Pergamon

Transpn Res.-A, Vol. 32, No. 5, pp. 311–322, 1998

Printed in Great Britain
0965-8564/98 $19.00+0.00

PII: S0965-8564(98)00003-2

MODELING THE ENDOGENEITY OF LANE-MEAN SPEEDS AND LANE-SPEED DEVIATIONS: A STRUCTURAL EQUATIONS APPROACH

VENKATARAMAN SHANKAR
Washington State Department of Transportation, Room 2B, Transportation Building, Box 47329, Olympia, WA 98504, U.S.A.

and

FRED MANNERING*
Department of Civil and Environmental Engineering, 121 More Hall, Box 352700, University of Washington, Seattle, WA 98195, U.S.A.

(Received 22 April 1997; in revised form 2 December 1997)

Abstract—This paper attempts to macroscopically address endogeneity issues related to lane-mean traffic speeds and lane-speed deviations. Methodologically, we seek to provide a better understanding of mean speeds and speed deviations across the lanes of a multilane highway. In so doing, the work may eventually be applied to better understand highway safety and the effects that lane-mean and lane speed deviations have on highway safety. We propose a structural model that relates mean speed and speed deviations by lane and is contemporaneously influenced by environmental, temporal, and traffic flow factors. Spot speed and vehicle classification data measured by lane in both the eastbound and westbound directions of Interstate 90 (I-90) in Washington State are used to develop the empirical relationships. The findings show that lane-mean speeds are endogenously related with adjacent lane speeds and exogenously related with associated environmental, traffic flow and temporal factors, while lane-speed deviations are endogenously related not only with adjacent lane speed deviations but also, through forward causality, lane-mean speeds and exogenously related with environmental, traffic flow and temporal factors as well. The approach shows significant promise in unraveling cause–effect relationships affecting macroscopic traffic flow continuums. © 1998 Elsevier Science Ltd. All rights reserved

Keywords: structural equations, three-stage least squares, speed deviation

1. INTRODUCTION

Prior speed–flow relationship studies have focused on single-regime 体制 or multi-regime functional relationships that were generally univariate 单元 or bivariate 二元 in nature. Linkages between speed and flow were generally studied over different traffic density ranges. Engineering intuition suggests that such approaches offer only a limited understanding of the underlying processes governing speed–flow relationships. Particularly in the context of intelligent transportation systems (ITS) where the use of technological components will likely result in fundamental shifts of assumed speed–flow relationships. In the presence of ITS, it is important that the causality underlying the processes affecting traffic speed–flow relationships and consequently safety be uncovered, because systemic affects associated with such technologies are potentially wide-ranging and often simultaneous.

Prior theories and empirical validations have established speed–flow relationships that are unidirectional and regime-based (see for example, Greenshields, 1935; Edie, 1961; May and Keller, 1968). Suggestions on structural modeling (i.e. a simultaneous equations approach), with its potential to provide an improved understanding of the interrelationships among the contemporaneous influences of lane-mean speeds, lane-speed deviations, environmental conditions, geometric elements, vehicle types, and temporal and seasonal factors, have been conceptual for the most part. Instead, significant effort has been focused on the use of independent ordinary

*Author for correspondence. Fax: 001 206 543 1543; e-mail:flm@u.washington.edu

or non-linear least squares estimation (see for example Easa and May, 1980). Use of independent regression equations that separately estimate speed and flow-related parameters without accounting for the contemporaneous correlation of the disturbances will cause the respective estimated parameters to be biased and inconsistent (Greene, 1993). Apart from the specification aspects mentioned above relating to the causal modeling of traffic speed and flow, little evidence is available on modeling frameworks that simultaneously incorporate the influence of environmental, geometric, temporal and traffic flow factors. Some efforts in this area have focused on the impact of weather (Ibrahim and Hall, 1994) and geometrics (Iwasaki, 1991), while others have focused on the temporal variations in traffic flow (see for example Brilon and Ponzlet, 1996).

The attempt of this research is to combine the need for a complete model that is comprehensive in factors identified in previous research with the need for an estimation framework that is structural in nature. It should be noted here that the focus of the research is on the structural relationship between lane-mean speeds (i.e. time-mean speeds) and related lane-speed deviations and the traffic characteristics, environmental conditions, and temporal and seasonal factors. As such, the investigation will focus on the contemporaneous inter-relationships at a given location in a given time period.

The paper begins by providing an overview of the modeling approach and estimation technique. This is followed by a description of the data-collection site and the presentation of model-estimation results. Finally, conclusions and recommendations are provided.

2. MODELING APPROACH

Our intent is to develop a model of mean speeds and speed deviations (measured over some time interval) for each lane of a multilane roadway at a macroscopic level.* Turning first to lane-mean speeds, from a structural point of view, it is important to note that the mean speed in each lane will not only be a function of traffic characteristics in the lane, but also a function of the mean speeds in the adjacent lanes. This suggests an equation system in which lane-mean speeds are determined simultaneously across the roadway′s lanes. In a similar fashion, speed deviations in each lane will be dependent on speed deviations in adjacent lanes. Lane-speed deviations will also be a function of the lane′s mean speed and the mean speeds in adjacent lanes. Because of this inter-relationship, lane-speed deviations must also be determined in a simultaneous equation system with mean speeds entering the equation system in a recursive fashion.

