國立成功大學

交通管理科學系

博士論文

以 ARX 模型預測公共自行車系統之運量—比較人口因子及旅次因子之差異性

Using Autoregressive with Exogenous Variable Models for the Ridership Predictions of Bicycle Sharing Systems: Comparison between Population- and Trip-based Data

研究生：劉昭堂

指導教授：胡守任　博士

共同指導教授：魏健宏　博士

中華民國 一百一十一年 七 月

National Cheng Kung University

Department of Transportation & Communication

Management Science

Doctoral Dissertation

Using Autoregressive with Exogenous Variable Models for the Ridership Predictions of Bicycle Sharing Systems: Comparison between Population- and Trip-based Data

by

Chao-Tang Liu

Advisor: Dr. Shou-Ren Hu
Co-Advisor: Dr. Chien-Hung Wei

July 2022

國立成功大學

博士論文

以ARX模型預測公共自行車系統之運量—比較人口因子及旅次因子之差異性

Using Autoregressive with Exogenous Variable Models for the Ridership Predictions of Bicycle Sharing Systems: Comparison between Population- and Trip-based Data

研究生：劉昭堂

本論文業經審查及口試合格特此證明

論文考試委員：[illegible] 鄭永祥
[illegible] 胡大瀛
溫傑華 魏健宏
李威勳 胡守任

指導教授：胡守任 魏健宏

單位主管：鄭永祥

(單位主管是否簽章授權由各院、系(所、學位學程)自訂)

中 華 民 國 111 年 7 月 21 日

摘要

公共自行車系統的發展與都市地區的實質特徵具高度的相關性，因都會區中的每個區域具有不同社會經濟與人口組成特性，這些特質對於公共自行車使用量及系統規模的增減有不同的影響，因此如何有效評估特定區域內公共自行車系統的營運績效，對於都市地區整體運輸系統的發展相當重要。彙整過往的研究發現，多數研究均以區域內「人口數」作為衡量公共自行車系統績效的重要因子之一，惟本研究認為「旅次」資料更適合用於預測公共自行車系統的使用量，因為在都市運輸規劃中「旅次」數量，方代表區域內人口的潛在活動強度，對於預測公共自行車系統的使用量及績效，更具代表性；另本研究亦從公共自行車歷年使用量及系統規模擴展資料，用以評估區域內之租賃站，於未來應增加或減少，以有效控制經營成本。

本研究彙整2009年至2017年高雄市行政區的相關資料(包含:社經資料、人口統計及旅次起訖資料)及高雄公共自行車系統(CityBike)的運量、租賃站數及車站容量等資料，並以租借站區域(站級資料)及行政區(地區資料)為單位進行資料分類，使用時間序列分析法並考量外生變量下，進行模型建構。研究結果顯示，「旅次」資料相對於人口數對於公共自行車的使用量之預測較為準確，且可判別租借站借還運量與不同旅次目的起訖量(例如：家-工作旅次或家-學校旅次等)的關聯性。另有關系統擴展議題，本研究從「CityBike系統單一租賃站1平方公里內租賃站容量」該因子之實證研究結果發現，並非所有的區域會隨著 CityBike 系統站數的擴增，而增加使用量，部份區域呈現負向的成長。本研究相關成果，可以提供公共自行車系統的營運者，準確評估既有系統的使用量，並在有限的預算之下，以整體系統最大化效益進行設備投資與增減租賃站之評估，期能實現公共自行車系統永續經營之目標。

關鍵字：公共自行車、社會經濟特性、旅次特性、外生輸入自我迴歸模式

ABSTRACT

The successful development of bicycle-sharing systems (BSS) has been influenced by the socioeconomic characteristics and geographical attributes of the metropolitan areas they are set up in. Trip generation and attraction volumes, which represent the true flow of resident activity within a specific area, may influence BSS ridership, particularly for people using a BSS for first- or last-mile services. However, most studies have thus far used population and other socioeconomic data to investigate BSS ridership and have not considered trip attributes. Although population-related attributes may influence BSS ridership, they cannot account for the spatial distributions of vehicular or passenger trips between specific origins and destinations. In addition, this study also explored BSS ridership data and system scale over the years to evaluate whether BSS rental stations in a specific area should be increased or decreased for cost control.

In contrast with previous studies, this study collected nine years of BSS ridership data, number of rental stations, station capacity and related socioeconomic characteristics regarding the CityBike system in Kaohsiung City, Taiwan, including users' trip attributes. The autoregressive with an exogenous variable (ARX) model was used to analyze the factors influencing the ridership of the CityBike system. The results indicated that trip attributes are more relevant than population data in predicting BSS ridership and determining correlations between renting and returning in the CityBike system and the original destination depending on the trip's purpose (such as journeys from home to work or school).

Concerning the system expansion issue, by investigating the effect of the "capacity of CityBike station in 1 km buffer" variable, this study found that some districts in urban areas are oversupplied with stations during specific periods, thereby decreasing the system's overall efficiency. Accurate predictions of BSS ridership over time can enable the allocation of limited resources to establish new stations or improve infrastructure for future sustainable BSS development.

Keywords: Bicycle-sharing system, Socioeconomic characteristic, Trip attribute, Autoregressive with exogenous variable model

致謝

博士學位，是我個人自我理想的實現，終於在今年7月順利通過審查，取得學位。在職期間接連攻讀碩、博士學位，箇中滋味如人飲水，冷暖自知，遙想甫取得博士入學資格時的興奮，彷彿昨日，當踏進成功大學的校門，期許自己在交通運輸領域，要有所專精與突破。

感謝在求學這一漫長路上，遇到的多位貴人。首先，感謝我的指導教授-胡守任教授，我們自淡江大學運管系「運輸工程」這門課相識，後續到成大就讀碩、博士時仍有幸成為老師的學生，如您常說，這是緣分。記得當時老師在運工課程的細心教學及對作業研究的嚴謹態度，後續到成大仍採用一致的標準，用心指導我。研究期間，因為自身專業知能上的不足，導致在論文審查及期刊投稿上跌跌撞撞了幾回，過程中老師始終鼓勵著我，耐心提點，並要求我恪守學術研究上的嚴謹態度，更鼓勵我走向國際，參與國際性交通運輸研討會，拓展視野，引領我渡過重重難關，心中的感謝無以言表，期許自己未來在運輸領域的表現及研究上的成果，能與老師看齊，對社會做出貢獻。

感謝魏健宏教授，擔任我的共同指導教授，在我論文最關鍵的時刻，給予我中肯的修正建議，對於相關行政程序上也多方提醒，讓我在最後口試階段，順利完成各項作業。感謝胡大瀛教授、鄭永祥教授及李威勳教授，在外審前的審查，皆細心並明確點出我論文缺漏及不足之處，並提供可行方向，讓論文內容更加紮實。感謝邱裕鈞教授、溫傑華教授及鍾智林教授播冗出席外審會議，花費諸多時間與我問答確認內容章節，並提出學術與實務上不同面向之看法，讓論文更加完善。感謝劉瀚聰教授，是我成大碩、博班的學長，自我就讀以來，無論任何時間，都不厭其煩解答我的疑問，並適時提供的支援，是我最佳的戰友。感謝系辦李淑秋助教、吳雅婷組員、林鳳靖助理及莊曜聰先生於論文口試期間，各項行政事務的協助與處理。

感謝我的父母及妹妹，給我生活上全力的支援，讓我無後顧之憂。特別感謝我的太太，靜琳，在我最辛苦的撞牆期，全心全意地陪伴著我、包容我的情緒、體諒我的辛勞，在精神上與生活中給予最大支持，並一直默默地守候在我身邊，不棄不離，共同攜手走到我人生重要的里程碑，愛妳。

再次感謝曾經幫助過我的每一位，也感謝高雄捷運股份有限公司各級長官對於員工進修的支持。最後，期許畢業後的我，秉持研究初衷，繼續向前。

劉昭堂　謹誌于成功大學交通管理學系 2022 年 7 月

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1 INTRODUCTION

1.1 Motivation

Rapid urbanization and economic development have led to significant population growth and vehicle ownership in most metropolitan areas worldwide. However, our dependence on motorized transport has resulted in severe traffic congestion, energy consumption, and air pollution. Bicycle-sharing systems (BSSs) are an effective way of mitigating the negative effects of high motor vehicle usage because they increase accessibility to Mass Rapid Transit (MRT) stations and bus stops for households. A BSS is a crucial part of a well-designed urban public transportation network. However, in order for BSSs to develop sustainably, the relationships between a BSS's ridership and socioeconomic characteristics should be analyzed in order to determine the optimal placement of future BSS stations.

Among the related socioeconomic characteristics, trip volumes, which represent the true flows of resident activity in a specific area, may influence BSS ridership, particularly for travelers using a BSS for first- or last-mile services. However, previous studies have relied on population and other socioeconomic data to investigate BSS ridership without considering trip attributes. Because population data is easily accessible, the statistical data can be obtained from the local government by calculating people's birth rate and mortality rate and it was not necessary to develop a mathematical model to obtain it. Normally, trip data in an area is obtained in one of two ways: 1) census, or 2) transportation demand forecasting models. Collecting census data requires a significant investment of time and labor to conduct questionnaires or home visits before the researcher analyzes those data for each area and divides them into different catalogs for trip purposes. Using transportation demand forecasting models means that it is necessary to aggregate the relevant influencing variables and use different models for estimation purposes (i.e., regression models and gravity models). Therefore, previous studies considered the completeness and convenience of the data, and the population factor was usually adopted for BSS ridership forecasting. The number of trips per person per day differs depending on the population of an area because people make trips for different purposes. People typically make two trips per day; that is, they leave and return home. However, they may make more trips for other purposes, such as attending classes, going out to eat, or exercising. Population-related attributes may influence BSS ridership, but they cannot account for the spatial distributions of vehicular or passenger trips between specific origin–destination pairs. Trip attributes and socioeconomic characteristics are the basic data used for predicting the total figures of BSS ridership and local residents' "true" behaviors. In addition, many cities around the world are establishing new BSSs or expanding their existing bicycle-sharing networks. However, Médard de Chardon et al. (2017) uncovered no direct relationship between ridership and the number of BSS stations. In other words, BSS ridership does not necessarily increase with the number of installed rental stations.

Therefore, this research is unique, to the best of our knowledge, in that it includes station and district-level data, including spatial variables (i.e., station capacity, capacity in the buffer range, and POI), socioeconomic variables (i.e., demographics), and infrastructure variables (i.e., bike lane lengths, number of lamps, road lengths), and the trip data is treated as an independent variable instead of the population factor. In particular, trip attributes (i.e., trip purpose, trip assignment, number of trip generations, and number of trip attractions) are considered the independent variables in the developed models. District-level data provide a comprehensive picture of demographic and economic characteristics, which can be combined with station-level data to generate information that can be used in modeling ridership evolution or expanding an existing BSS.

1.2 Research Scope

Previous studies have identified key factors (e.g., demographic and socioeconomic variables and spatial-temporal factors) that have different degrees of influence on the ridership of a BSS. However, some studies found that the population factor had no effect on the BSS ridership or did not consider it to be the main influence factor (Bachand-Marleau et al. 2012; Gebhart and Noland 2014; Scott and Ciuro 2019). However, these studies did not incorporate trip attributes into their models when estimating BSS ridership rates, and most of them used short-term (one-month to two-year datasets) or small-scale datasets to verify the capability of developed models in identifying key factors.

For this research, trip volume at the starting point is defined as trip generation, and passenger volume at the endpoint is defined as a trip attraction. Based on the 2009 census data, we collected trip data from each district of Kaohsiung City and classified them into student trip and worker trip demand. These trip attributes were considered to be key variables influencing BSS arrivals and departures in this research. To fill the gap in the literature, this research investigated the effect of trip attributes on BSS ridership in areas. In contrast to previous studies, this research aimed to explore trip attributes and socioeconomic variables that influence CityBike ridership in different district types by accounting for both station-level and district-level data collected from 2009 to 2017 in the developed models. During the modeling process, this research used trip factor instead of population factor and compared the impact and accuracy of BSS ridership forecasting.

The expansion of the BSS network is the other key issue of this research. Although previous studies have found that adding new stations in areas with a high density of bike stations can increase BSS ridership (Rixey 2013; Wang and Lindsey 2019), this research aims to determine whether there is a limit to the number of stations before BSS use levels off and causes an imbalance impact on supply and demand in those areas. Therefore, this research also focuses on the factors of BSS network expansion (e.g. station capacity and station capacity within a 1-km buffer) and conducts a cross-sectional analysis using a long-term statistical dataset.

The traditional cross-sectional regression model only forecasts or micro-analyzes BSS ridership over a specific short-term period or area (station-based). In order to evaluate whether or not people use the BSS system, or to predict BSS ridership, the logit model, negative binomial regression model, and time-series models (i.e., fixed-effect model and random-effect model) are most often adopted. However, these studies did not consider overall environmental (district-based) impacts on BSS ridership. This research collected data from 2009 to

2017 that had trend characteristics but did not have seasonality or cyclicity. Time-series methods, such as ARMA or ARIMA models, are not appropriate to use in this case, but the auto regressive (AR) model could be adopted. The AR model only uses the historical data effect on the current status but does not consider the effect of other exogenous variables (i.e., demographic, trip attributes, and spatial-temporal factors), which may cause parameter estimate bias. In order to reduce this deviation and raise the estimating accuracy of exogenous variables on BSS ridership, we combined traditional hierarchical regression analysis and considered the impact of the time-series, using the ARX model to predict BSS ridership and improve forecasting accuracy.

In the planning stage, the government needs to maximize the BSS system benefits within its limited budget in order to ensure the system's sustainable development. Therefore, the number of BSS stations that need to be constructed will be key. In this study, the regression model was adopted to analyze the effects between the key influencing factors and either BSS ridership or location in the planning stage. In order to achieve sustainable BSS development, BSS ridership is the main area that must be analyzed. The researchers need to find out the influencing factors and adjust operating strategies accordingly. In order to understand the impact of key factors on BSS usage and to predict BSS ridership, the time-series methods were adopted in the operational stage. In order to understand the impact of each data cluster on a BSS's ridership, the relative data of the CityBike system were collected monthly from all districts in Kaohsiung City, Taiwan between 2009 and 2017, including data on residents' trip attributes. The CityBike system of Kaohsiung City in southern Taiwan was chosen as an empirical study, and policy implications can be addressed based on the results. This research explores the potential effects of trip attributes in addition to other socioeconomic characteristics, on CityBike's ridership.

1.3 Objective

This research is concerned with the CityBike ridership prediction and the system's performance problems. The main objectives of the research are as follows.

(1) *To explore trip attributes and key socioeconomic variables that influence CityBike journeys to and from a station in different districts by accounting for both station-level and district-level data.*

(2) *To develop an integrated model based on regression-based modeling (hierarchical regression model, HRM, and autoregressive with exogenous variables model, ARX) that can solve the CityBike ridership prediction and system performance problems over the long term.*

(3) *To evaluate the performance of CityBike in each district and to determine whether or not to expand the scale of CityBike in the target area.*

The parameters related to the number of CityBike stations, such as "capacity of station" and "capacity of CityBike station in 1 km buffer," are analyzed and explored to provide political strategies for BSS operators or the government.

1.4 Research Methodology

This research consists of a CityBike ridership prediction and system performance evaluation with masses of data. Therefore, a linear regression model would be the most suitable methodology for modeling a variety of relationships between a dependent variable (DV) and multiple independent variables (IV) (Rawlings et al, 2001, Washington et al., 2003). The main objective of linear regression is to identify the relationship between a DV and one or more IVs.

In previous studies, many variables have been used to predict BSS ridership, with population one of the most used factors for usage forecasting. However, this research found that these impact factors have not been classified or ranked in order of importance in previous studies. In order to understand previous study theoretical assumptions and to prioritize the influence of several predictor variables in sequence, we wanted to determine the order of importance of influencing factors (i.e., population, trip attributes, and other influencing factors) in BSS ridership. This research used hierarchical regression to test such specific, theory-based hypotheses (Cohen 2003) and as a benchmark to compare the other developed model. In hierarchical regression, the main step focused on the variation in predictability associated with predictor variables entered later in the estimate process over and above that contributed by predictor variables entered earlier in the process. Cohen (2003) noted when there are several influence factors (predictors), those factors can specify the order in which they are to enter the equation. This ordering creates a hierarchy. When the ordering of variables is based on a theoretical model, the procedure is usually known as hierarchical regression. Specifically, hierarchical regression refers to the process of adding or removing predictor variables from the regression model in a series of steps. This research adopted this theory to formulate our developed model for predicting and considering the ordering of IVs based on past studies. It must be noted that the hierarchical regression model used in this research is not the hierarchal linear model (HLM), which is also referred to as "multi-level modeling;" this method is one of the families of analyses known as mixed-effect modeling or mixed models and is mostly used when the data have a nested structure.

In this research, ridership data of the CityBike are collected as monthly arrivals and departures for each station in all districts between 2009 and 2017. The collected data can be considered longitudinal as they are both quantitative and continuous. On the other hand, the collected dataset is essentially time-series data in nature and is also a panel dataset. The AR and the ARX models are generally used as regressive models for ridership prediction. Since this research considers different data types and data characteristics, different regression-based methods can be applied to compare and improve the prediction accuracy of the CityBike ridership data and to identify the key factors influencing system expansion.

1.5 Research Content with Research Flow Chart

A flowchart of this research is shown in Figure 1-1. There are five chapters in this research, and the main contents are as follows.

1. Chapter 1 introduces the motivation, scope, and objectives of the research.
2. Chapter 2 presents background information on BSS development and estimation issues and reviews the relevant literature.
3. Chapter 3 introduces the applied methods, the problem statement, model assumptions, and the model framework.

4. Chapter 4 outlines the background information on CityBike, as well as data collection, and model calibration, and presents the results of an empirical study conducted under real-world conditions.
5. Chapter 5 presents the conclusions, limitations, and suggestions of the research.

The dissertation is organized as shown in Figure 1-1 and is described as follows. The introduction includes the research motivation, research scope, and the research objectives. Based on the research purpose, related literature on BSS development, influencing variables and socioeconomic characteristics, ridership estimation, and prediction methods are reviewed. Based on the introduction and literature review, two regression mathematical models are then adopted to forecast ridership and find the key variables that will decide whether the CityBike system should increase the number of stations or not. The results of the empirical study are also sequentially revealed in different district categories. Finally, the conclusions and suggestions are outlined in the final chapter.

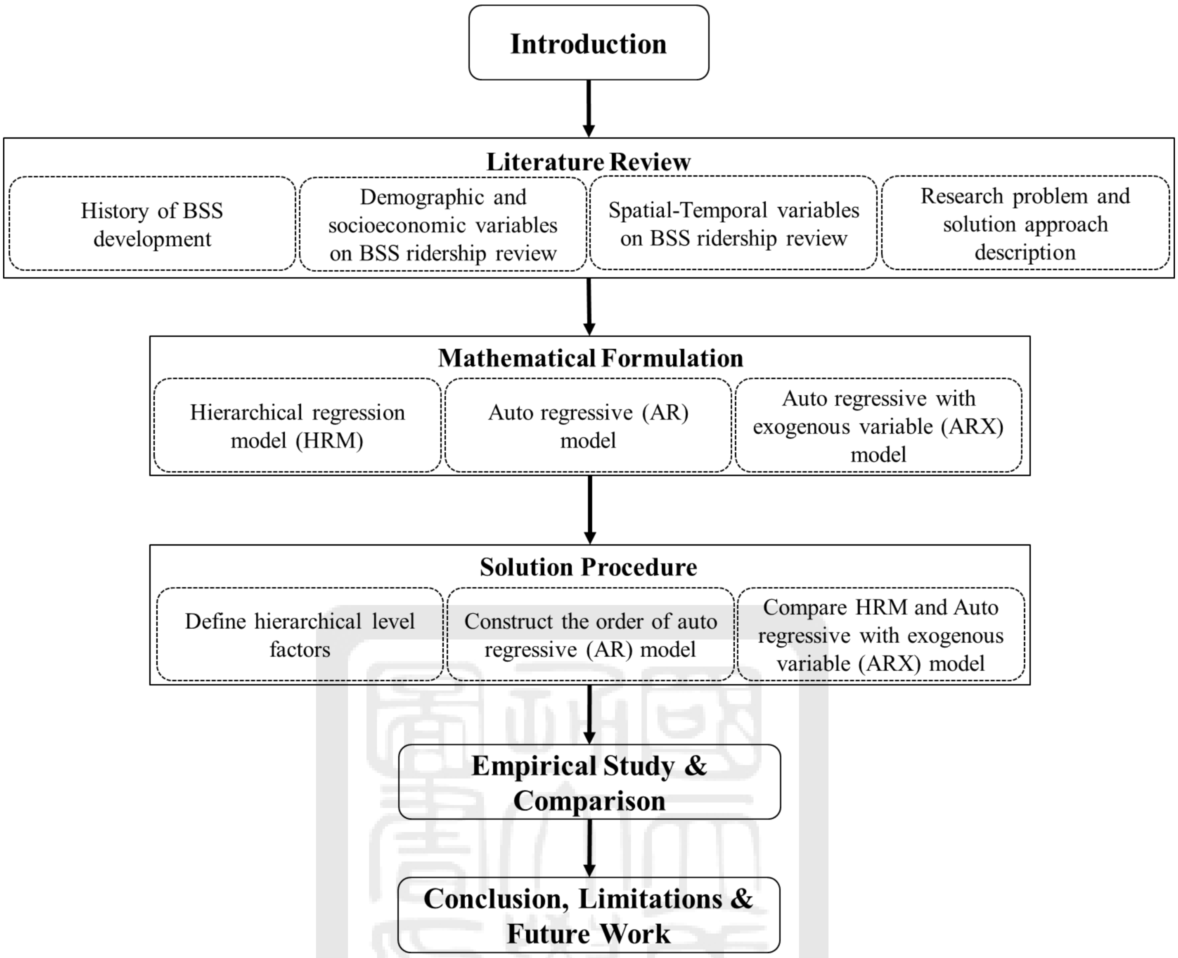


Figure 1-1. Flow chart of the dissertation

Chapter 2 LITERATURE REVIEW

This chapter introduces the relevant literature related to BSS development and the effects of key spatial-temporal factors on usage forecast. The modeling approaches proposed in related studies are discussed via a literature review, which covers both domestic and overseas studies. The chapter is further divided into three sub-sections. Section 2.1 summarizes the development of BSS; Section 2.2 discusses the key factors for BSS ridership forecasts and how spatial-temporal factors can affect BSSs; Section 2.3 summarizes modeling approaches for BSSs Section 2.4 summarizes the literature review.

2.1 BSS History and Development

Developing public transportation systems, such as the metro, bike-sharing systems, and buses aims to reduce private vehicle usage. The researchers used the mode choice analysis method to determine which transport mode would be the best choice. Mode choice analysis is one of four steps in the transportation forecasting model, the other three being trip generation, trip distribution, and route assignment. The modeler can use mode choice analysis to determine what transport modes will be used based on trip distribution results. The mode choice model is formulated by the users' choice of which transport mode to take (e.g., railway, metro, bicycle-sharing system, private vehicles, motor scooters, etc.). The researchers needed to decide which variables are relevant to the decision-making process. Several indicators are usually adopted into the models, such as walk time, travel cost, in-vehicle time, waiting time, and change times between other transport modes. The model considers the input variables about each possible transport mode that the user has available for the journey and gives the proportion of users that would use each transport mode. The implementation of new public transportation policies would change people's mode choices. Their choices and mode shift process reflect the development of existing public transportation systems. In previous studies, questionnaires and mathematical models (i.e., binomial logit model, the mixed logit model, the multinomial logit model, and the structural equation model) were usually adopted to solve this problem. The researchers used the aforementioned methods (Hensher and Reyes, 2000; Cervero, 2002; Lee et al., 2003; O'Fallon et al., 2004; Li et al., 2010; Diana, 2010; Ji et al. 2017; Bai et al., 2020; Holmgren and Ivehammar, 2020) to analyze people mode choices under different transport policies and time period effects. Private bicycles or bike-sharing systems were usually one of the transport modes in the models.

BSSs began to appear during the 1960s, although the literature surrounding BSSs has only grown considerably in the past two decades (DeMaio and Gifford 2004; DeMaio 2009; Fishman et al. 2013; Shaheen et al. 2013; Fishman 2016). Fishman (2016) reviewed research on bike-sharing schemes from North America, Asia, Europe, and Australia in 2013, focusing on the systems' history, growth, citizen journeys, and demographics. After the Second World War, many cities became dependent on private vehicles, which negatively impacted the environment in terms of congestion, air pollution, and safety issues (Handy et al., 2014). Consequently, these impacts inspired a huge interest in cycling options among urban, governmental, and non-governmental operators that combined a payment system, cycling infrastructure, and tracking technology to increase BSS growth.

For example, Faghih-Imani et al. (2017a) examined competition between CitiBike and private cars in New York City based on comparisons of journey times. They developed a panel mixed multinomial logit model in order to identify and understand the factors that affect journey times in order to improve CitiBike's service. Saltykova et al. (2022) examined the possibility of a BSS in Chengdu, China, acting as a substitute for private vehicles and other public transportation modes, which could reduce fuel consumption and mitigate CO_2.

The negative environmental impacts have greatly increased interest in bicycle travel and have led to a significant increase in BSSs across Europe, Asia, and the Americas. For example, the UK's Nation Cycling Strategy (NCS) aimed to increase bicycle usage fourfold from a 1996 baseline figure. The target was included in the Transport White Paper published in Summer 2004 and requested local councils to build more convenient, attractive, and secure environments for people to journey to work or school. Despite differences in the development and purpose of BSSs worldwide, they nonetheless increase public bicycle usage by integrating with other transportation systems to provide convenient and attractive alternative means for local citizens to get around.

The components and concepts of BSSs are simple to understand. A BSS is an alternative transport system provided for local citizens or tourists to make short-range trips in an environmentally friendly manner. People can also use bicycles without the construction costs and custodial responsibility of a BSS, and only need to pay the rental costs. The flexibility of BSSs allows users to pick up and return public bicycles at unmanned BSS stations.

In addition to being beneficial for their users, BSSs also provide social and transportation-related strategies. For instance, BSSs offer a low-carbon solution for making first- or last-mile trips (Nikita, 2018), the first and final miles referring to the short distance between home and the workplace or other transport hub.

Shaheen et al. (2010) noted several benefits to constructing BSSs, including (a) increased mobility options, (b) cost savings from modal shifts, (c) lower implementation and operational costs (i.e., in contrast to shuttle services), (d) reduced traffic congestion, (e) reduced fuel usage, (f) increased use of public transit and alternative modes (i.e., rail, buses, taxis, carsharing, ridesharing), (g) increased health benefits, and (h) greater environmental awareness.

Many researchers have conducted in-depth studies of BSSs from 1965 to 2012 (DeMaio and Gifford 2004; DeMaio 2009; Parkes et al. 2013; Fishman et al. 2013; Shaheen et al. 2013; Fishman 2016) and have shown that bicycle-sharing development took place over four generations. The first generation of bicycle-sharing programs began on July 28, 1965, in Amsterdam, with the initial idea of simply using ordinary bikes, painted white, which people could ride to their destination and then leave for the next user. However, this led to many bikes being stolen. Other cities in Europe that initiated a first-generation bicycle-sharing program included La Rochelle in France, which started its Vélos Jaunes program in 1974, and Cambridge in the UK, which set up its Green Bike scheme in 1993. The Green Bike scheme in Cambridge failed quickly, as nearly 300 of the shared bicycles were stolen. However, the Vélos Jaunes program in La Rochelle had the full support of the local community and became the first successful bicycle-sharing program in France.

In 1994, North America's first BSS was established in Portland, USA. The Yellow Bike system provided 60 bikes without locks for people to use from Pioneer Square in the city center. After the successful launch of the Yellow Bike program, the city of Boulder in Colorado developed the Green Bike Program with 130 bicycles that were provided for free use and were maintained by a group of volunteers consisting of local high school students. However, this program was eventually scrapped because of bike theft.

Because of the theft problems inherent in many first-generation BSSs, the government of Denmark advanced a bicycle service that was different from previous bicycle-sharing programs. This second generation of BSSs began in Denmark in 1995. During this period, bicycle design was improved, advertising plates appeared on solid rubber tires, and a coin-deposit system was set up whereby people could only pick up or return the bikes at specific locations. The Bycyken system, set up in Copenhagen, Denmark, was the world's first large-scale BSS. The system consisted of 1,100 bicycles with bike racks. Users had to deposit 20 Danish kroner to unlock the bike, which would be refunded when the bike was returned. Several other cities in Europe set up similar schemes, including Aarhus, Sandnes in Norway, and Helsinki in Finland. In 1996, the Yellow Bike Project was launched in the twin cities of Minneapolis and St. Paul. The project placed 150 bikes at specific locations around the city and used a coin-deposit system for rental. Over the next five years, other cities developed similar coin-deposit BSSs, which included the Olympia Bike Library in Olympia, Washington (1996); Yellow Bike in Austin, Texas (1997); Freewheels in Princeton, New Jersey (1998), and Decatur Yellow Bikes in Decatur, Georgia (2002) (Shaheen et al. 2010).

