Citation: WANG Ziyi, LI Jian. A Scenario-Oriented Macro and Micro Data Nested Analysis Method for Customized Bus Travel Demand [J]. Urban Transport of China, 2022 (6).

The English version here is roughly translated and has not been reviewed or edited. Please contact us if you need a finely translated and edited version.

A Scenario-Oriented Macro and Micro Data Nested Analysis Method for Customized Bus Travel Demand

WANG Ziyi[1], LI Jian[1,2]

1. Key Laboratory of Road and Traffic Engineering of Ministry of Education, Tongji University, Shanghai 201804, China;
2. Urban Mobility Institute, Tongji University, Shanghai 200092, China

Abstract: Analyzing the potential travel demand for customized buses is the basis of the service promotion. However, the existing relevant research focuses on the macro demand characteristics or mode choices' willingness under specific travel scenarios, which is difficult to generalize the various application scenarios and corresponding potential demand characteristics of customized bus systems. Combining the respective advantages of large data feature mining and small sample individual behavior, this paper puts forward a scenario-oriented macro and micro nested analysis method for customized bus travel demand. To estimate the customized bus travel demand in different scenarios, big data mining is utilized to reveal the macro characteristics of travel demand, and a small sample survey can help obtain the mode choices' willingness of micro individuals and analyze the nested connection through consistent travel scenarios. Then, Xiamen is selected as an example to verify the applicability of the method. The results show that, according to the characteristics of potential demand, various travel scenarios can be divided into four types including "high demand, high mode-shift rate", corresponding to differentiated customized public transportation marketing strategies. **DOI:** 10.13813/j.cn11-5141/u.2022.0605-en

Keywords: urban public transportation; customized bus; travel demand; scenario oriented; macro and micro data nested analysis

0 Introduction

Customized bus (CB) is an emerging form of demand-responsive public transportation that uses internet platforms to collect passengers' travel needs and plan routes and schedules based on demand characteristics [1–2]. CB offers personalized travel services that cater to residents' individualized needs, while online demand collection and reservation booking through mobile internet have improved response efficiency and service quality. In recent years, especially during the COVID-19 epidemic, CB in many cities have achieved ideal passenger flow levels [3–5], supplementing traditional buses and reflecting the advantages of flexibility, comfort, and traceable passenger flow [6–7]. Developing CB can activate and enrich the public transportation service market, enhancing its attractiveness to travelers, particularly when the passenger flow of urban buses is slow to recover.

To promote the development of CB, it is essential to analyze the potential travel demand for such services. Existing research on CB demand analysis has focused on macro feature analysis based on travel big data [8–13] and micro behavior willingness based on sample surveys [14–15]. However, these studies have limitations. Macro feature analysis often discusses urban travel demand as a whole or only focuses on commuting demand, ignoring the spatial and temporal demand characteristics of CB in different scenarios. On the other hand, micro behavior willingness analysis is limited by the sample size of the survey and only analyzes the behavioral mechanisms and travel willingness of specific travel scenarios, making it difficult to grasp the demand scale and spatial distribution of CB. The commercial promotion of CB needs to give full play to the advantages of travel big data and small sample behavior survey, clarify the potential demand scale and spatial-time distribution of current travel groups under different scenarios from the urban level, as well as the acceptance of different types of CB modes, then combine the conclusions above.

This paper proposes a scenario-oriented macro and micro nested analysis method for the potential travel demand of CB. First, segment travel scenarios according to travel attributes, and build a scenario-oriented travel demand analysis framework; Secondly, mine the macro characteristics of travel demand based on travel big data, obtain the choice willingness of micro individuals based on small sample survey; Then realize the nested connection of the above two conclusions through consistent travel scenario in order to study potential demand of CB under various scenarios, providing decision-making reference for the government's public resource allocation and CB enterprise operation.

Received: 2022-04-29

1 Literature review

1.1 Analysis of customized bus travel demand

Travel big data (or “macro data”) and survey data (or “micro data”) are two common data sources used for analyzing travel demand. Due to differences in collection purpose, methods, and information contained, the former is mostly used for macro feature mining while the latter is applied for analyzing micro individual demands and mode choice intentions. Each type of data contains its own information boundaries, leading to limitations in studies focusing on CB demand.

Macro feature mining based on travel big data is currently the main method for assessing potential CB demand. This method extracts a certain type of travel OD information based on resident travel data and assesses travel demand based on spatial-temporal clustering features. Some studies [8–9] use IC card data to reconstruct trip chains and identify commuting OD, with Chen Xi et al. [8] identifying CB demand hotspots using spatial clustering methods and Guo Rongge et al. [9] establishing utility indicators to explore CB service routes. Some studies [10–12] use taxi GPS data to mine potential passenger groups and OD hotspots, while Yu Qing et al. [13] discuss CB service route design methods based on mobile phone signaling data. However, the above data records do not include individual attribute information, making it difficult to analyze the shifting willingness of travel groups to CB, resulting in low credibility of research conclusions.