The structural equation system for lane-mean speeds and lane-speed deviations can be written as follows: For lane-mean speeds, over some time interval, the equation system is,

$$\begin{aligned} u_1 &= \alpha_1 + \beta_1 X_1 + \lambda_1 Z_1 + \theta_1 \bar{u}_1 + \varepsilon_1 \\ u_2 &= \alpha_2 + \beta_2 X_2 + \lambda_2 Z_2 + \theta_2 \bar{u}_2 + \varepsilon_2 \\ \vdots \quad & \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ u_n &= \alpha_n + \beta_n X_n + \lambda_n Z_n + \theta_n \bar{u}_n + \varepsilon_n \end{aligned} \tag{1}$$

where u_{n} is the mean speed in lane n, X_{n} is a vector of exogenous variables influencing the mean speed in lane n, Z_{n} is a vector of endogenous variables influencing the mean speed in lane n (i.e. traffic flow characteristics that may be influenced by lane-mean speeds such as proportion of total roadway traffic in the lane), $\overline{u_n}$ is a vector of mean speeds in lanes adjacent to lane n, α_n, β_n, λ_n, and θ_n are vectors of estimable coefficients, ε_n is a disturbance term. Similarly, lane-speed deviations, over some time interval, can be written as†

*Modeling at the macroscopic level limits our ability to investigate vehicle-level interactions in a direct manner. From a modeling perspective, this avoids problems with individual vehicle interactions (i.e. vehicles ahead and adjacent to the individual vehicle speed being modeled, if a microscopic approach had been use). However, as will be shown later in the paper, suggestive results on vehicle-level interactions can be obtained by appropriate inclusion of geometric-, seasonal-traffic-related factors.

†Note that our equations model lane-speed deviations as dependent variables, which are functions of lane-mean speeds. The reverse relationships between lane-mean speeds and lane-speed deviations is not specified. This is because there is not basis for assuming lane-mean speeds are influenced by speed deviations. This was borne out during some preliminary estimation runs that found lane-speed deviations to be statistically insignificant when in eqn (1).

$$\begin{aligned}
\sigma_1 &= \rho_1 + \eta_1 V_1 + \tau_1 Y_1 + \gamma_1 \overline{u_1} + \omega_1 \overline{\sigma_1} + \nu_1 \\
\sigma_2 &= \rho_2 + \eta_2 V_2 + \tau_2 Y_2 + \gamma_2 \overline{u_2} + \omega_2 \overline{\sigma_2} + \nu_2 \\
\vdots \quad & \quad \vdots \quad \vdots \;\; \vdots \quad \vdots \;\; \vdots \quad \vdots \;\; \vdots \quad \vdots \;\; \vdots \quad \vdots \\
\sigma_n &= \rho_n + \eta_n V_n + \tau_n Y_n + \gamma_n \overline{u_n} + \omega_n \overline{\sigma_n} + \nu_n
\end{aligned} \tag{2}$$

where σ_n is the standard deviation of speed in lane n, V_n is a vector of exogenous variables influencing the standard deviation of speed in lane n, Y_n is a vector of endogenous variables influencing the standard deviation of speed in lane n (i.e. traffic flow characteristics that may be influenced by lane-speed deviations such as proportion of total roadway traffic in the lane), $\overline{u_n}$ is a vector of mean speeds in lane n and in other lanes, $\overline{\sigma_n}$ is a vector of the standard deviation of speeds in lanes adjacent to lane n, ρ_n, η_n, τ_n, γ_n, and ω_n are vectors of estimable coefficients, ν_n is a disturbance term.

To estimate eqns (1) and (2), three-stage least squares (3SLS) is appropriate. This approach allows for simultaneous estimation of coefficients using information from the equation system. By so doing, it ensures that coefficient estimates are generally more efficient (asymptotically) than alternative simultaneous-equation estimation approaches such as the indirect least-squares (ILS), two-stage least squares (2SLS), and limited-information maximum likelihood (LIML).* An alternative estimation approach is full-information maximum likelihood (FIML), but because the asymptotic variance–covariance matrices of FIML and 3SLS can be shown to be equal, the choice of 3SLS is acceptable. The 3SLS estimation procedure is conducted by first getting two-stage least squares (2SLS) estimates of the equation system which are calculated using instruments (endogenous variables regressed against all exogenous variables). The 2SLS estimates are then used to estimate the equation system′s disturbances which are subsequently used to estimate the contemporaneous variance–covariance matrix of disturbances. Finally, generalized least-squares (GLS) is applied to estimate model coefficients using the estimated contemporaneous variance–covariance matrix of disturbances as a basis. See Greene (1993) for a complete description of the procedure.

3. EMPIRICAL SETTING

The study area is a rural location on Interstate 90 (I-90) located some 50 kilometers east of Seattle. This location on I-90 is in the Cascade Mountains with an elevation 975 m above sea level. The climate is harsh with an average of 215 cm of rainfall and 1140 cm of snowfall annually. In general, this portion of I-90 has significant variations in speeds (i.e. high lane-speed deviations), due to the combined impact of vehicle mix, inclement 恶劣 weather, seasonal effects (e.g. variations in traffic volume, precipitation 沉淀, and ambient 周围 temperatures), and challenging roadway geometrics. These speed variations significantly contribute to the likelihood and severity of accidents on this portion of I-90 (see Shankar *et al.*, 1995, 1996).

