The Effect of Ridesharing on Congestion: Evidence from Didi Chuxing in China *

Wansong Yang†
University of Macau

December, 2021

Abstract

As an important derivative of today's sharing economy, ridesharing has been exerting profound influence on society and environment. The question whether ridesharing really reduces congestion and improves urban traffic condition has been concerned by several scholars in past years. However, there is no literature investigating impact of ridesharing on congestion in China even though China has the largest ridesharing market around the world. In this study, we will employ the dataset of Didi Chuxing (Didi), a leading ridesharing platform in China to empirically test the overall effect of ridesharing on congestion under the framework of instrumental variable method. We get numerous interesting conclusions. First, ridesharing in China does improve the traffic congestion in urban areas. Second, this effect is significant if we consider the congestion throughout the day, but it doesn't survive in peak hours. Third, ridesharing can be treated as a strictly exogeneous variable to determine the congestion level. In other words, statistically speaking, traffic congestion doesn't change Chinese people's decision for ridesharing.

Keywords: Ridesharing, Congestion, Didi Chuxing, Instrumental Variable, Exogeneous Variable

*This paper is the final project of undergraduate level *Urban Economics* course at the University of Macau in Fall 2021

†Undergraduate Student in the Dpartment of Economics, University of Macau

1 Introduction and Literature Review

Sharing economy (SE), which refers to *"a socio-economic system enabling an intermediated set of exchanges of goods and services between individuals and organizations which aim to increase efficiency and optimization of sub-utilized resources in society"* (Cohen and Munoz, 2016), has become more and more popular in recent years. Some advantages like convenience, flexibility, and even social benefits motivate people to participate and maintain their peer-to-peer relationships in SE (Yang et al., 2017). Meanwhile, SE is substantially sculpting people's way of living by deemphasizing the ownership and reshaping their consumption behaviors (de Leeuw and Gössling, 2016). Among a wide range of SE activities, accommodation and transportation become the most predominant sectors (Hossain, 2020).

Transportation sharing, also known as ridesharing, refers to the phenomenon that at least two ridesharing participants share a vehicle (Furuhata et al., 2013). The earliest ridesharing can date back to the car-sharing club in World War II in United States to cut down the recourses consuming during the wartime (Chan and Shaheen, 2012). Recently, with integration with the Internet and social networking, platform-based, on-demand and real-time matching have become new characteristics of ridesharing service (Belk, 2014). This new trend has created a series of well-known ridesharing platforms like Uber, Lyft, and Careem. Ridesharing in China has been developing especially rapidly, even though it appeared a little later relatively to western developed countries. According to the data in 2020, over 36.19% of netizens in China often travel by ridesharing and the proportion of ridesharing passenger volume to traditional taxi passenger volume is about 56.7% (of Information Industry, 2021). Didi Chuxing, known as Chinese version of Uber, is one of the leading ridesharing companies in China. It launched its ridesharing business in June, 2015, and it has even become the largest mobile transportation platform around the world since 2021 after developing for several years with more than 550 million users and 31 million drivers (Staff, 2021).

As ridesharing becomes more and more influential, a multitude of scholars have been concerning its impact over the environment, consumers' behaviors, and traffic condition (Fellows and Pitfield, 2000) (Guo et al., 2019). Although a lot of existing literature points out with the widespread of ridesharing, traffic congestion could be improved (Hossain, 2020), the situation is not so simple like that, because intuitively, ridesharing can not only reduce the number of drivers to mitigate the roads' congestion, but also attract more non-drivers because of its convenience and economic benefits, which will rather intensify congestion. Thus, whether ridesharing intensifies congestion depends on the relative magnitude between the number of vehicles on road increased by participations of non-driving people and the number of vehicles reduced by ridesharing adoptions from drivers (Alexander and González, 2015). Li et al. (2022) define these two diverse effects as *"efficiency-enhancing"* effect and *"demand-including"* effect respectively. Although the direction and mechanisms of these two effects is clear, there is an absence of perfect empirical evidence regarding what's the overall effect of ridesharing phenomenon, in face of several challenges as follows to conduct the accurate estimation of the pure ridesharing impact.