CB mode choice intention analysis based on small sample survey data focuses on specific populations’ willingness to choose CB. Firstly, RP and SP surveys are conducted on commuting groups, followed by statistical analysis methods such as discrete choice models to discuss specific groups’ willingness to choose CB and the mechanism behind it, achieving research on individual-level behavior and attitudes [14–15]. However, it is difficult to conduct surveys on different travel groups using traditional questionnaire forms and difficult to expand survey results to the city level, hindering the broad promotion of research conclusions.

It should be noted that most studies discuss CB travel demand as a whole [10–13], or only focus on commuting [8–9, 14–15], lacking consideration of various practical service scenarios for CB and differences in travel demand between different scenarios. In fact, they do not analyze relevant demand from the perspective of CB diversification characteristics, which is not conducive to guiding CB practice.

1.2 Method of scenario oriented

The concept of scenario-oriented (or “scenario-driven”) originated from marketing. Marketing research has shown that consumer behavior often occurs in specific scenarios of daily life, and consumers’ perception of services is strongly related to these scenarios. Therefore, distinguishing different scenarios when designing products or services can more accurately meet consumers’ personalized needs in different scenarios [16]. In recent years, the method of scenario-driven has gradually expanded from the marketing process to product designing, smart manufacturing, and management system development, becoming an innovative development model. The concept of scenarios has also been generalized from consumer and application scenarios to demand scenarios.

The method of scenario-driven has been applied in the development of transportation systems, such as the development of big data platforms based on different data application scenarios [17], the design of signal control simulation systems for future urban transportation scenarios [18], and the research on autonomous driving systems in specific environmental scenarios [19]. However, existing methods for analyzing travel demand for new formats, such as CB, still use a top-down analysis model for traditional transit, and lack consideration of the demand differences caused by diversified travel service products from the perspective of multiple travel scenarios. In fact, introducing scenario-driven thinking and conducting scenario-based differential analysis of CB demand is an important foundational work for promoting public transportation priority using market-oriented approaches and aligns with the direction of dynamic, refined, and accurate governance [20].

1.3 Method of nested analysis

The analysis concept of macro-micro data nesting (also known as “integration of big and small data”) has emerged with the widespread application of big data, and it can specifically address the following shortcomings of big data: first, big data belongs to “weak designed” data [21], which does not serve scientific research specifically [22], and its information content is limited with observation boundaries; second, the lack of individual attribute data in big data leads to emphasize correlation rather than causality [23], making it difficult to verify causal relationships [24]; finally, big data is not a complete sample data [22–24], and cannot represent the entire social population. The macro-micro data nesting approach can leverage the advantages of big and small data, and through the construction of information keys, achieve effective matching of big and small data, while also verifying micro-behavioral mechanisms through macro feature correlation analysis [21–24].

The method of nested analysis has been widely applied in disciplines such as sociology and geography. Sun Shichao [25] used the IC card number as the linking key to construct a nested connection between macro transit travel data and micro individual survey data, thus achieving a comprehensive measurement of user behavior and attitude features, and completing user market segmentation in two dimensions. J. Blumenstock et al. [26] conducted a sample survey on telephone users in Rwanda, using the measured household asset information as a label to construct a supervised learning model for predicting household asset status based on anonymous phone call records big data, and analyzed the overall poverty situation in the area. Huang Wei et al. [27] established a mapping relationship between partial user mobile signaling location data and their true individual residence, and constructed a supervised learning model based on signaling big data to predict residential location, and analyzed the distribution of population in residential areas. Qin Xiao et al. [24] proposed three method frameworks for combining big and small data: EDCF, MFVE, and AADSF, to discuss the inter-influence mechanisms between urban space and resident behavioral activities. Applying the above framework, Qiu Jing et al. [28] discovered two types of spatial patterns in the main urban area of Changzhou city based on WeChat hot spot big data: balanced and unbalanced employment-residence spaces. They then built regression analysis models based on POI data and micro survey data to analyze the causes of the differences between the two types of spaces.

Taking the analysis of potential CB travel demand as an example, this paper combines traditional micro-sample survey methods to conduct a sample survey on corresponding travelers, which can supplement individual behavioral intention evidence for macro feature conclusions based on travel data analysis, and overcome the one-sidedness of analysis based on a single data source.

2 Framework of scenario-oriented macro and micro data nested analysis

2.1 Basic logic

Travelers in cities meet their diverse travel needs through various existing transportation modes. With the emergence of CB services, some travelers may switch from original mode to CB, and different travel scenarios may lead to differences in travelers' potential demand for CB. In various scenarios, the potential demand for CB depends on two factors: the demand (and characteristics) of the original travel mode, and the willingness or mode-shift rate of travelers.

Based on the above logic, combining with scenario-oriented macro-micro data nested analysis method, this paper proposes a framework for analyzing potential demand for CB (see Figure 1), which includes four steps: Research object selection and data collection, scene segmenting, macro and micro data analysis, and nested analysis.

