A Structural Equations Approach for Modeling the Endogeneity of Lane-mean Speeds

Qinglong Lu

Technical University of Munich

Department of Civil, Geo and Environmental Engineering

Chair of Transportation Systems Engineering

Qinglong.lu@tum.de

Outline

1. Motivation
2. Lane-mean speeds modeling
3. Empirical Study
4. Discussion & Conclusion

Motivation

1. **Endogeneity problem in lane-mean speed modeling.** Solutions:
   - Three-stage least squares (3SLS).
   - Full-information maximum likelihood (FIML).

2. **Previous models (Shankar and Mannering) solved the endogeneity problem, but**
   - did not account for the influence of downstream speed,
   - the data are aggregated in a too long time window (1 hour),
   - only considered two-lane highway,
   - investigated a long road segment.

3. **Provide a better understanding of mean speeds across the lanes of a multi-lane highway.**
   - Car-following behavior
   - Lane-changing behavior

Lane-mean speeds modeling

1. **Sectionalizing**
   - Shorter segments can improve the modeling accuracy.
   - The influence of the downstream traffic will be considered.

2. **Simultaneously incorporate the influence of**
   - environmental (e.g. weather),
   - geometric (e.g. curved radius),
   - temporal (e.g. peak time),
   - and traffic flow factors (e.g. traffic volume).