To model lane-mean speeds and lane-speed deviations at this location, data were collected using magnetic loop detectors. Interstate 90, at this location, is a three-lane divided freeway in each direction with the eastbound alignment on a 1.5% upgrade and the westbound alignment on a 2.5% downgrade. Eastbound and westbound traffic data were collected by lane. Data on spot speeds by lane, vehicle classification by lane, were gathered in the fall of 1994 and the winter, spring and summer months of 1995. Speed data were collected in speed bins of 10 m.p.h, aggregated over one hour.† Classification of vehicle types was based on four wheelbase classes of up to

*The 3SLS procedure is more efficient than single-equation methods such as ILS, 2SLS, and LIML, when the variance–covariance matrix is not diagonal. This will be the case when there is contemporaneous correlation among disturbance (i.e. the unobserved factors affecting mean speed in one lane are correlated with those unobserved factors that affect mean speed in other lanes). If these unobserved factors are not correlated (i.e. the case of a diagonal variance–covariance matrix), it can be readily shown that 3SLS reduces to 2SLS.

†Aggregation of speed data over one hour is likely to mask some underlying variation in the speed distribution; however, the level that is afforded at micro-speed data such as 5 or 20 s data is not likely to significantly alter 改变 the structure of the cause–effect relationship between speed and speed deviation. Any additional insight into the cause–effect relationship could stem from the stochasticity of peak hour flows. As will be demonstrated later, the stochasticity of peak hour flows and its impact on speed–speed deviation relationships will be captured adequately by indicator variables acting as surrogates 代理人 for peak hour phenomena thus eliminating potential omitted variable biases. The authors do acknowledge that micro-speed data does provide insight into merge and weave phenomena and shock wave-related incremental impacts on traffic flow continuums, but point out that the use of such data is different, namely to investigate ‘resulting conditions’ stemming from inconsistencies in traffic flow.

Table 1. Summary of lane-mean speeds and lane-speed deviations

		Hourly grouped speeds					
		Grouped lane-mean speed (m.p.h.)			Grouped lane-speed deviation (m.p.h.)		
Direction	Location	Mean	Minimum	Maximum	Mean	Minimum	Maximum
Eastbound	Right lane	70.193	31.250	76.760	7.164	4.440	16.150
	Middle lane	75.612	32.580	79.820	5.548	3.780	13.860
	Left lane	78.012	34.880	90.000	4.858	0.000	21.680
Westbound	Right lane	72.986	40.470	79.430	7.000	4.580	15.640
	Middle lane	76.441	43.570	81.940	5.756	3.000	14.210
	Left lane	78.830	40.000	86.670	5.310	0.000	28.720

26, 26 to 39, 39 to 65, and 65 to 114 ft. Lane-by-lane data were collected for spot speeds and vehicle classifications in both eastbound and westbound directions. Table 1 shows computed lane-mean speeds and lane-speed deviations by lane using one hour time periods.

Tables 2 and 3 show the results of the 3SLS estimation of grouped lane-mean speeds at the study location. Tables 4 and 5 show the results of the 3SLS estimation of grouped lane-speed deviations. For estimation purposes, the logarithm of the lane-mean speed was used as the dependent variable in the lane-mean speed model system. As seen in the tables, exogenous variables significantly determining lane-mean speed and lane-speed deviation include time-of-day, time-of-week, and seasonal, indicators. Vehicle mix and the distribution of traffic across the lanes were also found to be significant determinants of lane-mean speed and lane-speed deviation.* All estimated coefficients were found to be of plausible sign. For the eastbound direction, the system R^2 for the lane-mean speed model was 0.8629 and 0.3288 for the lane-speed deviation model. For the westbound direction, system R^2 was 0.9232 and 0.3087 for the lane-mean speed and lane-speed deviation models, respectively.† The interpretation of the estimation results is provided below.

3.1. 3SLS estimation of lane-by-lane grouped mean speeds‡

A brief summary of the important findings of the estimation results for lane-by-lane grouped mean speed models (Table 2 for eastbound and Table 3 for westbound) are presented below.

Variable: Lane traffic-flow indicator (flows less than 75 veh h^{-1})
Finding: Positively affects right-lane-mean speeds in the eastbound direction

This finding is intuitive in that it illustrates driver tendency to drive the allowable safe speed under near free-flow conditions. Under near free-flow conditions, the visual constraints posed by the presence of adjacent vehicles are removed thereby allowing lane-mean speeds to increase significantly beyond normal operating speeds (around the speed limit.) The effect appears to be significant in the eastbound direction only and it is likely that the downgrade effect for the westbound direction annuls the significance of low volumes on lane-mean speeds in the right lane.

Variable: Truck percentage in right lane§
Finding: Negatively affects lane-mean speeds in both eastbound and westbound directions

This finding reflects the impact of truck percentage on speed–flow distributions. Under general conditions, with no constraints on flow levels and accounting for the effect of all other factors, increasing truck percentage will tend to decrease lane-mean speeds.