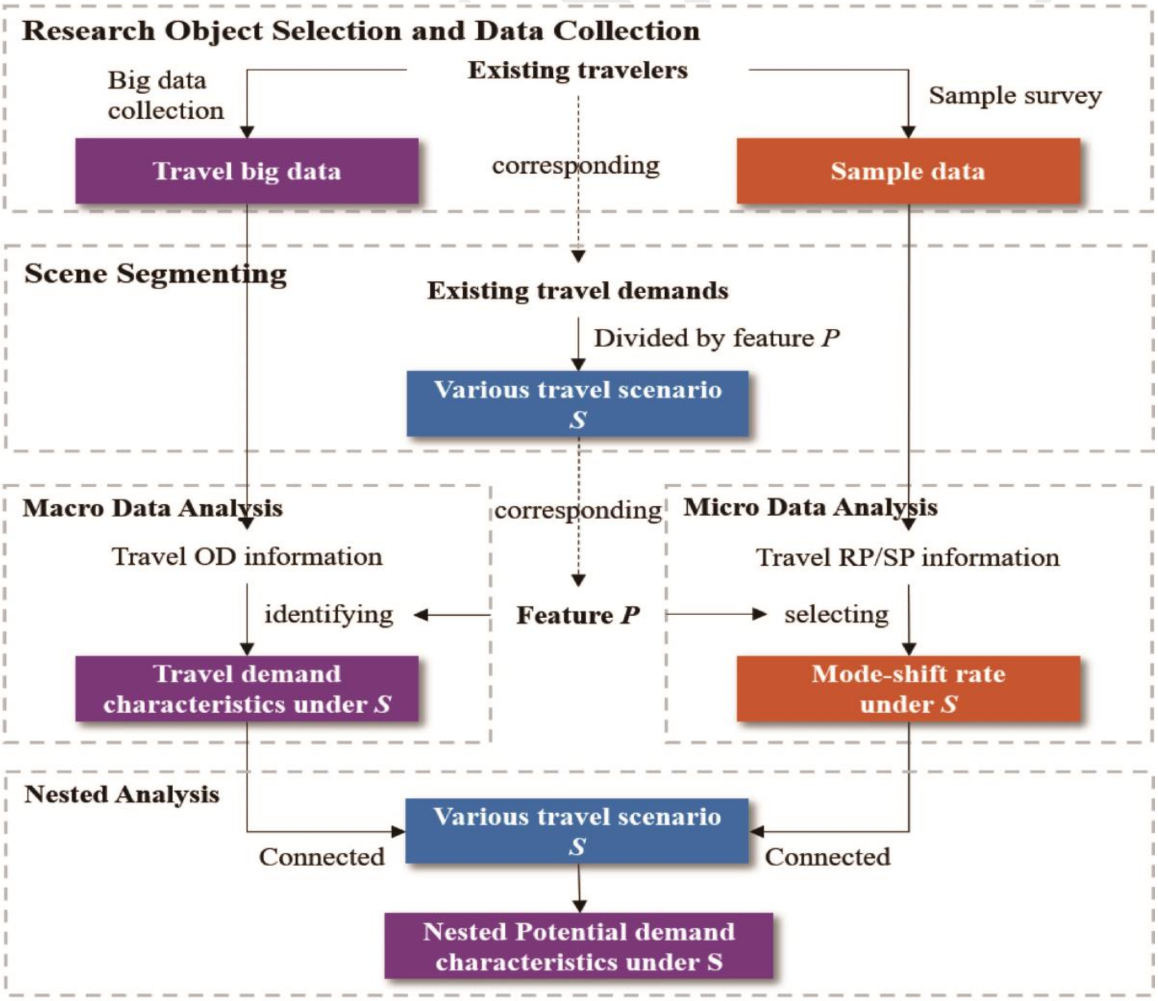

Fig.1 Scenario-oriented macro and micro data nested analysis framework for potential travel demand of customized bus

2.2 Research object selection and data collection

Select the existing travelers of a certain mode as the research object and collect data. The travel big data of individual travel activities (see Table 1), usually collected using public transit IC cards, mobile digital devices, etc.; sample survey data is obtained based on a sample survey of the research group. It is recommended to choose research subjects with relatively low data acquisition difficulty.

The method framework proposes the following requirements for the two types of data: macro data should be able to extract complete travel OD information from the raw data; micro data should focus on sampling methods, sample size, and survey content to improve the credibility of small-sample survey data.

1) Sampling methods

Understand the overall distribution of the target population's attributes (such as gender, age, etc.) and conduct stratified sampling based on these characteristics. If cost constraints prevent stratified sampling, simple random sampling can be used instead, and the consistency between the sample's attribute distribution and the overall distribution can be controlled to improve the representativeness of the sample.