Bicycle rental stations with coin-deposit lock systems proved to be reliable, dependable, and resistant to theft. However, this second generation of BSSs did not impose any time limitations on a user, and customer anonymity often meant that people could keep the bikes as long as they wanted for no more than the initial deposit, making it difficult for the operators to track them down. These first- and second-generation bicycle-sharing programs provided people with alternative transport methods, but the programs were not reliable or widespread enough to convince people to switch from cars. The weaknesses inherent in the second-generation bicycle-sharing programs provided the impetus for the development of a third generation, which integrated electronic lockers, bicycle locks, telecommunication systems, smart cards, mobile phones, and onboard computers. In 2005, a BSS in Lyon with 1,500 bikes called Velo'v became the catalyst for a new wave of third-generation BSSs. By the end of 2008, there were over 90 third-generation programs in use around the world. DeMaio (2009) and Fishman (2016) showed that bicycle-sharing programs had a remarkable effect on cycling populations, as well as leading to an increase in public transport use, decreased greenhouse gas emissions, and health benefits. BSSs improve connectivity to other transportation systems, such as railways, bus stops, and MRT stations because the systems provide first- or last-mile services that decreased private vehicle use. According to a survey conducted in Paris, the proportion of residents who did not normally use private vehicles increased from 29% to 49% in the first two years after the Velo'v system began. More than 20% of survey respondents used Velo'v to reach or return from other transportation modes and used them to begin and end their multi-legged journeys.

In a survey of the SmartBike system in Washington, D.C., 16% of participants said they would give up their private vehicles and would use SmartBike for their trips (Shaheen et al. 2010). Regarding greenhouse gas mitigation, Montreal's BSS, BIXI, was estimated to have saved over three million tons of greenhouse gas since the system's inception in May 2009.

In Asia, BSSs began with the third generation. Through the use of information technology (IT), Asia is now a significant growth market for bicycle-sharing programs. The first bicycle-sharing program was the TownBike system, established in Singapore in 1999 and ending in 2007. The second bicycle-sharing program was Taito Bicycle in Taito, Japan, established in 2002. This program consisted of 130 bicycles and 12 rental stations and was funded by the local municipal government. Many large-scale bicycle-sharing programs are currently operating successfully throughout Asia, including CityBike (2009) in Kaohsiung City, Taiwan, YouBike

(2009) in Taipei City, Taiwan, and the Hangzhor Public Bike system (2008) in Hangzhou, China.

The fourth-generation BSSs focused on improving system efficiency, operating sustainability, and usability. These new systems started using cutting-edge IT technology such as mobile systems and GPS to create new business and operating models. Fourth-generation BSSs included Obike in Singapore and T-Bike in Tainan, Taiwan, which improved upon bike distribution, rental station installation, and original destination tracking. Onboard GPS provides BSS bicycle positioning information, and users could inquire about real-time rental station information, hire a bike and pay, all through the mobile app. BSS operators could also provide points of interest and potential customer services by cooperating with businesses located near the BSS stations.

One of the barriers that led to a decline in bike-sharing services was accessibility to docking stations. To overcome this, BSSs can either increase the number of docking stations or simply change to a dockless system. Usually, limited space in a city constrains the number of docking stations that can feasibly be installed. Therefore, dockless BSSs were constructed to expand the system's range.

In 2016, two companies, Ofo and Mobike developed an innovative dockless BSS (called Mobike) in Shanghai, China. By the end of March 2017, the fleet size of dockless bikes had reached 450,000. The new dockless BSS combined mobile payments and GPS tracking, increasing the overall flexibility and efficiency of the system. With the Mobike system, an individual will use a smartphone app to locate a bike and browse the subscription and payment procedures. Once they find a bike, they can unlock and use the bike by scanning its QR code. The onboard GPS then collects large-scale riding trajectory data from the embedded GPS device, which can then allow researchers to analyze and improve the system (Shen et al. 2018; Shaheen and Cohen 2019; Chen et al. 2020). Table 2-1 shows the characteristics of the different BSS generations.

Table 2-1. Characteristics for different BSS generation

Generation Step	City, Years, System	Components	System Characteristics
First generation	Amsterdam, 1965, White Bikes; La Rochelle, 1974, Vélos Jaunes; Cambridge, 1993, Green-Bike; Portland, 1994, Yellow Bike; Colorado, 1994, Green Bike;	Bicycles	1. Distinct bicycle (by color) 2. Unlocked bicycle 3. Without docking stations 4. Free of charge
Second generation	Copenhagen, 1991~1995, Bycyken; Minneapolis and St. Paul, 1996, Yellow Bike Projects; Olympia, 1996, Olympia Bike; Austin, 1997, Yellow Bike; Princeton, 1998, Freewheels; Decatur, 2002, Decatur Yellow Bikes	1. Bicycles 2. Dock stations	1. Distinct bicycle (by color or special design) 2. Bicycle with lock 3. Specific docking stations 4. Coin-deposit system, free of charge
Third generation	Singapore, 1999, TownBike; Taito, 2002, Taito Bicycle Lyon, 2005, Velo'v system	1. Bicycles 2. Dock stations 3. Kiosk 4. User interface	1. Distinct bicycle (by color, special design and advertisement on body) 2. Bicycle with lock 3. Specific docking stations with kiosk 4. Using smart card for renting 5. Free in first 30 minutes or longer
Fourth generation	Montreal, 2005, BiXi; Hangzhor, 2008, Hangzhor Public Bike system;	1. Bicycles 2. Dock stations or dockless	1. Distinct bicycle (by color, special design, and advertisements on body) 2. Bicycle with lock 3. Specific docking stations with 4 G

	Kaohsiung City, 2009, CityBike system Taipei City, 2009, YouBike Shanghai, 2016, Mobike (dockless)	3. Kiosk 4. User interface 5. Electronic services system	technology 4. Using smart cards for renting 5. Real-time information and GPS technology 6. Mobile system 7. Optimal dispatch

Over the past decade, bike-sharing programs have increased significantly around the world. As of July 2022, 1,880 cities have BSSs in operation, with up to 8.96 million bikes in use, but also up to 1413 cities had closed local BSS. (Meddin and DeMaio 2022).

"Convenience" and "docking stations close to work" were the two main factors that motivated people to become bike-sharing program members, and the most common trip purpose was commuting to/from work (Fishman et al. 2013; Fishman 2016). Many city government agencies, including those in Washington D.C., London, and North America, conducted surveys among their local BSS members and non-members. More than 50% of survey respondents indicated that convenience, speed, and ease of use were all motivating factors in their BSS use. However, there are still several unrevealed issues regarding bike-sharing programs that need to be investigated further, especially key influencing factors on BSS ridership, system performance, and integration between bike-sharing programs and other public transport models.

2.2 Key Influencing Factors on BSSs Usage

People are interested in using BSSs as an alternative first- and last-mile mode of transport. BSSs offer multiple benefits such as convenience, good accessibility, good health, and connectivity to other transportation networks. For example, Cheng and Lin (2017) used a questionnaire and the mixed logit model to examine how metro stations can expand their service coverage for passengers by implementing a public bicycle-sharing system (PBSS) in the vicinity. As BSSs continue to be set up in cities around the world, studies are increasingly looking at what key factors would influence system usage and performance. The basic approach for analyzing BSSs ridership contains the following issues: BSS infrastructure (i.e., number and capacity of BSS stations), transportation network infrastructure (i.e., bike lane and major road length), land use, demographic and socioeconomic factors (i.e., population, household), environment (number of restaurants, hotels, places of interest, schools, and businesses) and other transportation stops. Relevant studies of those key factors are described below.

2.2.1 Demographic and Socioeconomic Variables

Population is usually one of the most important factors affecting local citizens' activities (Ewing and Cervero 2001). Several studies (Cervero and Radisch 1996; Handy, et al. 2002; Saelens et al. 2003; Dill and Voros, 2007; Lindsey et al. 2007; Parkin et al. 2008; Hankey et al. 2012) have demonstrated strong correlations between sociodemographic characteristics and non-motorized traffic volumes. According to the findings of these studies, follow-up research on BSSs considered population factors as one of the main influencing factors on BSSs ridership (Hampshire and Marla 2012; Cui et al. 2014; Faghih-Imani et al. 2014; Faghih-Imani and Eluru 2016a, 2016b; Noland et al. 2016; Wang et al. 2016; Faghih-Imani et al. 2017a; Faghih-Imani et al. 2017b; Wang and Lindsey 2019; Guidon et al. 2020; Morton et al. 2021). In previous studies, population density, age, gender, ethnicity, and household density were also potential indicators in evaluating the impact of BSS ridership (Cervero 1996; Dill and Voros, 2007; Bachand-Marleau et al. 2012; Hampshire and Marla 2012; Rixey 2013; Cui et al. 2014; Faghih-Imani et al. 2014; Médard de Chardon and Caruso, 2015; Faghih-Imani and Eluru 2016a, 2016b; Noland et al. 2016; Wang et al. 2016; Wang and Lindsey 2019). Household income or personal net income and members of BSS variables are also influencing factors in BSS research. Bachand-Marleau et al. (2012) and Riexy (2013) identified that middle-income users and annual BSS members positively influenced BSS ridership.

Although population is one of the main factors affecting the rate of public transport, it is not the most influential factor. Scott and Ciuro (2019) studied the interactions between BSSs ridership and population, weather, and temporal factors in Hamilton, Ontario. Weather and temporal variables were found to be a significant influence on ridership, but the effect of population was insignificant. The authors mentioned that students made up a large proportion of bicycle-sharing users but had not been included in population variables (possibly because students' accommodation may not have been captured by the local census). Previous studies usually used population or population density as an indicator for evaluating BSS ridership, while other studies would use working population or job density (Parkin et al. 2008; Hampshire and Marla 2012; Cui et al. 2014; Faghih-Imani et al. 2014; Faghih-Imani and Eluru 2016a, 2016b; Noland et al. 2016) or the number of schools as indicators (Faghih-Imani et al. 2014; Faghih-Imani and Eluru 2016a, 2016b; Faghih-Imani et al. 2017a, 2017b; Wang et al. 2016). For example, the working population and school population account for 68% and 15% of the total population of Kaohsiung City, respectively. These groups are the ideal users of BSS first- or last-mile services. Overall, previous studies confirmed population and employment as the key variables influencing BSS ridership. The impacts of population and employment are positive, but those variables' effects may differ from trip variables' effects. Previous studies did not comprehensively evaluate the influences of working and school trips on BSS ridership. Therefore, this study attempts to obtain different results and strategies by evaluating these factors.

2.2.2 Spatial Variables

Spatial variables such as BSS-related infrastructure, land use, POI in a buffer range, and built-up environment, have a significant influence on BSS ridership and other non-automotive transportation forms. Because of the different degrees of industrial, commercial, tourism, and public transport development (i.e., bus, metro) in a district, the number of factories, companies, hotels, and parks in such locations will influence the number of trips made to these areas (Boarnet and Crane 2001; Parkin et al. 2008; Krizek et al. 2009; Bachand-Marleau et al. 2012; Fishman et al. 2012; Cui et al. 2014; dell'Olio et al. 2014; Faghih-Imani et al. 2014; Marqués et al. 2015; Mateo-Babiano et al. 2016; Noland et al. 2016; Rowangould and Tayarani 2016; Faghih-Imani and Eluru 2016a, 2016b; González et al. 2016; Noland et al. 2016; Wang et al. 2016; Faghih-Imani et al., 2017a, 2017b; Wang et al. 2018; Majumdar and Mitra 2018; Guidon et al. 2020; Morton et al. 2021).

Boarnet and Crane (2001) pointed out the extremely complex relationships between urban form, land use, and travel behavior. They developed an ordered probit regression model which considered the effect of land use on non-work-related automobile trips in Orange Country Los Angeles, and San Diego. The empirical results of land use on travel behavior were shown to be very sensitive. Parkin et al. (2008) constructed an aggregate regression model using data from the 2001 UK Census to find out which influence factors (sorted into either socioeconomic, physical, or transport system variables) would affect the proportion of bicycle journals to work. The results of the logistic model revealed several key influential factors. Regarding socioeconomic variables, a higher proportion of males or white residents had a positive impact that cycle for journey to work raise, but a higher number of cars per employee and lower incomes cause the opposite result.

Regarding physical and transport system variables, increased commuting distances, high traffic volumes, poorly maintained highways, hilly terrain, and rainy days all negatively impacted the proportion of people who cycled to work. An aggregate forecasting regression model showed significant trends in car ownership and bicycle-related infrastructure in a proportion of cyclists. Krizek et al. (2009) used census data from the Minneapolis-St. Paul area from 1990 to 2000. The authors agreed to satisfy the three conditions proposed by John Start Mill in 1843 to seek the true causality between user behavior and bicycle facilities. The first condition is concomitant variation—the extent to which a cause (X) leads to an effect (Y), the second is causing an event where the effect must occur in sequence, and the third is that the investigated variables should be the only possible causal explanation. In the first, the authors identified and described the facilities around the research area, including on-street cycle lanes, off-street bicycle paths, and other related facilities. Secondly, they adopted buffering techniques and traffic analysis zones (TAZs) to measure the effect between the research area and the traffic purposes. In this study, the empirical study area located within 1.6 km of a facility is Buffer 1, and the area extending to 2.4 km is Buffer 2. The results showed the target areas in the facility buffer zones had a statistically significant increase in shared bicycle traffic, particularly when new facilities are constructed in areas that show an increase in bicycle-sharing. Cui et al. (2014) pointed out that existing bicycle infrastructures were still not developed enough for cyclists, particularly safe riding spaces. The authors also noted that most previous studies did not account for the effects of different regions (i.e., urban, suburban, and rural) on bicycle usage. Therefore, the authors used the Household Travel Survey from the Baltimore-Washington D.C. region in the US to develop a series of models to verify what key factors would influence bicycle usage in different regions. A spatial lag model (SLM) approach was adopted in the models to explore land use, built-up environment, demographics, socioeconomics, and traffic conditions in relation to bicycle ridership in 1,151 Statewide Modeling Zones (SMZs). The results showed that bicycle ridership in the SMZs increased with a higher number of households, population, household workers, zero-worker households, transit accessibility, school enrollment, industrial employment, retail employment, and other forms of employment. By contrast, bicycle ridership decreased with higher densities of single drivers, average congestion speed, and average freeway mileage. Bicycle facilities are therefore a key factor that influences citizens' aspirations for bicycle commuting.

Bachand-Marleau et al. (2012) mention that although shared bicycle systems are growing in number, little is known about why people use them. Therefore, the authors used a detailed online survey containing data from BIXI (Montreal's shared bicycle system) which was conducted in Montreal, Canada, in the summer of 2010 2010. The survey data included demographic data, travel behavior, and spatial information. The authors adopted a binary logistic model to determine the proximity of homes to docking stations and bicycle theft prevention that influence the BIXI system ridership.

Rixey (2013) investigated the effects of demographic and urban environmental factors near bike-sharing stations on the bike-sharing ridership of three BSS systems (Capital Bikeshare, Nice Ride, and B-Cycle). The results confirmed that population density, retail job density, cyclists, pedestrians, transit commuters, median income, education, and the presence of bikeways all had a positive effect on the three BSSs. Non-white population and rainy days had a negative effect.

Faghih-Imani et al. (2014) also revealed that people were more likely to use the BIXI system in good weather. The number of restaurants, businesses, universities, and stations, as well as station capacity, bicycle facilities, length of minor roads, population density, and job density all had a positive relationship with the arrivals and departures of the BIXI system over a specific period. Inversely, bicycle use decreased over the weekend and with increasing distance from the central business district (CBD).

El-Assi et al. (2015) conducted a comprehensive spatial analysis to determine the influences of sociodemographic attributes, land use, and the built-up environment on the ridership of Bike Share Toronto. The results of the empirical study revealed the significant influence of road network configuration (i.e., intersection density and spatial dispersion of stations), bike infrastructure (i.e., bike lanes, paths), and land use on the demands placed on the system.

Faghih-Imani and Eluru (2016a) mentioned the unobserved factors of BSS estimating models that influenced the DV (BSS ridership) and which also significantly impacted the IV (BSS infrastructure). The authors considered that this situation violated a basic assumption of econometric modeling (i.e., the error component of the model is not correlated with any exogenous variables). Therefore, they posed a multi-level econometric framework that combined with a measurement equation for BSS ridership prediction and usage equation to reduce estimation bias. The model adopted a repeated observation-based panel multi-level mixed ordered logit model and used data compiled from BIXI between April and August 2012. The proposed econometric models—three simple ordered logit models (3OL), two DV panel mixed ordered logit models (2PMOL), two DV panel mixed multi-level ordered logit models (2PMMOL), and three-dimensional panel multi-level mixed ordered logit models (3POMMOL)—were used, with the results of the model fit measures providing strong evidence to support their assumptions that ignoring unobserved factors during the installation process would affect model estimates. The results of the empirical study revealed an overestimation of BSS infrastructure impacts in models that neglect the installation process.

Faghih-Imani and Eluru (2016b) used spatial lag models to examine New York City’s bicycle-sharing (CitiBike) usage with urban environmental variables and determined that the length of bicycle routes, number of subway stations, areas of parkland near CitiBike stations have a significant positive impact on daily customer numbers but not annual customer numbers.

Wang et al. (2016) collected 13 independent variables and separated them into four categories (sociodemographic, built environment, transportation infrastructure, and economic activity) to identify any correlations with the Nice Ride BSS in Minnesota. Eight of the top ten most frequented Nice Ride stations are located near the Minneapolis CBD, a major retail hub area, and campus. In order to determine the main influencing factors, they formulated log-linear OLS regression and negative binomial regression models to determine any correlations between the usage of the Nice Ride system's stations and the explanatory variables. The results of the development models with a high goodness of fit revealed positive correlations in the two models in terms of the percentage of white residents, station location, paved trail in the station area, public transport stops, and economic activities as density factors. In a case study regarding the failure of the BSS in Pronto, Seattle, Sun et al. (2018) found users tend to use the BSS more at stations that have more bus stops nearby, but users often shifted to using bus services on rainy days or during peak hours.

Faghih-Imani et al. (2017b) collected arrival and departure data from the station level of two cities (Barcelona and Seville, Spain) and developed a mixed linear model to examine the influence of bicycle infrastructure, sociodemographic characteristics, and land use characteristics on bicycle-sharing arrivals and departures. Their results revealed that bicycle infrastructure variables, station density variables, and capacity per unit area variables have a strong positive influence on bicycle-sharing usage. Mattson and Godavarthy (2017) focused their research on smaller and more successful BSSs. The Great Rides Bike Share in Fargo, North Dakota launched with 11 stations and 101 bikes in 2015. The authors examined two years' worth of usage data and confirmed the impacts of weather, temporal, and spatial variables on BSS use. However, the main influencing key factors were the presence of a college campus, the location of stations on the college campus, and reduced barriers to the use of the BSS for college students.

Wang and Lindsey (2019) pointed out that previous study designs had been cross-sectional and were therefore unable to establish causality. The authors collected a six-year panel dataset of Nice Ride members' bike-share trips from 2010 to 2015 in Minneapolis-St. Paul and adopted a fixed-effect Poisson model to verify the significant effects of changes in accessibility on the frequency of individual members' usage. They found the distance variable had a negative impact on the frequency of use, and that the effects of increasing bike-share accessibility were greater (installing a new station in areas with a higher density of BSSs) in areas with Nice Ride services.

Guidon et al. (2020) used linear and spatial regression models to investigate arrivals and departure demand with Smide (an e-bike-sharing system in Switzerland) to predict possible expansion into a new city. The results revealed population, employment, restaurants, bars, and distance to central locations were the most important predictors that can have positive influences on Smide system usage demand.

Morton et al. (2021) adopted an SLM to investigate the interaction between the London Bicycle-Sharing Scheme's (LBSS) spatial demand and key factors, including the built environment, residential demographics, and workplace populations. Built environment variables, cycling infrastructure, railway stations, parks, university facilities, the density of shops, and the proximity of conventional roads near bike stations correlated strongly with LBSS trip generation volumes.

Zhao et al. (2021) used negative binomial regression models to examine the effect of the built environment on public bicycle system usage in Hexi District, Nanjing. The authors investigated 15 types of POI data and found residence, employment, entertainment, and metro stations had significant correlations with public bicycle system usage. The results of the empirical study indicated that POIs such as residence, employment, entertainment, restaurant, bus stop, metro station, amenities, and schools within a 300-meter buffer had significant positive effects on public bicycle system ridership, while POIs such as shopping, parks, attractions, sports, and hospitals within a 300-meter buffer had significant negative effects.

2.2.3 Temporal Variables

Temporal variables such as time windows, days (weekday vs. weekend), and seasons (summer vs. winter) also have a significant influence on BSS ridership. Faghih-Imani et al. (2014) collected the data compiled minute-by-minute from 410 stations on the BIXI website between April and August 2012 and used a linear mixed model to show a connection between arrivals and departures to and from a BIXI station with three groups of dependent variables (i.e., weather, temporal variables, and spatial variables). The authors found that usage was greatest during afternoons and evenings and was greater during weekdays rather than weekends when fewer people would be going to work.

Gebhart and Noland (2014) collected hourly CitiBike usage data and found a significant difference between peak and off-peak periods during different seasons. In a follow-up study, Noland et al. (2016) adopted negative binomial regression to examine the influence of bicycle infrastructure, population and employment, land use mix, and transit by season, weekday/weekend, and user type. Their results showed great interaction between population density, weekdays, and weekends, with the influence being much stronger at the weekend.

Faghih-Imani and Eluru (2016a) developed models that analyzed BIXI usage from April to August 2012. Their empirical study showed that time of day and weekend variables had a significant impact on BSS ridership. Locals tended to use BIXI more often during weekdays than at weekends as it is more often used for daily activities (i.e., to go to work or school) during the week than for recreational purposes at the weekend. The authors also observed that people mainly used the bikes during the afternoons and less during the evenings. For instance, workers used BIXI as a feeder transport mode for a short trip (home-work trip or to a restaurant) during the afternoon. However, during the evening peak hours, most of the BIXI users were tourists, non-member users, and students.

Faghih-Imani and Eluru (2016b) considered that arrivals and departures at one BSS station could potentially correlate with bicycle flow rates for neighboring stations in the same way that arrival and departure rates during one time period are influenced by activity at neighboring stations during an earlier period. The authors studied the spatial and temporal effects on BSS ridership by using spatial panel models. The authors noted the hourly arrivals and departures from CitiBike stations in New York and found a strong relationship between spatial-temporal variables and arrival and departure rates. For both annual members, and daily customers, time variables corresponding to morning, midday, afternoon, and evening have statistically significant efforts on arrivals and departures, particularly during the afternoon.

Faghih-Imani et al. (2017a) attempted to determine whether a BSS system or a taxi was better for people making short-range journeys by investigating the differences in observed travel times by CitiBike and taxis in 2014. The authors developed a panel mixed multinomial logit model to identify and understand the key factors that could influence differences in journey times in order to assist the operators in improving CitiBike's service. Their results showed that for journeys of less than three kilometers made on weekday mornings and afternoons, CitiBike was either faster or was competitive with taxis. Faghih-Imani et al. (2017b) also found that the lowest demand for BSSs in Barcelona was during the late-night period, while the lowest arrival demand for BSSs in Seville was during the late night and morning.

Sun et al. (2018) tried to understand the reason for the failure of the Pronto BSS in Seattle. The authors investigated the effects of land use, roadway design, elevation, bus services, weather, and temporal factors on three-hour-long Pronto bike-sharing system data. They developed a generalized additive mixed model (GAMM) to resolve the temporal autocorrelations and nonlinear seasonality. The results of their empirical study showed that university students tended to use the bike-sharing system in neighborhoods with higher household densities and a higher percentage of residential land use on both weekdays and weekends.

Gao and Lee (2019) proposed a moment-based model and a new hybrid approach that combined a fuzzy C-means (FCM)-based genetic algorithm (GA) with a backpropagation network (BPN) to effectively forecast demand for the Capital Bikeshare system to improve user satisfaction and find out the key factors that influence user behavior. The results revealed that more people used the Capital Bikeshare system in the summer and fall, while public transport was favored more in winter and spring. There was also a different trend on weekdays and weekends. Because more people used the Capital Bikeshare system to commute to work, demand increased earlier in the morning and during the evening peak period on weekdays, while demand was at its highest during the afternoons at the weekend.

2.2.4 Other Influencing Variables

In addition to the spatial-temporal variable effects on BSS ridership, other variables were often used in previous studies. The weather factor was one of the most considered variables for BSSs. In the past, when the researchers used weather variables, they usually noted variables such as rainy days, temperature, humidity, and wind strength, and they normally collected short-term or mid-term period data (usually more than one month and less than two years) (Faghih-Imani et al. 2014; El-Assi et al. 2015; Faghih-Imani and Eluru 2016a, 2016b; Mattson and Godavarthy 2017; Médard de Chardon et al. 2017; Sun et al. 2018; Gao and Lee 2019; Zhao et al. 2021). The classification of weather variables was usually divided into either hourly (every hour, morning and afternoon periods), daily (weekdays, weekends, and holidays), or seasonal (spring, summer, fall, and winter) periods according to the needs of spatial variables.

2.3 Modeling Approach

BSS ridership has been widely analyzed in the field of transportation. Past studies provide the recent developments in the area of ridership prediction and influencing factors definition. In relative literatures, the major types of modeling approach to BSS ridership prediction, i.e., regression base model and time-series methods. Therefore, in this section, recent researches on BSS ridership prediction are reviewed and briefly classified.

In the field of BSSs, ridership prediction and influencing factors are commonly formulated as linear regression programs. The most common methodology adopted to study continuous dependent variables (e.g., BSS arrival and departure usage) is the linear regression model, despite it not being appropriate in its traditional form for studying all data types of observations.

Rixey (2013) used statistical software to determine the bivariate correlations between each independent variable and the dependent variable and examined which variables should be contained for the regression process. Multivariate linear regression was improved to establish a predictive model of BSS ridership in the three input systems (Capital Bikeshare, Denver B-Cycle, and Nice Ride MN systems). Rixey aggregated a dataset from three BSS systems rather than one, improving the robustness of the regression results.

Faghih-Imani et al. (2014) used a multilevel approach to examine the influence variables (i.e. temporal effects, bike infrastructures, and the built environment) and their attributes on arrival and departure flows at the station level. In this study, the traditional linear regression model is not appropriate to study data with multiple repeated observations. The researchers observed the hourly BIXI station usage at the same time of day for each station and adopted a multilevel linear regression model that recognized the dependencies associated with the flow variable generated from the same BIXI station. Specifically, a linear mixed modeling approach was established on the linear regression model while incorporating the effects of the repeated observations from the same BIXI station. The linear mixed model would be a simple linear regression model in the absence of any station-specific effects. A traditional cross-sectional linear regression model would ignore the internal correlations through multiple repeated measurements meaning that the model results would be inefficient, and the parameter estimates would contain bias. In order to estimate the impacts of exogenous factors on BSS usage, Faghih-Imani et al. (2017b) developed a mixed linear model to estimate the effects of bicycle infrastructure and demographic characteristics on BSS station arrivals and departures in Barcelona and Seville. Their mixed model allows them to simultaneously incorporate different correlation structures into one model and to simultaneously evaluate their influence. Several studies modeling BSS demand used regression models that consider special dependence (Faghih-Imani and Eluru, 2016; Guidon et al., 2020).

Researchers such as Gebhart and Noland (2014), Noland et al. (2016), Wang et al. (2016), and Zhao et al. (2021) developed negative binomial regression models to examine the effects of the built environment, demographic characteristics, and transit accessibility on BSS usage. This was because the collected data of those studies did not pass the normality test and strongly rejected the hypothesis that the collected data are normally distributed. Therefore, models cannot be estimated by ordinary least squares (OLS) regression. The collected data of those studies can be regarded as count data, meaning that negative binomial regressions are suitable for examining them. The negative binomial distribution model is an extended version of the Poisson model. One of the strongest assumptions is that the variance and mean are equal. If the count data is fitted by an ordinary Poisson regression model and it is found that the data is over-dispersed, the negative binomial distribution model would recommend refitting the data.

The logit regression model was also adapted to predict BSS ridership. A binary logistic model is a type of logit regression in which the DV is binary. Bachand-Marleau et al. (2012) determined the influencing factors that encouraged people to use the BIXI system and influenced their usage frequency. A binary logistic model was developed to answer the research questions. In this case, the dependent variable is the previous use of a BIXI (yes–no). Next, linear regression was applied to our subsample of BIXI users to identify factors that affected the frequency of use of shared bicycles. BSS ridership prediction involves time-period issues, such as peak/off-peak and day types, part of previous studies also formulated the models based on logit model structures. Faghih-Imani et al. (2017a) developed a multivariate analysis using a random utility framework in the form of a panel mixed multinomial logit model that identified and revealed the factors that the effects of different travel time can help operators to enhance their BSS service.

The spatial lag regression model was also adopted to reveal the relationship between BSS ridership and spatial autocorrelation (Faghih-Imani and Eluru, 2016b; Morton et al., 2021). The model considered the DV of a target area with the other areas associated with it. The data, which to some extent are often geographical in nature, often have spatial autocorrelations. One of the methods for solving this spatial dependence issue is to directly formulate the model with autocorrelation, which would be achieved using autocorrelated time-series data. The spatial lag model is one of the models in which a dependent variable is predicted by using the value of the dependent variable of an observation's "neighbors".

Box and Jenkins (1987) defined a time-series as a sequence of observations taken sequentially in time. Numbers of datasets appear as a time-series situation, such as the daily quantity of goods in a company, a person's monthly income, or the annual deaths from car accidents in a particular country. A time-series has a distinctive characteristic that neighboring observations are dependent on. This characteristic of dependence in observation data has a practical interest. In order to understand the interaction, a number of methods for time-series data were developed, the concepts of the methods will be briefly reviewed below.

