JUAN SEBASTIÁN MIRANDA PATIÑO – 201516821

ECONOMÍA URBANA

UNIVERSIDAD DE LOS ANDES

2022

1 DESCRIPRION OF INPUTS

1.1.Description of the data sources and raw data

Table 1.1 Raw data information.

Data.source	Page	Data.Files	Location	Provided	Cited
"Baidu Maps"	5	Not available	data/to clean/	FALSE	TRUE
"2010- 2015 Beijing Household Travel Surveys	19	Not available	data/to clean/	FALSE	TRUE
2010- 2017" "Statistical Yearbook of Chinese Citie	A14	Not available	data/to clean/	FALSE	TRUE
"2016 Beijing Transportation Annual Report"	A24	Not available	data/to clean/	FALSE	TRUE

1.2.Description of the analytic data

Table 1.2: Analysis data information

Analytic.data	Location	Description
BaseSamp.dta	\Subways and Road Congestion\Data	Data for generate most results
ExtendSample.dta	\Subways and Road Congestion\Data	Data for Figure 6
HhdVKTSample.dta	\Subways and Road Congestion\Data	Data for Table 5
IndModeSample.dta	\Subways and Road Congestion\Data	Data for Table 5
PublicTransits.dta	\Subways and Road Congestion\Data	Data for Appendix Table B.4.
RushNonrushHours.dta	\Subways and Road Congestion\Data	Data for Figure 6
SubwayLines.dta	\Subways and Road Congestion\Data	List of subway lines

WithAgainstTraffic.dta	\Subways and Road Congestion\Data	Data for Figure 6
2017_part18.dta	\Subways and Road Congestion\Data	Data for Appendix Figure A.4
linkInfo_new.dta	\Subways and Road Congestion\Data	Data for Appendix Figures A.5 and B.6.
HourlySample.dta	\Subways and Road Congestion\Data	Data for Appendix Table B.5.
UrbanTransitReports.dta	\Subways and Road Congestion\Data	Data for city-level public transit analysis.
Sample_ControlCities.dta	\Subways and Road Congestion\Data	Control cities analysis.
Sample_Treated Cities.dta	\Subways and Road Congestion\Data	Treated cities analysis.

1.3.Description of the code scripts

Table 1.3 Codes files information.

File_name	Location	Inputs	Outputs	Description	Primary_type
_Setup.do	\Subways and Road Congestion\Codes\Data cleaning	DOES NOT APPLY	DOES NOT APPLY	The necessary packages are installed to be able to reproduce the code.	Construction
ExtractControlCities.do	\Subways and Road Congestion\Codes\Data cleaning	Speed.dta	Sample_ControlCities.dta	Maintain 2.5 km road sections for control lines, drop road segments with no road name, and local streets. In addition, only observations are kept during peak hours, weekends and holidays, and finally duplicates are eliminated.	Cleaning/ construction
ExtractTreatedLines.do	\Subways and Road Congestion\Codes\Data cleaning	TreatedLines_SelectedRoads.dta	Sample_TreatedLines.dta	Maintain 2.5 km road sections for treated lines, drop road segments with no road name, and local streets. In addition, only observations are kept during peak hours, weekends and holidays, and finally duplicates are eliminated.	Cleaning/ construction
Gen_BaseSamp.do	\Subways and Road Congestion\Codes\Data cleaning	City_AllCasesWkly_ExclTestRides.	BaseSamp.dta	Additional cleanup in the cross City_AllCasesWkly_ExclTestRides.do database as well as	Cleaning/ construction

				variable creation, labeling, and selection.	
Gen_ExtendSamp.do	\Subways and Road Congestion\Codes\Data cleaning	CrossCityPolybyHour_ExclTestRides_AllCasesWkly.dta	ExtendSample.dta	Treated road segments are incorporated. also of generating, selecting and labeling variables with geographical characteristics	Cleaning/ construction
Gen_HhdVKTSample.do	\Subways and Road Congestion\Codes\Data cleaning	Data_TravelSurvey	HouHhdVKTSample.dta HourlySample.dta	Creation of the sections of the control and treaty cities, cleaning of speed data, creation of treaty groups and creation of hourly samples are incorporated.	Cleaning/ construction
Gen_IndModeSample.do	\Subways and Road Congestion\Codes\Data cleaning	Data_TravelSurvey	IndModeSample.dta	Compare Beijign household travel account data from 2010 to 2015, extract household characteristics information, metro distance information, transportation options and individual choice information for travel	Construction
Gen_HhdVKTSample.do	\Subways and Road Congestion\Codes\Data cleaning	Sample_TreatedLines.dta Sample_ControlCities.dta CrossCity_AllCasesWkly_ExclTestRides.dta	HourlySample.dta	Combination of commuter survey of Beijing households from 2010 to 2015, information on car ownership, use and characteristics, and a sample of domestic vehicle use is generated.	Construction
Fig1.do	\Subways and Road Congestion\Codes	UrbanTransitReports.dta	TablesFigures/Fig1.pdf	It produces Figure 1 which shows the growth of subway stations and the number of passengers between 2000 and 2020 in China.	Analysis
Fig2.do	\Subways and Road Congestion\Codes	BaseSamp.dta	TablesFigures/Fig2.pdf	Produces the dynamic effects of launches from subway stations.	Analysis
Fig3.do	\Subways and Road Congestion\Codes	BaseSamp.dta	TablesFigures/Fig3.pdf	Allows plotting of Wald statistics from estimates with placebo opening dates vs actual opening week.	Analysis
Fig4.do	\Subways and Road	BaseSamp.dta	TablesFigures/Fig4.pdf	Gives graphical evidence of the stacked	Analysis