First of all, it's not easy to measure the scale of ridesharing and the extent of the traffic congestion. Some existing studies take the entry of ridesharing platform into local market as a dummy variable and then employ the difference-in-differences strategy to estimate the causal effect (Li et al., 2022). But the entry dummy may not be a perfect measure for the ridesharing volume, making their estimation not precise. In this study, based on a new dataset of Didi Chuxing from China, we employ a new variable, Gross Merchandise Volume (GMV) of 93 prefecture-level cities in China which calculates the total turnover of Didi's ridesharing business to measure the size of ridesharing. Compared with entry dummy, GMV is definitely much more desirable.

Secondly, a potential endogeneity issue must be taken into consideration which comes from the reverse causality between ridesharing and congestion. Specifically, it's reasonable to consider that the congestion also affects people's ridesharing decision, and less ridesharing volume will be expected given higher congestion level, due to higher time cost from congestion. This assumption has been consolidated by theoretical framework (Xu et al., 2015). Although the existing empirical analysis (Li et al., 2016) has taken unemployment rate as an IV to control the potential biased estimation, this IV isn't so strong as expected. In this study, a new and much stronger IV, telecommunication per capita will be employed.

The rest of the paper will be organized as follows: We will describe our dataset in detail in the second part, and then the estimation method will be given subsequently in the third part. Supplementary analysis and robustness test of our econometric model will be conducted in the fifth part, and finally we will summarize our findings.

2 Data and Descriptive Analysis

Didi Chuxing, known as Chinese version of Uber, is one of the most prominent ridesharing platforms in China, and it has become the largest travel service platform around the world since 2021. In this study, we mainly focus on the data of 93 cities in China, and the Gross Merchandise Volume (GMV), which measures the total turnover of Didi's ridesharing business in a specific region, will be taken as a proxy variable to stand for the ridesharing scale of these concerning cities. From *Didi Urban Development Index Report* (Lu et al., 2019), we can find the GMV data of 93 cities in 2017. In this report, GMV is a value that has been treated by standardization with maximum value equals to 10 and the minimum value equals to 0. The bigger the value of GMV, the larger the scale of the ridesharing will be in this city.

The Congestion Delay Index and Real Speed throughout the day and in peak hours can be used to measure the congestion level of given Chinese cities for different time periods. These data are issued by AMAP (Gaode Map), one of the most prevailing map service and auto-navigation suppliers in China. We collect these congestion data from *2017 Traffic Analysis Report of Major Cities in China* (AMAP, 2018). The Congestion Delay Index of AMAP is calculated as follows:

$$Congestion\ Delay\ Index = \frac{Real\ Travel\ Time}{Free\ Flow\ Travel\ Time}$$

It's reasonable to apply Congestion Delay Index to evaluate the urban congestion level, because if the proportion of the time delayed during the travel is bigger, higher level of congestion will be implicated. Meanwhile, for higher congestion level, the lower actual speed will be expected.

The Big Data Analysis Department of AMAP also calculated the Congestion Delay Index and Actual Speed regarding the duration both throughout the day and in peak hours respectively. The whole day refers the period from 6:00 to 22:00, while the peak hours include both morning peak hours (from 7:00 to 9:00) and evening peak hours (from 17:00 to 19:00). The summary statistics of these four variables which will act as dependent variables in succeeding analysis is shown in Table 1.