Previous research has involved panel data or spatial effects issues. The time-series-based methods, such as fixed effects, random effects, and mixed effects models, were adopted to examine the relationship between variables and BSS usage. Mattson and Godavarthy (2017) used a one-way random effects model to examine influencing variables on Great Rides Bike Share ridership. El-Assi et al. (2015) and Médard de Chardon et al. (2017) combined the advantages of the fixed effects and random effects models to create a mixed effects model to determine the variables that influence BSS ridership. The fixed effects model assumed that the explanatory variable has a fixed relationship with the response variables with all observations. Fixed effects of variable are a constant across individuals, such as gender, age, or ethnicity, which do not change over time. In other words, any change will cause the same effects on an individual. In a fixed effects model, random variables are treated as though they were non-random, or fixed. The opposite of the fixed effects model is the random effects model, which assumes that explanatory variables have fixed relationships with the response variables across all observations, but that these fixed effects may vary from one observation to another.

2.4 Comments on the Reviewed Literature

Table 2-2 provides a comparison of the key previous literature of this research, which shows the key influencing factors and methodologies from the previous studies and describes the differences between this research and previous studies.

Table 2-2. Summary of the key literature of BSS

Author(s)	**Dependent variable**	**Independent variables**		**Solution method and data range**	**Contributions**
		Key Demographic variable	**Key Socioeconomic variables, Environment variables, and others**		
Bachand-Marleau et al. (2012)	● BIXI system arrivals and departures year rate ● All stations	● Gender	● Station less than 500 m from home or destination ● Household income ● Member of BSS ● Spatial variable (distance from home to downtown)	● Survey and binary logistic model ● January to December 2010	● Determine spatial factor that influence BSS ridership ● Theft prevention
Rixey (2013)	● Capital Bikeshare, Nice Ride, and B-Cycle systems monthly arrivals and departures ● All stations	● Number of populations (Non-white population) ● Number of Jobs	● Bachelor’s degree ● Number of stations in 200-6400 m buffer ● Bicycle facility	● Multivariate regression model and Bivariate regression ● October 2010, to September 2011	● Determine key factors that influence BSS Determine network effect on bicycle-sharing system ridership
Gebhart and Noland (2014)	● Capital Bikeshare system trips per hour ● Average trip duration per hour ● All stations	● None	● Weather and temperature	● a negative binomial model ● September 15, 2010, to December 31, 2011	● Determine key factors that influence Capital Bikeshare usage ● Determine Metro serves as a backup option when weather conditions are

					unfavorable for bicycling
Faghih-Imani et al. (2014)	• BIXI system the 5-min arrival and departure rate • All stations	• Population density • Job density	• Station capacity • Number of stations or capacity in 250 m buffer • Weather, temperature, and temporal effects • Bicycle facility and road infrastructure • Land use factors	• Linear mixed regression model • April to August 2012	• Micro analysis in station-base • Determine BIXI station size and location decision
El-Assi et al. (2015)	• Bike Share Toronto daily arrivals and departures • All stations	• Population density • Employment density	• Number of stations or capacity in 200 m buffer • Weather and temperature • Bicycle facility • Land use factors	• Distribution lag model • Multi-level/Linear Mixed Effects model • January to December 2013	• Determine key factors that influence BSS ridership
Faghih-Imani and Eluru (2016a)	• BIXI system daily arrivals and departures • All stations	• Population density • Job density	• Station capacity • Number of stations or capacity in traffic analysis zone (TAZ) • Weather and temperature • Bicycle facility and road infrastructure in TAZ • Land use factors	• Mixed ordered logit model • April to August 2012	• Evaluation in TAZ • Determine key factors that influence BIXI usage • Determine bicycle infrastructure influence in BSS initial stage
Faghih-Imani and Eluru (2016b)	• CitiBike system hourly arrivals departures • Single station	• Population density • Job density	• Station capacity • Number of stations or capacity in 250 m buffer	• Spatial lag model (spatial autoregressive), spatial error model	• Determine key factors that influence CitiBike usage

			● Weather, temperature, and temporal effects ● Bicycle facility and road infrastructure ● Land use factors	(spatial autocorrelation) ● September 2013	● Determine spatial and temporal interaction of BSS station's demand
Noland et al. (2016)	● CitiBike system arrivals departure of weekday, weekend, and holiday ● All stations	● Number of Population ● Number of employments	● Station capacity ● Land use factors	● Negative binomial regression ● February, July, and November of 2014	● Determine key factors that influence CitiBike usage in a season of the year and weekday/weekend ● CitiBike usage forecasting in 2015
Wang et al. (2016)	● Nice Ride station daily trip original (departure) and trip destination (arrival) ● All stations	● Percentage of White/Caucasian resident ● Percentage of residents younger than 5 years or older 64 years ● Total jobs within 30-min transit accessibility	● Distance to nearest lake, river, central business district (CBD), park, station ● Station at campus ● Number of business ● Land use factors	● Log-linear and negative binomial regression models ● January to December 2011	● Determine sociodemographic, build environment, and transportation-related infrastructure on Nice Ride system daily trips ● Optimize bike-share operations, locate new stations, and evaluate the potential of new bike-share

					programs
Faghih-Imani et al. (2017a)	● CitiBike system hourly arrivals and departures	● Population density ● Job density	● Station capacity ● Bicycle facility and road infrastructure ● Temporal factors (AM/PM, day type) ● Land use factors	● Panel mixed multinomial logit model. ● January to December 2014	● Comparison between CitiBike and private vehicles using ● Determine the factors that affect the travel time for improving the CitiBike service
Faghih-Imani et al. (2017b)	● BSS of Barcelona and Seville arrival and departure rate	● Population density	● Station capacity ● Percentage of point of interest (recreation, restaurant, hotel) ● Temporal factors (AM/PM, day type) ● Land use factors	● Mixed linear model and binary logit model ● May 1 to September 20, 2009	● Determine key factors that influence BSS ridership ● Improving bike rebalancing (refilling and removal)
Mattson and Godavarthy (2017)	● Great Rides daily ridership	● Population density	● Station capacity ● Weather and temperature ● Temporal factors (day type, season) ● Land use factors	● One-way random effects model ● January 2014 to December 2015	● Determine BSS successful factor in a small city ● Determine impacts of weather, temporal, and spatial variables
Médard de Chardon et al. (2017)	● Monthly average trip per day per bike (TDB) of 75 BSSs around the world	● Population	● BSS-related elements (dock type, seasonality, number of bikes/stations, ratio of bicycles to stations, station	● Robust regression model and fixed-effect model ● 904 months of data	● Determine network effect on BSS ridership ● Using TDB as an

			empty, station density) ● Weather ● Member of BSS ● Other transportation modes	from 75 BSSs	indicator to evaluate BSS performance
Sun et al. (2018)	● Pronto's bike trip data	● Household density ● Employment density	● Weather ● Temporal factors (AM/PM, day type) ● Land use factors	● Generalized additive mixed model (GAMM) ● October 2014 to August 2016	● Determine encourage, or discourage factors on BSS trip generation/attraction at the station level ● Unsuccessful BSS analysis
Gao and Lee (2019)	● Capital Bikeshare arrivals and departures	● Population size	● Weather ● Temporal factors (hourly, day type, monthly, season)	● Fuzzy C-means (FCM)-based genetic algorithm ● January 2011 to December 2012	● Determine temporal and weather attribution in Capital Bikeshare usage forecasting
Scott and Ciuro, (2019)	● Ontario's bike-share arrivals and departures	● Population in 200 m buffer (age 15-64)	● Weather ● Temporal factors (season, day type)	● Random intercept multi-level model ● April 1, 2015, to March 31, 2016	● Determine weather, temporal, hub attributes, and a one-day lag on Ontario's bike-share ridership ● Confirm population did not influence

					Ontario’s bike share usage
Wang and Lindsey (2019)	● Nice Ride Minnesota weekly usage by annual member	● Population density ● Job density	● Spatial variable (network distance to the nearest bike stations) ● Bicycle facility	● Fixed effects Poisson model ● 2010 to 2015	● Determine the accessibility effect on the frequency of individual members’ use of Nice Ride.
Guidon et al. (2020)	● Smide (e-bike-sharing system in Switzerland) arrivals and departures	● Population size	● Number of workplace ● Spatial factors (distance to main infrastructures, public transport) ● Land use factors	● Linear and spatial regression models ● October to December 2018	● Estimate bicycle-sharing demand and predict expansion to a new city
Morton et al. (2021)	● London Bicycle-Sharing Scheme arrivals and departures	● Number of residents ● Number of workers	● Temporal factors (hourly) ● Spatial factors (distance to bike lane, park, Railway station)	● Spatial lag model ● January to December 2016	● Determine built environment, demographics, and temporal attribution in London Bicycle-Sharing usage.
Zhao et al. (2021)	● Public bicycle (in Hexi District of Nanjing) trip generation, trip attraction, trip distribution, and trip duration	● Number of households ● Number of enterprises	● Station capacity ● Number of stations in 250 m buffer ● Weather and temperature ● Temporal factors (day types)	● Negative binomial regression model ● November 2015	● Determine the relationship between public bicycle trip generation,/attraction and build environment, weather, population density factors.

Hu and Liu (2022)	● **CityBike system monthly arrivals and departures** ● **All stations**	● **Worker population** ● **Student population** ● **Worker trip demand-generation/attraction** ● **Student trip demand-generation/attraction**	● **Station capacity** ● **Number of stations or capacity in 1 kilometer** ● **Bicycle facility and road infrastructure** ● **Land use factors**	● **ARX model based on the hierarchical level** ● **January 2009 to December 2017**	● **Determination of trip attribution in the CityBike usage forecasting** ● **Raising the accuracy of CityBike usage forecasting with long-term data** ● **Determination of the CityBike system size in individual district**

A summary of the related literature for BSS research is described as follows.

1. Refer previous studies, there were two types of BSS ridership prediction models. Firstly, when the dependent variables were aggregated in hours, days, or months, a logit model or negative binomial model was adapted for analysis. The authors can understand the relationship between the influencing variables and BSS ridership in different time slots (e.g. peak hour vs. off-peak hour; morning vs. evening) and day types (e.g. weekly vs. weekend). However, those methods were not suitable for analyzing a long-term dataset, because it is hard to collect short-term period data. Secondly, regression-based models were usually adopted for analysis, such as a multivariate regression model and bivariate regression, linear mixed regression model, and spatial lag model. Those models can analyze long-term data and multiple types of variables (i.e., continuous or categorical variables) and consider the time-series nature. In order to find out the BSS ridership trend, the authors usually collected mass and variant variables over the long term. Therefore, these models were not suitable for short-term data analysis.
2. In previous research, regardless of the methods used to analyze BSS ridership or performance, the population variable is essential data. However, population cannot represent the real travel behavior in a particular area; only people's true behavior can influence BSS arrivals and departures. To fill this research gap, this study used the trip variable to replace the population variable as it can obtain real-world effects on BSS ridership prediction based on the travel behaviors of the citizens. Since spatial and temporal factors were usually adopted for BSS issues in past studies, this study also collected the same or similar spatial-temporal variables from the target areas into the developed models to ensure that minimal bias was generated in the modeling process.
3. From the operator or government perspective, the key factors of BSS ridership prediction that need to be considered are also different depending on the regional scale. Most previous studies had focused on one single rental station or small-scale BSS system, and only collected the data

near the station or in a limited buffer range, such as 200–500 meters around the station (Rixey 2013; Faghih-Imani et al. 2014; El-Assi et al. 2015; Faghih-Imani and Eluru 2016b; Scott and Ciuro 2019; Zhao et al. 2021). This study attempts to divide the research areas into four district categories of scale to define the key factors influencing the particular areas. The data would be analyzed at the station-level and district-level to comprehensively evaluate BSS performance.

4. Previous studies analyzed the relationship between BSS ridership and influencing factors over the short term and on a small scale; part of previous studies focused on the effects of spatial factors, and the other focused on temporal factors. The analysis of temporal factors was usually based on different time periods, such as weekdays/weekends or seasons (Faghih-Imani et al. 2014; Faghih-Imani et al. 2017a; Faghih-Imani et al. 2017b; Mattson and Godavarthy 2017; Sun et al. 2018; Scott and Ciuro 2019; Morton et al. 2021; Zhao et al. 2021). There were no studies of long-term perspectives to observe the interactions between BSS ridership and these factors. This study attempts to analyze nine years of related data from both cross-sectional and longitudinal perspectives.
5. In previous studies, the analysis between BSS ridership and key influencing factors was usually conducted over a specific, normally short-term period. There were a few studies that discussed optimal BSS size, but they only focused on a single station (Faghih-Imani et al. 2014). This is because there is no long-term dataset with which researchers can observe the variation in BSS stations within a buffer range in the target areas. A long-term dataset would be able to provide a sufficient number of cross-sectional samples to observe the relationship between BSS station variation and BSS ridership.

Chapter 3 MODEL FORMULATION

In this chapter, we develop the mathematical formulations to predict CityBike ridership. Chapter 3 is organized as follows: Section 3.1 describes the problem statement and assumptions made in this research; Section 3.2 describes the model design for BSS ridership forecasting; Section 3.3 describes the mathematical formulation of this research.

3.1 Problem Statements and Assumptions

Past studies have only modeled short-term arrival and departure rates at BSS stations (i.e., minutes or hours of a day) as functions of the environment (i.e., a 250–500 m buffer surrounding a BSS station), demographics, population, land use, POI, system capacity, and socioeconomic characteristics at the station level. However, the use of small-scale or single-station data is insufficient for evaluating levels of BSS ridership over time. Previous studies have also not considered trip attributes and district-level features in the modeling process. Therefore, this research is unique, to the best of our knowledge, in that it includes station and district-level data, including spatial variables (i.e., station capacity, capacity in the buffer range, and POI), socioeconomic variables (i.e., demographics), and infrastructure variables (i.e., bike lane lengths, number of lamps, road lengths), and the trip data is treated as an independent variable instead of the population factor.

Kaohsiung City was divided into different verification district categories in terms of scale, and we used HRM and ARX models to forecast the CityBike ridership. The assumptions of the developed models in this research are described as follows.

1. In previous research, population has been one of the most important influencing factors, with a strong correlation for BSS ridership prediction. According to Cohen’s (2003) definition of hierarchical regression, population factor would be input into the first level of HRM, while other socioeconomic characteristics would be sorted into the second and third

levels. In contrast to previous studies, trip attributes were the key variables influencing the CityBike ridership in this research. To evaluate the effects of trip attributes on the CityBike ridership, trip generation, and attraction flow classified by trip purpose (i.e., work-oriented or school-oriented) were used instead of population by the models developed in this research.

2. Previous studies have evaluated the spatial effects, modeled as a vector of influence factors (including number of POI, BSS capacity, and number of BSS stations), on BSS ridership near rental stations with a limited buffer zone (Kittelson et al. 2003; Cui et al. 2014; Faghih-Imani et al. 2014; Faghih-Imani and Eluru 2016a, 2016b). Kittelson et al. (2003) and Faghih-Imani et al. (2014) assumed that 250–500 m is a reasonable distance for people walking to and from a rental station. In this research, the factors assumed to affect CityBike usage were standardized into square-kilometer zones for district-level data, and rental station capacity and the number of rental stations near each rental station within a 1-km buffer zone were collected at the station level. According to the ridership data of the CityBike system, CityBike users typically travel more than 250 m (KRTC 2016). Chang and Lin (2013) reported that Kaohsiung City residents will tolerate a 12.26-minute walk to the nearest transport station (assuming a 5 km per hour walking speed, which is approximately equal to a distance of 1,266 meters). Therefore, a 1-km buffer zone for the collected data was considered reasonable for the developed models.

3.2 Model Design

In this section, the modeling process is presented based on previous definitions and assumptions. The main steps are described as follows.

1. In previous BSS studies, many impact factors were determined and used to predict BSS ridership. However, to the best of our knowledge, no studies have classified and defined the priority order of impact of those factors on BSS ridership. This research used HRM to analyze and verify the priority ranking of these factors. According to Cohen (2003), hierarchical regression

must be based on the theoretical basis of previous research. This research found that the population factor is indeed one of the main factors affecting BSS ridership after reviewing past studies and understanding that socioeconomic and spatial-temporal factors also have significant effects on BSS ridership. Once the studies were reviewed, this research compiled a total of 27 influencing factors, including population (worker and student populations), trip attributes (worker and student generation/attraction), and other factors that have been confirmed to impact BSS ridership by previous studies. The HRM was adopted to classify those factors into three levels to examine the degree of influence at each level on the BSS ridership forecast in terms of citywide, urban district, suburban district, and individual district.

2. This research collected CityBike ridership data between 2009 and 2017. Due to the time-series nature of these data, this research considered the possible impact of past historical BSS ridership on the current state. Therefore, before formulating the ARX model, this research used the AR model to define the order of historical ridership data on current ridership. The autocorrelation function and partial autocorrelation function (Box and Jenkins 1987; Kendall and Ord 1990) are two major tools used to define the order of an autoregressive model.
3. When the AR mode defines the order of the model, besides the series correlation of the ridership data, CityBike ridership would be affected by several exogenous variables. Therefore, the ARX model was adopted to forecast CityBike usage. The ARX model considers time-series characteristics of the ridership data and combines the influence of other exogenous variables to forecast CityBike usage. At the same time, because this research obtained the hierarchy (ranking) of each factor affecting the CityBike ridership from the HRM, the advantage of the HRM was incorporated into the ARX model and used to estimate CityBike usage. The ARX model with hierarchical characteristics not only classifies the influencing factors but also considers the time-series nature of data (including the impact of historical data on current data). This research also

compared the HRM with the ARX model to confirm whether the predicted CityBike ridership is more accurate.

4. In order to confirm the impact of the number of CityBike rental stations on CityBike ridership, this study used cross-sectional data to understand the relationship between the number of CityBike rental stations and people's travel demand in different periods. That could provide the operators or government adjust the construction policies and strategies for CityBike operations.

3.3 Model Structure and Formulation

This section introduces the developed models for the CityBike ridership predictions. The basic concepts of the aforementioned models are introduced as follows.

3.3.1 Description of the Developed Models

To examine the efficiency of different factors and time intervals, the HRM, AR, and ARX model's specifications were set up to control for trip attributes and socioeconomic factors as much as possible. A flow chart of the indexing scheme of the different models is shown in Figure 3-1, and each of the models is described in Table 3-1.

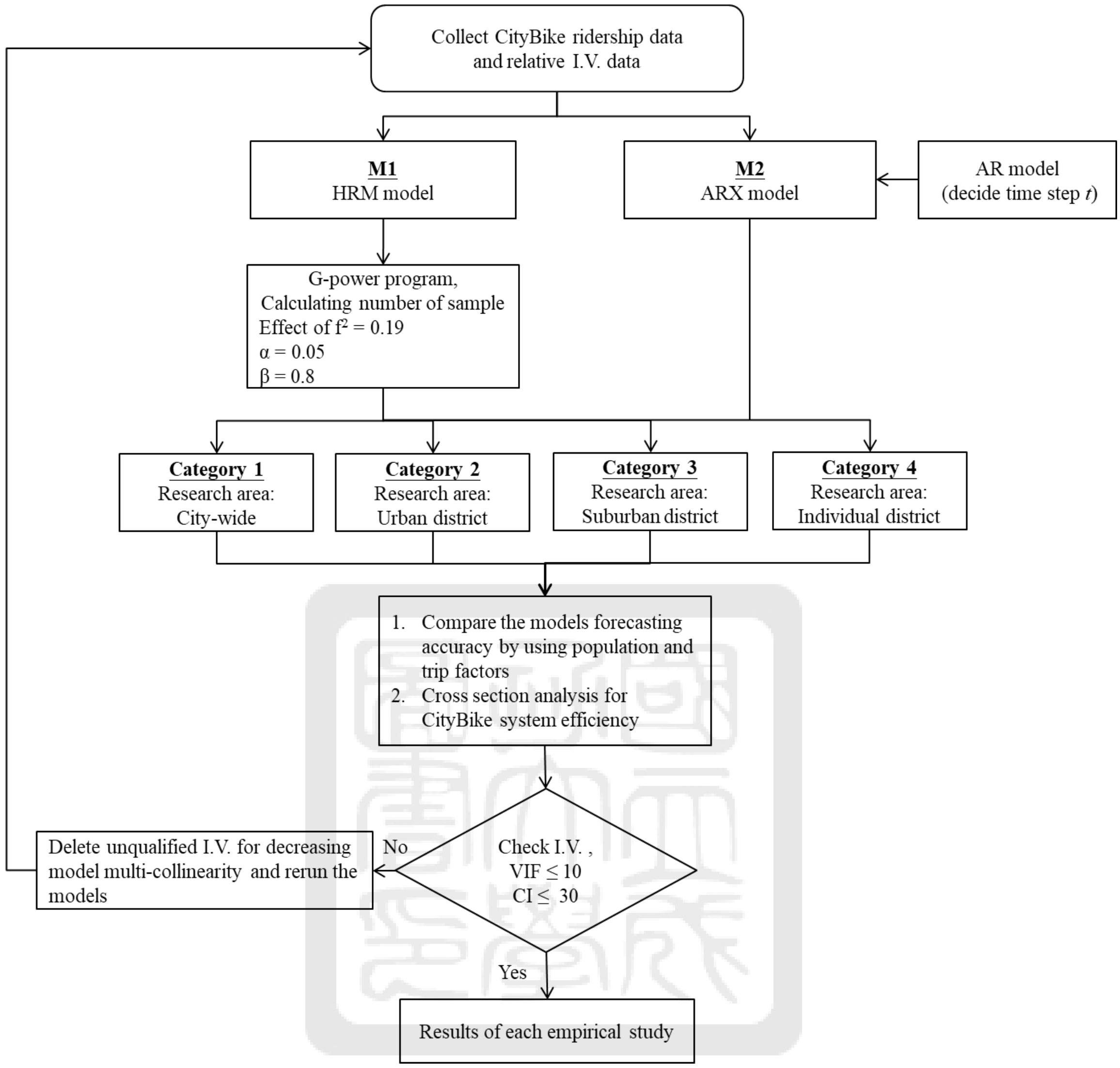


Figure 3-1. Flow chart of indexing scheme of the developed models.

Table 3-1. Description of the developed models

Model	Independent Variables	District categories	Method	Sample type
M1	1. Trip generation and attraction 2. Population	1. Citywide 2. Urban districts 3. Suburban districts 4. Individual districts	Hierarchical regression	Panel data
M2			ARX-Autoregressive With exogenous variable + Hierarchical regression	

The following criteria were used when choosing variables.

1. In order to verify the hypothesis of this study, population factors and trip attributes (i.e., trip purpose and the number of trip generations and attractions) were assumed to be individual IVs among the collected variables and were put into the model separately in the modeling process.
2. The student and worker populations were added to the models simultaneously to examine the influence of the variables on CityBike departures or arrivals. The same steps were compiled in trip variables.
3. Figure 3-1 shows that multicollinearity was detected during the modeling process because some of the variables' variance inflation factors (VIFs) were greater than 10, and the condition index (CI) was greater than 20. Belsley et al. (1980) stipulated that a CI of less than 30 is acceptable, between 30 and 100 is moderate, and more than 100 is critical. In this study, variables were removed from the parameter estimation if their CI exceeded 30 in a developed model. The VIFs of the remaining variables were confirmed to be less than 10. Effective explanatory variables were considered significant at the 0.1 level of significance.

As shown in Table 3-1, M1 and M2 are used to represent different models and populations, and the trip factors are adopted into the models for comparison. Model M1 without the trip attribute and time-lagged term effect was used as a benchmark for comparison with M2. In this research, variables at both the station and district levels included spatial variables (i.e., station capacity, capacity in the buffer range, and POI), socioeconomic variables (i.e., land use and demographics), and infrastructure-related variables (i.e., bikeway length, number of lamps, road length) were collected. The collected variables were incorporated into the developed models to comprehensively investigate the key factors affecting the CityBike ridership.

Two models (M1 and M2) were developed and tested under four district categories. These four district categories differed in terms of scale, and were named as (1) citywide—the 19 districts that have CityBike stations grouped into one area for model testing; (2) urban districts—the 12 districts located in the center of Kaohsiung City, and which have the CityBike stations grouped into one area for model testing; (3) suburban districts—the seven districts located on the outskirts of Kaohsiung City, and which have the CityBike stations grouped into one area for model testing, and (4) individual districts—the 19 districts of Kaohsiung City that have CityBike stations tested separately in the models. In Kaohsiung City, CityBike has different functions depending on the scale. Citywide, the CityBike system serves as one part of the public transportation system in Kaohsiung City. The number of stations and their locations affect the performance of the overall system. In urban districts, CityBike stations were usually constructed near institutions, schools, shopping malls, offices, and other points of interest, and to serve as a feeder mode to citizens during their daily lives (Department of Transportation, Taipei City Government 2016; Transportation Bureau of Kaohsiung City Government 2009). In suburban districts, most CityBike stations are constructed near points of interest for touristic purposes because commercial activities and public transportation are not as well developed as in the more urban areas. This study attempts to examine the various influencing factors on CityBike ridership at different district scales in Kaohsiung City.

3.3.2 Dependent Variables Hierarchical Division

In general regression analysis, all explanatory variables were included in the models, and the overall effect on the DV was examined. However, we could not obtain results for specific variables. Hierarchical regression can be used to determine whether explanatory variables can explain statistically significant variations in a DV after considering all other variables (Wampold and Freund 1987; Cohen 2003). Hierarchical regression was used to determine whether the inclusion of more variables increases the percentage of explained variance in the DV.

In the past, when researchers had several potential IVs without an overall theory to guide selection, traditional regression methods (e.g., multiple regression and stepwise regression models) were usually adopted for analysis. This was because those methods use computers to select the variables, relieving the researchers of the responsibility of making decisions regarding their logical or causal priority or relevance before the analysis. However, interpreting the findings may not be made any easier. In a hierarchical regression, when researchers believe that a more orderly advance in the behavioral sciences is likely to occur, the researchers armed with theories provide an *a priori* ordering that reflects causal hypotheses rather than when computers order independent variables post and for a given sample. The researchers may understand or hypothesize that some groups of variables occur logically, causally, or structurally before other variables, and do not have a basis for ordering variables within the group. In the solution process, hierarchical regression adds a set of variables into the regression process to determine how much effect the prediction of DV has over and above the contribution of the previously included IVs. In the simplest equation form, the IVs are entered cumulatively in a specified sequence, while the R^2, partial regression, and correlation coefficients are determined as each independent variable joins the others. A series of hierarchical procedures for IVs consists of a series of n regression analyses and equations, each with one more variable than its predecessor. However, some researchers used sets of independent variables in their models. The choice of a particular cumulative sequence of independent variables is made in advance and is dictated by the purpose and logic of the research (Cohen, 2003).

In this research, in order to understand the effects of trip, population, and other factors' attributes on the CityBike ridership, the key step in the hierarchical regression process was to divide the IVs into multiple levels based on their characteristics and importance before being input into the model sequentially. Firstly, the effects of the variables collected in this research had been identified based on BSS usage in previous studies. Population was usually considered the main factor in BSS ridership prediction. Therefore, based on Cohen's (2003) research, this research considered population and trip variables to be the most important factors and ordered them into the first level of the hierarchical regression process. Secondly, previous researchers also collected environmental factors (e.g., number of POI, industries, BSS capacity) and also used them to predict BSS ridership, even though some factors did not significantly affect the models. This research considered those factors to be of secondary importance and ordered them into the second level of the hierarchical regression process. Other factors had not been fully adopted into the predicting model and we ordered them into the third level of the hierarchical regression process. The three levels consisted of (1) first-level demographic and trip purpose variables; (2) second-level environmental variables; and (3) third-level "other" variables. By conducting this step, the influence of factors of three levels on the arrival and departure rates of a BSS could be identified. The collected variables in this research are shown in Table 3-2.

Table 3-2. Summary of the variables

Variables	Hierarchical level	Expected sign.
Dependent Variable (DV)		
CityBike departures	-	-
CityBike arrivals	-	-
Independent Variable (IV)		
Worker population	first level	positive
Student population	first level	positive
Worker trip attraction	first level	positive
Worker trip generation	first level	positive
Student trip attraction	first level	positive
Student trip generation	first level	positive
Capacity of station	second level	positive
CityBike station capacity in a 1-km buffer	second level	positive
Bikeway Length	second level	positive
Number of streetlamps	second level	positive
Length of major road	second level	positive
Number of parks	second level	positive
Number of companies	second level	positive
Number of factories	second level	positive
Number of hotels	second level	positive
Number of markets	second level	positive
Number of schools	second level	positive
Number of public transportation yard	second level	positive
Tourist	third level	positive
Income	third level	negative
Private vehicle	third level	negative
Land value	third level	negative

Table 3-2 presents the factors associated with the three levels of IVs and their expected positive and negative correlations with BSS ridership. In this research, the data set consisted of nine years of monthly data for the 185 CityBike stations operating in Kaohsiung City. Because each rental station began operating on different dates, a total of 11,474 samples were collected. In addition, this research collected the CityBike samples of related data based on the four district categories (i.e., diverse spatial scales), and most data were standardized to 1 km^2.

3.3.3 Notation of Development Models

Before constructing the basic formulations, the notations and terms used in the models are listed in Table 3-3.