	Congestion\Codes			DID specsification for the weeks pre and post treatment.	
Fig5.do	\Subways and Road Congestion\Codes	Data/BaseSamp.dta	TablesFigures/Fig5A.pdf	Divide Figure 4 by time periods for the year 2017.	Analysis
Fig6.do	\Subways and Road Congestion\Codes	ExtendSample.dta	TablesFigures/Fig6.pdf	It produces the heterogeneous effects by characteristics of the road segment.	Analysis
Fig7.do	\Subways and Road Congestion\Codes	ExtendSample.dta	TablesFigures/Fig6.pdf	It illustrates the welfare impact for people who continue to commute by car as well as for those who commute by subway.	Analysis
Tab2	\Subways and Road Congestion\Codes	CityChars.dta	TablesFigures/Tab2B	Summarizes the descriptive statistics of the reference data	Analysis
Tab3	\Subways and Road Congestion\Codes	BaseSamp.dta	TablesFigures/Tab3.tex	Shows base estimates	Analysis
Tab4	\Subways and Road Congestion\Codes	BaseSamp.dta	TablesFigures/Tab4.tex	Shows the robust estimates of the model.	Analysis
Tab5	\Subways and Road Congestion\Codes	BaseSamp.dta	TablesFigures/Tab5_Bstrp.dta	Displays descriptive statistics regarding the individual and domestic VKT mode of transport options	Analysis
AppA_FigA4	\Subways and Road Congestion\Codes	AppA_FigA4	AppA_FigA4.pdf	Generates Appendix Figure A.4.	Analysis
AppA_FigA5	\Subways and Road Congestion\Codes	AppA_FigA4	TablesFigures/AppA_FigA5.pdf	Generates Appendix Figure A.5.	Analysis
AppB_FigB1FigB2FigB3TabB1	\Subways and Road Congestion\Codes	BaseSamp.dta	AppB_SeasonalityCityChars_H.pdf	Generates Appendix Figure B.1, Figure B.2, Figure B.3, and Appendix Table B.1	Analysis
AppB_FigB4	\Subways and Road Congestion\Codes	BaseSamp.dta	AppB_FigB4_D.pdf	Generates Appendix Figure B.4.	Analysis
AppB_FigB5TabB2	\Subways and Road Congestion\Codes	BaseSamp.dta	AppB_FigB5.pdf	Generates Appendix Figure B.5 and Table B.2.	Analysis

AppB_Fig B6	\Subways and Road Congestion\Codes	BaseSamp.dta	AppB_CasebyCase2.pdf	Generates Appendix Figure B.6	Analysis
AppB_Fig B7	\Subways and Road Congestion\Codes	BaseSamp.dta	AppB_FigB7_Tab.tex	Generates Appendix Figure B.7	Analysis
AppB_Tab B3	\Subways and Road Congestion\Codes	ExtendSample.dta	AppB_TabB3.tex	Generates Appendix Table B.3.	Analysis
AppB_Tab B4	\Subways and Road Congestion\Codes	PublicTransits.dta	AppB_TabB4.tex	Generates Appendix Table B.4.	Analysis
AppB_Tab B3	\Subways and Road Congestion\Codes	HourlySample.dta	AppB_Tab5.tex	Generates Appendix Table B.5.	Analysis

2 ASSIGN A REPRODUCIBILITY SCORE.

As specified by ACRe Guidlines, the reproducibility level is 6 (L6). This taking into account that the analysis and cleaning code are available and complete. In the same way, it has all the data files for data analysis, however, it does not have the raw data. So one level improvement is being able to add partially (L7) or completely (L8) this raw data.

	Analysis Code		Analysis Data		CRA	Cleaning Code		Raw Data		CRR
	P	C	P	C		P	C	P	C	
L1: No materials	–	–	–	–	–	–	–	–	–	–
L2: Only code	✓	✓	–	–	–	–	–	–	–	–
L3: Partial analysis data & code	✓	✓	✓	–	–	–	–	–	–	–
L4: All analysis data & code	✓	✓	✓	✓	–	–	–	–	–	–
L5: Reproducible from analysis	✓	✓	✓	✓	✓	–	–	–	–	–
L6: All cleaning code	✓	✓	✓	✓	–	✓	✓	–	–	–
L7: Some raw data	✓	✓	✓	✓	–	✓	✓	✓	–	–
L8: All raw data	✓	✓	✓	✓	–	✓	✓	✓	✓	–
L9: All raw data + CRA	✓	✓	✓	✓	✓	✓	✓	✓	✓	–
L10: Reproducible from raw data	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

[a] **Computationally Reproducible from Analytic data (CRA):** The output can be reproduced with minimal effort starting from the *analytic* datasets.

[b] **Computationally Reproducible from Raw data (CRR):** The output can be reproduced with minimal effort from the *raw* datasets.