Table 1: Descriptive Statistics (Dependent Variables)

Variables	N	Mean	Standard Deviation	Min	Max
Congestion	93	1.511	0.0924	1.260	1.717
Peak Congestion	93	1.701	0.139	1.326	2.067
Speed (Km/h)	93	29.69	3.724	23.42	41.28
Peak Speed (Km/h)	93	26.44	3.675	21.12	39.20

Moreover, we also select 6 variables as control variables for the following econometric model. Among these control variables, we collect data of population (Population), proportion of employees in the secondary industry (Industry), passenger volume of public transportation (Passenger Volume)[1], and number of primary and middle schools (School) from *2018 China City Statistical Yearbook* (of Statistics, 2018). Data of density of road network in built district (Road Density) can be found from *2017 China Urban Construction Statistical Yearbook* (of Housing and Development, 2018), while data of free-flow speed can be gathered from Lu et al. (2019).

Telecommunication refers to the telecommunication service per capita. This data can also

Table 2: Descriptive Statistics (Independent Variables)

Variables	N	Mean	Standard Deviation	Min	Max
GMV	93	0.973	1.594	0.00301	10
Free-Flow Speed (Km/h)	93	44.61	3.938	37.33	56.30
Road Density (km/km^2)	93	6.182	2.258	0.320	14.57
Population (Million People)	93	6.520	4.109	0.590	33.90
Passenger Volume (100 Million People)	93	5.045	6.237	0.330	33.90
School (1,000)	93	0.945	0.635	0.163	4.072
Industry	93	0.481	0.139	0.0812	0.798
Telecommunication (1,000 Yuan)	93	1.591	1.471	0.408	7.962

be found from *2018 China City Statistical Yearbook* (of Statistics, 2018). We will use this variable as an instrumental variable (IV) in the subsequent analysis. The descriptive statistics of above variables are shown in Table 2 above.

[1]Passenger Volume of Public Transportation data of Dongguan City and Zhongshan City in 2017 are missing. What we used here is the estimated value from ARIMA model based on the data of previous years.

3 Econometric Model

3.1 Basic Model

We consider the following simple econometric model, which is the basic model for our subsequent analysis.

$$Congestion_i = \alpha GMV_i + \mathbf{X}_i^{'}\beta + \delta + \epsilon_i$$

As shown in this model, the congestion is the dependent variable and GMV which stands for the ridesharing scale of each city is the primary independent variable we concern with. **X** is the vector of 6 control variables, and they are the most important determinants for the urban traffic congestion based on the related literature.

Falcocchio and Levinson (2015) argue the disparity between the demand for and supply of transportation facilities primarily causes traffic congestions. Accordingly, we employ passenger volume of public transportation to measure the travel demand and road density to evaluate the supply of transportation facilities. Although the passenger volume of public transportation can only precisely gauge the demand for public transportation, due to the lack of other more accurate variables, we can reasonably utilize it as a proxy for total transportation demand for both public and private facilities. Besides, McClintock (1925) pointed out congestions also come from some special elements in roads that obstruct the free flow. To control innate issues of roads in each city, we use free-flow speed to attribute how smooth the road is originally. Moreover, industrial structure is another determinant of traffic congestion based on the empirical evidence from China (AMAP, 2018), and the proportion of employees in secondary industry is used to represent local industrial structure. Plus, according to La Vigne (2007), areas surrounding schools tend to be under greater pressure of traffic congestion because of too many car transportations of children to and from schools. We keep this in mind and consider number of schools in our empirical model. We anticipate there will be more serious traffic congestions for more schools in urban areas.

3.2 Instrumental Variable

The simple OLS estimation could be risky because of the potential endogeneity issue. By intuition, it is easy to imagine that at the same time when the ridesharing exerts influence on the traffic congestion, the congestion could also have an effect on people's decision for ridesharing. That is, for higher congestion level on roads, taking a vehicle will cause a higher time cost. People may change their choice to travel by ridesharing and they might take the subway or bicycles. This is a classical Reverse Causality issue, which will generate a biased estimation. An IV must be employed to ensure unbiasedness of the model.