2) Sample size

If the sampling rate of resident travel survey cannot be reached by using small sample survey, the sample size required for simple random sampling can be calculated according to the accuracy requirements. For the common normal population, when the total number of subjects is > 5 000 and the confidence level is 95%, the error range of ±10% can be met when the number of samples reaches 96, ±5% can be met when the number of samples reaches 384, and ±3% can be met when the number of samples reaches 1 067 [29]. The specific sample size is determined according to the required accuracy requirements and costs.

3) Survey content

Various attribute characteristics (including travel purpose, travel time, travel distance, etc.) of the study object using the original transportation mode should be collected, as well as their willingness to shift from the original mode to CB under the above travel conditions. Considering that CB has not been widely popularized, the concept of CB should be objectively introduced before the SP survey to reduce the error caused by insufficient cognitive level.

2.3 Scene segmenting

Based on the attributes related to travel (recorded as "*P*"), all original travel demands can be segmented into several travel scenes (recorded as "*S*"). In this method, one or more attributes, such as travel purpose, travel time period, and travel distance, can be selected to form *P*. Considering that the willingness of travelers to switch modes is usually influenced by multiple factors [30–31], selecting more travel attributes to form *P* may reduce errors in subsequent analyses

Tab.1 Types of travel big data

Mode	Type of travel big data
Transit	IC card data
Shared bike	Shared bike order data Shared bike GPS data
Taxi (Including Online-hilling)	Taxi Order data Taxi GPS data
Shared car	Shared car GPS data

of mode choice. However, overly complex *P* may result in non-representative scene divisions and increase practical difficulties. In addition, CB services still have public service characteristics, and not all user demands can be considered in the early potential demand estimation process.

This paper suggests that *P* be determined based on the types of CB services planned by the city (such as routes serving hub transfer and leisure trip), and scene division should be conducted to achieve targeted analysis. If conditions permit, pre-surveys can also be conducted to analyze the main influencing factors of travelers' willingness to choose CB services and determine *P*.

2.4 Macro and micro data analysis

Macro data analysis is based on the spatial-temporal features of OD obtained from travel big data, using reference attribute features *P* to infer the travel scenario *S* that individual trips belong to. Then, through aggregation and other methods, the demand for travel under the original mode (and the features) is extracted and calculated in the travel scenario *S*.

Micro data analysis is based on the survey of travel attributes of research subjects, using reference attribute features *P* to select and extract samples that meet the travel conditions of the travel scenario *S*. Then, the mode-shift rate of travelers from the original mode to CB is calculated and aggregated under the travel scenario *S*.

2.5 Nested analysis

Nested analysis is used to connect the conclusions of macro and micro data analysis based on the same travel scenario *S*. With the above process, we can explore the differentiated demand for CB under various specific travel scenarios.

3 Case study

3.1 Research object and data collection

Take Xiamen as an example to verify the proposed method framework. Select local taxi order data for macro analysis.

This data includes the city's online taxi and cruise taxi order records from May 31 to June 9, 2019 (5 days for working days and 5 days for holidays). The OD data format is shown in Table 2. The amount of data after cleaning is about 4.6 million.

Conduct sampling survey on local taxi passengers and obtain small-scale sample data for microscopic analysis. The questionnaire was distributed to the target group through the "Questionnaire Star" platform in July 2021. According to the data requirements in the method framework, the design content of the questionnaire is shown in Table 3. Finally, 200 valid samples were recovered.

Tab.2 OD data format of taxi orders

Field name	meaning	Field name	meaning
ID	Unique order ID	ETIME	Order end time
DATE	Order date	ELON	Longitude of travel destination
STIME	Order start time	ELAT	Latitude of travel destination
SLON	Longitude of travel origin	PASS_MILE	Travel distance
SLAT	Latitude of travel origin		

Tab.3 Questionnaire contents

Type	Content	Example
RP	Socio-economic attributes, including age, gender, monthly income, et.	Your gender?
	Perception of CB	How do you know about CB?
	The details of the last taxi trip, including the start and end of the trip, the purpose of the trip, the time period of the trip, et.	Please check the travel record of the last time you used a taxi and answer: The origin of this trip? The destination of this trip? The purpose of this trip? (Commuting, going to and from school, seeing a doctor, entertainment, scenic spot travel, hub transfer, and others); The time period of this trip? (Morning peak on weekdays, off-peak on weekdays, holidays)
-	Introduction of CB	Custom bus is a new bus service that can be booked. You can enter your planned travel time, starting and ending points in advance on the dedicated APP to complete the ticket purchase; According to your travel requirements, the platform will match passengers with similar trajectories, set bus routes and stops, and complete the transfer at the nearest stop to your travel start and end on time. The CB adopts the strict "one person, one seat" system, uses 10 to 20 people medium and small buses, and is equipped with Wi-Fi, USB sockets, air conditioning and other equipment. There are few stations along the line, and the price is between the public transport service and taxi service. It is comfortable, affordable and green.
SP	Based on the recent experience of using taxi to travel, the willingness to shift to CB	If you have CB modes to choose during this trip, would you like to use CB instead?