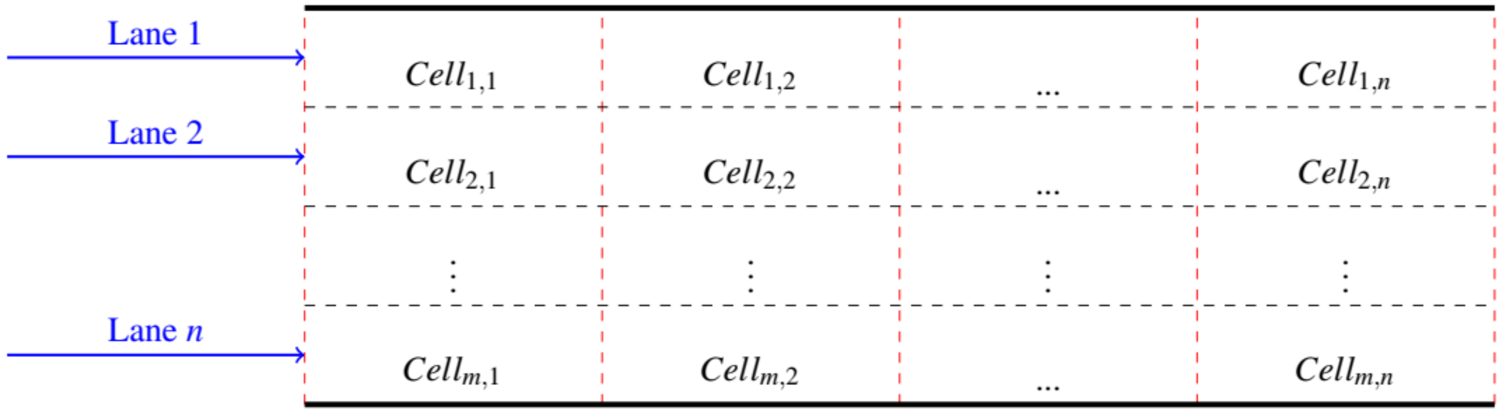


Figure 1: Cells structure of a road segment

Lane-mean speeds modeling

3. Structural equations system for lane-mean speeds (on segment *j*)

$$
\begin{cases}
u_{1,j} = \alpha_{1,j} + \boldsymbol{\beta}_{1,j} \cdot \boldsymbol{X}_{1,j} + \boldsymbol{\lambda}_{1,j} \cdot \boldsymbol{Z}_{1,j} + \boldsymbol{\theta}_{1,j} \cdot \boldsymbol{v}_{1,j} + \eta_{1,j} \cdot u_{1,j+1} + \varepsilon_{1,j} \\
u_{2,j} = \alpha_{2,j} + \boldsymbol{\beta}_{2,j} \cdot \boldsymbol{X}_{2,j} + \boldsymbol{\lambda}_{2,j} \cdot \boldsymbol{Z}_{2,j} + \boldsymbol{\theta}_{2,j} \cdot \boldsymbol{v}_{2,j} + \eta_{2,j} \cdot u_{2,j+1} + \varepsilon_{2,j} \\
\vdots \\
u_{i,j} = \alpha_{i,j} + \boldsymbol{\beta}_{i,j} \cdot \boldsymbol{X}_{i,j} + \boldsymbol{\lambda}_{i,j} \cdot \boldsymbol{Z}_{i,j} + \boldsymbol{\theta}_{i,j} \cdot \boldsymbol{v}_{i,j} + \eta_{i,j} \cdot u_{i,j+1} + \varepsilon_{i,j} \\
\vdots \\
u_{m,j} = \alpha_{m,j} + \boldsymbol{\beta}_{m,j} \cdot \boldsymbol{X}_{m,j} + \boldsymbol{\lambda}_{m,j} \cdot \boldsymbol{Z}_{m,j} + \boldsymbol{\theta}_{m,j} \cdot \boldsymbol{v}_{m,j} + \eta_{m,j} \cdot u_{m,j+1} + \varepsilon_{m,j}
\end{cases}
$$

Where i is the lane number, j is the segment number. $u_{i,j}$ is the lane-mean speed.

$\boldsymbol{X}_{i,j}$ is a vector of exogenous variables.

$\boldsymbol{Z}_{i,j}$ is a vector of endogenous variables.

$\boldsymbol{v}_{i,j}$ is a vector of mean speeds in adjacent lanes in the same segment.

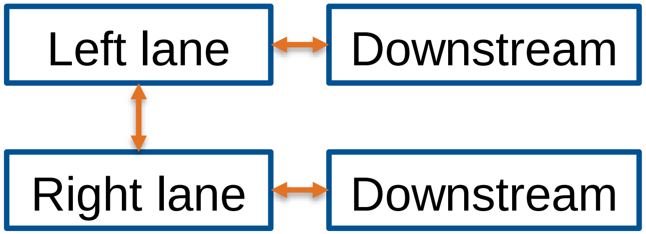
$u_{i,j+1}$ is a vector of downstream mean speeds in the same lane in the same segment.

$\alpha_{i,j}$, $\boldsymbol{\beta}_{i,j}$, $\boldsymbol{\lambda}_{i,j}$, $\boldsymbol{\theta}_{i,j}$, $\eta_{i,j}$ are vectors of estimable coefficients. $\varepsilon_{i,j}$ is a disturbance term.

Lane-mean speeds modeling

4. **Endogeneity problem ($Z_{i,j}$, $v_{i,j}$)**

e.g. For a two-lane highway, speed of the left lane and speed of the right lane has influence on each other, so when we model the mean speed of the left lane, the speed of the right lane will be a regressor, vice versa.



Lane-mean speeds modeling

- Solution (3SLS)

  Stage 1: Gets the two-stage least squares (2SLS) estimates of the model system;

  - Stage 1 of 2SLS:

$$\hat{\boldsymbol{Z}}_{i,j} = \hat{\boldsymbol{\beta}}_{i,j}^{(\boldsymbol{Z})} \boldsymbol{X}_{i,j} + \hat{\boldsymbol{\alpha}}_{i,j}^{(\boldsymbol{Z})}$$

$$\hat{\boldsymbol{v}}_{i,j} = \hat{\boldsymbol{\beta}}_{i,j}^{(\boldsymbol{v})} \boldsymbol{X}_{i,j} + \hat{\boldsymbol{\alpha}}_{i,j}^{(\boldsymbol{v})}$$