Here, we choose telecommunication service per capita as our IV. This idea is obtained from Chase (2015)'s argument that SE is driven by three key features: platform leverage, peer-to-peer interaction, and underutilized resources. Telecommunication service per capita can be a good reflection of the power of ridesharing platform and the strength of peer-to-peer interactions of each city. Then, according to the framework of Angrist and Pischke (2009), we consider the following First Stage Regression:

$$GMV_i = \pi^F Z_i + \mathbf{X}_i^{'}\beta^F + \delta^F + \epsilon_i^F$$

Where Z refers to the telecommunication per capita, and X is the vector of control variables. The telecommunication service per capita computed in money value measures the total amount of telephone call volume, short message service, and internet usage.

For one thing, since people need to ask for ridesharing services via Didi's online platform and they need to have communications with drivers through the cellphone, only netizens are possible to take Didi's ridesharing. Undoubtedly, a large proportion of Chinese people can't use the Internet particularly the old, and this ratio is especially high in small and relatively undeveloped cities. The value of telecommunication service per capita is calculated by us manually through dividing telecommunication service by local population, so for cities with higher telecommunication service per capita, they tend to have higher proportion of netizens and higher scale of ridesharing. We can suppose that the telecommunication per capita is a key determinant to the ridesharing scale, with $Cov(GMV, Z) \neq 0$ and the first stage exists. This is just the intuition, and we will conduct more detailed weak IV test in subsequent analysis.

For another, it's reasonable to assume that telecommunication service per capita has no direct effect on the traffic congestion with $Cov(Z, \epsilon|X) = 0$ and the exclusion restriction condition can be satisfied.

$$Congestion_i = \alpha^S \widehat{GMV}_i + \mathbf{X}_i^{'}\beta^S + \delta^S + \epsilon_i^S$$

According to Angrist and Pischke (2009), we follow the 2SLS procedure to estimate the fitted value of GMV from the First Stage Regression and then consider the Second Stage Regression above.

4 Results

4.1 OLS Regression

We run the simple OLS regression of four different dependent variables with respect to the GMV, independent variable and other control variables. The OLS results together with Breusch-Pagan statistic are shown in table 3. The result of Breusch-Pagan tests indicates that there is no heteroscedasticity issue except the last regression where we want to interpret the variation of peak-hour speed, and we substitute its standard errors by robust standard errors.

We mainly focus on the statistical inference of the partial effect of GMV. From the results, we can see the ridesharing has positive influence to alleviate the traffic congestion throughout the day.

That is, the effect of GMV is negative to the Congestion Delay Index and positive to the Actual Speed throughout the day and it is statistically significant. However, for the peak-hour case, although the sign of coefficients provides us the similar indications with the whole-day situation, unfortunately, they are not statistically significant.

Table 3: OLS Output

Variables	(1) OLS Congestion	(2) OLS Speed	(3) OLS Peak Congestion	(4) OLS Peak Speed
GMV	-0.0173**	0.362**	-0.0168	0.230
	(0.00840)	(0.168)	(0.0142)	(0.300)
Population	-0.00674*	0.143*	-0.00357	0.0884
	(0.00371)	(0.0745)	(0.00629)	(0.0883)
Industry	-0.271***	5.519***	-0.262***	4.159***
	(0.0518)	(1.039)	(0.0877)	(1.374)
Road Density	-0.00787**	0.191***	-0.00762	0.143
	(0.00303)	(0.0608)	(0.00513)	(0.0903)
Passenger Volume	0.00976***	-0.201***	0.0144***	-0.217**
	(0.00245)	(0.0492)	(0.00415)	(0.0943)
Free-Flow Speed	-0.00588***	0.802***	-0.00804***	0.746***
	(0.00173)	(0.0347)	(0.00293)	(0.0429)
School	0.0344	-0.743*	0.0276	-0.654
	(0.0217)	(0.436)	(0.0368)	(0.507)
Constant	1.931***	-9.483***	2.174***	-8.784***
	(0.0802)	(1.610)	(0.136)	(2.132)
Observation	93	93	93	93
R-squared	0.582	0.896	0.469	0.801
Breusch-Pagan	0.22	2.47	0.53	3.94
P-Value (B-P test)	0.6384	0.1157	0.4665	0.0472

Standard errors in parentheses for OLS 1 to 3
Robust standard errors in parentheses for OLS 4
*** p<0.01, ** p<0.05, * p<0.1

In conclusion, this simple OLS regression implies the ridesharing does improve the traffic congestion in China during whole-day period. But the ridesharing has no effect on the congestion in peak hours.