Table 3-3. Notations of the developed models

Symbol	Definition	Model
Dependent variable		
$Ridership^{(i)}{}_{s,d,t}$	The number of ridership at the *s-th* station and in the *d-th* district type at time *t*	HRM/AR/ARX
Independent variables		
$Ridership^{(i)}{}_{s,d,t-1}$	the number of ridership at the *s-th* station and in the *d-th* district type at time *t-1*	HRM/AR/ARX
$1stLIV_{s,d,t}$	The first-level independent variable at the *s-th* station and in the *d-th* district type at time *t*	HRM/ARX
$2ndLIV_{s,d,t}$	the second-level independent variable at the *s-th* station and in the *d-th* district type at time *t*	HRM/ARX
$3rdLIV_{s,d,t}$	the third-level independent variable at the *s-th* station and in the *d-th* district type at time *t*	HRM/ARX
Parameters		
β_0	constant coefficient	HRM/ARX
β_1	the coefficients of the first-level independent variables	HRM/ARX
β_2	the coefficients of the second-level independent variables	HRM/ARX
β_3	the coefficients of the third-level independent variables	HRM/ARX
β_4	the coefficients of the number of *t*-1 phase ridership	HRM/ARX
$\varepsilon_{s,d,t}$	the random error term at the *sth* station and in the *dth* district type at time *t*	HRM/ARX

Sets		
X	the matrix that includes all the independent variables	HRM/ARX
X^T	the transposed matrix that includes all the independent variables	HRM/ARX
Y	the matrix that includes observed ridership at the *s-th* station and in the *d-th* district type at various time points	HRM/ARX

3.3.4 Hierarchical Regression Model (HRM model, M1 model)

The HRM proposed in this research is described in Eqs. (3.1) to (3.4).

$$Ridership^{(i)}{}_{s,d,t} = \beta_0 + \beta_1 \cdot 1stLIV_{s,d,t} + \varepsilon_{s,d,t}, \tag{3.1}$$

$$Ridership^{(i)}{}_{s,d,t} = \beta_0 + \beta_1 \cdot 1stLIV_{s,d,t} + \beta_2 \cdot 2ndLIV_{s,d,t} + \varepsilon_{s,d,t}, \tag{3.2}$$

$$Ridership^{(i)}{}_{s,d,t} = \tag{3.3}$$

$$\beta_0 + \beta_1 \cdot 1stLIV_{s,d,t} + \beta_2 \cdot 2ndLIV_{s,d,t} + \beta_3 \cdot 3rdLIV_{s,d,t} + \varepsilon_{s,d,t},$$

$$\varepsilon_{s,d,t} \sim NID\ (0, \sigma^2) \tag{3.4}$$

where β_0 is a constant coefficient; β_1, β_2, and β_3 are the coefficients of the IVs; $1stLIV_{s,d,t}$, $2ndLIV_{s,d,t}$, and $3rdLIV_{s,d,t}$ are the first-level IVs, second-level IVs, and third-level IVs at the *s-th* station and in the *d-th* district type at time *t*, respectively; and ε_t is the random error term at time *t*.

Moreover, *i* represents the CityBike arrivals and departures, respectively denoted by 1 and 2; *s* is the serial number of a CityBike station; *t* represents the t-th month of the research period; and d includes four types of city scales (i.e., citywide, urban districts, suburban districts, and individual districts, denoted as 1 to 4, respectively). In the regression model, $\varepsilon_{s,d,t}$ is the only random errors term. These assumptions imply that the $Ridership^{(i)}{}_{s,d,t}$ also has common variance σ^2 and to be independent in pairs. In order to have tests with significance, the random error terms are assumed to be a normal distribution, which indicates the $Ridership^{(i)}{}_{s,d,t}$ also has a normal distribution. The assumption of random error terms is frequently formulated as Eq. (3.4).

In Eqs. (3.1) to (3.3), the LSM is also adopted to solve the values of each parameter. The error term $\varepsilon_{s,d,t}$ is assumed to be independent with common variance σ^2, and the IVs are also assumed to be measured without error. The LSM is applied to solve this regression model which can then be formulated as follows.

$$SS(Res) \tag{3.5-1}$$

$$= \sum \left(Ridership^{(i)}{}_{s,d,t} - \widehat{Ridership}^{(i)}{}_{s,d,t} \right)^2$$

$$= \sum \left(Ridership^{(i)}{}_{s,d,t} - \hat{\beta}_0 - \hat{\beta}_1 \cdot 1stLIV_{s,d,t} \right)^2$$

$$SS(Res) \tag{3.5-2}$$

$$= \sum \left(Ridership^{(i)}{}_{s,d,t} - \widehat{Ridership}^{(i)}{}_{s,d,t} \right)^2$$

$$= \sum \left(Ridership^{(i)}{}_{s,d,t} - \hat{\beta}_0 - \hat{\beta}_1 \cdot 1stLIV_{s,d,t} - \hat{\beta}_2 \cdot 2ndLIV_{s,d,t} \right)^2$$

$$SS(Res) \tag{3.5-3}$$

$$= \sum \left(Ridership^{(i)}{}_{s,d,t} - \widehat{Ridership}^{(i)}{}_{s,d,t} \right)^2$$

$$= \sum \left(Ridership^{(i)}{}_{s,d,t} - \hat{\beta}_0 - \hat{\beta}_1 \cdot 1stLIV_{s,d,t} - \hat{\beta}_2 \cdot 2ndLIV_{s,d,t} - \hat{\beta}_3 \cdot 3rdLIV_{s,d,t} \right)^2$$

$$\boldsymbol{X} = \begin{bmatrix} 1 & 1stLIV_{1,1,1} & 2ndLIV_{1,1,1} & 3rdLIV_{1,1,1} \\ 1 & 1stLIV_{2,1,2} & 2ndLIV_{2,1,2} & 3rdLIV_{2,1,2} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1stLIV_{k,1,n} & 2ndLIV_{k,1,n} & 3rdLIV_{k,1,n} \end{bmatrix} \tag{3.6}$$

$$\boldsymbol{Y} = \left(Ridership^{(i)}{}_{1,1,1}\ Ridership^{(i)}{}_{2,1,2}\ Ridership^{(i)}{}_{3,1,3} \cdots Ridership^{(i)}{}_{k,1,n} \right)^T \tag{3.7}$$

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X^T X})^{-1} \boldsymbol{X^T Y} \tag{3.8}$$

$$\boldsymbol{\varepsilon} = \boldsymbol{Y} - \boldsymbol{X} \cdot \boldsymbol{\beta} \tag{3.9}$$

$$\boldsymbol{\beta} = (\boldsymbol{\beta_0} \quad \boldsymbol{\beta_1} \quad \boldsymbol{\beta_2} \quad \boldsymbol{\beta_3})^{\boldsymbol{T}} \tag{3.10}$$

In Eq. (3.5-1) to (3.5-3), where β_i is the coefficient of the *i*-th input variable, the LSM method is adopted for calibrating the coefficients of the HRM. $\boldsymbol{X}$ in Eq. (3.6) is a matrix that includes all the IVs listed in Table 3-3 of the CityBike stations at each time point, where the index n is the number of time points.

$\boldsymbol{Y}$ is the observed ridership at the *sth* station and in the *dth* district type at various time points in Eq. (3.7). The values of $\boldsymbol{\beta}$ can be estimated using the LSM as in Eq. (3.5). The values of "X," "$\boldsymbol{\beta}$," "ε," and "$\boldsymbol{Y}$" can be calculated using Eqs. (3.3) to (3.10), where the index "n" is the number of time steps.

3.3.5 Autoregressive Model (AR Model) and Autoregressive with Exogenous Variable Model (ARX Model, M2 model)

This research denote the value of a process at equally spaced times, $t, t-1, t-2, \dots\dots t-p$ by $Ridership^{(i)}{}_{s,d,t}, Ridership^{(i)}{}_{s,d,t-1} \dots\dots . Ridership^{(i)}{}_{s,d,t-p}$. And define $\widetilde{Ridership}^{(i)}{}_{s,d,t} = Ridership^{(i)}{}_{s,d,t} - \mu$ is the series of deviations from μ. The equation of the AR process of order p is as follows.

$$\begin{aligned}\widetilde{Ridership}^{(i)}{}_{s,d,t} & \\ &= \phi_1 \widetilde{Ridership}^{(i)}{}_{s,d,t-1} + \phi_2 \widetilde{Ridership}^{(i)}{}_{s,d,t-2} + \cdots \\ &+ \phi_p \widetilde{Ridership}^{(i)}{}_{s,d,t-p} + \alpha_t\end{aligned} \tag{3.11}$$

Because the AR model is autoregressive, the variable $Ridership^{(i)}{}_{s,d}$ of Eq. (3.11) is regressed on the previous value of itself. This research used ACF and PACF to define the order of the AR model for the CityBike system. The equations for determining the order of the AR model are as follows.

$$\begin{aligned}\mu = E\left[Ridership^{(i)}{}_{s,d,t}\right] & \\ &= \int_{-\infty}^{\infty} Ridership^{(i)}{}_{s,d} p\left(Ridership^{(i)}{}_{s,d}\right) dRidership^{(i)}{}_{s,d}\end{aligned} \tag{3.12}$$

$$\sigma_z^2 = E\left[\left(Ridership^{(i)}{}_{s,d,t} - \mu\right)^2\right] \tag{3.13}$$

$$= \int_{-\infty}^{\infty} \left(Ridership^{(i)}{}_{s,d,t} - \mu\right)^2 p\left(Ridership^{(i)}{}_{s,d}\right) dRidership^{(i)}{}_{s,d}$$

$$\gamma_k = cov\left[Ridership^{(i)}{}_{s,d,t}, Ridership^{(i)}{}_{s,d,t+k}\right] \tag{3.14}$$

$$= E\left[\left(Ridership^{(i)}{}_{s,d,t} - \mu\right)\left(Ridership^{(i)}{}_{s,d,t+k} - \mu\right)\right]$$

$$\rho_k = \frac{E\left[\left(Ridership^{(i)}{}_{s,d,t} - \mu\right)\left(Ridership^{(i)}{}_{s,d,t+k} - \mu\right)\right]}{\sqrt{E\left[\left(Ridership^{(i)}{}_{s,d,t} - \mu\right)^2\right]}\left[\left(Ridership^{(i)}{}_{s,d,t+k} - \mu\right)^2\right]} \tag{3.15}$$

$$= \frac{E\left[\left(Ridership^{(i)}{}_{s,d,t} - \mu\right)\left(Ridership^{(i)}{}_{s,d,t+k} - \mu\right)\right]}{\sigma_z^2}$$

$$\rho_k = \frac{\gamma_k}{\gamma_0} \tag{3.16}$$

$$\widetilde{Ridership}^{(i)}{}_{s,d,t-k}\widetilde{Ridership}^{(i)}{}_{s,d,t} \tag{3.17}$$

$$= \phi_1 \widetilde{Ridership}^{(i)}{}_{s,d,t-k}\widetilde{Ridership}^{(i)}{}_{s,d,t-1} + \phi_2 \widetilde{Ridership}^{(i)}{}_{s,d,t-k}\widetilde{Ridership}^{(i)}{}_{s,d,t-2} + \cdots + \phi_p \widetilde{Ridership}^{(i)}{}_{s,d,t-k}\widetilde{Ridership}^{(i)}{}_{s,d,t-p} + \widetilde{Ridership}^{(i)}{}_{s,d,t-k}\alpha_t$$

$$\gamma_k \widetilde{Ridership}^{(i)}{}_{s,d,t-t} = \phi_1\gamma_{k-1} + \phi_2\gamma_{k-2} + \cdots + \phi_p\gamma_{k-p}, \quad k>0 \tag{3.18}$$

$$\rho_k \widetilde{Ridership}^{(i)}{}_{s,d,t} = \phi_1\rho_k + \phi_2\rho_{k-1} + \cdots + \phi_p\rho_{k-p}, \quad k>0 \tag{3.19}$$

$$\rho_j = \phi_{k1}\rho_{j-1} + \phi_{k2}\rho_{j-2} + \cdots + \phi_{k(k-1)}\rho_{j-k+1} + \phi_{kk}\rho_{j-k} \quad (3.19\text{-}1)$$

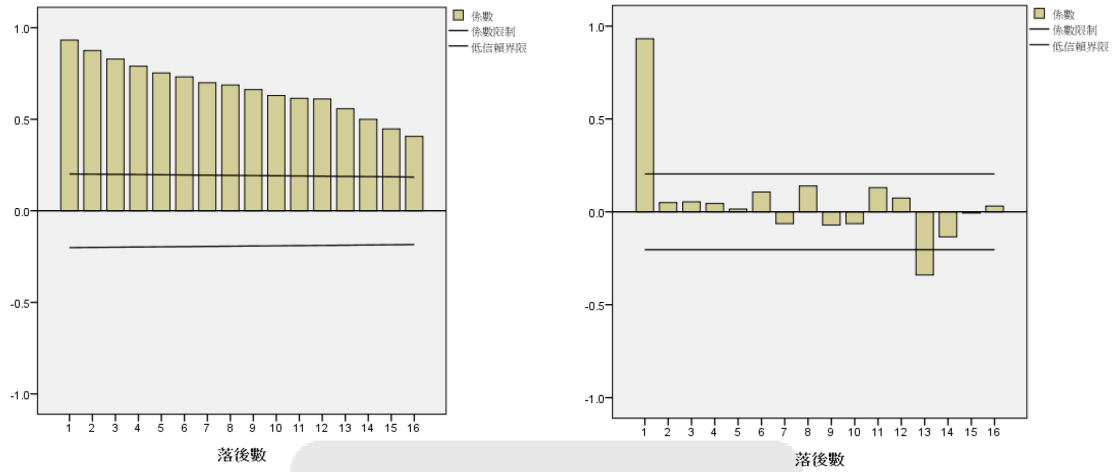
$$j = 1,2, \dots \dots, k$$

Because of the time-series nature of the CityBike ridership data, the ridership prediction was conducted using the AR model. A critical feature of time-series models is an assumption of statistical equilibrium forms. One beneficial assumption is "stationarity." Generally, we can describe this by calculating the mean, variance, or autocorrelation function for a stationarity time-series. A stationary process is based on the assumption that the process is in a statistical equilibrium state. Therefore, for a discrete process to be strictly stationary, the joint distribution of any set of observations must be unaffected by shifting all the times of observation forward or backward. The AR model assumes that a DV has a linear dependence on the DV's previous values plus a stochastic error term. In this research, the ridership data from all the research districts were examined using the autocorrelation function and the PACF (Box and Jenkins 1987; Kendall and Ord 1990). In an autoregressive process, we do not know the p_{th} order of the AR model that needs to be defined from the dataset and decided by the number of IVs of the multiple regressions. The partial autocorrelation function (PACF) is a statistical tool that explores correlation, whereas an AR(p) process has an autocorrelation function that is finite in extent and the partial autocorrelations are zero beyond lag p. The PACF can be represented in terms of p non-zero functions of the autocorrelation. The ACF in a stationary process has an important recurrence relation. The results of four district categories shown in Figures 3-2 to 3-5 indicated that the first-order model, AR (1), can adequately estimate the CityBike ridership at the 0.01 level of significance. The model AR (1) is written in Eq. (3.20).

$$Ridership^{(i)}{}_{s,d,t} = \beta_0 + \beta_1 \cdot Ridership^{(i)}{}_{s,d,t-1} + \varepsilon_{s,d,t}, \quad (3.20)$$

Category 1-citywide

ACF PACF



Category 2- urban districts

ACF PACF

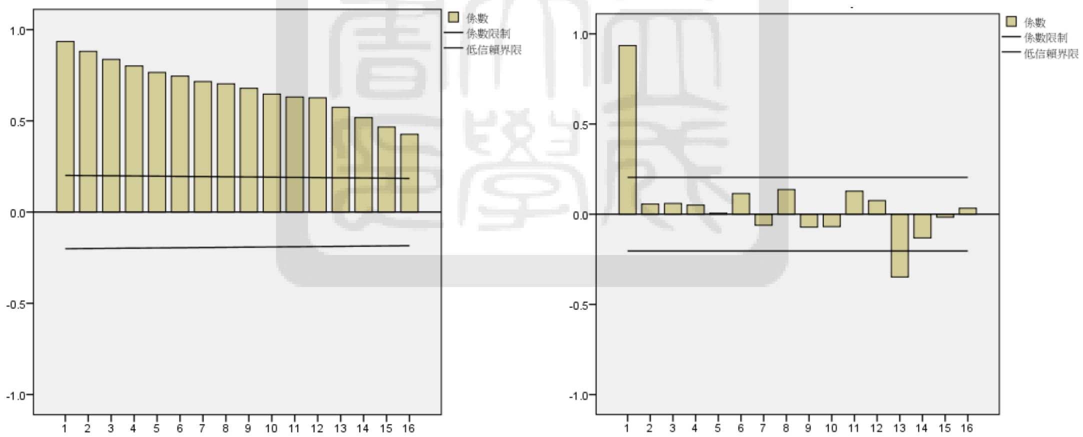
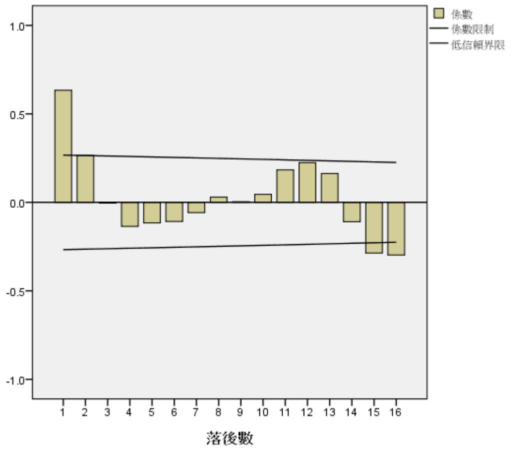


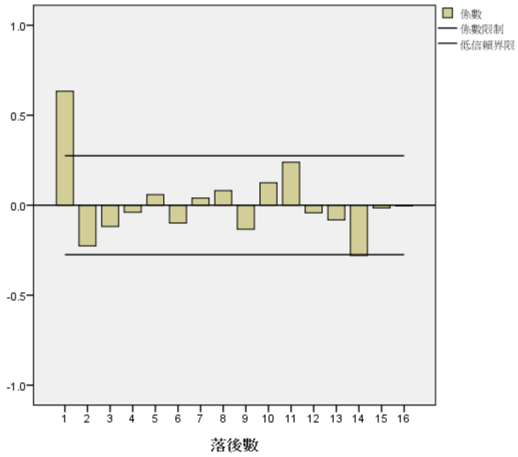
Figure 3-2. ACF and PACF of CityBike departure of scenarios 1~2.

Category 3- suburban districts

ACF

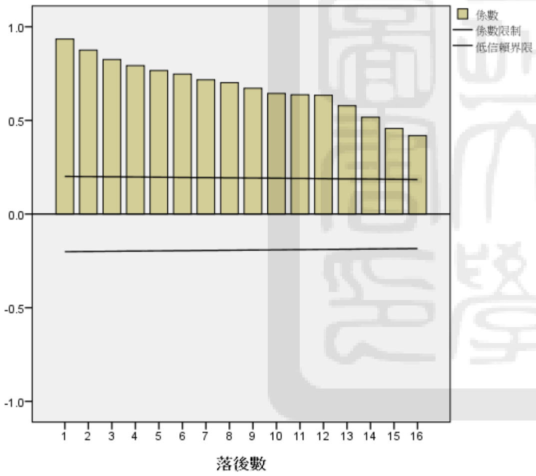


PACF



Category 4- Sanmin District

ACF



PACF

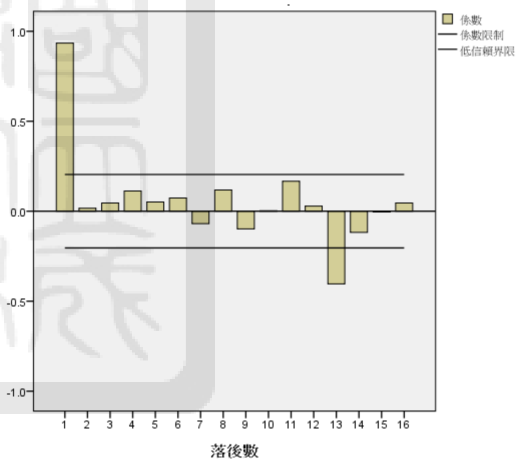
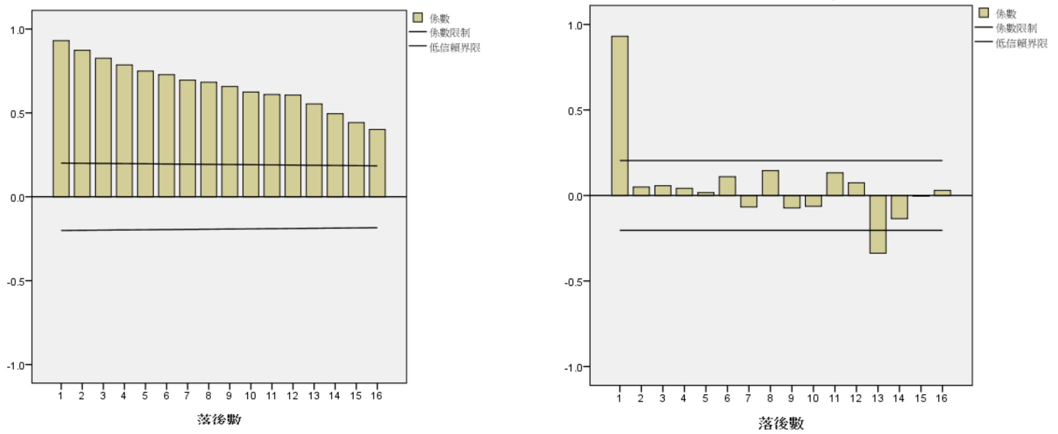


Figure 3-3. ACF and PACF of CityBike departure of scenarios 3~4.

Category 1-citywide

ACF PACF



Category 2- urban districts

ACF PACF

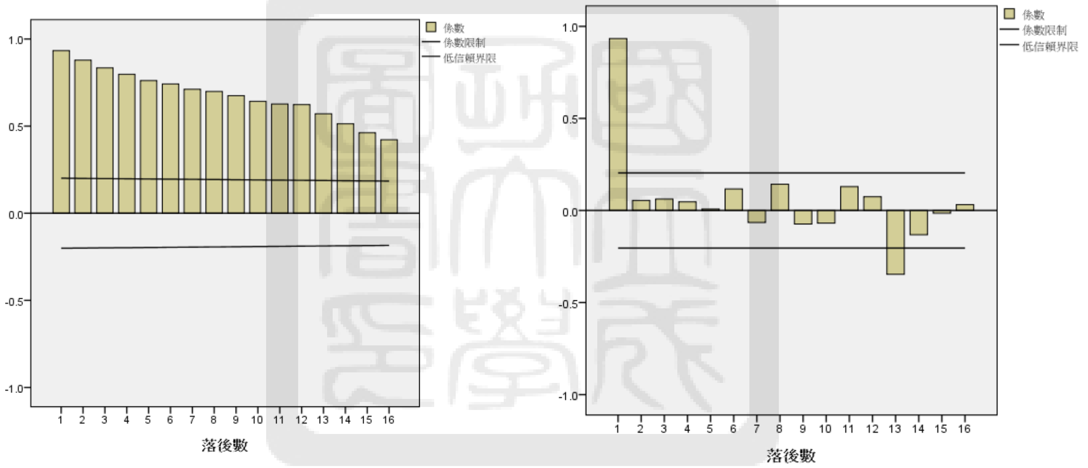
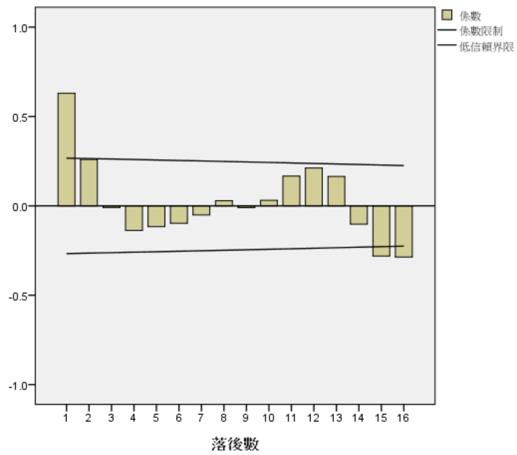


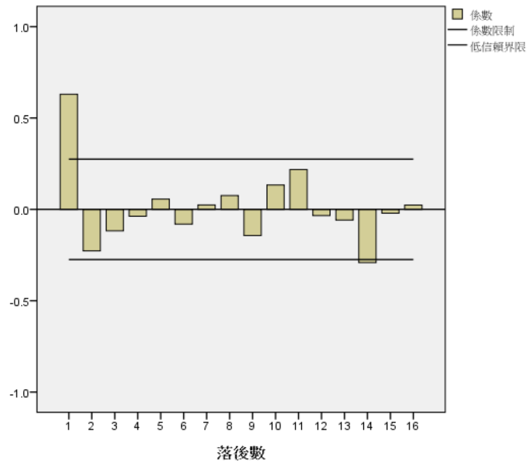
Figure 3-4. ACF and PACF of CityBike arrival of scenarios 1~2.

Category 3- suburban districts

ACF

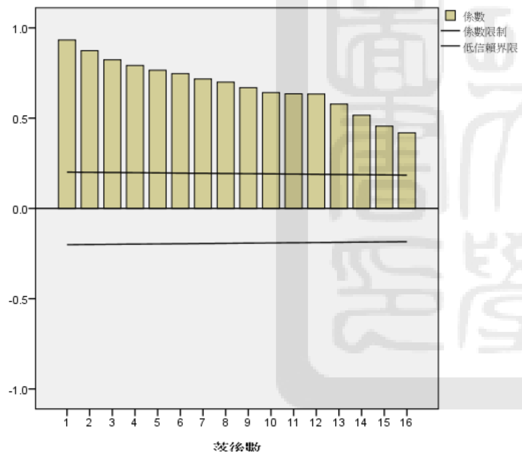


PACF



Category 4- Sanmin District

ACF



PACF

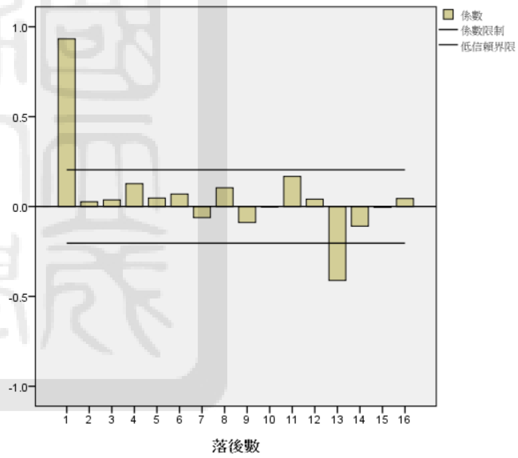


Figure 3-5. ACF and PACF of CityBike arrival of scenarios 3~4.

In Eq. (3.20), where β_1 is the coefficient of the IV, $Ridership_{t-1}$, represents the ridership data at time ($t - 1$). In Eq. (3.20), β_0 is a constant term, and ε_t is a random error term at time t. Referring to Figures 3-2 to 3-5, there were correlations between the ridership variable of a series and others from the same series separated by time t in citywide, urban, and individual districts, but not suburban districts. For the other variables, most of them were first-order autocorrelation. However, there was no autocorrelation for the tourist variable because the variation in tourists was affected by the different locations of POIs and the travelling period.

The AR model predicts the current ridership of a BSS on the basis of historical ridership data. However, if significant changes associated with socioeconomic characteristics of the built environment occur, the AR model may produce erroneous ridership predictions because it relies solely on previous ridership information. Under such circumstances, Eqs. (3.1) to (3.3) and (3.20) can be combined into one ARX model (Box and Jenkins 1987; Kendall and Ord 1990; Yun et al. 2012; Sarwar et al. 2017), where the X in ARX represents the exogenous inputs, which are the same as the IVs in Eqs. (3.1) to (3.3). The ARX (1) model for BSS ridership prediction is expressed in Eqs. (3.21) to (3.23).

$$Ridership^{(i)}{}_{s,d,t} = \beta_0 + \beta_1 \cdot 1stLIV_{s,d,t} + \beta_4 \cdot Ridership^{(i)}{}_{s,d,t-1} + \varepsilon_{s,d,t}, \quad (3.21)$$

$$Ridership^{(i)}{}_{s,d,t} = \quad (3.22)$$

$$\beta_0 + \beta_1 \cdot 1stLIV_{s,d,t} + \beta_2 \cdot 2ndLIV_{s,d,t} + \beta_4 \cdot Ridership^{(i)}{}_{s,d,t-1} + \varepsilon_{s,d,t},$$

$$Ridership^{(i)}{}_{s,d,t} =$$

$$\beta_0 + \beta_1 \cdot 1stLIV_{s,d,t} + \beta_2 \cdot 2ndLIV_{s,d,t} + \beta_3 \cdot 3rdLIV_{s,d,t} + \beta_4 \cdot \quad (3.23)$$

$$Ridership^{(i)}{}_{s,d,t-1} + \varepsilon_{s,d,t},$$

In Eqs. (3.21) to (3.23), β_i is the coefficient of the *i*-th input variable and ε_t is a random error term. The LSM is a standard method for calibrating the coefficients and error covariance of an ARX model's random error terms. Specifically, $\boldsymbol{X}$ in Eq. (3.24) is a matrix that includes all the IVs listed in Table 3-3 of the CityBike stations at each time point, where the index *n* is the number of time points. $\boldsymbol{Y}$ is the observed ridership at the *sth* station and in the *d-th* district type at various time points in Eq. (3.25). The values of $\boldsymbol{\beta}$ can be estimated using the LSM as in Eq. (3.28). The values of X, $\boldsymbol{\beta}$, ε, and $\boldsymbol{Y}$ can be calculated using Eqs. (3.24) to (3.28), where the index "n" is the number of time steps.

$$\boldsymbol{X}=\begin{bmatrix} 1 & 1stLIV_{1,1,1} & 2ndLIV_{1,1,1} & 3rdLIV_{1,1,1} & {Ridership^{(i)}}_{1,1,0} \\ 1 & 1stLIV_{2,1,2} & 2ndLIV_{2,1,2} & 3rdLIV_{2,1,2} & {Ridership^{(i)}}_{2,1,1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1stLIV_{k,1,n} & 2ndLIV_{k,1,n} & 3rdLIV_{k,1,n} & {Ridership^{(i)}}_{k,1,n-1} \end{bmatrix} \tag{3.24}$$

$$\boldsymbol{Y} = \left({Ridership^{(i)}}_{1,1,1}\ {Ridership^{(i)}}_{2,1,2}\ {Ridership^{(i)}}_{3,1,3} \cdots {Ridership^{(i)}}_{k,1,n}\right)^T \tag{3.25}$$

$$\widehat{\boldsymbol{\beta}}=(\boldsymbol{X^T X})^{-1}\boldsymbol{X^T Y} \tag{3.26}$$

$$\boldsymbol{\varepsilon}=\boldsymbol{Y}-\boldsymbol{X}\cdot\boldsymbol{\beta} \tag{3.27}$$

$$\boldsymbol{\beta}=(\boldsymbol{\beta_0}\quad \boldsymbol{\beta_1}\quad \boldsymbol{\beta_2}\quad \boldsymbol{\beta_3}\quad \boldsymbol{\beta_4})^{\boldsymbol{T}} \tag{3.28}$$

In this chapter, we showed how the BSS ridership forecasting problem for the CityBike system, considering that population, trip attributes, and relative factors are formulated by HRM and ARX models. This research considered trip factors to be a potential predictor of BSS ridership and ordered the factors determined by previous studies in priority of importance. Therefore, this research used the HRM to define the importance of input factors and compare the differences between population and trip factors for the CityBike ridership. Different from previous studies, the nine years of CityBike data collected for this research are of a time-series nature, which necessitated the ARX model and the hierarchical characteristics of HRM in order to analyze the effects. The research results of the two models are described in the next chapter under different district categories.