3.2 Potential travel demand analysis

1) Scene segmenting

Divide taxi travel scenarios. In consideration of the limitation of the number of samples and the difficulty of identifying the OD scenario, “travel purpose” is selected as the attribute feature *P*, and five travel scenarios (namely commuting, hub transfer, leisure trip, traveling to scenic spot and seeking medical treatment) are set up, corresponding to the potential service scenario of CB.

2) Macro data analysis

The travel purpose of individual OD in the taxi order data is inferred, and then the taxi travel demand characteristics under various travel scenarios are obtained. Considering that the existing travel order data often does not directly contain travel purpose information, this paper identifies and verifies the travel purpose of OD based on land use information and the space-time characteristics of OD. The technical route is shown in Figure 2. The main steps are as follows: based on the local POI data, identify the nature of the traffic area land [32] and match it with the aggregated OD data; Select some ODs with strong spatial aggregation (to avoid the impact of a large number of scattered demands on scenario inference), and judge and verify the corresponding travel purpose based on time-varying characteristics and land use characteristics, as shown in Table 4.

3) Micro data analysis

Based on the survey data, the mode-shift rate of taxi passengers to CB is calculated under various travel scenarios. The calculation formula is

$$P_i = \frac{S_i^w}{S_i^t},$$

P_i is the mode-shift rate of taxi passengers to CB under the fourth travel scenario. S_i^t is the total number of respondents whose travel demand meets *i* travel scenario; S_i^w is the number of respondents who choose “willing” or “very willing” to shift from the original mode to CB. The calculation results are shown in Figure 4 (the dotted line represents the average level of the mode-shift rate).

4) Nested analysis

Figure 5 shows the relationship between the mode-shift rate from taxi to CB (horizontal axis) and the original demand for the travel mode (vertical axis) in various travel scenarios. Based on a certain threshold for the two indicators, the plane is divided into four parts, corresponding to different types of potential demand for CB in different travel scenarios:

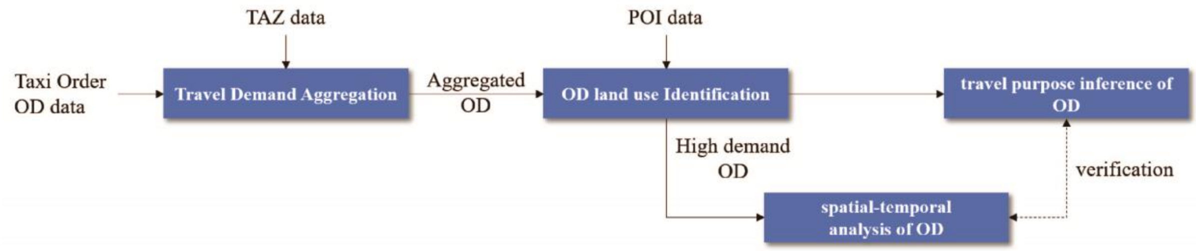

Fig.2 Travel purpose inference of taxi OD data

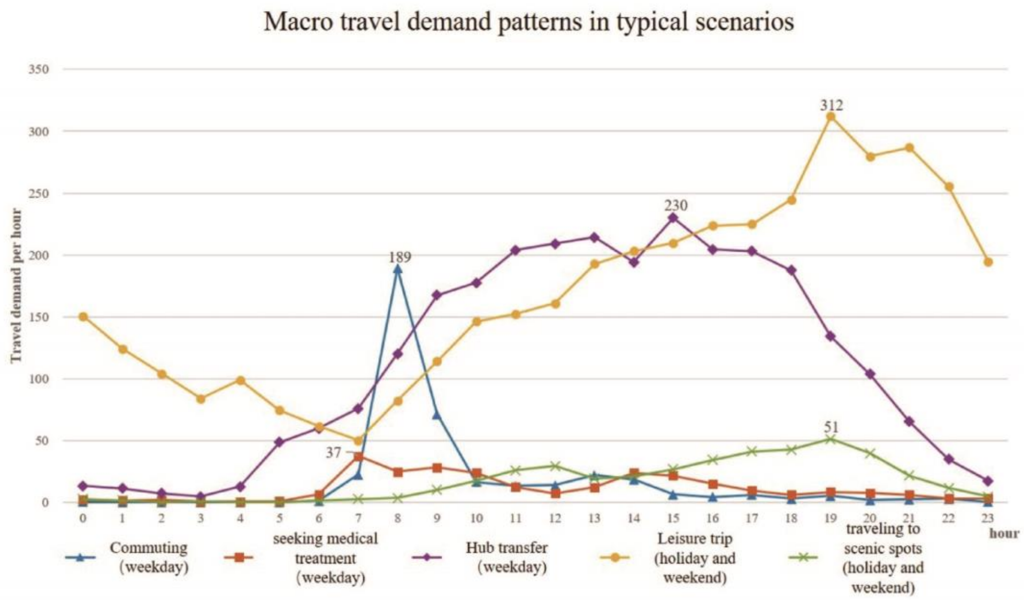

Fig.3 Macro characteristics of taxi travel demand under each travel scenario

Tab.4 Travel purpose inference

Type number	Land use characteristics	Temporal characteristics	Possible travel scenario
1	Land use of O/D is inferred as office zone, public service facilities or commercial zone by geo- information	Morning and evening peak characteristics in weekdays	Commuting
2	Land use of O/D is inferred as transportation hub by geo-information	-	Hub transfer
3	Land use of O/D is inferred as commercial zone or public service facilities by geo-information	Travel demand in holidays and weekends significantly higher than in work days	Leisure trip
4	Land use of O/D is inferred as scenic spots by geo- information	Travel demand in holidays and weekends significantly higher than in weekdays	Traveling to scenic spots
5	Land use of O/D is inferred as medical facilities by geo-information	-	Seeking medical treatment

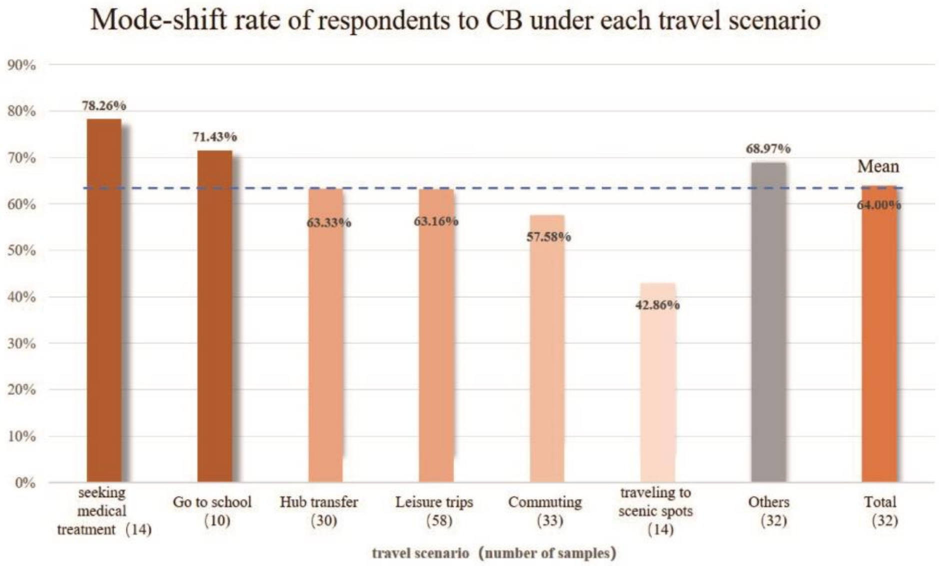

Fig.4 Mode-shift rate of respondents to customized bus under each travel scenario

Quadrant 1 represents scenarios of "high demand and high mode-shift rate". Taxi demand exhibits strong spatial and temporal clustering, and the proportion of passengers shifting to CB is high. Therefore, CB of this type have good operational potential and are suitable for priority promotion.

Quadrant 2 represents scenarios of "high demand and low mode-shift rate". Taxi demand exhibits strong spatial and temporal aggregation, such as commuting during rush hours, hub transfers, and leisure trips at night. However, the proportion of passengers shifting to CB is lower than the average level. Therefore, CB serving the above travel scenarios may not be attractive enough to original taxi passengers. If this type of CB is planned to be promoted, attention should be paid to service attributes such as ticket price and punctuality to enhance passenger flow attraction.

Quadrant 3 represents "low demand and low mode-shift rate" scenarios. Taxi demand does not exhibit strong spatial and temporal aggregation, and the willingness of passengers shifting to CB is also low, such as traveling to scenic spots. In this travel scenario, original taxi passengers are not the favorable service objects of CB.

Quadrant 4 represents "low demand, high mode-shift rate" scenarios. Taxi demand does not exhibit strong aggregation, but the willingness of taxi passengers to shift to CB is high,