  - Stage 2 of 2SLS:

$$u_{i,j} = \alpha_{i,j} + \boldsymbol{\beta}_{i,j} \cdot \boldsymbol{X}_{i,j} + \boldsymbol{\lambda}_{i,j} \cdot \hat{\boldsymbol{Z}}_{i,j} + \boldsymbol{\theta}_{i,j} \cdot \hat{\boldsymbol{v}}_{i,j} + \eta_{i,j} \cdot u_{i,j+1} + \varepsilon_{i,j}$$

$$\hat{u}_{i,j}^{(2SLS)} = \hat{\alpha}_{i,j}^{(2SLS)} + \hat{\boldsymbol{\beta}}_{i,j}^{(2SLS)} \cdot \boldsymbol{X}_{i,j} + \hat{\boldsymbol{\lambda}}_{i,j}^{(2SLS)} \cdot \hat{\boldsymbol{Z}}_{i,j} + \hat{\boldsymbol{\theta}}_{i,j}^{(2SLS)} \cdot \hat{\boldsymbol{v}}_{i,j} + \hat{\eta}_{i,j}^{(2SLS)} \cdot u_{i,j+1}$$

  Stage 2:

$$\varepsilon_{i,j} = u_{i,j} - \hat{u}_{i,j}^{(2SLS)}$$

Lane-mean speeds modeling

- Solution (3SLS)

  Stage 3: Generalized least squares (GLS) is used to compute parameter estimates again.

$$E(\boldsymbol{\varepsilon}_{i,j}\boldsymbol{\varepsilon}_{i,j}^T) = \boldsymbol{\Omega}$$

$$\boldsymbol{W}_{i,j} = (1;\boldsymbol{X}_{i,j};\hat{\boldsymbol{Z}}_{i,j};\hat{\boldsymbol{v}}_{i,j};u_{i,j+1})$$

$$\boldsymbol{B}_{i,j} = (\alpha_{i,j};\boldsymbol{\beta}_{i,j};\boldsymbol{\lambda}_{i,j};\boldsymbol{\theta}_{i,j};\eta_{i,j})$$

$$\boxed{\hat{\boldsymbol{B}}_{i,j}^{(3SLS)} = (\boldsymbol{W}_{i,j}^T\boldsymbol{\Omega}^{-1}\boldsymbol{W}_{i,j})^{-1}\boldsymbol{W}_{i,j}^T\boldsymbol{\Omega}^{-1}\boldsymbol{u}_{i,j}}$$

$$\hat{\boldsymbol{u}}_{i,j}^{(3SLS)} = \hat{\boldsymbol{B}}_{i,j}^{(3SLS)T}\boldsymbol{W}_{i,j}$$

Empirical Study

1. Scenarios

- Compare the presented model with the approach proposed in Shankar and Mannering (to value the influence of the downstream speeds).

- Compare model performance on different highway segments with different number of lanes but have similar lengths, e.g. 0.2 miles two-lane segment, 0.2 miles three-lane segment, 0.2 miles four-lane segment.

- Compare model performance on different highway segments with different lengths but have the same number of lanes, e.g. 0.1 miles two-lane segment, 0.2 miles two-lane segment, 0.3 miles two-lane segment.

Empirical Study

2. Data description

- All data used in this study come from Caltrans Performance Measurement System (PeMS). Figure 2 depicts the simplified drawing of the segments.
- Granularity in 5 minutes.