*For estimation purposes these variable were instrumented (see description in Tables 2–5) because of possible endogeneity. This is because changes in lane-mean speeds and/or lane-speed deviations can affect the distribution of traffic flow over the lanes. Thus changing values in the dependent variable could change values in the independent variable, which is violation of least-squares assumptions. Not correcting for this will results in biased and inconsistent coefficient estimates.

†Cross-validation of the models was conducted on post-data samples collected for different time periods. The findings indicated no changes in the R^2 of the equations. Also, the magnitudes of individual coefficient estimates did not change significantly.

‡The lanes are defined as right, middle, and left relative to the direction of travel.

§Trucks are defined as vehicles with wheelbases exceeding 65 ft.

Table 2. Three-stage least squares estimation of grouped lane-mean speeds for eastbound I-90

Variable[a]	Estimated coefficient	*t*-statistic
Equation 1: logarithm of right-lane mean speed (dependent variable)		
Constant	−0.1106	−2.6503
Lane traffic flow indicator (1 if traffic flow in right lane is less than 75 veh h^{-1}, 0 otherwise)	0.0021	2.7372
Truck percentage in right lane	−0.0292	−15.2267
High truck flow in right lane (1 if hourly truck flow is greater than 100 veh h^{-1}, 0 otherwise)	0.0030	6.8196
Relative truck flow indicator 1 (1 if truck percentage in right lane exceeds 60% and total traffic flow in right lane is less than 50 veh h^{-1}, 0 otherwise)	0.0047	5.8343
Relative truck flow indicator 2 (1 if truck percentage in right lane is less than or equal to 20% and total traffic flow in right lane exceeds 200 veh h^{-1}, 0 otherwise)	0.0034	5.3791
Logarithm of middle-lane mean speed	1.0107	104.7271
Time-of-day indicator 1 (1 if hour of observation is between midnight and 6:00 a.m., 0 otherwise)	−0.0030	−3.3621
Seasonal indicator 1 (1 if it is winter, 0 otherwise)	−0.0021	−3.5543
Seasonal indicator 2 (1 if it is spring, 0 otherwise)	−0.0010	−2.7644
Time-of-week indicator (1 if it is weekend, 0 otherwise)	0.0104	8.7886
Time-of-day indicator 2 (1 if it is p.m. peak hour, 0 otherwise)	0.0018	3.7352
Time-of-day indicator 3 (1 if it is a.m. peak hour, 0 otherwise)	−0.0014	−2.7911
Number of observations	2233	
R^2	0.9072	
Corrected R^2	0.9067	
Equation 2: logarithm of middle-lane mean speed (dependent variable)		
Constant	0.3628	12.0474
Logarithm of right-lane mean speed	0.4257	59.2642
Logarithm of left-lane mean speed	0.4960	80.9855
Hourly traffic flow in middle lane	−0.000014	−10.2548
Lane use distribution between middle lane and right lane (ratio of flows in middle lane to right lane)	−0.0010	−4.2564
Time-of-day indicator 4 (1 if it is night-time, 0 otherwise)	−0.0030	−3.5560
Time-of-week indicator (1 if it is weekend, 0 otherwise)	−0.0072	−13.2349
Number of observations	2233	
R^2	0.9022	
Corrected R^2	0.9019	
Equation 3: logarithm of left-lane mean speed (dependent variable)		
Constant	−0.6949	−12.0579
Truck percentage in left lane	0.0057	2.1422
Lane distribution between left lane and middle lane (ratio of flows in ~~middle~~ right lane to ~~right~~ middle lane)	0.0050	3.1882
Logarithm of middle-lane mean speed	1.1671	87.7616
Time-of-day indicator 4 (1 if it is night-time, 0 otherwise)	0.0050	3.0050
Number of observations	2233	
R^2	0.7961	
Corrected R^2	0.7958	
System R^2	0.8629	

[a]Variables in italics are instrumented because of possible endogeneity. This is done by regressing the variable against exogenous variables and using the regression-predicted values for the 3SLS estimation. Variables in bold are endogenous and part of the simultaneous equation estimation. Finally, trucks are defined as vehicles with wheelbases exceeding 65 ft.

However, as will be illustrated in the following discussions, certain truck percentage–flow combinations will create desirable conditions for traffic flow.

Variable: High truck flow in right lane
Finding: Increases right-lane mean speeds in eastbound and westbound directions

This finding suggests that when truck flow in the right lane exceeds a threshold of flow, lane-mean speeds will increase as a result of a combination of factors. Truck drivers driving in high truck volumes tend to 'draft' taking advantage of the relatively greater uniformity of vehicle type
均匀性

Table 3. Three-stage least squares estimation of grouped lane-mean speeds for westbound I-90