4.2 2SLS Results

The results of the 2SLS Regression are shown in Table 4, where we combine OLS and 2SLS estimation together. From the results of 2SLS, the effect of ridesharing on congestion level is still statistically significant if we consider whole-day situation, while it's not statistically significant in peak hours. Therefore, we still can't get any evidence that the ridesharing can exert influence on the traffic congestion in peak hours in China.

For the whole-day situation, the partial effect of ridesharing estimated by 2SLS is -0.0453, and it's higher than OLS estimator, which means that the effect coefficient is underestimated due

Table 4: 2SLS Output

Variables	(1) OLS Congestion	(2) OLS Peak Congestion	(3) 2SLS Congestion	(4) 2SLS Peak Congestion
GMV	-0.0173**	-0.0168	-0.0453**	-0.0400
	(0.00840)	(0.0186)	(0.0186)	(0.0301)
Population	-0.00674*	-0.00357	-0.00807**	-0.00466
	(0.00371)	(0.00524)	(0.00386)	(0.00624)
Industry	-0.271***	-0.262***	-0.244***	-0.239***
	(0.0518)	(0.0812)	(0.0551)	(0.0891)
Road Density	-0.00787**	-0.00762	-0.00845***	-0.00810
	(0.00303)	(0.00536)	(0.00310)	(0.00501)
Passenger Volume	0.00976***	0.0144**	0.0165***	0.0200***
	(0.00245)	(0.00589)	(0.00470)	(0.00760)
Free-Flow Speed	-0.00588***	-0.00804***	-0.00646***	-0.00851***
	(0.00173)	(0.00238)	(0.00179)	(0.00289)
School	0.0344	0.0276	0.0364*	0.0293
	(0.0217)	(0.0311)	(0.0221)	(0.0358)
Constant	1.931***	2.174***	1.947***	2.187***
	(0.0802)	(0.111)	(0.0821)	(0.133)
Observations	93	93	93	93
R-squared	0.582	0.469	0.527	0.452

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

to the potential endogeneity issue. This result quite fits our intuition because under the analytical assumption of reverse causality, decreased congestion level caused by ridesharing will further influence people's ridesharing decision and increase the scale of ridesharing. This rise of ridesharing will continue declining the congestion level. So, the first-round effect estimated by OLS will absolutely smaller than the final effect estimated by 2SLS.

5 Supplementary Analysis

5.1 Test for the Instrumental Variable

To check the robustness of our results, we run the first-stage regression shown in the Table 5. From the first-stage regression, we find that our IV, telecommunication service per capita, is strongly significant and highly correlated with GMV, with t statistic equal to 4.77. Then we follow Stock and Yogo (2005) to further conduct the weak identification test and get Cragg-Donald Wald F statistic equal to 22.748. This especially large statistic indicates our IV is valid. Meanwhile, since Anderson LM statistic is 19.635 (p-value is 0), the null hypothesis that the IV is under identified must be rejected. There is also no over identified issue because the number of

IV is exactly equal to the number of endogenous explanatory variable.

5.2 Test for the Endogeneity

We go along with the procedure of Hausman (1978) to test the whether the ridesharing scale is endogenous or not. We add the residual of the first-stage regression into the original OLS to test whether the residual term is significant in the new regression. The result is displayed in the fourth regression of Table 5. The t statistic of the residual term is 1.75 which is not highly significant. Additionally, when we apply Hausman test to compare if there is a systematic difference between OLS and 2SLS estimate, Hausman statistic is 2.986 (p-value is 0.0840) and we can't reject the null hypothesis with 5% significant level.