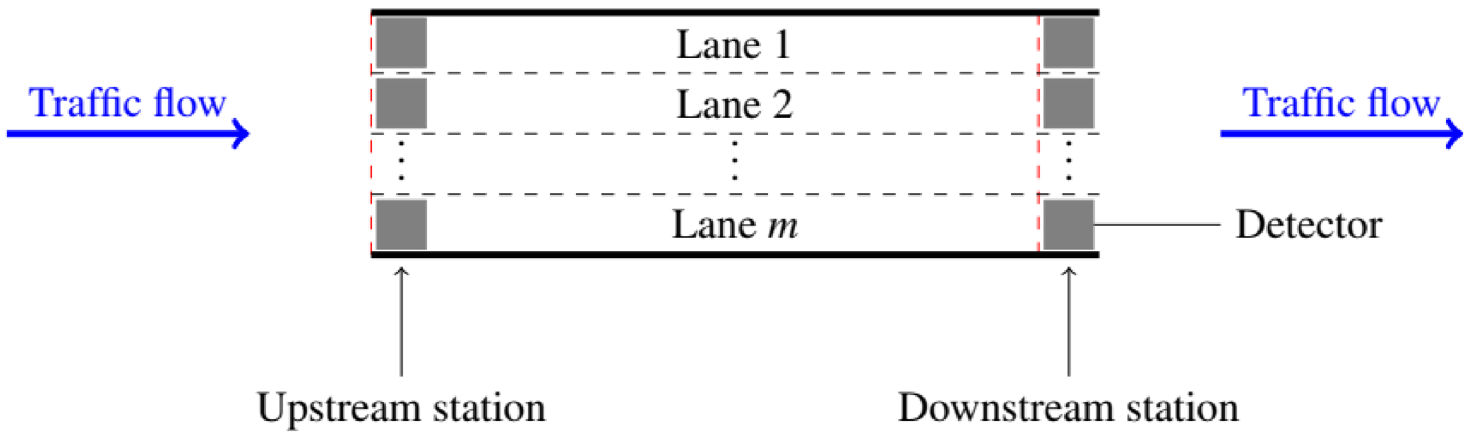


Figure 2: A segment and its attached information

Table 2: Variables considered for the tests

Factor	Variables	Remark
Traffic data	Mean speed of lane	
	Downstream speed	The mean speed collected by next station
	Flow of the lane	
	Low flow indicator	Equal to 1 if flow of lane $< 75veh/h$
	Ratio of flow	Ratio of flows in current lane to crucial adjacent lane
	Truck percentage	
	Truck percentage indicator 1	Equal to 1 if truck percentage $> 60\%$ and flow $< 50veh/h$
	Truck percentage indicator 2	Equal to 1 if truck percentage $\leq 60\%$ and flow $> 200veh/h$
	Truck speed	Equal to $truck\ VMT/truck\ VHT$ [a]
	High truck flow	Equal to 1 if truck flow $> 100veh/h$
Time of year (dummy)	Spring	
	Summer	
	Autumn	
	Winter (reference)	
Time of week (dummy)	Monday	
	Tuesday	
	Wednesday	
	Thursday	
	Friday	
	Saturday	
	Sunday (reference)	
Time of day (dummy)	Early morning	From 0:00 to 6:00
	AM peak	From 7:00 to 8:00
	PM peak	From 17:00 to 19:00
	Nighttime	From 19:00 to 24:00
	Other time (reference)	

[a] truck VMT is truck Vehicle Miles Traveled and truck VHT is truck Vehicle Hours Traveled, both of which are available in PeMS.