Chapter 4 EMPIRICAL STUDY AND RESULTS

In this section, the developed models are adopted in order to evaluate the performance of the target BSS. This research chose the CityBike system, located in Kaohsiung City, Taiwan for an empirical study. The CityBike system was the first BSS established in Taiwan and is the second-largest in the country. The outline of this empirical study of CityBike is as follows: Section 4.1 provides a background of the CityBike system, Section 4.2 describes the collected data of the CityBike system, and Section 4.3 presents the empirical results of the development models in different district categories.

4.1 Background Information

In Taiwan, the Taipei City First Generation Bike-Sharing System, which was launched in Taipei City Riverside Park on August 16, 1997, was a pilot BSS program providing bikes for recreational purposes (Figure 4-1). CityBike launched on March 1, 2009, in Kaohsiung City, while YouBike, the other BSS in Taipei, commenced its pilot program with 500 bicycles and 11 stations on March 11, 2009 (Figure 4-2). Both CityBike and YouBike are designed for various trip purposes, including work, recreation, and first- and last-mile journeys to other public transportation hubs.

Figure 4-1. Taipei City first generation bike-sharing system (Source: Taipei City Government, 2020).

Figure 4-2. Taipei City YouBike system (Source: YouBike Corporation, 2020).

Kaohsiung City is located in southern Taiwan and comprises 38 districts (Figure 4-3). The city is the third-largest in Taiwan, with a population of 2.74 million as of 2021. CityBike stations equipped with physical docking bays are found throughout the main CBD and in some outer districts. A total of 19 districts have bicycle rental stations, 12 of which are in the CBD, while the others are subdistricts (Figure 4-4). In the first stage of the CityBike operations, the Kaohsiung Environmental Protection Bureau installed 20 rental stations with 300 bicycles on March 1, 2009. In the second stage, additional 49 rental stations and 700 bicycles were added to the system on May 1, 2009. On August 18, 2011, the Kaohsiung Rapid Transit Corporation (KRTC) was appointed to operate and maintain the system. Finally, the system expanded again to 314 rental stations and 5,628 bicycles on March 25, 2020. After the contract between the government and the operator expired, the YouBike system was selected to replace CityBike on July 1, 2020. During the initial stages of the CityBike operation, only credit cards could be used for rentals, with complimentary use for the first hour and a charge of NT$10 for every 30 minutes thereafter. The Kaohsiung Environmental Protection Bureau also provided a NT$4 subsidy for two-way transfers between the Kaohsiung City metro system and CityBike. These strategies encouraged a gradual increase in CityBike ridership during its early stages of operation. From December 5, 2011, passengers could register iPASS cards (a type of e-ticket smart card used in Taiwan) for rentals, which led to substantial increases in CityBike ridership from January 2012 onwards. In 2018, other e-ticket smart cards such as EasyCard, HappyCash, and iCash were included as authorized payment methods on the CityBike system, and the fare structure was modified depending on the rental duration (Hu and Liu 2014; Kaohsiung Public Bike 2020).

Figure 4-3. Kaohsiung City districts (Reprinted from Kaohsiung City Government, 2020a)

Figure 4-4. City-bike stations in Kaohsiung City. (Source: Kaohsiung Public Bike, 2020)

Urbanization and economic development have led to significant growth in population and vehicle ownership in most metropolitan areas worldwide. Kaohsiung City is currently home to more than 3 million private vehicles, including cars and motor scooters. Because of the high vehicle usage rate, the BSS was introduced to improve accessibility to and from MRT stations and bus stops, to provide a greener transportation alternative for users, and serve as a convenient feeder mode to meet commuter needs. A comprehensive evaluation of the costs related to both operators and users needs to be conducted to ensure the financial sustainability of the CityBike system. Generally, in Taiwan, the budget for transportation development in a city's financial plan is somewhat lower for other public services. In Kaohsiung City, the total budget for transportation development is NT$1.976 billion, which is only 1.6% of the general budget. Since 2011, the Kaohsiung Environmental Protection Bureau (KEPB) has subsidized the KRTC by providing approximately NT$10–12 million in financial aid every year to support the CityBike system. However, this financial support is not guaranteed because the funds allocated for promoting green transport modes change each year. Thus, operating the CityBike system in a cost-effective manner is a key problem that must be addressed by both the government agencies and the KRTC.

Before 1946, Kaohsiung City had only ten districts. Later, as industrial development progressed, the population grew to one million, and it became the second-largest city in Taiwan in 1979. Currently, the city's vision is to get rid of its heavy industry by developing green environmental, cultural, technical, and natural elements. To ensure the city develops sustainability in the future, a seamless transportation system is an essential component and is supported by the government. Kaohsiung City is unique in that it has several forms of public transport, including a ferry, a metro, high-speed rail, a regional railway, light rail, city and intercity bus lines, and a BSS, all of which provide citizens with many different ways to carry out their daily commute. Therefore, determining how to seamlessly integrate these systems is an important issue.

The CityBike system was implemented for both recreational and commuting purposes and is a suitable feeder mode in the city for connecting with other transportation modes.

In general, bicycle rental stations were installed near Kaohsiung's MRT stations and popular tourist attractions, with their installation based on the professional judgment of engineers with no systematic evaluation of cost-effectiveness. This research considers different types of variables both at the station and district levels, including spatial variables (i.e., station capacity, capacity in the buffer range, and POI), socioeconomic variables (i.e., land use and demographics), and infrastructure variables (i.e., bike lane lengths, number of lamps, road lengths) to comprehensively investigate the key factors that may impact on the ridership of the CityBike system. In particular, trip attributes (i.e., trip purpose, trip assignment, number of trip generations, and attractions) are considered IVs in the developed models. The previous studies reviewed in Chapter 2 have revealed that population is one of the key factors that influence BSS ridership. This research collected data on population, trips, and the number of CityBike stations from the districts shown in Figures 4-5 and 4-6.

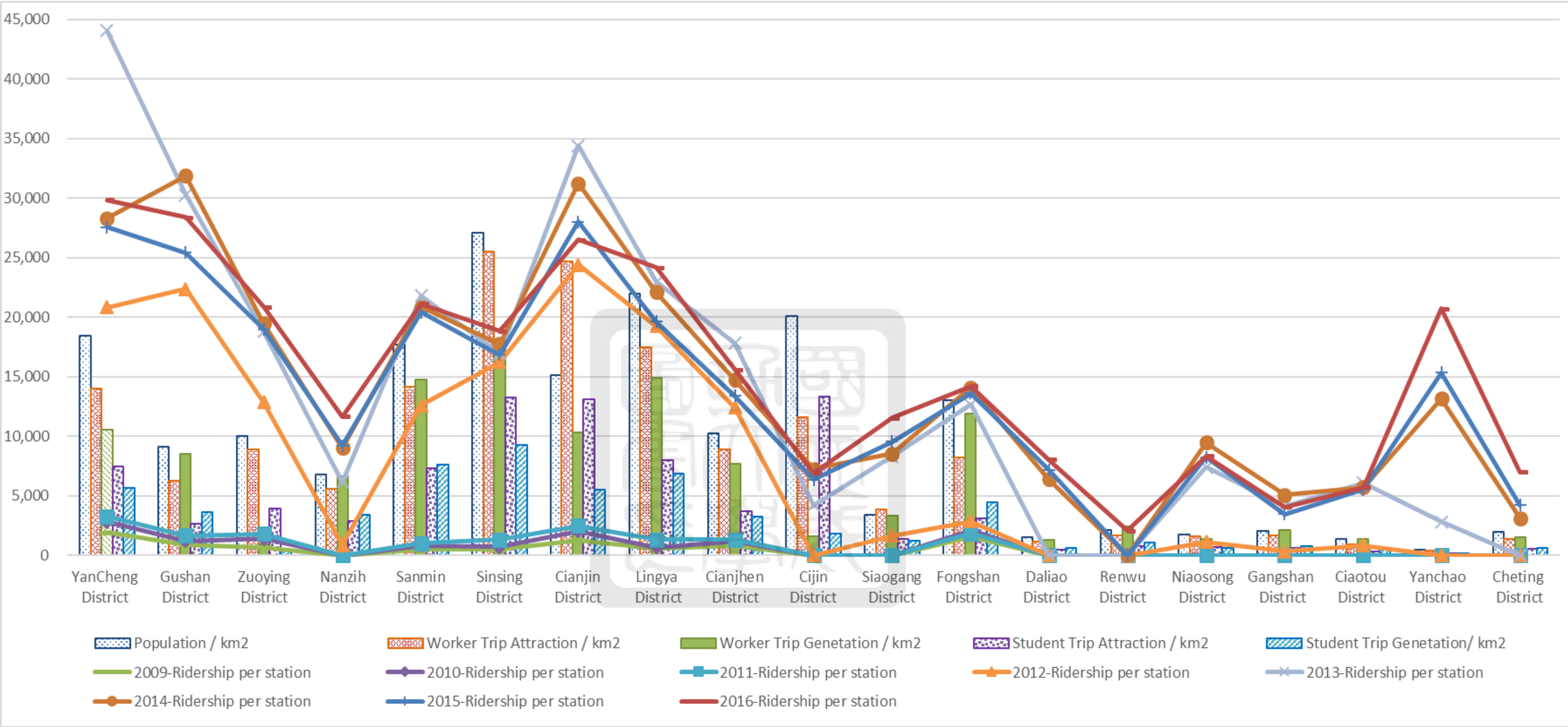


Figure 4-5. Distributions of population, trips and CityBike ridership of each district

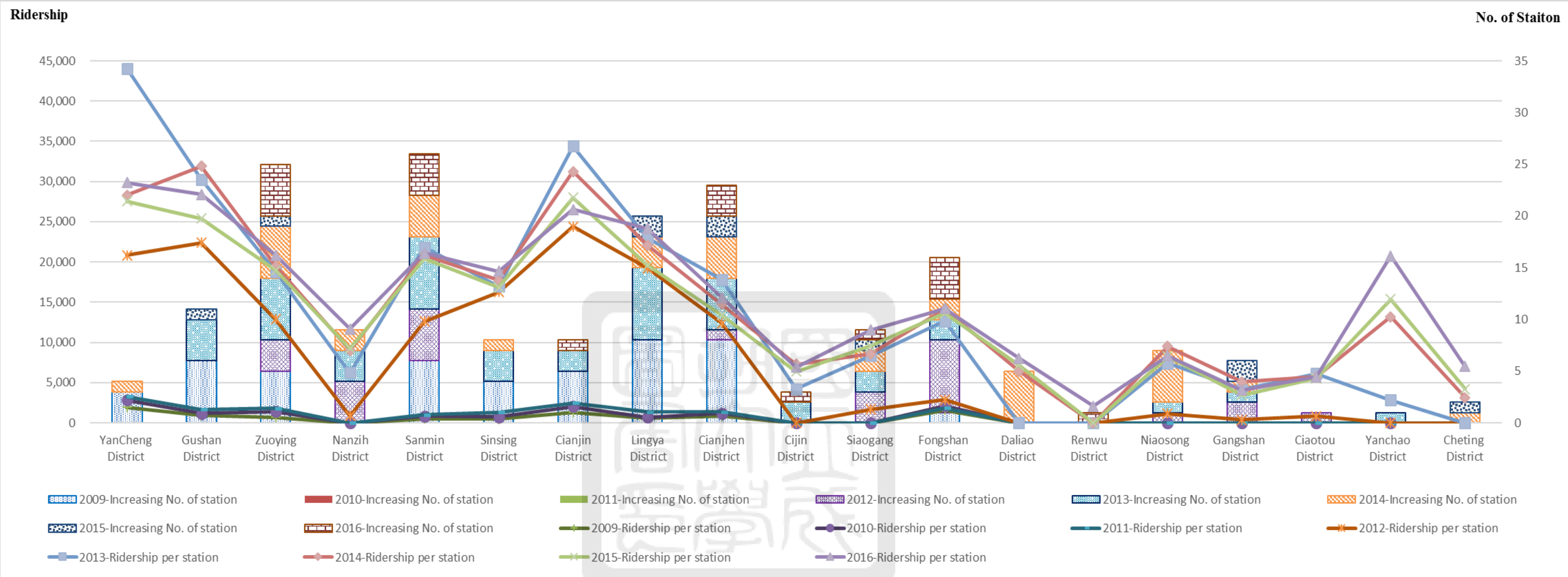


Figure 4-6. Distributions of CityBike ridership and number of stations.

Figure 4-5 presents statistics for population, home-based work trips, home-based school trips, and ridership in the districts of central Kaohsiung City. The population and trip indexes only showed positive correlations in the Samin, Sinsing, Cianjin, and Lingya Districts meaning that CityBike ridership was above average in the aforementioned districts. Although the socioeconomic indexes of Yancheng, Gushan, and Zuoying Districts were lower than in other districts, these districts, particularly Gushan District, had higher than average CityBike ridership, indicating that there are unrevealed interactions among trip attributes, socioeconomic characteristics, and CityBike ridership. For example, there are not enough bicycles to rent in BSS stations, which might cause a bias in BSS ridership. Chung and Huang (2015) examined disposition criteria and established supplementary criteria for Taipei's YouBike system. They found that YouBike stations confront issues such as fewer bikes or fewer available parking slots, which could otherwise unbalance rental demand. In this situation, users are unable to rent or return a bicycle to a BSS station, which causes estimated errors in BSS ridership prediction. In addition, because the municipal government closed or relocated some of the rental stations in Yancheng District, the ridership also varied significantly between 2012 and 2014. Figure 4-6 shows the ridership and number of stations in districts centralized in Kaohsiung City between 2009 and 2016. Gushan District has the higher ridership but also the second-lowest number of rental stations in urban district.

The number of rental stations in Cianjin District is similar to those in Nanzih District, Sinsing District, and Siaogang District, but ridership is much higher than in these other areas. The data seem to show that if an area increases its number of BSS stations annually, ridership will rise simultaneously. However, it is not necessarily true that ridership in a particular area is always positively related to the number of CityBike stations installed. There may be areas that have too many stations that adversely affect the overall performance of the BSS.

4.2 Data Collection

The literature review revealed that the factors contributing to BSS ridership include demographics, socioeconomic characteristics, and spatial variables (i.e., bicycle infrastructure, land use, POI, and types of public transport). Trip attributes (trip generation and attraction flow classified by trip purpose) are the particular factors collected in this research for comparison with the effect of population effect. This research collected relevant data from government offices and the KRTC, with the descriptive statistics for the explanatory variables presented in Table 4-1 and Appendices A-13 to A-14.

Table 4-1. Descriptive statistics of the collected data

Variables	Min	Max	Mean	Std. deviation	Data resource
Station-level data					
CityBike departures (per station/ per month)	0	9,889	1,315	1,231	KRTC
CityBike arrivals (per station/ per month)	0	10,013	1,317	1,218	
Capacity of station (units)	12	32	28	6	
CityBike station capacity in a 1-km buffer (units)	0	254	83	57	
District-level data					
Population (people/ per square kilometer)	459	28,205	13,326	6,677	DOBAS, TBKCG, DOSKCG
Worker population (worker/ per square kilometer)	311	18,601	9,016	4,458	
Student population (student/ per square kilometer)	59	3,466	1,840	822	
Tourist (people/ per square kilometer)	0	711,902	17,078	50,891	
Income (NT$/ year)	399,531	826,201	545,160	55,074	
Private vehicle (units / per square kilometer)	475	30,923	14,102	7,136	
Bikeway length (meter/ per square kilometer)	0	58,482	4,587	9,353	PBWKCG
Number of streetlamps (units / per square kilometer)	124	3,095	1,036	703	
Length of major road (meter/ per square kilometer)	6,280	35,116	17,305	6,555	LABKCG
Land value (NT$/ per square kilometer)	470	245,834	38,661	42,910	
Number of parks (units / per square kilometer)	0	12	4	2	DOSKC, MOTC
Number of companies (units / per square kilometer)	8	2,348	648	586	
Number of factories (units / per square kilometer)	2	29	10	8	
Number of hotels (units / per square kilometer)	0	85	12	22	
Number of markets (units / per square kilometer)	0	6	1	1	
Number of schools (units / per square kilometer)	0	4	2	1	
Number of public transportation yards (units / per square kilometer)	0	25	6	7	TBKCG
Trip attraction (trips/ per square kilometer)	1,152	84,034	33,983	22,034	DOSKC, MOTC (Census data)
Trip generation (trips/ per square kilometer)	1,275	58,913	26,760	16,728	
Worker trip attraction (trips/ per square kilometer)	147	13,782	5,612	3,668	
Worker trip generation (trips/ per square kilometer)	163	9,662	4,344	2,724	

Student trip attraction (trips/ per square kilometer)	399	26,555	11,571	6,612
Student trip generation (trips/ per square kilometer)	441	18,616	9,177	5,366
Simple size		11,474		

Note:
1. KRTC: Kaohsiung Rapid Transit Corporation
2. DOBAS: Department of Budget, Accounting, and Statistics
3. TBKCG: Transportation Bureau of Kaohsiung City Government
4. DOSKCG: Department of Statistics of Kaohsiung City Government
5. PBWKCG: Public Work Bureau of Kaohsiung City Government
6. LABKCG: Land Administration Bureau of Kaohsiung City Government
7. MOTC: Ministry of Transportation and Communications

Referring to Table 4-1 and Appendices A-13 to A-14, there were significant differences in the same variables in different district scales because of differences in the demographic and socioeconomic characteristics. The effects of these variables can impact CityBike ridership in different ways, depending on the district. Based on the summary of the literature review in Chapter 2, this research collected 27 indicators of demographics, trip factors, and related spatial factors for analysis. Previous studies (Rixey 2013; Faghih-Imani et al. 2014; El-Assi et al. 2015; Médard de Chardon 2015; Faghih-Imani and Eluru 2016b; Faghih-Imani et al. 2017b; Sun et al. 2018; Zhao et al. 2021) had focused on station-level data in small areas and short-term BSS ridership forecasting and did not take into account the impacts of socioeconomic factors on BSS ridership at a district level. In order to fill this gap in the literature, this research combined station-level, and district-level data, which were obtained from the KRTC and other Taiwanese government agencies, for a comprehensive analysis of the CityBike ridership. The basic concepts and terms related to the developed models for this research are defined below, and the model's development process will be based on these definitions.

1. Station-level and district-level data: Previous studies collected relevant BSS data obtained at the station level within a buffer of a nearby bike station (Rixey 2013; Faghih-Imani et al. 2014; El-Assi et al. 2015; Médard de Chardon 2015; Faghih-Imani and Eluru 2016b; Faghih-Imani et al. 2017b; Sun et al. 2018; Zhao et al. 2021). Only a few studies considered a comprehensive picture of target areas. This research collected relative demographic and socioeconomic characteristics of each district of Kaohsiung City as district-level data for the following reasons.
   (1) By dividing Kaohsiung City into "district" levels, sufficient samples from the CityBike system can be obtained for analysis. Subdividing

the areas further to, e.g., “village” level, would not yield enough samples for this research.

(2) Most of the census data or annual socioeconomics data in Kaohsiung City were collected by district. Only a few categories were collected from smaller areas (e.g., village level) and not on an annual basis.

Therefore, this research collected relative demographic and socioeconomic characteristics of each district of Kaohsiung City as district-level data, including trip generation and attraction, population, number of POI, bike infrastructure, and road infrastructure. The station-level data included CityBike usage, station capacity, and CityBike capacity within a 1-km buffer.

2. CityBike arrivals and departures: The monthly ridership data of all the CityBike stations were collected from 2009 to 2017, with arrivals (returning the bike) and departures (renting the bike) calculated separately. Those data are considered DVs and station-level data in the developed models in this research.

3. Trip attractions and generations: The population size does not necessarily represent the actual number of trips per day in a given area. The worker and student population of Kaohsiung City is approximately 1.92 million (Transportation Bureau of Kaohsiung City Government 2009; Department of Statistics 2016). Districts with higher worker and student populations will generate more home-based trips to workplaces and schools. Kaohsiung City residents generate 6.12 million trips per day, which equates to 2.21 trips per person per day. In the trip purpose index, the sum of home-based work and school trips comprised 52.6% of all trips, indicating that these trips make up the bulk of journeys undertaken by Kaohsiung City residents. To obtain the “true” influence of population on BSS ridership, a critical step is to consider trip attributes and BSS ridership as an interactive process. In this research, the number of trips made in each district was collected as an independent variable, which would be compared with the population effect. The number of trips in each district was modeled as an IV. Trip attributes, including the number of trip generations and attractions, the number of student trip generations and attractions, and the number of worker trip generations and attractions were obtained from the Kaohsiung City census data. These three trip types (student trip demand, worker trip demand, and other trip demand) accounted for 92.1% of all trips generated in Kaohsiung City. This research estimated the number of trip generations and attractions of the districts in Kaohsiung City from 2009 to 2016 based on the census data

and annual population growth rate. This research believed that the developed models using the "estimated" trip data were superior to the "register" population data on BSS ridership forecasting. Those data were not calculated by the travel demand model of traditional transportation planning models.

4. Population data: Previous studies used the total population of a specific area as an influencing factor. In this study, in order to compare the impact of the trip factors (i.e., number of student trip generations and attractions, number of worker trip generations and attractions) with population factors on CityBike usage, working population, and school population data were collected separately in target regions for modeling.
5. Socioeconomic data: The monthly related socioeconomic characteristics of the CityBike system were collected from the different organizations of the Kaohsiung City municipal government. These data are considered district-level data of the developed model in this research.
6. Model samples: In the hierarchical process, the multiple regression method is adopted to analyze the effect of independent variables with sufficient samples. Calculating how many samples need to be collected in the development model is a key step, and there are no similar studies that can be referred to. Therefore, this research refers to Cohen's (1988) research, under a certain level of type I error and type II error (α=0.05, β=0.8) and medium effect size (ES, ES=0.15, R^2=0.13). This research used G-power software based on previous settings and the number of variables used in the models to calculate the required sample size. For the ARX model, we used the time-series method and collected the monthly data of each variable as required samples.
7. District categories design: The collected variables of urban districts, suburban districts, and individual districts have different impacts on the CityBike ridership forecasting. In this research, in order to examine the impact of various factors on the CityBike ridership on a different scale, four district categories, which differed in terms of scale were set up and named as citywide, urban district, suburban district, and individual district.

4.3 Results and Discussion

4.3.1 Model Calibration

The developed models in this research were used to evaluate the district categories. Key statistical information, including the t-test results, estimated coefficients, significance test results, adjusted R^2, and Durbin–Watson statistics, are described and presented in the following tables. To determine the goodness of fit for the developed models, an F-test, the Akaike information criterion (AIC), and the Bayesian information criterion (BIC) was adopted. An F-test is typically used to analyze statistical models with more than one parameter to determine the suitability of some or all of the parameters in the model for data estimation. The AIC and BIC were used to determine the goodness of fit of the developed models. Specifically, the model with the smallest AIC or BIC value has the best fit (Akaike 1974; Schwarz 1978).

The empirical study results were analyzed using the HRM and ARX methods for the citywide, urban district, suburban district, and individual district levels. The factors related to arrivals and departures in the CityBike system were also investigated. The development models also compared population factors with trip factors on BSS ridership in different area scales. The results of the four district categories are shown in Tables 4-2 to 4-5.

Table 4-2. Models M1 and M2 estimation results of citywide, urban district, suburban district, and individual district (CityBike departures, IV: trip).

Model / IV	M1 Citywide	M1 Urban district	M1 Suburban district	M1 Sanmin District	M2 Citywide	M2 Urban district	M2 Suburban district	M2 Sanmin District
Constant	512.285	964.092	536.566	14640.941	-49574.829	-278409.363	2363.947	84289.017
	(0.538)	(1.136)	(1.095)	(4.489***)	(-4.082***)	(-4.691***)	(0.582)	(7.167***)
Lag 1	-	-	-	-	0.583	0.519	0.329	0.413
	-	-	-	-	(6.156***)	(6.805***)	(2.373**)	(5.152***)
Worker trip attraction	-	-	-	-				
	-	-	-	-				
Worker trip generation	-	-	-	-				-0.803
	-	-	-	-				(-6.998***)
Student trip attraction	-	-	-	-				
	-	-	-	-				
Student trip generation	-	-	-	-	0.460	0.306	0.088	
	-	-	-	-	(5.696***)	(4.752***)	(0.560)	
Capacity of station	-0.081	0.028	-0.005	-0.089				
	(-0.897)	(0.313)	(-0.061)	(-0.899)				
CityBike station capacity in a 1-km buffer	0.375	0.117	-0.013	0.063				-0.234
	(3.058**)	(1.123)	(-0.134)	(0.645)				(-2.638***)
Bikeway length	.0343		-0.167				0.109	
	(2.040**)		(-1.090)				(0.677)	
Number of streetlamps			-0.125					
			(-1.334)					
Length of major road	0.766	0.990	0.519				-0.267	
	(3.834**)	(4.483***)	(3.139**)				(-1.514)	
Number of parks	-0.066	0.664		0.726				
	(-0.400)	(3.619***)		(5.095***)				
Number of companies			0.291					
			(2.598**)					
Number of factories	-0.157	-0.317		-0.379	-0.302	0.036		-0.145
	(-1.432)	(-3.256**)		(-2.979***)	(-5.209***)	(0.757)		(-3.816***)
Number of hotels	-0.546	-0.416						-0.238
	(-3.137**)	(-2.181**)						(-4.344***)
Number of markets	-0.307	0.145						
	(-1.875*)	(0.749)						
Number of schools		-0.989				0.005		
		(-4.081***)				(0.083)		
Number of public transportation yard					-0.192	-0.038		
					(-4.529***)	(-1.043)		
Tourist	-0.282	-0.331		0.165	0.120	0.122	0.213	-0.003
	(-2.109**)	(-2.423**)		(1.830*)	(3.417***)	(3.473***)	(1.949*)	(-0.110)
Income	-0.010	-0.067	0.006	-0.720	-0.204	-0.091		-0.024
	(-0.099)	(-0.712)	(0.067)	(-4.055***)	(-4.645***)	(-2.652**)		(-0.697)
Private vehicle			-1.168		-0.158	-0.118	-0.389	-0.127
			(-7.246***)		(-3.212***)	(-2.233**)	(-2.274**)	(-1.353)
Land value	-0.244	-0.257	0.023	-0.523				
	(-1.384)	(-1.630)	(0.242)	(-3.089***)				
Adjusted R^2	0.222	0.311	0.604	0.349	0.940	0.937	0.53	0.935
F value	3.731***	5.299***	16.939***	7.596***	211.979***	179.005***	10.763***	171.682***
DW	1.491	1.503	1.775	0.903	1.735	1.763	1.778	1.642
AIC	1393.699	1366.264	1015.900	1088.114	1004.780	1020.148	502.233	1046.428
BIC	1422.997	1395.561	1038.885	1105.376	1022.731	1040.663	514.054	1066.942

Note: 1. *90% level of significance, ** 95% level of significance, *** 99% level of significance
2. *t*-statistics values are shown in parenthesis.

Table 4-3. Models M1 and M2 estimation results of citywide, urban district, suburban district, and individual district (CityBike departures, IV: population).