Variable[a]	Estimated coefficient	t-statistic
Equation 1: logarithm of right-lane mean speed (dependent variable)		
Constant	−0.4308	−14.3947
Truck percentage in right lane	−0.0144	−9.2982
High truck flow in right lane (1 if hourly truck flow is greater than 100 veh h^{-1}, 0 otherwise)	0.0017	2.7649
Logarithm of middle-lane speed	1.0895	157.8630
Seasonal indicator 1 (1 if it is winter, 0 otherwise)	0.0012	2.5684
Time-of-week indicator 1 (1 if it is weekend, 0 otherwise)	0.0046	4.8303
Time-of-day indicator 3 (1 if it is a.m. peak hour, 0 otherwise)	−0.0013	−2.5191
Number of observations	2230	
R^2	0.9472	
Corrected R^2	0.9470	
Equation 2: logarithm of middle-lane mean speed (dependent variable)		
Constant	0.1919	7.7056
Logarithm of right-lane mean speed	0.4539	62.1349
Logarithm of left-lane mean speed	0.5047	66.5657
Hourly traffic flow in middle lane	−0.000015	−9.5357
Lane use distribution between middle lane and right lane (ratio of flows in middle lane to right lane)	−0.0012	−4.4492
Time-of-day indicator 4 (1 if it is night-time, 0 otherwise)	−0.0036	−5.1740
Time-of-week indicator (1 if it is weekend, 0 otherwise)	−0.0030	−6.6528
Number of observations	2230	
R^2	0.9454	
Corrected R^2	0.9452	
Equation 3: logarithm of left-lane mean speed (dependent variable)		
Constant	−0.1134	−2.6376
Hourly traffic flow in left lane	0.000035	6.3908
Lane distribution between left lane and middle lane (ratio of flows in middle lane to right lane)	0.0040	2.6966
Logarithm of middle-lane mean speed	1.0321	104.1540
Time-of-day indicator 4 (1 if it is night-time, 0 otherwise)	0.0067	4.6959
Number of observations	2230	
R^2	0.8797	
Corrected R^2	0.8795	
System R^2	0.9232	

[a]Variables in italics are instrumented because of possible endogeneity. This is done by regressing the variable against exogenous variables and using the regression-predicted values for the 3SLS estimation. Variables in bold are endogenous and part of the simultaneous equation estimation. Finally, trucks are defined as vehicles with wheelbases exceeding 65 ft.

in the lane. This finding is consistent with the truck equivalency factors presented in the U.S. Highway Capacity Manual (Transportation Research Board, 1994).

Variable: Right-lane relative truck flow indicators (truck percentage exceeding 60% and total traffic flow less than 50 veh h^{-1} or truck percentage less than or equal to 20% and total lane flow exceeding 200 veh h^{-1})
Finding: Increases lane-mean speeds in the eastbound direction

This finding is illustrative of the significance of the impact of vehicle mix on traffic flow distribution. Under low or near free-flow conditions but with a high percentage of trucks, or under higher volume conditions but with a relatively low percentage of trucks, lane-mean speeds are found to increase because of the uniformity of vehicle type.

Variable: Adjacent lane-mean speeds
Finding: Increasing adjacent-lane speeds increase in-lane mean speeds in eastbound and westbound directions

These variables capture the endogenous lateral cause–effect relationships between adjacent lane speeds.* As will be evidenced in subsequent discussions, adjacent lanes tend to positively affect traffic speeds. The underlying process this factor captures is the need to drive faster to merge into adjacent lanes and also the psychological impact faster traffic in the adjacent lane has on drivers.

Variable: Time-of-day indicator (midnight to 6:00 a.m.)
Finding: Negatively impacts right-lane mean speeds in eastbound and westbound directions

This finding represents selection effects of drivers choosing the right lane for travel in the morning. Drivers who tend to use the right lane under free-flow conditions, as expected in the midnight to early morning hours, usually consist of slower passenger-car drivers or truck drivers. This portion of the population tends to have lower travel speeds.

Variable: Seasonal indicators (winter, spring)
Finding: Tend to decrease right-lane mean speeds in the eastbound direction and increase right-lane mean speeds in the westbound direction

These variables capture the effect of weather on right-lane operations. Particularly in this area of I-90 where snow and associated inclement conditions occur in winter and early spring, right lanes tend to operate at lower speeds due to vehicle chaining requirements and the deterrence of adverse driving conditions in the eastbound direction. The westbound direction seems to experience anomalous effects, however, but this is likely an artifact of the data, especially the positive effect of winter coupled with no significant effect for spring.

Variable: Time-of-week indicator (weekend)
Finding: Tends to increase right-lane mean speeds in both directions and decrease middle-lane mean speeds in eastbound and westbound directions

These variables represent the near free-flow conditions that exist on weekends, in addition to capturing the effect of uniformity of traffic mixes. Truck traffic in weekend periods is minimal and as evidenced before, with greater vehicle type uniformity, right-lane mean speeds are expected to increase. As opposed to a positive impact on right-lane speeds, weekend effects tend to decrease middle-lane speeds indicating that speeds will not be as high as expected in weekends.

Variable: Time-of-day indicators (p.m. and a.m. peak hours)
Finding: Right-lane speeds increase during the p.m. peak hour in the eastbound direction and decrease during the a.m. peak hour in eastbound and westbound directions

The peak hour variables capture the effect of several factors such as commute direction and vehicle mix uniformity. The lack of a significant p.m. peak hour effect in the westbound direction is likely an artifact of the data, and in generic situations likely will play a significant role in both directions.