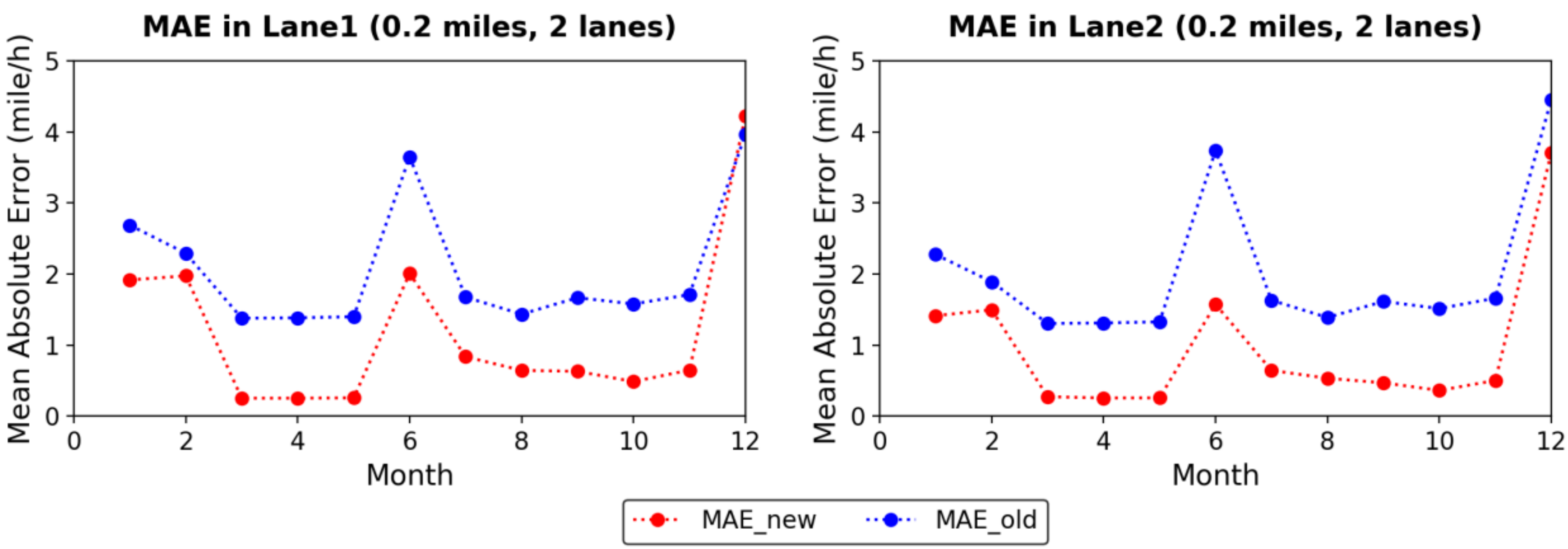
Empirical Study

3. Results

- The influence of the downstream speeds

  MAE_new: the improved model

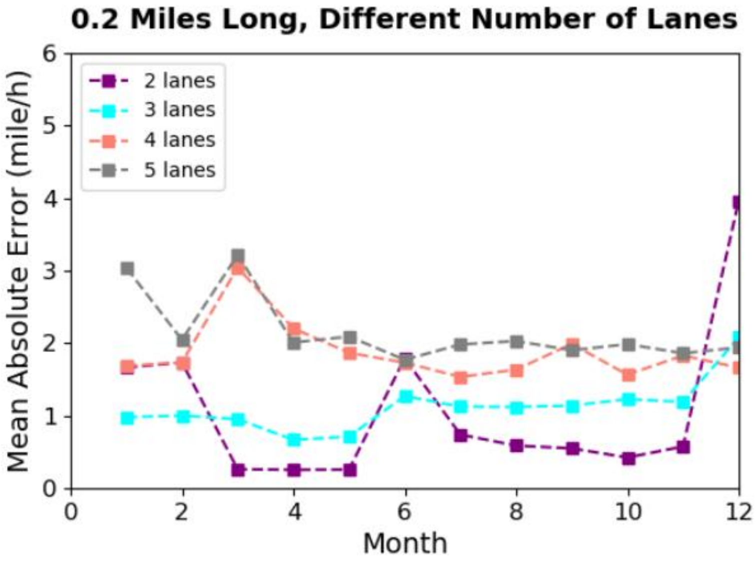
  MAE_old: Shankar and Mannering's model