Model / IV	M1 Citywide	M1 Urban district	M1 Suburban district	M1 Sanmin District	M2 Citywide	M2 Urban district	M2 Suburban district	M2 Sanmin District
Constant	275.552 (0.241)	1542.858 (1.688*)	1497.651 (3.403***)	126138.441 (18.899***)	-99320.154 (-6.258***)	-102125.42 (-2.75***)	2042.688 (0.444)	84565.738 (7.186***)
Lag 1	- -	- -	- -	- -	0.283 (3.753***)	0.547 (9.380***)	0.329 (2.373**)	0.411 (5.132***)
Worker population				-1.432 (-17.145***)		0.141 (3.045***)	0.088 (0.560)	
Student population	-0.092 (-0.377)	-0.119 (-0.603)	-1.030 (-6.722***)		0.916 (8.033***)			-0.806 (-7.017***)
Capacity of station	-0.076 (-0.831)	0.060 (0.620)	-0.066 (-0.757)	-0.002 (-0.036)				
CityBike station capacity in a 1-km buffer	0.362 (3.067***)	0.198 (1.786*)	-0.002 (-0.020)	0.038 (0.853)				-0.234 (-2.640***)
Length of bikeway	0.296 (1.420)		-0.310 (-1.751)				0.109 (0.676)	
Number of streetlamps			-0.264 (-2.724)					
Length of major road	0.823 (2.282***)	0.574 (2.240**)	0.172 (1.244)				-0.267 (-1.514)	
Number of parks	-0.026 (-0.133)	0.278 (1.639)		-0.286 (-3.263***)				
Number of companies			0.334 (2.737***)					
Number of factories	-0.161 (-1.454)	-0.233 (-2.192**)		-0.219 (-3.686***)	-0.373 (-6.862***)	0.029 (0.582)		-0.146 (-3.850***)
Number of hotels	-0.592 (-2.772***)	-0.154 (-0.658)						-0.238 (-4.354***)
Number of markets	-0.288 (-1.680*)	-0.326 (-1.855*)				-0.124 (-1.929*)		
Number of schools					0.294 (4.266***)	-0.009 (-0.138)		
Number of public transportation yard					-0.219 (-5.652***)	-0.001 (-0.034)		
Tourist	-0.276 (-2.043**)	-0.131 (-0.846)		-0.158 (-3.456***)	0.101 (3.188***)	0.139 (3.787***)	0.213 (1.949*)	-0.003 (-0.114)
Income	0.017 (0.142)	-0.154 (-1.385)	-0.058 (-0.656)	-0.344 (-4.059***)	-0.207 (-5.066***)	-0.038 (-1.135)		-0.023 (-0.683)
Private vehicle					-0.387 (-6.480***)		-0.390 (-2.279**)	-0.125 (-1.338)
Land value	-0.221 (-1.178)	-0.203 (-1.171)	-0.111 (-1.208)	0.289 (3.164***)				
Adjusted R^2	0.215	0.191	0.582	0.862	0.952	0.930	0.530	0.941
F value	3.401***	3.260***	15.533***	68.038***	237.738***	159.221***	10.763***	172.034***
DW	1.506	1.448	1.81	1.306	1.562	1.961	1.778	1.642
AIC	1395.537	1383.146	1021.099	954.213	983.091	1030.711	502.233	1046.242
BIC	1427.498	1412.433	1044.084	973.941	1003.606	1051.225	514.054	1066.757

Note: 1. *90% level of significance, ** 95% level of significance, *** 99% level of significance
2. *t*-statistics values are shown in parenthesis.

Table 4-4. Models M1 and M2 estimation results of citywide, urban district, suburban district, and individual district (CityBike arrivals, IV: trip).

Model / IV	M1				M2			
	Citywide	**Urban district**	**Suburban district**	**Sanmin District**	**Citywide**	**Urban district**	**Suburban district**	**Sanmin District**
Constant	515.858	895.056	467.659	14524.152	-321843.151	-111292.356	704.127	86103.105
	(0.539)	(1.050)	(0.944)	(4.403***)	(-7.921***)	(-2.218**)	(0.159)	(7.191***)
Lag 1	-	-	-	-	0.242	0.709	0.322	0.407
	-	-	-	-	(3.155***)	(11.160***)	(2.297**)	(5.055***)
Worker trip attraction	-	-	-	-		0.143		-0.816
	-	-	-	-		(2.452**)		(-7.030***)
Worker trip generation	-	-	-	-				
	-	-	-	-				
Student trip attraction	-	-	-	-	0.721		0.143	
	-	-	-	-	(8.372***)		(0.905)	
Student trip generation	-	-	-	-				
	-	-	-	-				
Station capacity	-0.075	0.032	-0.021	-0.089				
	(-0.822)	(0.361)	(-0.232)	(-0.891)				
CityBike station capacity in a 1-km buffer	0.353	0.110	-0.022	0.062		-0.038		-0.233
	(2.911**)	(1.046)	(-0.223)	(0.634)		(-0.513)		(-2.597**)
Bikeway length	0.344		-0.192				0.148	
	(2.031**)		(-1.226)				(0.897)	
Number of streetlamps			-0.120					
			(-1.256)					
Length of major road	0.761	1.004	0.508				-0.312	
	(3.776***)	(4.514***)	(2.999***)				(-1.733*)	
Number of parks	-0.072	0.668		0.720				
	(-0.435)	(3.619***)		(5.027***)				
Number of companies			0.331					
			(2.884***)					
Number of factories	-0.146	-0.311		-0.379	-0.354	-0.059		-0.151
	(-1.322)	(-3.178***)		(-2.965***)	(-6.895***)	(-1.051)		(-3.908***)
Number of hotels	-0.549	-0.436						-0.235
	(-3.127***)	(-2.272**)						(-4.287***)
Number of markets	-0.310	0.152				-0.192		
	(-1.875*)	(0.781)				(-2.507**)		
Number of schools		-1.018			0.116			
		(-4.172***)			(2.011**)			
Number of public transportation yard					-0.136	0.007		
					(-4.093***)	(0.179)		
Tourist	-0.282	-0.319		0.162	0.106	0.159	0.202	-0.003
	(-2.096**)	(-2.317**)		(1.783*)	(3.406***)	(4.288***)	(1.822*)	(-0.109)
Income	-0.010	-0.056	0.043	-0.702	-0.242	-0.056		-0.023
	(-0.105)	(-0.590)	(0.491)	(-3.935***)	(-5.748***)	(-1.521)		(-0.686)
Private vehicle			-1.165		-0.454		-0.359	-0.113
			(-7.067***)		(-7.058***)		(-2.122**)	(-1.195)
Land value	-0.243	-0.250	-0.004	-0.519				
	(-1.365)	(-1.577)	(-0.039)	(-3.049***)				
Adjusted R^2	0.208	0.301	0.586	0.343	0.953	0.927	0.518	0.934
F value	3.512***	5.109***	15.757***	7.403***	241.638***	150.713***	10.298***	168.262***
DW	1.494	1.496	1.780	0.898	1.596	2.076	1.766	1.656
AIC	1394.597	1367.177	1018.018	1090.109	982.066	1035.923	503.283	1050.06
BIC	1423.894	1396.475	1041.003	1107.370	1002.581	1056.438	515.105	1070.575

Note: 1.*90% level of significance, ** 95% level of significance, *** 99% level of significance
2. *t*-statistics values are shown in parenthesis.

Table 4-5. Models M1 and M2 estimation results of citywide, urban district, suburban district, and individual district (CityBike arrivals, IV: population).

Model / IV	M1 Citywide	M1 Urban district	M1 Suburban district	M1 Sanmin District	M2 Citywide	M2 Urban district	M2 Suburban district	M2 Sanmin District
Constant	258.261	1470.473	1408.122	127597.4	-48495.33	-103918.36	517.441	86479.481
	(0.225)	(1.579)	(3.179**)	(19.086***)	(-4.391***)	(-2.751***)	(0.112)	(7.232***)
Lag 1	-	-	-	-	0.402	0.649	0.322	0.403
	-	-	-	-	(5.310***)	(9.440***)	(0.322**)	(5.011***)
Worker population				-1.443		0.142	0.144	
				(-17.359***)		(3.028***)	(0.905)	
Student population	-0.101	-0.140	-1.034		0.630			-0.820
	(-0409)	(-0.702)	(-6.630***)		(6.391***)			(-7.071***)
Station capacity	-0.070	0.064	-0.081	-0.001				
	(-0.752)	(0.660)	(-0.902)	(-0.016)				
CityBike station capacity in a 1-km buffer	0.358	0.193	-0.010	0.037				-0.233
	(3.004**)	(1.724*)	(-0.101)	(0.407)				(-2.602**)
Length of bikeway	0.294		-0.341				0.147	
	(1.394)		(-1.891*)				(0.896)	
Number of streetlamps			-0.260					
			(-2.633**)					
Length of major road	0.823	0.590	0.165				-0.313	
	(3.255**)	(2.28**)	(-1.171)				(-1.735*)	
Number of parks	-0.029	0.272		-0.299				
	(-0.145)	(1.591)		(-3.409***)				
Number of companies			0.378					
			(3.036**)					
Number of factories	-0.151	-0.228		-0.218	-0.353	0.029		-0.152
	(-1.347)	(-2.123**)		(-3.684***)	(-5.914***)	(0.579)		(-3.937***)
Number of hotels	-0.600	-0.180						-0.237
	(-2.781**)	(-0.758)						(-4.323***)
Number of markets	-0.289	-0.326				-0.115		
	(-1.671*)	(-1.836*)				(-1.787*)		
Number of schools						-0.012		
						(-0.177)		
Number of public transportation yard					-0.218	-0.002		
					(-5.122***)	(-0.048)		
Tourist	-0.276	-0.120		-0.164	0.115	0.142	0.202	-0.004
	(-2.027**)	(-0.767)		(-3.603***)	(3.31***)	(3.805***)	(1.822*)	(-0.115)
Income	0.019	-0.141	-0.020	-0.323	-0.237	-0.038		-0.023
	(0.156)	(-1.253)	(-0.219)	(-3.834***)	(-5.303***)	(-1.131)		(-0.672)
Private vehicle					-0.240		-0.358	-0.113
					(-4.504***)		(-2.114**)	(-1.199)
Land value	-0.218	-0.192	-0.137	0.300				
	(-1.150)	(-1.097)	(-1.46)	(3.294***)				
Adjusted R^2	0.201	0.176	0.566	0.863	0.942	0.928	0.518	0.934
F value	3.205***	3.036**	14.638	68.769***	222.579***	154.886***	10.298***	169.031***
DW	1.51	1.445	1.811	1.294	1.665	1.964	1.766	1.655
AIC	1396.406	1384.633	1022.33	954.5	1000.821	1033.475	503.283	1049.649
BIC	1428.367	1413.931	1045.315	974.227	1018.771	1053.99	515.105	1070.163

Note: 1. *90% level of significance, ** 95% level of significance, *** 99% level of significance
2. *t*-statistics values are shown in parenthesis.

4.3.2 Results of Hierarchical Level in M1/M2

Most of the previous studies did not rank the variables affecting BSS ridership forecasting, and only put the effective factors into development models. In order to understand the importance and influence of each factor on BSS ridership and establish criteria, this research referred to Cohen's (2003) theoretical and practical concepts and used the hierarchical concepts of the HRM method to explore each factor. The factors were sorted into hierarchical levels (see Table 3-2) by the HRM and ARX models and the results are shown in Table 4-6. Specifically, this research listed the effects of population and trip factors separately.

Table 4-6. The results of IV in hierarchical level in M1/M2

Category 1: Citywide			First level		Second level		Third level	
			Adj. R^2	**F value**	**Adj. R^2**	**F value**	**Adj. R^2**	**F value**
HRM (M1)	IV population	DV departures	0.017	1.799	0.251	3.899***	0.037	3.401***
		DV arrivals	0.014	1.429	0.242	3.657***	0.037	3.205***
	IV trip	DV departures	0.008	0.824	0.256	3.827***	0.043	3.428***
		DV arrivals	0.015	1.541	0.236	3.575***	0.042	3.208***
ARX (M2)	IV population	DV departures	0.912	479.036***	0.01	209.944***	0.035	237.738***
		DV arrivals	0.909	467.054***	0.007	248.301***	0.03	222.579***
	IV trip	DV departures	0.911	475.237***	0.007	254.057***	0.026	13.729***
		DV arrivals	0.914	493.126***	0.002	196.120***	0.041	241.638***
Category 2: Urban district			**First level**		**Second level**		**Third level**	
			Adj. R^2	**F value**	**Adj. R^2**	**F value**	**Adj. R^2**	**F value**
HRM (M1)	IV population	DV departures	0.002	0.157	0.237	3.796***	0.038	3.260***
		DV arrivals	0.001	0.063	0.229	3.621***	0.032	3.036***
	IV trip	DV departures	0.011	1.175	0.335	5.640***	0.039	4.850***
		DV arrivals	0.014	1.499	0.286	5.202***	0.048	4.558***
ARX (M2)	IV population	DV departures	0.915	499.026***	0.008	177.798***	0.013	159.221***
		DV arrivals	0.913	487.579***	0.008	172.656***	0.014	154.886***
	IV trip	DV departures	0.923	554.158***	0.001	219.106***	0.019	179.005***
		DV arrivals	0.913	487.579***	0.008	172.656***	0.014	154.886***
Category 3: Suburban district			**First level**		**Second level**		**Third level**	
			Adj. R^2	**F value**	**Adj. R^2**	**F value**	**Adj. R^2**	**F value**
HRM (M1)	IV population	DV departures	0.532	105.999***	0.083	19.862***	0.007	15.533***
		DV arrivals	0.493	90.281***	0.105	18.455***	0.01	14.638***
	IV trip	DV departures	0.519	100.289***	0.11	21.039***	0.001	16.062***
		DV arrivals	0.266	33.723***	0.218	13.753***	0.009	10.443***
ARX (M2)	IV population	DV departures	0.472	22.321***	0.007	11.041***	0.105	10.763***
		DV arrivals	0.467	21.930***	0.011	11.021***	0.094	10.298***

	IV trip	DV departures	0.472	22.328***	0.008	11.047***	0.105	10.763***
		DV arrivals	0.467	21.921***	0.011	11.011***	0.095	10.298***
Category 4: Individual district (Sanmin District)			**First level**		**Second level**		**Third level**	
			Adj. R^2	**F value**	**Adj. R^2**	**F value**	**Adj. R^2**	**F value**
HRM (M1)	IV population	DV departures	0.484	79.651***	0.282	53.088***	0.108	68.038***
		DV arrivals	0.491	81.889***	0.28	54.309***	0.106	68.769***
	IV trip	DV departures	0.489	81.281***	0.283	54.648***	0.103	68.130***
		DV arrivals	0.491	82.122***	0.28	54.780***	0.104	68.600***
ARX (M2)	IV population	DV departures	0.911	476.176***	0.026	269.891***	0.003	172.034***
		DV arrivals	0.909	465.794***	0.027	265.847***	0.003	168.262***
	IV trip	DV departures	0.911	476.235***	0.026	268.976***	0.003	171.682***
		DV arrivals	0.909	465.996***	0.028	267.140***	0.003	169.031***

Note: 90% level of significance, ** 95% level of significance, *** 99% level of significance.

Referring to Tables 4-2 to 4-6, this research used the HRM and ARX models to examine the aggregated factors of hierarchical level in four categories. The results are described as follows.

1. The explanatory ability of the HRM was insufficient in Categories 1 to 4, particularly in Categories 1 and 2. Tables 4-2 to 4-5, show that the adjusted R^2 value is only 0.176 to 0.222. Additionally, when examining the explanatory power of the factors in the three hierarchical levels, Categories 3 and 4 were in line with the research hypothesis. Population and trip factors in the first level of the hierarchical level are the main influencing factors on the CityBike ridership, with the explanatory power being higher than the factors in the second and third levels. Categories 1 and 2 showed the opposite: CityBike ridership variation was affected by the factors in the second and third levels.
2. The explanatory ability of the ARX model was sufficient in Categories 1 to 4 and the factors' effects on each hierarchical level are also in line with the hypothesis of this research. The adjusted R^2 of population and trip factors in the first level reached 0.5 in the HRM and considering the time-series nature that added the historical data, the adjusted R^2 value reached 0.9 in the ARX model.
3. The results confirmed that the population and trip factors are the main influencing factors on the CityBike ridership and verified why the previous studies adopted the factors in their BSS predicting model. The adjusted R^2 of the other factors in both the second and third levels was less than 0.1.

The HRM can classify the factors by their importance based on hierarchical levels, but the explanatory power of the model and the BSS ridership prediction were insufficient. In order to fill this gap, based on the collected time-series data, the ARX model was adopted to incorporate historical data and to combine the hierarchical advantage of HRM to forecast the CityBike ridership. The ARX model not only improved the explanatory power (at the 0.01 level of significance) but also provided the importance of the hierarchical level of each factor for the CityBike ridership prediction. The results of these two models revealed historical ridership data and showed that population and trip factors were the most important factors for predicting BSS ridership.

4.3.3 Results of the First-Level Variables in M1/M2

In Models M1 and M2, most of the variables had the same effect on CityBike ridership. The variables' effects on different district categories would be either positive or negative for the following reasons. Firstly, this research designed the concept of district-level data in order to determine whether or not the variables had different effects on the four district categories. The effects of the variables, whether positive or negative, could allow the operators or local government to design suitable planning or strategies for BSS construction in different district categories, which is why four district categories were set up to verify the hypothesis in this study. Secondly, Model M2 considered both historical data and key variables for CityBike ridership prediction. Compared with the traditional regression model and cross-sectional characteristic, Model M2 considered the nature of the data time-series and analyzed it from a longitude-sectional perspective, which had better results and higher accuracy regarding CityBike ridership prediction when using the evaluating indicators. The variables' effects were verified by this model, and their effects on the three levels are described as follows.

For the first-level variables, and in contrast to previous studies, this research used trip data to verify that the number of CityBike arrivals and departures was related to the number of trip generations and attractions individually in different categories. Because model M2 incorporated trip data as an IV in the developed models, it was able to estimate the CityBike ridership more accurately than model M1. Referring to Tables 4-2 to 4-5, the goodness of fit of M1 was lower and R^2 less than 0.5 at an insignificant level in all four categories but particularly in the citywide and urban districts. Compared with M1, the goodness of fit of M2 had higher explanatory power at a 99% significance level in all categories and better statistical results with the evaluating indicators (i.e., Durbin–Watson statistics, AIC, and BIC). The results of the population and trip factors impact the CityBike ridership. The trip factors were correlated with the CityBike ridership at a 99% significance level in the four categories, but the population factor was not statistically significant.

The trip factors were positively correlated with the CityBike ridership in the citywide and urban districts, while the individual districts had positive or negative effects at the 99% significance level. For example, Scenario 4 of model M2 represents the research in Sanmin District in Table 4-2, where the worker trip generation attribute and CityBike departures were negatively correlated at the 99% significance level with R^2 reaching 0.935. Referring to the statistical data from 2009 to 2016, the number of CityBike stations in Sanmin District increased from 2 to 26 stations, and the average ridership increased from 69 per station per month to 2,091 per station per month, although the number of workers decreased from 244,359 to 237,385. According to the M2 results, operators or the government can understand the impacts of the trip data on CityBike ridership in different area scales in Sanmin District. Although the CityBike stations increased annually in Sanmin District, the decrease in the number of worker trips caused a significant negative impact on CityBike ridership, which suggests that CityBike now has sufficient rental stations. Determining whether to install more stations in each district will be discussed in Chapter 4.3.4. The results in most categories showed that trip data reached statistical significance. However, there was a negative, but statistically insignificant effect in the suburban district, since there were fewer rental stations in this area, which could not generate enough samples for estimation by the models. Therefore, the results of the most explanatory variables were insignificant, and their goodness of fit did not research at a certain level.

This research compiled trip data and classified the data into four categories (worker trip attraction, worker trip generation, student trip attraction, and student trip generation). Such trip classifications help to clarify which trip type is the key factor that affects CityBike arrivals or departures in different districts in the developed models. The tests for Scenario 4 yielded strong overall correlations between CityBike usage and the trip factors shown in Table 4-7. A causal relationship was revealed between BSS ridership in terms of the CityBike ridership and trip variables.

Table 4-7. Correlation between the CityBike ridership and trip variables.

Area	Ridership$^{(i)}$ i = 1, Departures i = 2, Arrivals	Variables	Coefficient, t value
Yancheng District	1	Student trip generation	-0.886, -1.253
	2	Student trip attraction	-0.699, -1.286
Gushan District	1	Worker trip generation	6.520, 7.639***
	2	Student trip attraction	20.673, 7.584***
Zuoying District	1	Worker trip generation	32.581, 4.574***
	2	Worker trip attraction	3.439, 4.418***
Nanzih District	1	Worker trip generation	3.547,5.001***
	2	Student trip attraction	8.123, 4.978***
Sanmin District	1	Worker trip generation	-4.856, -6.998***
	2	Worker trip attraction	-5.204, -7.030***
Sinsing District	1	Student trip generation	-11.654, -8.760***
	2	Student trip attraction	-12.309, -8.741***
Cianjin District	1	Worker trip generation	-0.566, -1.012
	2	Worker trip attraction	-0.234, -0.991
Lingya District	1	Worker trip generation	-2.594, -7.377***
	2	Student trip attraction	-4.969, -7.542***
Cianjhen District	1	Student trip generation	-10.914, -4.147***
	2	Worker trip attraction	-3.932, -4.027***
Siaogang District	1	Worker trip generation	5.709, 0.628
	2	Worker trip attraction	5.568, 0.648
Fongshan District	1	Worker trip generation	1.778, 6.180***
	2	Student trip attraction	7.230, 6.472***

Note: * 90% level of significance, ** 95% level of significance, *** 99% level of significance

Referring to Table 4-7, in most districts, the CityBike arrivals and departures were significantly correlated with the number of students and workers. In most districts, such as Gushan, Zuoying, Nanzih, Siaogang, and Fongshan Districts, CityBike ridership is positively correlated with the number of students or workers. This means that CityBike usage in these districts is not yet saturated and the CityBike system is currently able to meet the students' and workers' transportation needs.

In contrast, CityBike ridership is negatively correlated with the number of students and workers in the Yancheng, Sanmin, Sinsing, Cianjin, and Lingya Districts. For example, Sinsing District has the highest density of schools in Kaohsiung City, with 3.54 schools and 3,180 students per km^2. However, these students will not have a positive impact on CityBike ridership for two possible reasons. Firstly, the CityBike system scale cannot satisfy student demand for transportation, and secondly, the locations of the CityBike stations are often too far away from their schools or houses. These issues will need to be addressed by the operators and the municipal government in future works.

4.3.4 Result of the Second Level Variables in M1/M2

For the second-level variables, the results revealed the effects of a capacity issue and POI on the CityBike ridership prediction. There was a major finding in the "capacity of CityBike station in 1 km buffer" variable to support our hypothesis in this research. Normally, if there is an increase in the number of system stations within 1 km of a single station, the arrivals and departures should increase as well. When the effects of capacity and POI on the CityBike ridership in the second level of variables were observed, the "CityBike station capacity in a 1-km buffer" variable showed that ridership in terms of arrivals and departures typically increased with the number of stations within 1 km of a specific station.

However, as indicated in Tables 4-2 to 4-5 and Appendices A-1 to A-12, this positive relationship between station capacity in a 1-km buffer and ridership in the target buffer zone did not always hold for all districts, including the suburban districts, Sanmin District, Yancheng District, Zuoying District, Nanzih District, Gushan District, Sinsing District, and Cianjin District. For example, the coefficient trend was negative in Sanmin District, indicating that an excess of rental stations in a buffer zone reduces the marginal benefit of a single station because rider demand in a specific area is ultimately limited. In order to find out the main reasons for this situation, this research used the M1 model to gather cross-sectional data from the total number of samples and to observe the effect of this factor on ridership at different time periods in different regions. The empirical results are shown in Figures 4-7 to 4-9.

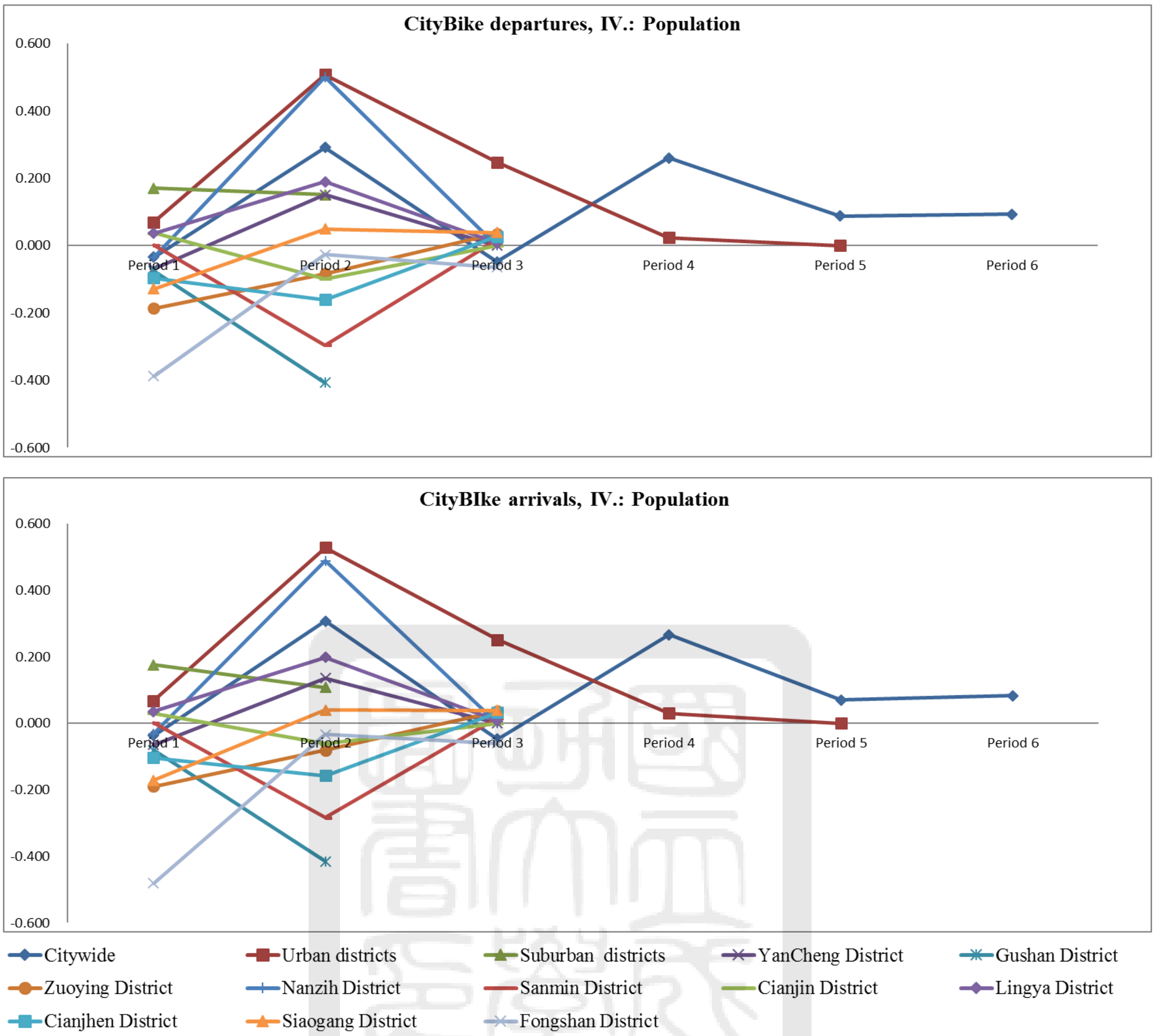


Figure 4-7. Comparison of regression coefficients of CityBike departures and arrivals with IV. population.

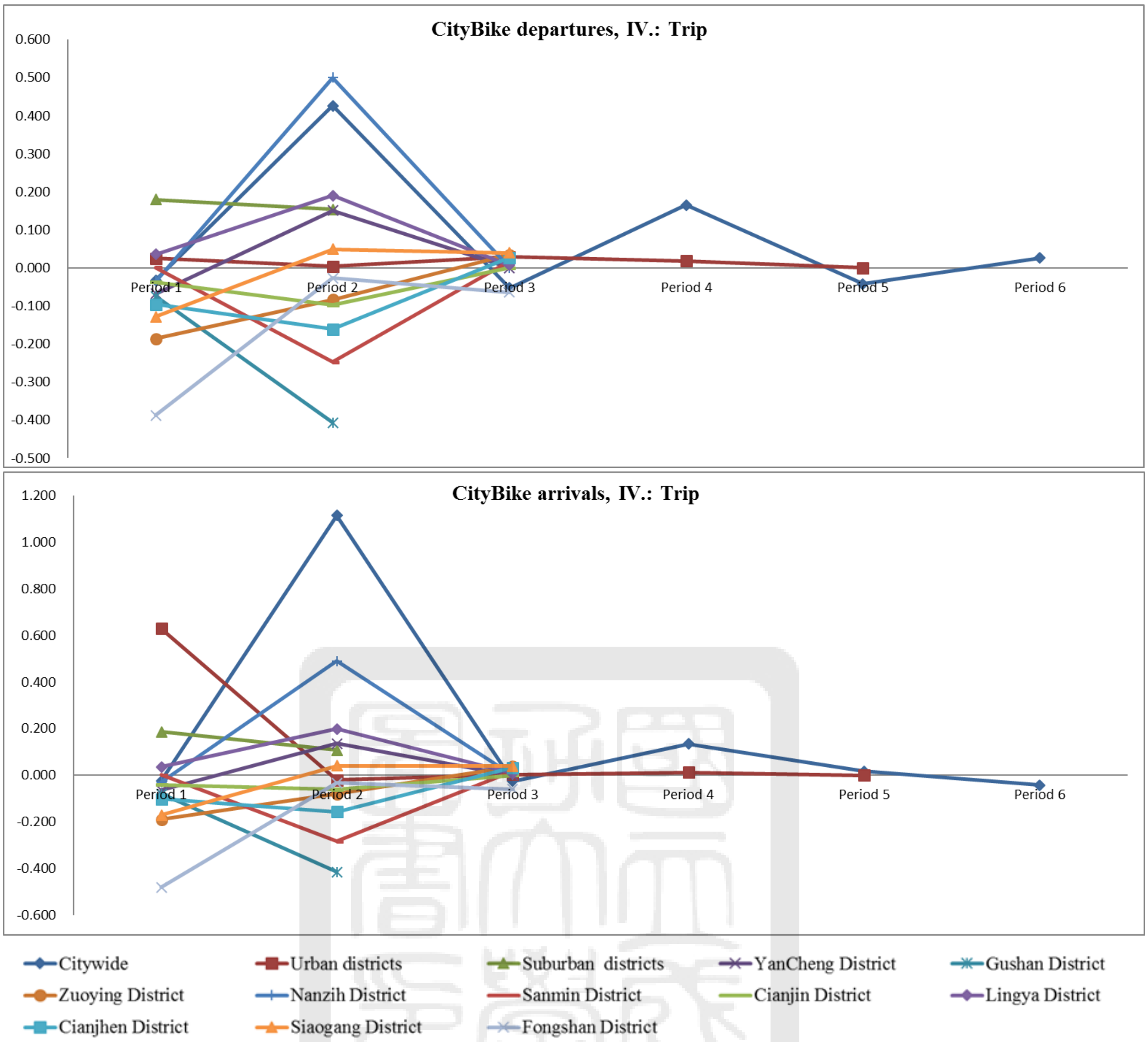


Figure 4-8. Comparison of regression coefficients of CityBike departures and arrivals with IV. trip.