Variable: Time-of-day indicator (night-time)
Finding: Tends to decrease middle-lane speeds and increase left-lane speeds in eastbound and westbound directions

This finding provides interesting insight into drivers′ perception of lane usage by time-of-day. Under night-time conditions, the use of the middle lane as a passing lane declines in favor of the left lane for drivers who tend to drive significantly faster than the average driver. Thus the night-

*Note that only the immediately adjacent lane has a statistically significantly impact on lane speeds (i.e. the left-lane speeds were not found to affect right-lane speeds).

time factor captures aggressive driving behavior and the locational occurrence of such behavior in a cross-sectional context.

Variable: Hourly traffic flow in middle lane
Finding: Tends to negatively impact middle-lane speeds in eastbound and westbound directions

This finding is consistent with flow-speed relationships observed in other empirical studies. Given that truck-related factors were not found to significantly affect middle-lane speeds, this finding indicates that as flow in the middle lane (as opposed to the right lane) increases it represents the gradual approach to congestion, and the consequent decrease in speeds.

Variable: Lane use distributions (ratio of in-lane to adjacent-lane flows)
Finding: Increase in ratio decreases middle-lane speeds in both directions while increasing left-lane speeds in eastbound and westbound directions

This finding illustrates the effect of congestion and the declining choice of the middle lane as a passing lane as a result of increasing congestion. As congestion levels are approached, the use of the middle lane changes from a passing lane to a capacity lane. Consequently, driver behavior appropriately reflects a tendency to slow down under increasing flows. However, as traffic flows in the middle and right lanes approach thresholds where lane speeds have to decrease to maintain safe operations, the use of the left lane as a passing lane increases thereby attracting faster drivers.

Variable: Truck percentage in left lane
Finding: Increasing truck percentage increases left-lane speeds in eastbound direction

This finding illustrates the primary effect of a passing lane on cross-sectional flow–speed relationships when from a capacity standpoint. A higher truck percentage in the left lane reflects truck drivers′ tendencies to pass slower traffic in order to accelerate up the steeper grade that is immediately upstream of the eastbound direction.

3.2. 3SLS estimation of lane-by-lane grouped speed deviations

A brief summary of the important findings of the estimation results for lane-by-lane grouped mean speed deviation models (Table 4 for eastbound and Table 5 for westbound) are presented below.

Variable: Speed deviations in adjacent lanes
Finding: In general adjacent lane-speed deviations positively affect in-lane speed deviations in eastbound and westbound directions

The adjacent lane-speed deviation variables captures the lateral lane effects across the roadway. However the impact includes the car-following response effect (not expected in adjacent lane speed effects) due to lane changes that adjacent-lane deviations bring. Greater deviations in the adjacent lane indicate to drivers more opportunities, although intermittent, for lane changing than a lower deviation would. Hence, the car-following driver response in the right lane is simultaneously being influenced by the opportunity for lane changing which causes the sub-conscious effect of higher fluctuation in in-lane speeds.

Variable: Lane-mean speeds*
Finding: In general in-lane mean speeds positively affect in-lane speed deviations while adjacent lane-mean speeds negatively affect in-lane speed deviations in eastbound and westbound directions

*Lane-mean speeds are instrumented variables in the speed deviation system. Predicted values from the lane-mean speed system were used in this 3SLS estimation.

Table 4. Three-stage least squares estimation of grouped lane-speed deviations for eastbound I-90

Variable[a]	Estimated coefficient	t-statistic
Equation 1: right-lane speed deviation (dependent variable)		
Constant	34.5707	7.6989
Speed deviation in middle lane	0.1996	2.5356
Logarithm of right-lane mean speed[b]	3.6272	2.6788
Logarithm of middle-lane mean speed[b]	−10.1006	−10.0808
Time-of-day indicator 1 (1 if hour of observation is between midnight and 6:00 a.m., 0 otherwise)	0.2384	3.4834
Time-of-day indicator 2 (1 if it is p.m. peak hour, 0 otherwise)	−0.0917	−1.4435
Seasonal indicator 1 (1 if it is winter, 0 otherwise)	−0.2234	−4.3246
Time-of-week indicator (1 if it is weekend, 0 otherwise)	−0.2969	−5.8715
Truck-to-passenger car flow ratio	−0.1238	−7.8162
Number of observations	2233	
R^2	0.2959	
Corrected R^2	0.2934	
Equation 2: middle-lane speed deviation (dependent variable)		
Constant	34.6818	10.6756
Speed deviation in right lane	−0.0516	−1.3422
Speed deviation in left lane	0.3791	13.6739
Logarithm of right-lane mean speed[b]	−31.0753	−11.7211
Logarithm of middle-lane mean speed[b]	11.2859	8.9724
Logarithm of left-lane mean speed[b]	12.0551	5.0897
Time-of-week indicator (1 if it is weekend, 0 otherwise)	0.4592	8.0088
Seasonal indicator 1 (1 if it is winter, 0 otherwise)	−0.1081	−2.3134
Time-of-day indicator 1 (1 if hour of observation is between midnight and 6:00 a.m., 0 otherwise)	0.2430	6.1934
Time-of-day indicator 2 (1 if it is p.m. peak hour, 0 otherwise)	−0.1829	−3.5588
Number of observations	2233	
R^2	0.3598	
Corrected R^2	0.3572	
Equation 3: left-lane speed deviation (dependent variable)		
Constant	22.9298	3.5773
Speed deviation in middle lane	1.0753	10.3436
Logarithm of middle-lane mean speed[b]	−29.5472	−13.8764
Logarithm of left-lane mean speed[b]	24.0178	11.6909
Passenger car percentage	−1.0800	−2.4323
Seasonal indicator 1 (1 if it is winter, 0 otherwise)	0.6884	6.5079
Time-of-day indicator 2 (1 if it is p.m. peak hour, 0 otherwise)	0.4557	3.3125
Number of observations	2233	
R^2	0.3285	
Corrected R^2	0.3267	
System R^2	0.3288	