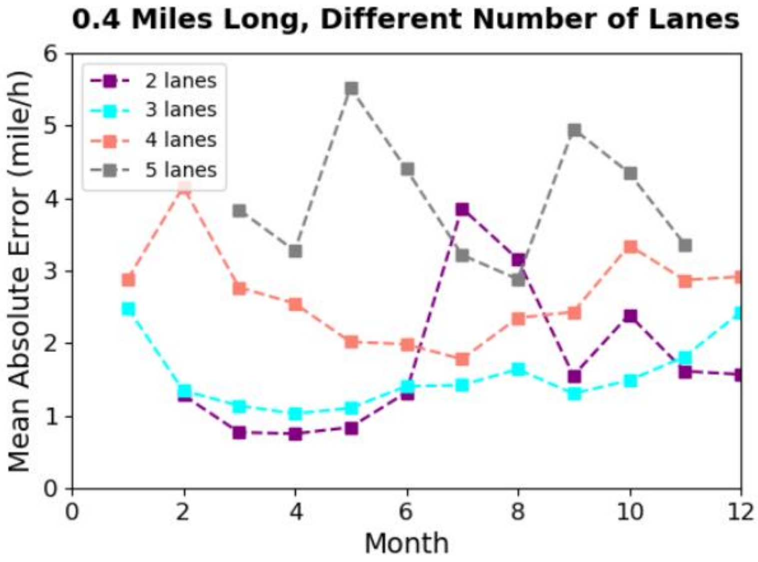


Empirical Study

3. Results

- Segments with different number of lanes

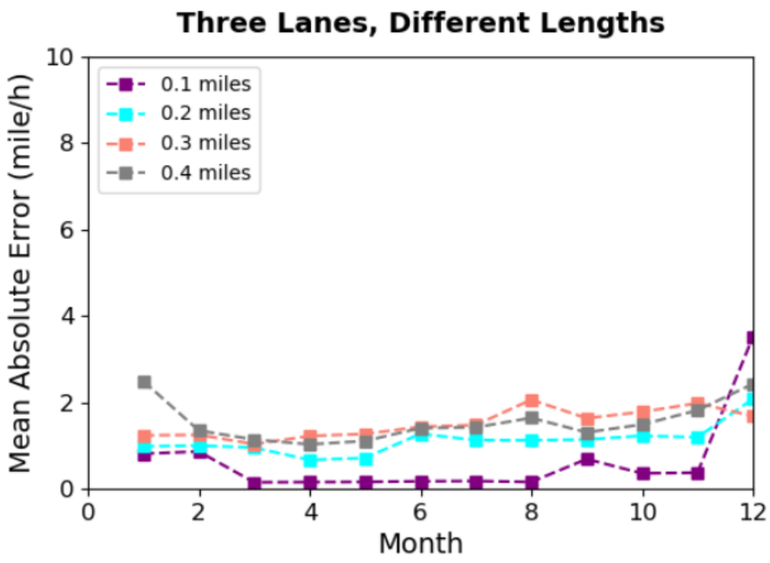


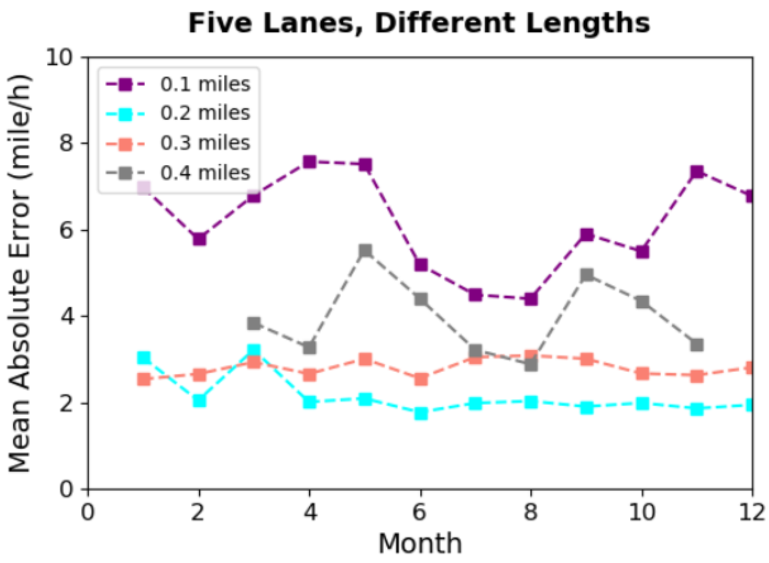


Empirical Study

3. Results

- Segments with different lengths





Discussion & Conclusion

1. Lane-mean speeds are affected by the adjacent-lane speeds.
2. The introduction of downstream speeds can improve the accuracy of lane-mean speed estimation.
3. This model can be applied in a small granularity (5 minutes here).
4. Model is more reliable in 3-lane segments. Model is more reliable in 0.2-mile segments.

Works for the future

- The effects of on-ramp stream and off-ramp stream on the speed of vehicles running on the main road are not considered.
- Research the relationship between lane-mean speeds and car following behavior as well as lane-changing behavior.
- Research traffic safety and accidents based on lane-mean speeds.

Thank you
for your attention!