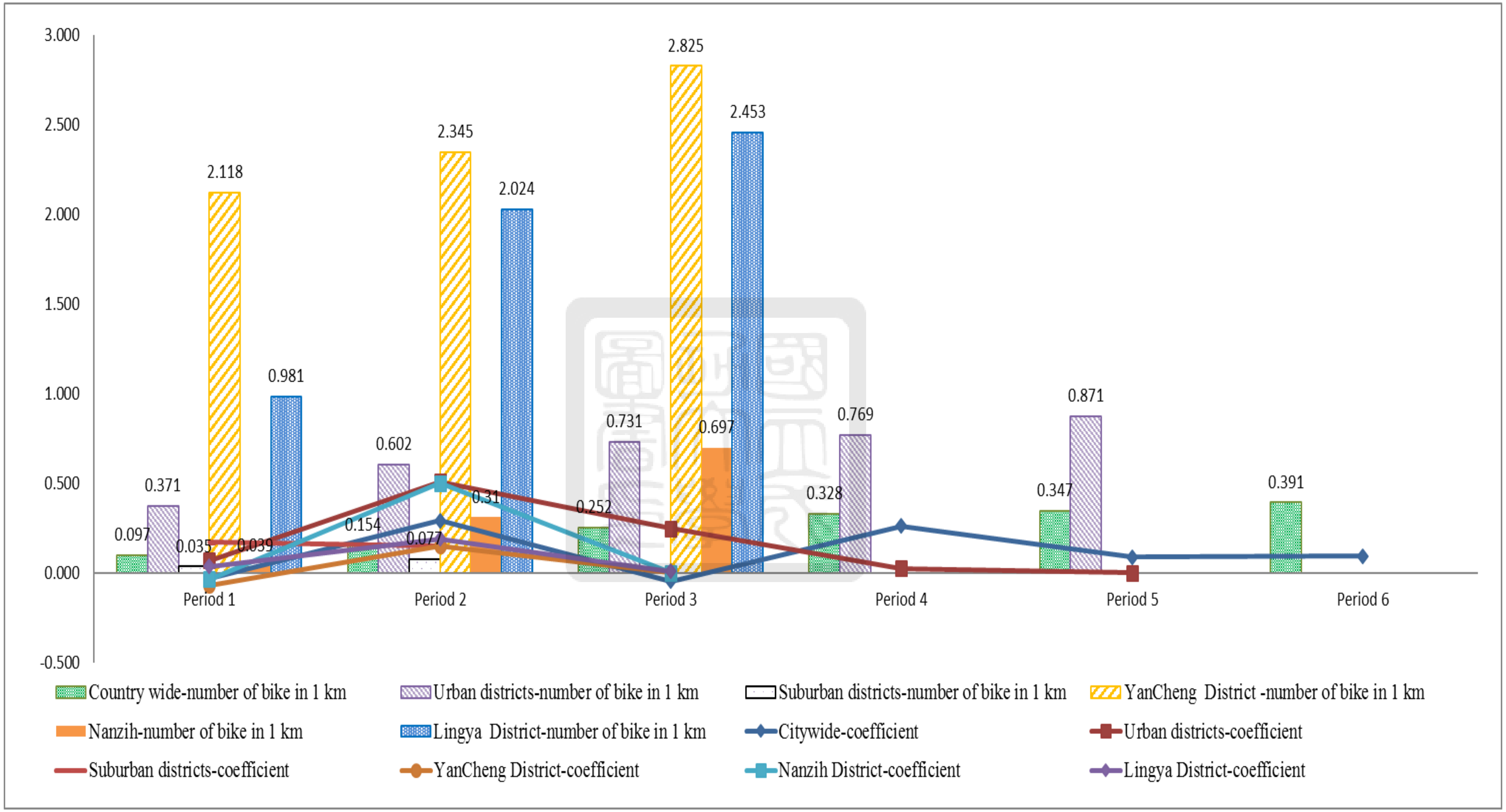


Figure 4-9. Comparison of regression coefficients and number of CityBike stations in 1-km buffer.

Figures 4-7 and 4-8 show the comparison between the number of stations in the 1 km buffer zone and the estimated coefficient. The estimated coefficient in most districts increased in Period 1 and Period 2 when the number of stations increased; ridership also increased during this time. However, in Period 3, the coefficients dropped significantly while the number of stations continued to increase, such as in the citywide scenario, the urban district scenario, and some of the individual districts (such as Yancheng District, Nanzih District, and Lingya District). The influence estimated coefficient of this variable decreased or changed from positive to negative between Period 2 and Period 3, which represented the impact on the CityBike ridership becoming weaker, or even negative. When the number of rental stations increased from Period 2 to Period 3, the estimated coefficient trend decreased, indicating that too many rental stations in a buffer area will reduce the marginal benefits of a single station since demand is limited.

The high peaks shown in Figures 4-7 and 4-8 determine the maximum number of stations per km^2 in different districts. For example, in Period 2 of Figure 4-9, it can be seen that the upper boundary of the number of stations in the 1 km buffer zone for Yancheng District is 2.345 rental stations per km^2. Table 4-8 summarized the upper or lower boundaries of the number of CityBike stations in the urban district.

Table 4-8. Model results and estimate upper/lower bound of the number of CityBike stations within 1 km^2

Area	**Estimated coefficients of the number of stations in 1 km^2 increase or decrease during Periods 2 and 3?**	**Add new CityBike stations or not?**	**Upper or lower bound of the number of stations in 1 km^2**
Yancheng District	Decrease	No	Upper bound, 2.345
Gushan District	Decrease	No	Upper bound, 0.712
Zuoying District	Increase	Yes	Lower bound, 0.851
Nanzih District	Decrease	No	Upper bound, 0.271
Sanmin District	Increase	Yes	Lower bound, 1.011
Sinsing District	Decrease	No	Upper bound, 3.790
Cianjin District	Decrease	No	Upper bound, 3.769
Lingya District	Decrease	No	Upper bound, 2.024
Cianjhen District	Increase	Yes	Lower bound, 0.837
Siaogang District	Increase	Yes	Lower bound, 0.132
Fongshan District	Increase	Yes	Lower bound, 0.411

To summarize the above results, there appear to be limits on the development of a BSS network in a given region, based on the previous test. Table 4-8 summarizes the estimated coefficient changes during Periods 2 to 3 and determines the most suitable areas to increase the number of rental stations. The upper or lower boundary of the number of CityBike stations in each district depends on the high peak of the estimated coefficient shown in Figures 4-7 and 4-8. The model results can provide the operators and the government with the ideal locations for constructing new stations. The average capacity within 1 km^2 of a single rental station should be calculated using a geographic information system, and the result can be used as the upper limit for the number of stations in a particular district. If the variable has a positive effect, the new rental station can be installed on a site within 1 km^2 of the upper boundary. If not, then no more CityBike stations should be constructed in that district. Specifically, increasing the number of BSS stations in a buffer area will not necessarily increase the ridership of a BSS.

The "bikeway length within 1 km^2" variable had a positive effect on the CityBike arrivals and departures in Cianjin District, but the variable had a negative effect on the suburban districts. According to official statistics, bikeway length increased from 727 to 1,288 m/km^2 from 2009 to 2016 in Kaohsiung City, while the number of CityBike stations increased from 46 to 185 during the same period. The simultaneous growth of bikeways and CityBike stations has had a positive effect on CityBike ridership. From an individual district perspective, Cianjin District has the highest station density (4.307 stations/km^2) and the second-longest bikeway length (20,379 m/km^2) in the CityBike network, which has had a positive effect on its ridership. The relatively low density and the relatively few stations in suburban districts (463 m/km^2, 0.08 stations/km^2) may negatively affect CityBike ridership, mainly because having exclusive cycle lanes can increase the willingness of residents to use CityBike. In other words, suburban districts continue to construct and extend bikeways because there are fewer CityBike rental stations, which makes it difficult to increase overall CityBike ridership. In Kaohsiung City, the municipal government has constructed bikeways along with public transportation systems (i.e., the metro, light rail, and bus systems) and POI (i.e., the Love River and Pier-2 Art Center), which might increase the willingness of people to use the CityBike. The variable of "length of major roads within 1 km^2" yielded the same effect in Cianjin District and Yancheng District as did the IV of bikeway length.

The "streetlamp" variable did not significantly affect CityBike ridership in most of the test areas, although it did in the suburban districts. Because the variation was less than 2% in all districts from 2009 to 2016, the Kaohsiung City municipal government did not consider the environmental demand of the CityBike system when constructing streetlamps; instead, it determined that the local citizens' needs and security were the main considerations. Therefore, this variable can be ignored in BSS ridership prediction models in future research.

However, the "number of parks" variable positively affected the CityBike arrivals and departures in urban districts, such as Yancheng District, which has the most parks among all the districts (12 parks/km^2). Additionally, the number of parks has increased annually since 2009. However, some urban districts, such as

Fongshan District, showed a negative relationship between the number of parks and the CityBike ridership. According to official statistics, these areas are not completely prepared for bicycle rental stations near parks and have not expanded the network annually. In addition, the variables of "number of companies," "number of factories," "number of markets," and "number of hotels" in most districts had a negative effect on the CityBike arrivals and departures. We inferred that rental stations installed near these activity points have been few in number because operators have neither applied to install a CityBike rental station near these points nor have paid for the construction costs. Therefore, residents in such areas have regarded the system as less convenient and are therefore less likely to use the CityBike. The installation cost of a single rental station is approximately NT$1.2 million, and the annual operating and maintenance cost is NT$10,000–NT$20,000. This is too expensive for the operators of different companies, factories, markets, or hotels.

The Kaohsiung City government considered fairness and meeting most residents' travel needs, as well as proximity to public transportation stations and POI, such as schools, to be the top priorities when constructing the CityBike stations, with the overall aim of achieving seamless transportation. The government did not prioritize the construction of stations close to private companies, factories, markets, or hotels because that would mostly benefit private enterprises rather than the public. Nanzih District is an exception because numerous domestic and international companies have constructed their factories in the area, making it a key destination for commuters. The municipal government considered employment development, salary levels, and transfers to the public transportation system (i.e., Kaohsiung MRT and urban bus stops) before opting to construct CityBike stations near the factories. Therefore, CityBike ridership in Nanzih District was positively correlated with the number of factories variable. In addition, the number of public transportation points increased slightly, and the CityBike ridership showed a downward trend in 2017, which negatively affected the CityBike arrivals and departures in citywide and urban districts. In Kaohsiung City, most public transportation stations, such as the metro, railways, light rail, high-speed rail, and bus transfer stations, are concentrated in urban areas, meaning that the CityBike system plays a feeder role in most public transportation systems in urban areas, providing mutual benefits for the transfer ridership of both public transportation systems and the CityBike system.

4.3.5 Result of the Third-Level Variables in M1/M2

For the third-level variables, the “tourist” variable had a positive effect in some areas such as the Yancheng, Gushan, Zuoying, and Lingya Districts. For example, the Kaohsiung City government constructed the Asia New Bay Area in Yancheng District in 2014. There are many remarkable exhibition buildings located here, such as the Kaohsiung Exhibition Center and the Pier 2 Art Center. In addition, Kaohsiung Light Rail Transit also goes through the entire area.

Yancheng District is popular with tourists, who increase the CityBike ridership. Based on this result, the government needs to develop a comprehensive plan for district development and create a friendly and attractive environment for everyone who uses the CityBike system.

The “income” and “private vehicle” variables negatively affected CityBike arrivals and departures. High-income households tend to use more convenient transportation modes, such as scooters, cars, or taxis. Similarly, households with more private vehicles are more likely to use private modes of transportation. Finally, land values did not significantly affect the CityBike ridership. When the government installs a rental station, the land on the construction site must be levied. Therefore, the land value of the area affects the ease of levying. However, between 2009 and 2017, land values in each district fluctuated, making it difficult to determine the influence of land value on CityBike ridership.

4.3.6 Results of CityBike Ridership Prediction in M1/M2

This research collected relative data between 2009 and 2017 to calibrate model M1 (HRM) and model M2 (ARX model). These models were used to predict ridership in 2017 as a performance evaluation. Figures 4-10 to 4-17 present actual ridership data for 2009 to 2017 in test areas as well as the CityBike arrivals and departures prediction results for models M1 and M2 in the four categories in 2017.

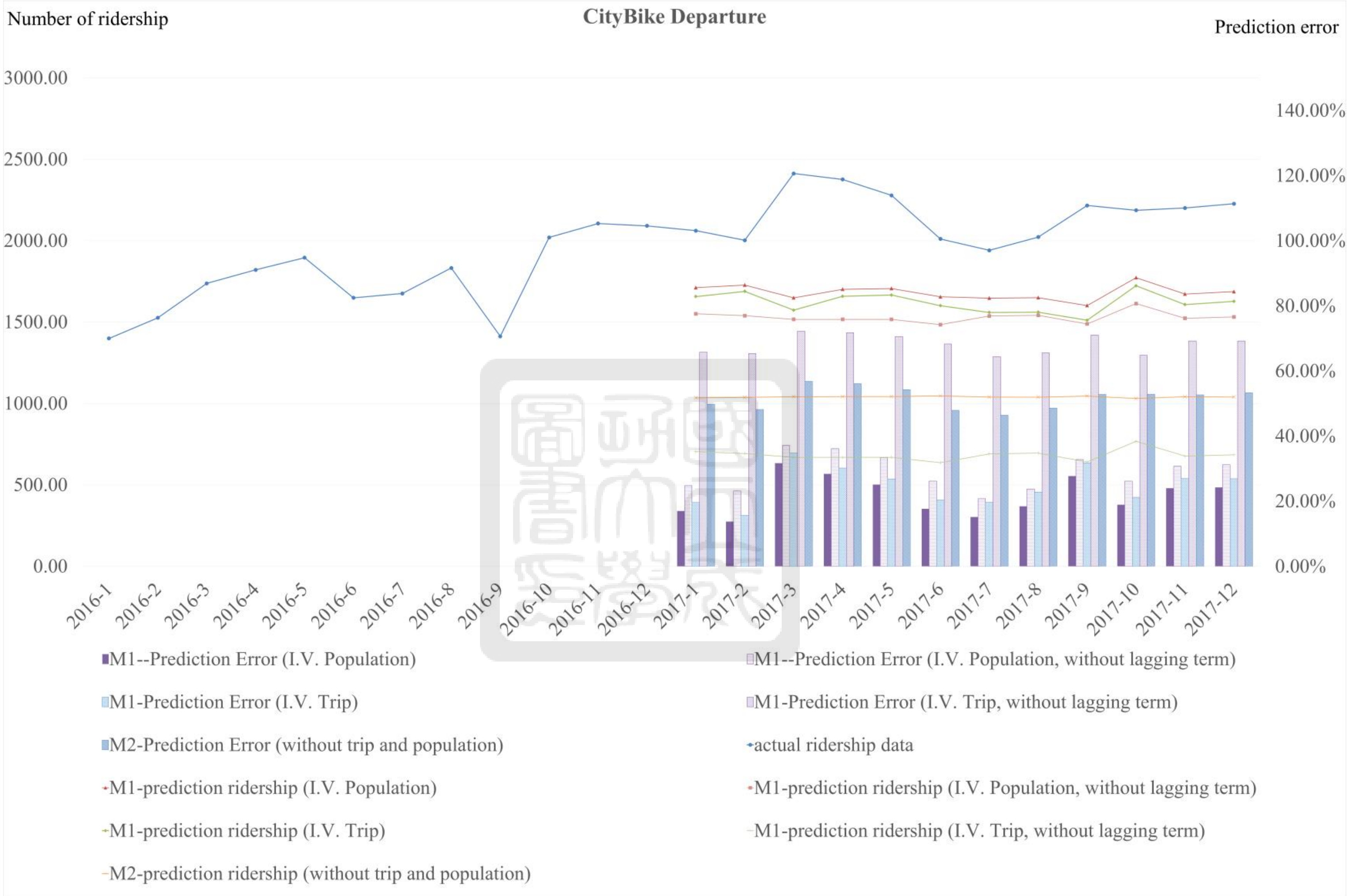


Figure 4-10. Comparison of predicted and actual CityBike departures in citywide district, 2017.

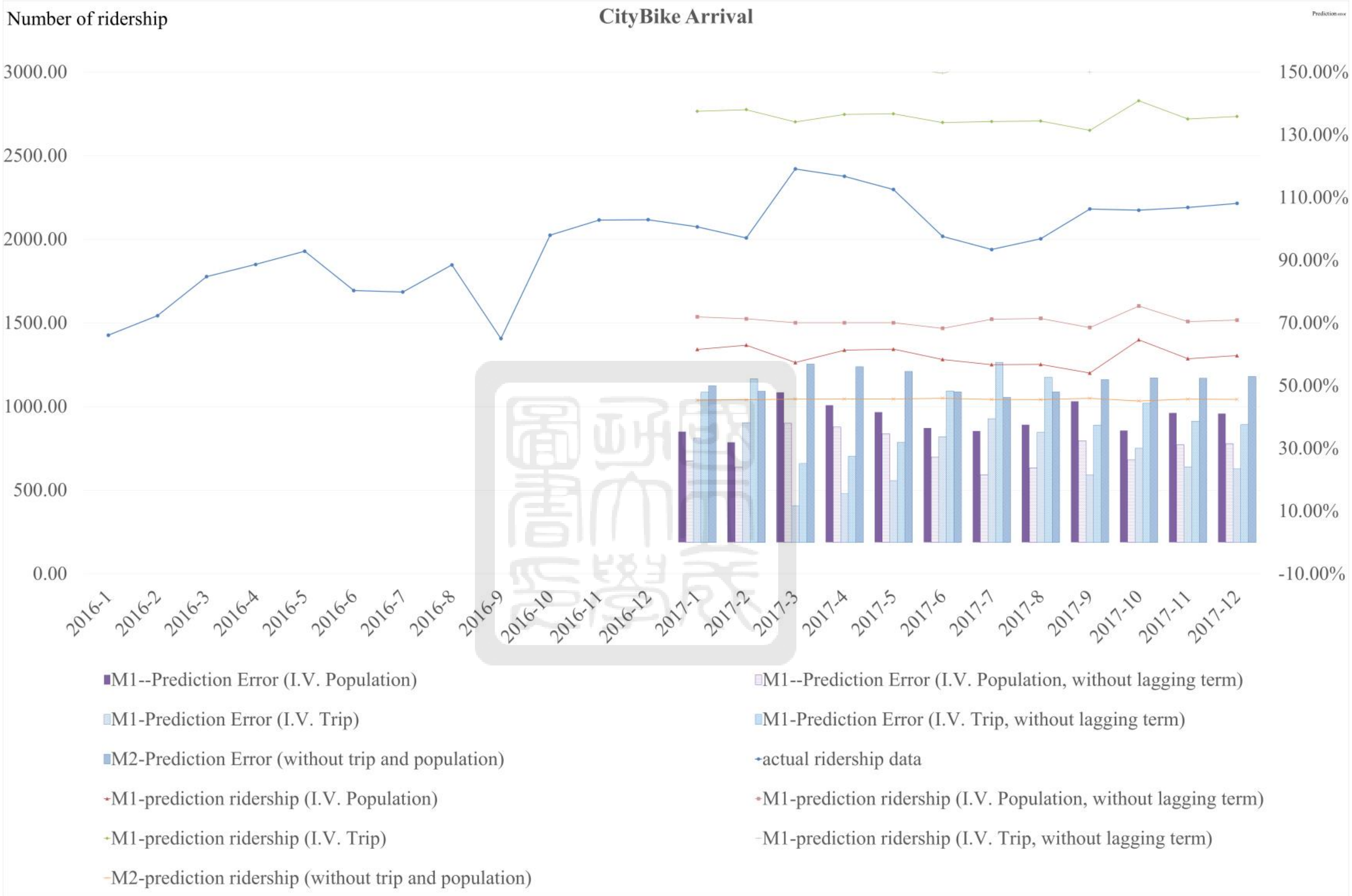


Figure 4-11. Comparison of predicted and actual CityBike arrivals in citywide district, 2017.

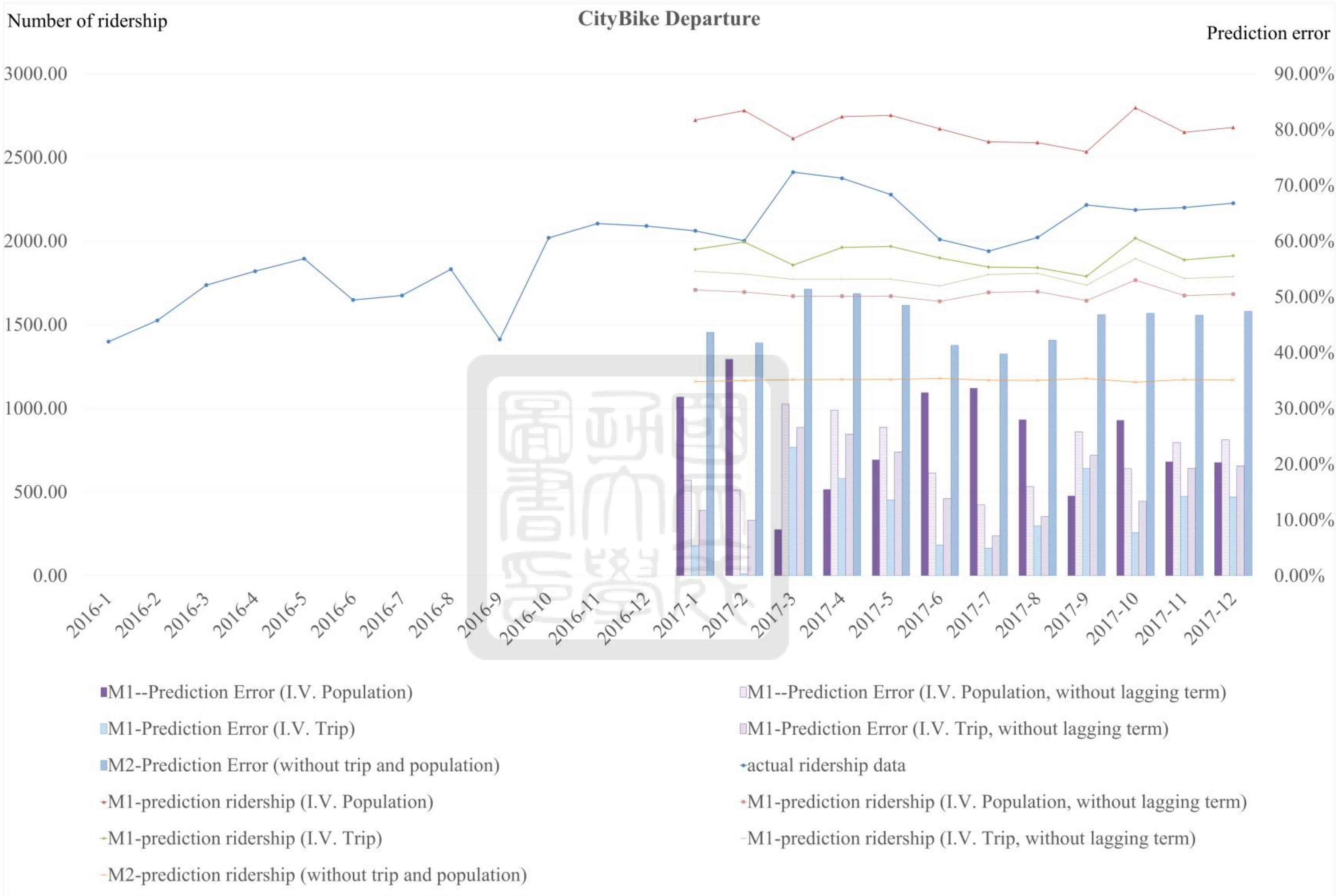


Figure 4-12. Comparison of predicted and actual CityBike departures in urban district, 2017.

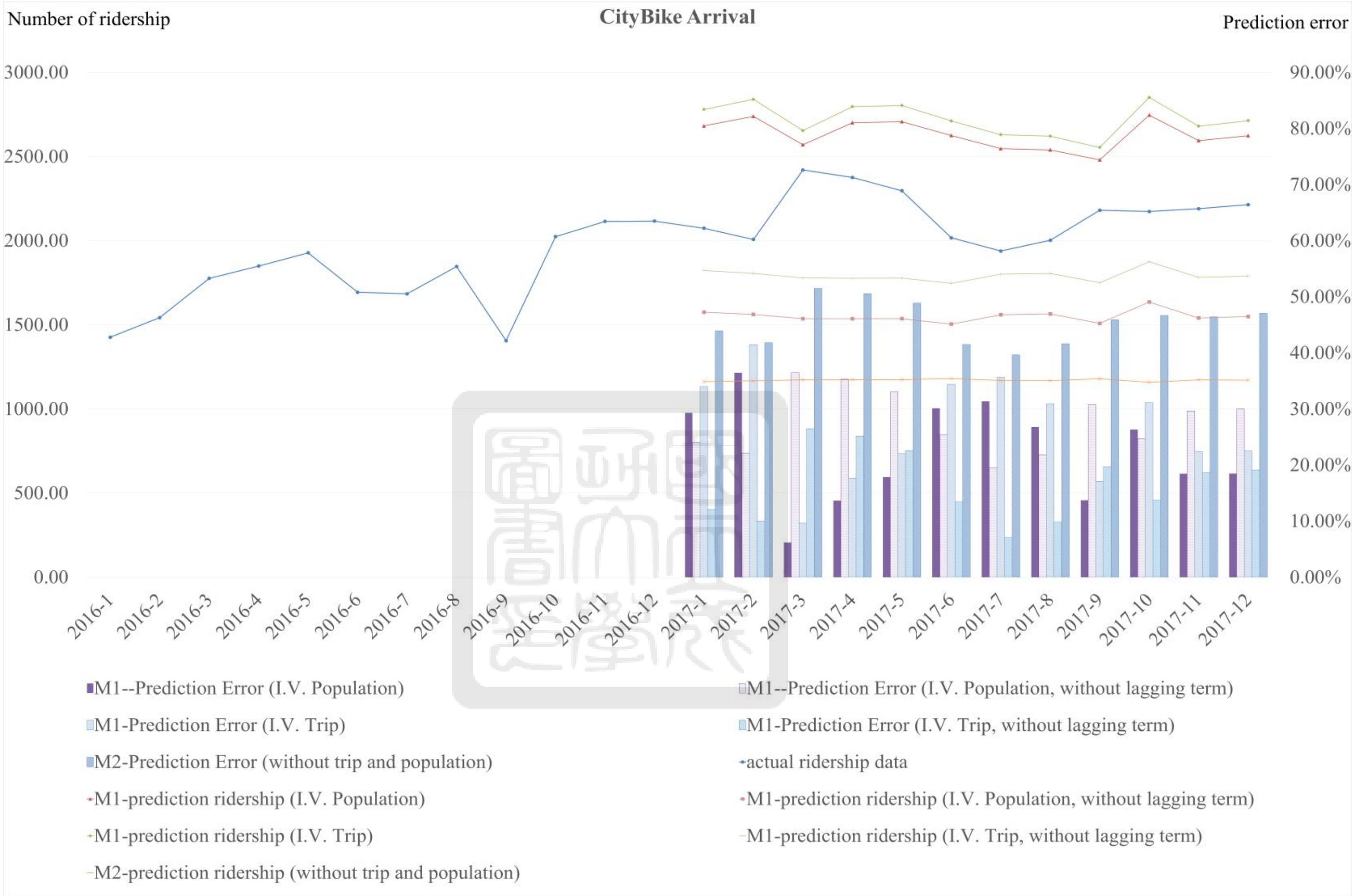


Figure 4-13. Comparison of predicted and actual CityBike arrivals in urban district, 2017.