[a]Variables in italics are instrumented because of possible endogeneity. This is done by regressing the variable against exogenous variables and using the regression-predicted values for the 3SLS estimation. Variables in bold are endogenous and part of the simultaneous equation estimation. Finally, trucks are defined as vehicles with wheelbases exceeding 65 ft.
[b]Lane-mean speeds are instrumented variables in the speed deviation system. Predicted values from the lane-mean speed system were used in this 3SLS estimation.

This finding is intuitive and consistent with the relationships drawn in previous studies between the coefficient of dispersion and mean speeds (see for example May 1990). The negative impact of adjacent-lane speeds on in-lane deviations indicates that drivers tend reduce their deviations as adjacent lane speeds go up in order to make their lane changing operations safer.

Variable: Time-of-day indicators (early morning and p.m. peak)
Finding: In general early morning effects cause an increase in in-lane deviations in the eastbound direction while p.m. peak hour effects cause a decrease in in-lane speed deviations in the eastbound direction. In the westbound direction the p.m. peak hour effect is insignificant while the a.m. peak hour effects are reversed.

Table 5. Three-stage least squares estimation of grouped lane-speed deviations for westbound I-90

Variable[a]	Estimated coefficient	t-statistic
Equation 1: right-lane speed deviation (dependent variable)		
Constant	−0.5669	−0.3505
Speed deviation in middle lane	0.8431	37.1615
Logarithm of right-lane mean speed[b]	9.7616	9.7441
Logarithm of middle-lane mean speed[b]	−9.0162	−8.7505
Time-of-day indicator 1 (1 if hour of observation is between midnight and 6:00 a.m., 0 otherwise)	−0.1023	−2.0570
Time-of-week indicator (1 if it is weekend, 0 otherwise)	−0.1552	−3.8241
Number of observations	2230	
R^2	0.3965	
Corrected R^2	0.3951	
Equation 2: middle-lane speed deviation (dependent variable)		
Constant	2.1069	08975
Speed deviation in right lane	1.1797	31.0057
Speed deviation in left lane	−0.0332	−1.6305
Logarithm of right-ane mean speed[b]	−12.3373	−3.6331
Logarithm of middle-lane mean speed[b]	9.5426	7.1049
Logarithm of left-lane mean speed[b]	1.6071	0.5003
Time-of-week indicator 1 (1 if it is weekend, 0 otherwise)	0.1856	3.5658
Seasonal indicator 1 (1 if it is winter, 0 otherwise)	0.0380	1.5280
Time-of-day indicator 1 (1 if hour of observation is between midnight and 6:00 a.m., 0 otherwise)	0.1429	2.4480
Number of observations	2230	
R^2	0.3460	
Corrected R^2	0.3436	
Equation 3: left-lane speed deviation (dependent variable)		
Constant	29.3353	5.8855
Speed deviation in middle lane	0.3794	5.2335
Logarithm of middle-lane mean speed[b]	−33.2354	−15.0579
Logarithm of left-lane mean speed[b]	26.9555	11.1272
Seasonal indicator 1 (1 if it is winter, 0 otherwise)	0.4621	4.2952
Number of observations	2230	
R^2	0.2682	
Corrected R^2	0.2669	
System R^2	0.3087*	

[a]Variables in italics are instrumented because of possible endogeneity. This is done by regressing the variable against exogenous variables and using the regression-predicted values for the 3SLS estimation. Variables in bold are endogenous and part of the simultaneous equation estimation. Finally, trucks are defined as vehicles with wheelbases exceeding 65 ft.
[b]Lane-mean speeds are instrumented variables in the speed deviation system. Predicted values from the lane-mean speed system were used in this 3SLS estimation.

The 'midnight to early morning' variable, as discussed previously in its effects on lane speeds, captures driver response under near-free-flow conditions. Depending on whether it is an upgrade (eastbound) or a downgrade (westbound), driver response in car following is expected to change. In the p.m. peak hour, traffic flow increases to levels that warrant use of the middle and left lanes from a capacity standpoint, and coupled with the greater uniformity in vehicle mix, the net effect on speed deviations in the right lane is a decline.