Table 5: Regressions for Supplementary Analysis

Variables	(1) 1st Satge GMV	(2) 2SLS Congestion	(3) OLS Congestion	(4) OLS Congestion
GMV		-0.045**	-0.017**	-0.045**
		(-2.44)	(-2.06)	(-2.51)
Telecommunication (IV)	0.399***			
	(4.77)			
Residual				0.036*
				(1.75)
Control Variables	Included	Included	Included	Included
Constant	0.190	1.947***	1.931***	1.947***
	(0.21)	(23.71)	(24.07)	(24.40)
Observations	93	93	93	93
R-squared	0.815	0.527	0.582	0.596
Cragg-Donald Wald F statistic	22.748			
Anderson LM statistic	19.635			
Hausman Statistic		2.986		
P-value (Hausman)		0.0840		

t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

We can conclude from Hausman test that the endogeneity issue is not so serious, and we can treat the ridesharing scale as exogeneous variable if we consider its effect on congestion in China. This finding is appealing because it implies that traffic congestion almost doesn't have an effect on Chinese people's decision to take ridesharing. In other words, Chinese people generally will keep on travelling by ridesharing even though time cost is increased by traffic congestion.

6 Conclusion

This paper mainly examines the overall impact of ridesharing on congestion under China's background. As far as we know, compared to past empirical analyses, this study primarily focuses on China's situation and employ the most precise proxy variable, GMV of Didi Chuxing in 2017 to measure ridesharing level for the first time. In faced of potential endogeneity issue, we find an especially strong IV, telecommunication service per capita and employ 2SLS estimation to make the estimation consistent. We use different variables, Congestion Delay Index and Actual Speed, as dependent variables to measure the congestion level both throughout the day and during peak hours and run both OLS and 2SLS regression to detect the effect of GMV on these dependent variables. Finally, we do the Hausman test to detect whether the ridesharing is endogenous with respect to traffic congestion in Chinese market. From the empirical evidence, several interesting implications can be derived as follows:

First, in China, the ridesharing can only improve the whole-day traffic condition, but it does nothing for peak-hour congestions. That is, both OLS and 2SLS estimation show ridesharing has negative effect on congestion delay and positive impact on actual speed throughout the day, based on which we can conclude that ridesharing does mitigate traffic congestion for whole-day duration. However, we don't have the similar inference if we consider congestion in peak hours because the estimator is statistically insignificant.

Second, the endogeneity of ridesharing is so weak that we can treat is as strictly exogeneous variable when we test its impact on traffic congestion, which implies that the congestion level doesn't affect the ridesharing decision of Chinese people throughout the day. Although there is a bit difference in magnitude, statistically significant estimation results from both OLS and 2SLS motivate us to conduct the Hausman test to investigate whether the endogeneity issue exists or not. After confirming the IV is especially strong, we follow the procedure of Hausman (1978) and discover that there is no systematic difference between OLS and 2SLS estimation in statistical sense, indicating no endogeneity issue actually exists.

However, this study still has a series of shortcomings which need to be improved by subsequent research. On one hand, the fundamental weakness is about the dimension of our dataset. Due to the limitation of data sources, we can only get the cross-sectional data of the ridesharing business of Didi Chuxing in 2017. Our conclusions will be much more convincing if a higher-dimension panel dataset can be employed. But the analysis in this paper is the best work we can do. On the other hand, although the GMV data of Didi Chuxing is reasonable to measure the ridesharing scale of each city, it's not perfect since there are still other famous ridesharing platforms widely used by Chinese people like Dida Chuxing and Hello Chuxing. Further studies had better take the ridesharing scale of multiple platforms into account to make the estimation much closer to the real-world situation.