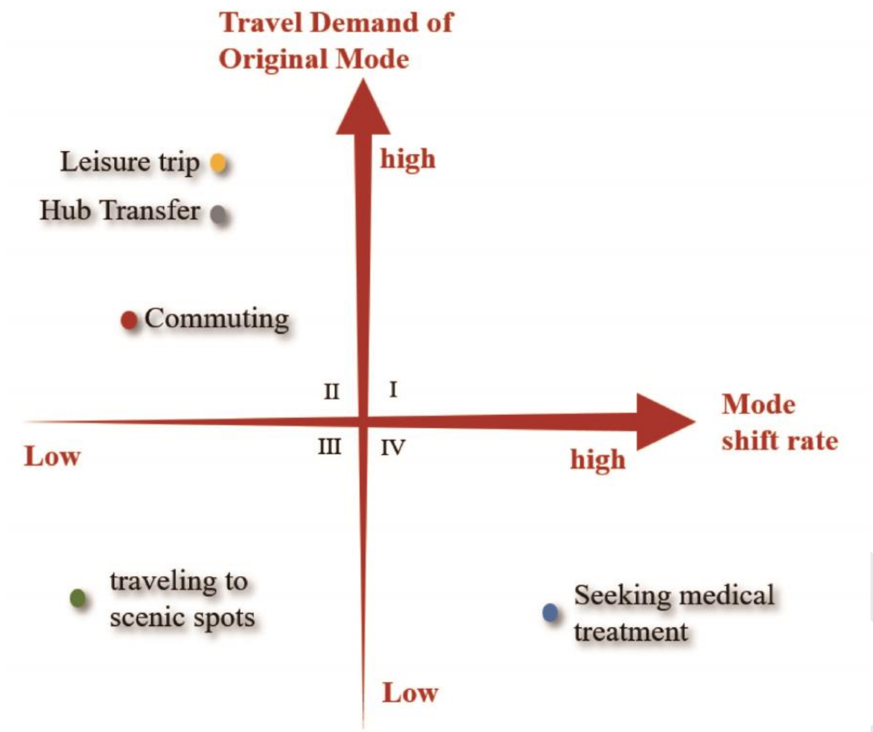

Fig.5 Conclusion of macro and micro data nested analysis

such as seeking medical treatment. Low travel demand may reflect two situations: the absolute number of such trips is small; There are already travel services similar to customized public transport (such as conventional buses), which makes a large number of travelers not choose taxi. If CB serving this type of scenario are planned to be promoted, further research on the current mode of transportation for corresponding travelers should be conducted to avoid unfavorable competition between CB and existing travel modes.

4 Conclusion

This paper proposes a scenario-oriented macro-micro data nested analysis method for potential demand for CB. Based on travel attributes, the travel scenarios are segmented, and an analysis framework guided by specific scenarios is constructed. The advantages of big and small data are utilized to mine the macroscopic features of travel demand, and the willingness of individual passengers is obtained through small-sample surveys. The conclusions of the two aspects are nested through consistent travel scenarios to analyze CB travel demand in different scenarios. Xiamen is selected as an example to verify the operability and data applicability of the method. The results show that various travel scenarios can be classified into four types of “high demand and high mode-shift rate” based on potential demand characteristics, and corresponding differentiated CB market promotion strategies can be formulated.

This article still has certain limitations that need to be addressed. Firstly, the travel scenarios are presented in a qualitative manner without sufficient quantitative analysis. Future research could incorporate quantitative indicators to facilitate more accurate alignment and connection of macro and micro data, leading to a more comprehensive understanding of CB’s operational potential. Secondly, the sample size of the case study is limited by uncontrollable factors, including the ongoing pandemic. Consequently, only taxi passengers were used as a nested analysis example, and the micro-investigation sample size is limited, which restricts the generalizability of the results. Future research could expand the sample size and consider diverse travel groups, such as transit passengers and shared bicycle users, to provide more comprehensive insights into CB’s potential demand.

References

[1] RYLEY T J, STANLEY P A, ENOCH M P, et al. Investigating the contribution of demand responsive transport to a sustainable local public transport system [J]. Research in Transportation Economics, 2014, 48 (12): 364–372.
[2] LIU T, CEDER A A. Analysis of a new public transport-service concept: customized bus in China [J]. Transport Policy, 2015, 39 (4): 63–76.
[3] Beijing Daily Client. How Can Public Transportation in Big Cities Attract Passenger Flow? Customized Buses Have Been Opened in 29 Provinces Including Beijing and Hebei [EB/OL]. (2021-11-12) [2022-1-25]. https://baiji-ahao.baidu.com/s?id=1716225464413604888&wfr=spider&for=pc.
[4] ZHAO X G, ZHANG S J, KONG G Q, et al.Consideration on the value of the online customized bus to urban traffic management: taking the quality public transport service of “Shenzhen E Bus” as an example [J]. Urban Public Transport, 2018 (7): 42–46.
[5] Hangzhou Public Transport Group. Help to Build ‘Digital Intelligence Hang-zhou · Livable Paradise’ [EB/OL]. (2021-02-02) [2022-3-18].