Variable: Seasonal indicator (winter)*
Finding: In general winter effects tend to decrease speed deviations in the eastbound direction with reversed effects in the westbound direction

Winter effects capture the effects of driver behavior under inclement conditions. Although speeds tend to decline under inclement conditions, driver behavior is altered to the extent that significantly more attention is paid to the driving task. Drivers tend to maintain constant

*In the absence of microscopic weather data, lateral lane effects are captured by the seasonal indicator. While this does not cause an omitted variable bias, real-time microscopic weather information will provide interesting insights into the impacts of factors such as precipitation vs snow pileup, and rainfall vs pavement drainage on driver behavior.

headways, and minimize lane changing operations with the net effect being an associated decline in in-lane deviations*

Variable: Time-of-week indicator (weekend)
Finding: Decreases right-lane speed deviations in eastbound and westbound directions while increasing speed deviations in the middle lane

The finding on this variable illustrates selectivity in the driving population that chooses the right lane on weekends. As mentioned previously, perhaps, this class of drivers not only maintain lower speeds but also lower deviations because they are risk averse. Given that traffic volumes in the middle lane during weekends are expected to be minimal, the effects are reversed.

Variable: Truck-to-passenger car ratio
Finding: Decreases right-lane speed deviations in the eastbound direction with no significant effect in the westbound direction

Increasing truck-to-passenger car ratio effects reiterate the impact of vehicle-mix uniformity and 'truck drafting phenomena' on reduction of speed deviations. In the presence of significant upgrades, there is self-selection of the right lane by heavier traffic. In the presence of a downgrade, as evidenced in the westbound direction, this need is not compelling.

Variable: Passenger car percentage
Finding: Negatively affects eastbound speed deviations in the left lane alone

This finding illustrates locale-specific effects related to grades. The eastbound direction, which experiences significant upgrades, also experiences a greater distribution of truck traffic across the cross-section. The passenger car variable captures this effect and corroborates the impact of uniformity of vehicle mix on traffic flow dispersion.

4. CONCLUSIONS AND RECOMMENDATIONS

Endogenous relationships within lane speeds and between lane speeds and speed deviations were found to be statistically valid. The westbound and eastbound directions of our study site experienced dissimilar effects related to grade, time-of-day and time-of-week characteristics. On the other hand, the endogenous relationships in large part are similar, with estimated coefficients of like sign, means and standard errors. Our findings show that in-lane speeds are affected only by adjacent-lane speeds and in-lane speed deviations are affected progressively by adjacent lane speed deviations and in addition, in-lane and adjacent-lane speeds. Coupled with findings on the contemporaneous impact of temporal and vehicle-mix factors, such inferences corroborate the need for a comprehensive investigation into lane-mean speed and lane-speed deviation relationships. The data we used were limited (i.e. a single site) in that they did not allow us to explore variations in geometric characteristics, functional classifications, and other factors that might vary from site to site. Further insights could be gained from a more diverse data set that encompasses various regions and roadway functional classes.

The findings gathered from this paper appear promising for further application of the structural equations methodology to macroscopic traffic–flow modeling. It is quite possible that dynamic effects could be uncovered to a greater extent with more microscopic data by incorporating predetermined lane-mean speed and lane-speed deviation variables in the specifications. Such a study could have objectives relating to the unraveling of incremental dynamics in traffic flow under smaller time windows and greater seasonal, vehicle-mix constraints. While the present paper offers generic insights, understanding the cause–effect relationships between lane-mean speed and lane-speed deviations under such constraints could enrich our knowledge of driver response under

*Such behavior is more prevalent among drivers who choose to use the right and middle lanes. On the contrary, for effects on left-lane speed deviations, the self-selection of riskier drivers in left-lanes will likely cause an increase in speed deviations.

specific conditions. Such knowledge will be beneficial to the design and planning of advanced traffic management systems intended for the improving traffic flow and safety.

REFERENCES

Brilon, W. and Ponzlet, M. (1996) Variability of speed–flow relationships on German autobahns. Presented at the 75th Annual Meeting of the Transportation Research Board, Washington, D. C.

Easa, S. M. and May, A. D. (1980) Generalized procedure for estimating single- and two-regime traffic flow models. *Transportation Research Record* **772**, 24–37.

Edie, L. C. (1961) Following and steady-state theory for non-congested traffic. *Operations Research* **9**, 66–76.

Greenshields, B. D. (1935) A study in highway capacity. *Highway Research Board Proceedings* **14**, 448–477.

Greene, W. (1993) *Econometric Analysis*. Macmillan Publishing, New York, NY.

Ibrahim, A. T. and Hall, F. L. (1994) Effect of adverse weather conditions on speed–flow–occupancy relationships. *Transportation Research Record* **1457**, 184–191.

Iwasaki, M. (1991) Empirical analysis of congested traffic flow characteristics and free speed affected by geometric factors on an intercity expressway. *Transportation Research Record* **1320**, 242–250.

May, A. D. and Keller, H. E. M. (1968) Evaluation of single- and two-regime traffic flow models. *Fourth International Symposium on the Theory of Traffic Flow Proceedings*, Karlsruhe, Germany.

May, A. D. (1990) *Traffic Flow Fundamentals*. Prentice Hall, Englewood Cliffs, NJ.

Shankar, V., Mannering, F. and Barfield, W. (1995) Effect of roadway geometrics and environmental factors on rural accident frequencies. *Accident Analysis and Prevention* **27**, 371–389.

Shankar, V., Mannering, F. and Barfield, W. (1996) Statistical analysis of accident severity on rural freeways. *Accident Analysis and Prevention* **28**, 391–401.

Transportation Research Board (1994) *Highway Capacity Manual*. Special Report 209, National Research Council, Washington, DC.