References

Lauren P Alexander and Marta C González. Assessing the impact of real-time ridesharing on urban traffic using mobile phone data. *Proc. UrbComp*, 15:1–9, 2015. 1

AMAP. 2017 traffic analysis report of major cities in china, 2018. 2, 3.1

Joshua D Angrist and Jörn-Steffen Pischke. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press, 2009. 3.2

Russell Belk. You are what you can access: Sharing and collaborative consumption online. *Journal of business research*, 67(8):1595–1600, 2014. 1

Nelson D Chan and Susan A Shaheen. Ridesharing in north america: Past, present, and future. *Transport reviews*, 32(1):93–112, 2012. 1

Robin Chase. Sharing economy—reconstructing the new model of future business. *Hangzhou, Zhejiang: Zhejiang People's Publishing House*, 2015. 3.2

Boyd Cohen and Pablo Munoz. Sharing cities and sustainable consumption and production: towards an integrated framework. *Journal of cleaner production*, 134:87–97, 2016. 1

Tim de Leeuw and Tobias Gössling. Theorizing change revisited: An amended process model of institutional innovations and changes in institutional fields. *Journal of Cleaner Production*, 135:435–448, 2016. 1

John C Falcocchio and Herbert S Levinson. *Road traffic congestion: a concise guide*, volume 7. Springer, 2015. 3.1

NT Fellows and David E Pitfield. An economic and operational evaluation of urban car-sharing. *Transportation Research Part D: Transport and Environment*, 5(1):1–10, 2000. 1

Masabumi Furuhata, Maged Dessouky, Fernando Ordóñez, Marc-Etienne Brunet, Xiaoqing Wang, and Sven Koenig. Ridesharing: The state-of-the-art and future directions. *Transportation Research Part B: Methodological*, 57:28–46, 2013. 1

Yue Guo, Xiaotong Li, and Xiaohua Zeng. Platform competition in the sharing economy: Understanding how ride-hailing services influence new car purchases. *Journal of Management Information Systems*, 36(4):1043–1070, 2019. 1

Jerry A Hausman. Specification tests in econometrics. *Econometrica: Journal of the econometric society*, pages 1251–1271, 1978. 5.2, 6

Mokter Hossain. Sharing economy: A comprehensive literature review. *International Journal of Hospitality Management*, 87:102470, 2020. 1

Nancy Gladys La Vigne. *Traffic congestion around schools*. US Department of Justice, Office of Community Oriented Policing Services ..., 2007. 3.1

Ziru Li, Yili Hong, and Zhongju Zhang. An empirical analysis of on-demand ride sharing and traffic congestion. In *Proc. International Conference on Information Systems*, 2016. 1

Ziru Li, Chen Liang, Yili Hong, and Zhongju Zhang. How do on-demand ridesharing services affect traffic congestion? the moderating role of urban compactness. *Production and Operations Management*, 31(1):239–258, 2022. 1

Ming Lu, Waiting Wang, Hongwei Xu, Xin Feng, and Yuhong Chen. Towards an efficient and livable city: Didi "urban development index" report, 2019. 2, 2

Miller McClintock. *Street traffic control.* McGraw-Hill Book Company, 1925. 3.1

Ministry of Housing and Urban-Rural Development. 2017 china urban construction statistical yearbook, 2018. 2

Ministry of Information Industry. China's sharing economy development report, 2021. 1

National Bureau of Statistics. 2018 china city statistical yearbook, 2018. 2, 2

CNBC Staff. Didi chuxing: 2021 cnbc disruptor 50, 2021. 1

James Stock and Motohiro Yogo. Asymptotic distributions of instrumental variables statistics with many instruments. *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg*, 6:109–120, 2005. 5.1

Huayu Xu, Jong-Shi Pang, Fernando Ordóñez, and Maged Dessouky. Complementarity models for traffic equilibrium with ridesharing. *Transportation Research Part B: Methodological*, 81: 161–182, 2015. 1

Shuai Yang, Yiping Song, Sixing Chen, and Xin Xia. Why are customers loyal in sharing-economy services? a relational benefits perspective. *Journal of Services Marketing*, 2017. 1