P1

Hello everyone, my name is Qinglong, a master’s student at TUM. I’m happy to be here to present our project on the equations approach for modeling the endogeneity of lane-mean speeds.

P2

My presentation will consist of four parts. The first part is the motivation of this research. The second part is the methodology about how to use equations system to model lane-mean speeds. The third part will show the performance of this model. Finally, conclusions we can get from this research will be given.

P3

When modeling the lane-mean speeds, we should take care of the endogeneity problem. Endogeneity means the mean speed in each lane will not only be a function of traffic characteristics in it, but also a function of the mean speeds in the adjacent lanes. There are some methods available to solve this problem, such as three-stage least squares and full-information maximum likelihood. Shankar and Mannering had considered this problem in their researches of lane-mean speeds estimation before. However, they did not consider the influence of downstream speed. And the mean speeds in their model were aggregated in one hour which is too long. And they did not show the applicability of the model in other scenarios, so we would like to improve that model to avoid these problems and to make it more reliable. We want to provide a better understanding of mean speeds across the lanes of a multi-lane highway. In doing so, the work may eventually be applied to better understand multi-lane traffic, like car following and lane changing behavior.

P4

In our approach, a road will be composed of shorter segments. We think that using shorter segments can improve estimates whilst also allowing the influence of downstream traffic to be considered. We propose a structural model in which lane-mean speeds are influenced by environmental conditions, geometric elements, temporal, and traffic flow factors at the same time, because the simultaneous equations approach has the potential to provide an improved understanding of the interrelationships among these factors.

P5

In the equations system, the lane-mean speeds of each lane will be the responses, while the factors mentioned just now will be the independent variables. Some of them are exogenous variables which means they are not intercorrelated with the responses, while others are intercorrelated with the responses. And we can see from the equations system that the speed of a lane may be the response in one equation, but is a regressor in another equation. This is the endogeneity problem we need to solve.

P6

For example, for a two-lane highway, the speed of the left lane and the speed of the right lane have an influence on each other, so that when we model the speed of the left lane, the speed of the right lane will be a regressor, and vice versa.

P7

There is a method called three-stage least squares that can be used to solve this problem. In three-stage least squares, the first stage is to get the two-stage least squares estimates of the model system. In two-stage least squares, each endogenous variable will be regressed by all exogenous variables and the regression values will be used to replace the endogenous variables. Then each equation will be estimated by using the original least squares. The second stage of the three-stage least squares is to use the estimates from the first stage to calculate the disturbance of the system.

P8

Finally, generalized least squares is used to compute the parameter estimates.

P9

In order to find out in which conditions our model can perform best, in the experiments we calculate the results of both of the presented model and the model proposed by Shankar and Mannering, and do a comparison between them to assess if our model achieves an improvement. Furthermore, we also compare the model performance in different experiment setups, which are different number of lanes and different lengths of sections.

P10

All data used in this study come a system for transportation management in California. We find some road segments which have two successive sections with the same number of lanes and similar lengths so that we can apply our model on them. For example, if the first section is 200 meters long and has 3 lanes, then the second section should also be about 200 meters long and has 3 lanes. And the data are aggregated in 5 minutes.

P11

Following the advice of Shankar and Mannering’s research, we use their variables listed in this table in our experiments. Flow indicators, ratio of flow in the current lane, truck percentage and so on. In their research they have given an explanation about why they chose these variables, and we are not going to explain it here.

P12

We use mean square error as the assessing metric in our experiment. And we found that the improved structural equations approach is superior at estimating speed compared to the previous system. This figure is an example for the experiment in 0.2 miles long sections with 2 lanes. What we should notice is that the main difference between these two systems is whether downstream speeds are considered. So we can say the downstream speed plays a positive role in modeling lane-mean speeds.

P13

In order to find out in how many lanes the approach can have a better performance, we analyzed the results of two series, a 0.2-mile series and a 0.4-mile series. As can be seen from this figure, generally speaking, segments with fewer lanes receive better results. However, the curves of 2-lane segments are more fluctuating than those of 3-lane segments.

P14

Similarly, for exploring how long a segment can be to have the best system performance, we analyze the results of two series, a 3-lane series and a 5-lane series. We can see from the figure that 0.1-mile segments have the best result in the 3-lane series, but have the worst results in the 5-lane series. Combining results from both plots, we can conclude that the approach is more reliable in 0.2-mile segments.

P15

In this research, we can reach the following conclusions. The first was already obtained in previous studies. And after researching the first scenario, we know the introduction of downstream speeds can benefit the lane-mean speeds estimation. In addition, the equations approach can be applied in a small granularity. But there are still some problems with our study, we did not consider the influence of on-ramp stream and off-ramp stream on the traffic flow on the main road. Future studies can research the relationship between lane-mean speeds and car-following behavior as well as lane-changing behavior, and also the effects of lane-mean speeds on traffic safety problems. That’s the end of my presentation, thank you for your attention!