https://www.hzbus.com.cn/hzbus/detail/36/4176
[6] Chen Baobin, Hu Xiaodan, Wang Xiuming, Ke Boxin, Liu Shitang, Huang Guosu. Development of Customized Public Transport in Xiamen and its Role in the Epidemic Situation [C]//Traffic Governance and Space Reconstruction—Proceedings of 2020 China Urban Transportation Planning Annual Conference, 2020: 2148–2154. DOI: 10.26914/c.cnkihy.2020.032282.
[7] LIU F, YANG X Y. Situational prejudgment and policy suggestions on development of new transport formats after COVID-19 epidemic [J]. Transport Research, 2020, 6 (3): 39–47.
[8] CHEN X, WANG Y H, DAI Z, et al. Research on demand identification for customized bus based on multi-source mobility data [J]. Big Data, 2020, 6 (6): 111–124.
[9] GUO R G, GUAN W, HUANG A L, et al. Exploring potential travel demand of customized bus using smartcard data [C]//IEEE.2019 IEEE Intelligent Transportation Systems Conference. New York: IEEE, 2019: 2645–2650.
[10] LYU Y, CHOW C Y, LEE V C S, et al. CBplanner: a bus line planning framework for customized bus systems [J]. Transportation Research Part C: Emerging Technologies, 2019, 101 (4): 233–253.
[11] LYU Y, CHOW C Y, LEE V C S, et al.T2CBS: mining taxi trajectories for customized bus systems [C]//IEEE.2016 IEEE Conference on Computer Communications Workshops. New York: IEEE, 2016: 441–446.
[12] CHEN C, ZHANG D, LI N, et al. CB-planner: planning bidirectional night bus routes using large-scale taxi GPS traces [J]. IEEE Transactions on Intelligent Transportation Systems, 2014, 15 (4): 1451–1465.
[13] YU Q, ZHANG H, LI W, et al. Mobile phone GPS data in urban customized bus: dynamic line design and emission reduction potentials analysis [J]. Journal of Cleaner Production, 2020, 272: 122471.
[14] LI J, LYU Y, MA J, et al. Factor analysis of customized bus attraction to commuters with different travel modes [J]. Sustainability, 2019, 24: 1–13.
[15] CAO Y, WANG J. The key contributing factors of customized shuttle bus in rush hour: a case study in Harbin City [J]. Procedia Engineering, 2016, 137: 478–486.
[16] LI G Y, LIU L. Scene digitization: building a scene driven development model [J]. Tsinghua Business Review, 2021 (6): 87–91.
[17] TAN Z Z, QIU J D, YI S, et al. Scenario driven data architecture design of transport big data platform [J]. Traffic & Transportation, 2020, 33 (s2): 62–66.
[18] ZHANG L L, WANG L. Design of integrated simulation platform for traffic signal control based on scene driven [J]. Journal of Chongqing Jiaotong University (Natural Science), 2020, 39 (1): 1–7.
[19] QIU S L, QIAN L J, LU J H. Research on AVPS based on scenario driven and hierarchical planning [J]. China Mechanical Engineering, 2019, 30 (19): 2364–2371.
[20] WANG G T, SHAN X N, ZHANG H, et al.Digital transformation in urban transportation governance [J]. Urban Transport of China, 2022, 20 (1): 1–9.
[21] FAN X G. Digitization and resolving the dilemma of empirical sociology research methods [J]. China Social Science Review, 2020 (3): 66–75.
[22] LAZER D, HARGITTAI E, FREELON D, et al. Meaningful measures of human society in the twenty-first century [J]. Nature, 2021 (595): 189–196.
[23] TANG W F. Big data and small data: discussion on research methods of social science [J]. Journal of Sun Yat-Sen University (Social Science Edition), 2015, 55 (6): 141–146.
[24] QIN X, ZHEN F. Combination between big data and small data: new methods of urban studies in the information era [J]. Scientia Geographica Sinica, 2017, 37 (3): 321–330.
[25] SUN S C. An approach to classify the public transit users based on macro&micro nested data [J]. Journal of Transportation Systems Engineering and Information Technology, 2018,18 (3): 108–112.
[26] BLUMENSTOCK J, CADAMURO G, ON R. Predicting poverty and wealth from mobile phone metadata [J]. Science, 2015, 350 (6264): 1073–1076.
[27] HUANG W, SUN S C, SUN N. Improved method of population distribution identification based on cellular signaling data [J]. Urban Transport of China, 2021, 19 (1): 95–101.
[28] QIU J, QIN X, ZHEN F. Research on job housing balance based on combination of big and small data: a case study on Changzhou, China [J]. Modern Urban Research, 2020 (6): 56–63.
[29] Creative Research Systems. Sample size calculator [EB/OL]. [2022-07-29]. https://www.surveysystem.com/sscalc.htm.
[30] GAO Y P, GAO H, PENG X C, et al. Research on influencing factors of customized bus user's continuous choice behavior [J].Journal of Wuhan University of Technology (Transportation Science & Engineering), 2021, 45 (4): 639–642.
[31] JIAN Y C, ZHANG J, JIN S. Analysis of influencing factor of residents' willingness to custom bus travel [J]. Journal of Wuhan University of Technology (Transportation Science & Engineering), 2019, 43 (2): 247–252.
[32] HU X M, LI X L, WEI F. A POI data-based study of the urban functional areas of Chongqing and their mix degree recognition [J]. Journal of Southwest University (Natural Science Edition), 2021, 43 (1): 164–173.