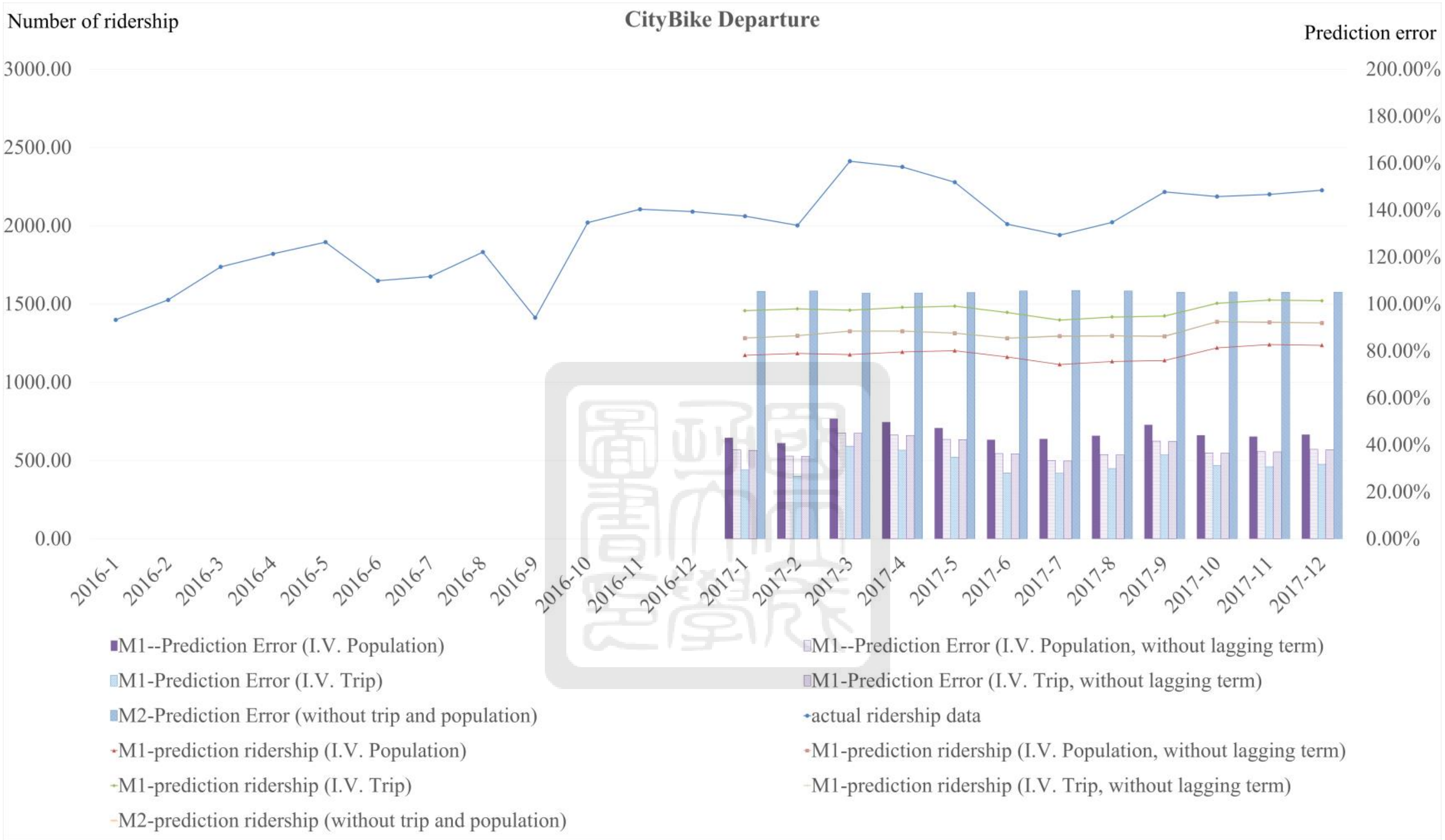


Figure 4-14. Comparison of predicted and actual CityBike departures in suburban district, 2017.

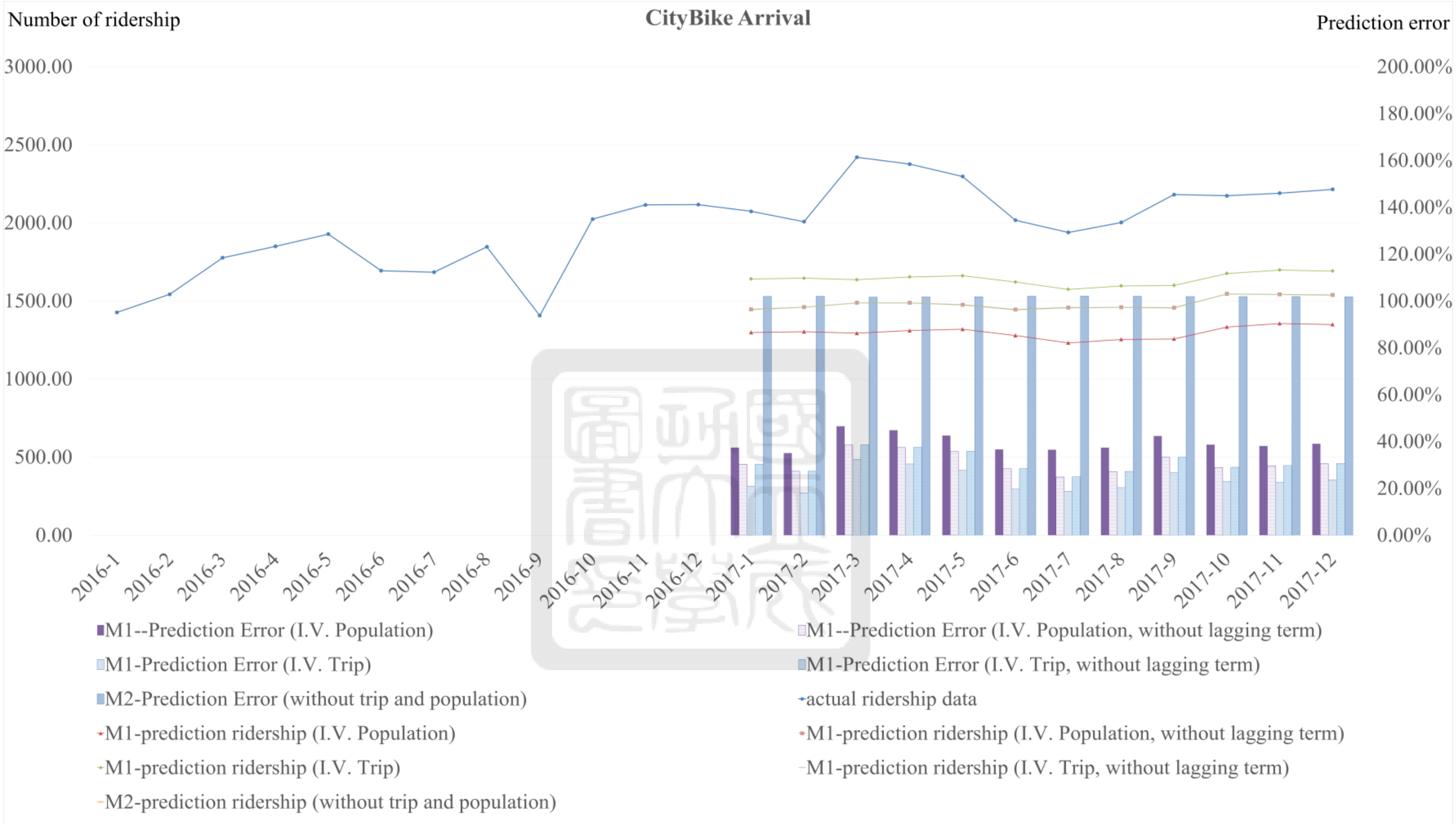


Figure 4-15. Comparison of predicted and actual CityBike arrivals in suburban district, 2017.

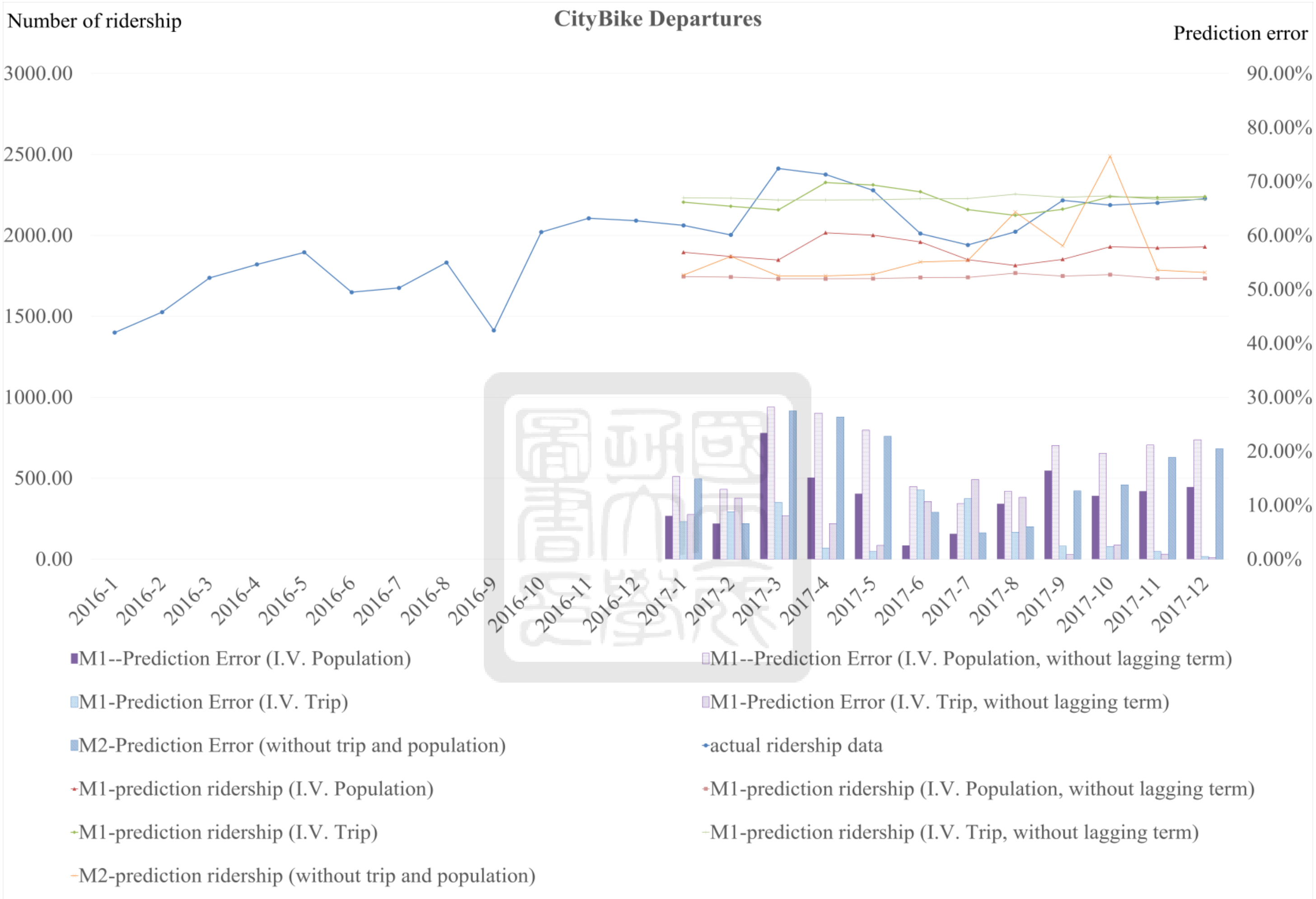


Figure 4-16. Comparison of predicted and actual CityBike departures in Sanmin District, 2017.

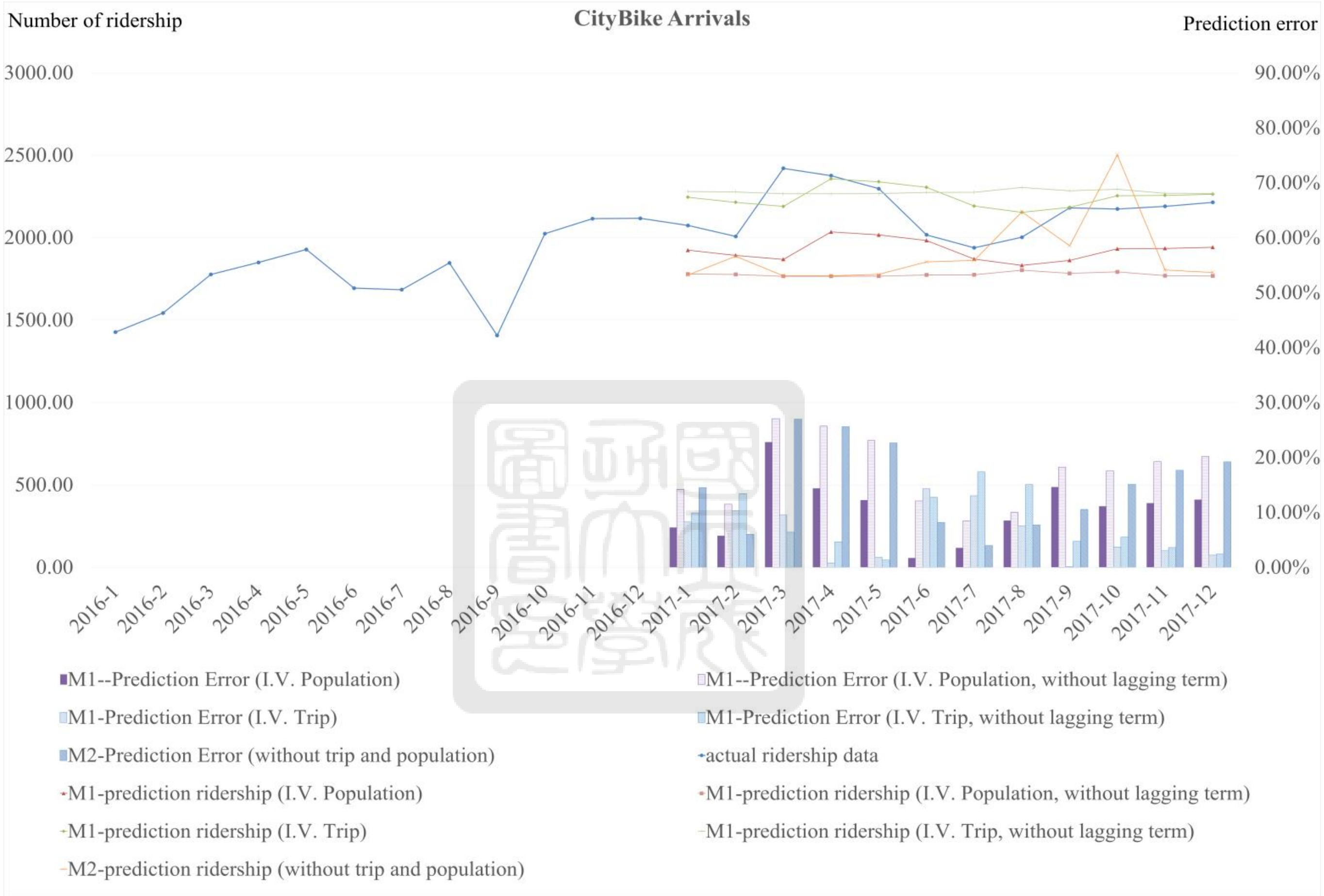


Figure 4-17. Comparison of predicted and actual CityBike arrivals in Sanmin District, 2017.

Compared with model M2, model M1 yielded less accurate predictions of CityBike arrivals and departures at the citywide, urban district, suburban district levels, and most individual districts. For example, Figures 4-16 and 4-17 present actual ridership data for 2009–2017 in Sanmin District as well as the prediction results of models M1 and M2 for 2017. The prediction error of models M1 and M2 in Sanmin District is shown in Table 4-9.

Table 4-9.Results of CityBike ridership prediction error in Sanmin Ditrict

Model	IV.	DV.	Prediction error
M1	-		15.30%
M2	Trip	Departures	5.49%
M2	Population		11.43%
M1	-		14.84%
M2	Trip	Arrivals	6.22%
M2	Population		10.46%

The average prediction errors for CityBike departures for model M2 with the trip attribute, model M2 with the population attribute, and model M1 were 5.49%, 11.43%, and 15.30%, respectively. The average prediction errors of CityBike arrivals for model M2 with the trip attribute, model M2 with the population attribute, and model M1 were 6.22%, 10.46%, and 14.84%, respectively. The collected data showed high levels of autocorrelation (up to 0.9 at the 99% significance level). Appendix B-1 to B-22 presents actual ridership data for 2009–2017 in the other individual districts as well as the prediction results of models M1 and M2 for 2017. For other districts, the average prediction errors of the CityBike ridership ranged between 11.21% and 31.93% in model M2 with trip attributes, ranged between 21.83% and 45.14% in model M2 with population attributes, and reached 66.89% in model M1.

To reduce the prediction error of the CityBike ridership, the AR model adopted the first-order lagged term by using the autocorrelation function and the PACF that were estimated from the historical CityBike ridership data, which was incorporated into the ARX model. If the lagged term is not considered in the ARX model, the prediction error will increase. The results with or without lagged term in Sanmin District and other district categories are shown in Tables 4-10 and 4-11.

Table 4-10. Results of CityBike ridership prediction error with/without lagged term in Sanmin District.

Model	**IV.**	**Lagged term**	**Prediction error**
M1	-	No	14.84%~15.30%
M2	Trip/Population	Yes	5.49%~6.22%
M2	Trip/Population	No	6.55%~8.10%

Table 4-11. Results of CityBike ridership prediction error with/without lagged term in City wide, urban district, and suburban district.

Model	**IV.**	**Lagged term**	**Prediction error**
M1	-	No	45.47%~105.20%
M2	Trip/Population	Yes	11.21%~31.93%
M2	Trip/Population	No	16.51%~68.23%

Similar to the results for Sanmin District, the prediction error of the ARX model with trip attributes ranged between 5.49% and 6.22%, the models without the lagged term showed errors ranging between 6.55% and 8.10%, and the prediction error of the HRM without trip attributes and the lagged term ranged between 14.84% and 15.30%. In the citywide, urban districts, and suburban districts, the prediction error of the ARX model with trip attributes was between 11.21% and 31.93%, while the model without the lagged term was between 16.51% and 68.23%; the prediction error of the HRM without trip attributes and the lagged term was 45.47% to 105.20%.

During most months, the prediction errors in the ARX model with trip attributes for the CityBike ridership were lower than those in the ARX model with population attributes and multiple regression models, and the R^2, AIC, and BIC of the ARX model with trip attributes maintained a certain level. For example, the CityBike ridership in Sanmin District changed considerably in March 2017, with ridership increasing by 20.4% compared with the previous month, but the average variation rate in the CityBike ridership was 5.39% in 2017. Under this situation, the prediction error was 23.40% in model M2 and 27.48% in model M1 with the population attribute, but only 10.58% in model M2 with the trip attribute.

According to the aforementioned results, models that considered trip factors and time-series for usage prediction performed better than the traditional regression models. The trip data can reflect the true passenger behavior and can more accurately predict BSS ridership. Important statistical information such as adjusted R^2, AIC, and BIC showed that Model M2 had a significantly better goodness-of-fit than Model M1. At the same time, all factors for the models were compliant with VIF and DW indicators. Therefore, when predicting the CityBike ridership, model

M2 outperformed model M1 with respect to trip attributes, and model M2 with trip attributes also outperformed model M2 with population attributes.

4.3.7 Policy Implication

To the best of our knowledge, this research is the first to use trip data to investigate the long-term causal relationship between trip generation/attraction and BSS arrivals/departures. The two models, M1 and M2, were developed to evaluate the effects of key factors on the CityBike ridership forecast and system expansion. The resulting observations provided policy implications for the operator or government are described below.

1. This study adopted trip factors to replace population factors and considered time-series nature in order to obtain the highest possible level of accuracy when predicting CityBike ridership. Many previous studies have used population factors and relative methodologies to forecast BSS ridership over the short term or in smaller areas. However, they did not evaluate or compare prediction accuracy with the current state. This study has verified that the trip factor influencing BSS ridership can obtain better prediction results. In the future, operators or governments should collect trip data within the area either monthly or annually in order to evaluate usage prediction or potential BSS expansion.
2. The model's results indicated that population and trip factors are the main influencing factors on the CityBike ridership, as the explanatory power is higher than the other factors. Table 4-7 summarizes the interactions between trip factors and CityBike ridership. There are two main reasons why some districts have trip factors with negative impacts. Firstly, the number of CityBike rental stations in those areas cannot satisfy the demand for transportation among the student body, and secondly, the locations of the CityBike stations are often too far away from their schools or houses. These results suggest that the operators or municipal government must adjust the number of rental stations or locations to maximize the CityBike system performance. On the other hand, in districts with a positive impact on trip factors, the CityBike system is successfully meeting demand and attention can instead be focused on stabilizing the current system.
3. The "Station capacity" and "CityBike station capacity in a 1-km buffer" variables serve as an example. Because travel demand is limited in some districts, constructing too many CityBike rental stations in a buffer range reduces the marginal benefit of the CityBike system and does not increase the CityBike ridership. This phenomenon was seen in Yancheng, Gushan,

Nanzih, Sinsing, Cianjin, and Lingya Districts. Those variables, defined as environmental variables in this research with positive effects in the ARX model, could help the government and BSS operators determine where to install new rental stations or remove unnecessary stations in order to efficiently meet the travel demands of each district. The findings can also serve as guidance for the government and BSS operators on appropriate investment decisions in the context of limited resources, ultimately enabling the BSS to achieve financial sustainability. By contrast, the aforementioned BSS station adjustment constituted a major strategy with a limited budget. If the operator had a large enough budget (e.g., YouBike 2.0 in Taiwan), they could set up BSS rental stations without restrictions, which would massively increase YouBike 2.0's ridership. This is because the YouBike 2.0 system can construct rental stations near metros, bus stops, schools, business districts, and other POIs without any budget limitations. However, in this situation, the results of this study could not be extrapolated.

4. The variables are defined as environment variables such as bikeway length and the number of parks, in this study with positive effects in the ARX model. The results could help the government and BSS operators decide what infrastructure to install in order to increase ridership. For example, in the "bikeway length within 1 km^2" variable, the model estimation result revealed that increasing bikeways in urban districts encouraged more people to use the CityBike than similar actions in suburban districts. Based on these results, the government should initially be constructing BSS infrastructure in urban districts despite their limited budget. Most bikeways in suburban districts were constructed near tourist areas for people who own bikes. Therefore, the government could consider constructing fewer bikeways in suburban districts since there are fewer CityBike stations (0.08 stations/km^2) and it would therefore not be convenient for people to rent bikes in these areas.
5. The results of this study indicated that constructing CityBike stations near parks will increase CityBike ridership since people usually rent bikes to ride to the parks for exercise and recreation. However, some districts, such as Fongshan District, showed a negative impact between the number of parks and CityBike ridership because there were no CityBike rental stations in the vicinity. The municipal government needs to follow the criteria of the BSS construction policy but cannot choose the place as their wish. The best locations for building rental stations and increasing CityBike ridership will be near public areas (e.g. parks or squares). In addition, the government is limited by budget constraints and the criteria of BSS construction in that

CityBike rental stations should usually be close to public transport systems or in more populated areas. To expand the CityBike system, the operators and government should firstly consider simplifying the rental station equipment and reducing construction costs, which would allow the budget to stretch to the construction of CityBike stations near private enterprises, such as companies, factories, or hotels. Secondly, the operators or the municipal government could cooperate with private institutions to maintain one single rental station, which would save them money.

6. The other finding in the second-level variables, the variables of "number of companies," "number of factories," "number of markets," and "number of hotels" had a negative effect on the CityBike ridership in most districts. The operators or the local government should therefore consider constructing CityBike stations in locations where they could be more easily rented, used, and returned.
7. Because the government has not provided a detailed social economics database, BSS time division ridership, or O-D pair data, the researchers only obtained by assuming or mathematical models estimating. For academic research and planning design, we suggest the following data should either be made public or be calculated by the government: (1) BSS ridership volume in hours, (2) BSS O-D pairs data, (3) trip attraction and generation per district per month.
8. Based on the models' results, this study proposes the following BSS station installation instructions that could consider being added to the CityBike system's installation laws and regulations made by the Kaohsiung City Government (2020b).
   (1) The government needs to select districts where the population is growing annually.
   (2) BSS rental stations need to be built near the POI with a walkable distance.
   (3) The government needs to construct BSS stations near bikeways within a suitable distance.
   (4) The government needs to prioritize bikeway construction in urban districts over suburban districts.
   (5) The government needs to analyze BSS ridership every year and adjust the number of BSS stations accordingly.

Chapter 5 CONCLUSIONS AND FUTURE WORK

This dissertation investigated trip attributes, population, and socioeconomic variables on CityBike ridership forecasting, with hierarchical regression, AR, and ARX models applied to solve the problem. The mathematical formulation combined the advantage of the HRM with the ARX model and revealed the key variables influencing different categories. The conclusions, limitations, and future works for this research are discussed below.

5.1 Conclusions

Most previous BSS-related studies focused on different aspects such as ridership (i.e., arrivals and departures), bicycle network and system design, location choices, and bicycle redistribution. Regarding BSS ridership, previous studies adopted typical methods, such as regression models, logit models, and time-series methods, in order to determine the effects of demographics, socioeconomic characteristics, infrastructure, the built environment, POI, system capacity, and land use on BSS ridership. Specifically, previous studies have typically treated population as an influential factor in the short term or on a small scale. However, the population variable can often represent neither real citizen activity nor the spatial relationships of BSS ridership in a specific area, which may result in biased model estimations. To fill this gap in the BSS-related literature, this research collected and incorporated trip data into the modeling process, and the empirical results confirmed that our ARX model provides accurate estimates of the CityBike ridership under different testing categories. A summary of the conclusions and findings is shown below.

1. The empirical research results revealed a strong correlation between trip attributes and BSS ridership. In the hierarchical modeling process, trip attributes are also the major influencing factors in BSS ridership forecasting. In addition, the CityBike ridership estimated by the developed ARX model with trip attributes was more accurate than that estimated by the ARX model with population attributes, and the HRM. Trip generation and attraction of a region had negative effects on CityBike ridership, such as in Yancheng, Sanmin, Sinsing, Cianjin, and Lingya Districts. There were two possible reasons for this: firstly, demand from students and workers is

exceeding supply, and secondly, the locations of the CityBike stations may be too inconvenient.

2. Different from previous studies that focused on small-scale datasets, this study created a type of “district-level data” to evaluate the relationship between BSS ridership and other influencing factors at different district scales. That information could not be gleaned from previous studies. The results of this study showed that the same factors had different effects on different district scales, which would alter the plans or policies that local operators or municipal governments would make for a particular city.
3. This research combined the advantages of the HRM and ARX models, which sorted the importance of each factor at the hierarchical level for the CityBike forecasting. The results determined that historical ridership data, population, and trip factors are the most important factors for BSS ridership forecasting. The second most important factors are spatial variables, such as POI and environmental factors, while factors that are not directly related to BSS ridership, such as land price and private vehicle ownership, are less important.

5.2 Limitations of the Research

1. The collected trip data (i.e., student trip demand or worker trip demand) of this research were obtained from the census data of the Kaohsiung City government in 2009. The census included classification data of different trip purposes for each district in Kaohsiung City. However, because censuses are not performed every year, this research used census data from 2009 and then used the population growth rate of each district to estimate trip data for each year.
2. There are not enough CityBike stations in seven of the suburban districts in Kaohsiung City (namely Daliao, Renwo, Niaosong, Gangshan, Ciaotou, Yanchao, and Cheting Districts), to provide enough samples for BSS ridership forecasting. This research, therefore, considered these seven suburban districts to be one area in order to obtain more samples for analysis.

5.3 Future Work

The demographic and socioeconomic characteristics of modern, well-developed cities do not change much over the short term. Small-scale, short-term BSS data may clarify variations in ridership at a single rental station, but in order to assess a BSS's performance over the long term in the name of sustainable development, the evolution of BSS ridership should be analyzed by incorporating trip-related data and other key factors.

There are several directions in which this research could be extended in the future.

1. The importance and priority of key factors were defined in the developed model in this research. Other developed BSSs can refer to the designed hierarchical level in this research and can construct their own forecasting models. Exploring the effect of each variable and performing sensitivity analysis on BSS ridership can be a logical next step. Moreover, unobserved factors that predict BSS ridership can be incorporated into the models of more comprehensive research on BSS ridership predictions. The main objectives of future research are as follows.
   (1) Home-based "other" trips comprised approximately 30% of all trips (i.e. home-based recreation trips), but this research did not consider those trips in the models. Including these trips could alter the variable coefficients in the models, so the researchers will need to clarify and classify those trips into the relevant categories in future research.
   (2) The research could divide some IVs into sub-categories. For example, the school variables contain elementary schools, middle schools, and universities. Students at these institutions would have different travel behaviors, which is also something that can be studied in the future.
2. This research confirmed that some districts had constructed too many CityBike rental stations within a buffer range. Future research could determine how to adjust excessive station numbers in areas, or within a 1-km buffer zone.
3. The dataset resembled panel data with time-series and cross-sectional characteristics. Before verifying the models, the panel data must in stationary status and need to calculate by unit root test method. Panel data models can be used to measure fixed effects and random effects within data in future research. This study also focused on the effects of variables on BSS ridership, which are traditional statistical methods that are easy to analyze. If the researchers want to predict BSS ridership more accurately, they can use both AI and machine-learning methods to mine data for more information.

REFERENCES

1. Akaike, H. (1974). “A new look at the statistical model identification.” IEEE Transactions on Automatic Control, Vol. 19, Issue 6, 716-723.
2. Bachand-Marleau, J., Lee, B.H.Y. and El-Geneidy, A.M. (2012). “Better understanding of factors influencing likelihood of using shared bicycle systems and frequency of use.” Transportation Research Record, Vol. 2314-09, 66－71.
3. Bai, L., Sze, N.N., Liu, P., and Haggart, A.G. (2020), “Effect of Environmental Awareness on Electric Bicycle Users’ Mode Choices.” Transportation Research Part D: Transport and Environment, Vo1.82, 102320.
4. Belsley, D.A., Kuh, E., and Welsch, R.E. (1980). “Regression Diagnostics: Identifying influential data and sources of collinearity.” New York:John Wiley and Sons.
5. Boarnet, M. and Crane, R. (2001). “The influence of land use on travel behavior: specification and estimation strategies.” Transportation Research Part A: Policy and Practice, Vol. 35, Issue 9, 823–845.
6. Box, G.E.P. and Jenkins, G.M. (1987). “Time Series Analysis: Forecasting and Control, 5th Edition.” New York:John Wiley and Sons.
7. Cervero, R. (1996). “Mixed land-uses and commuting: evidence from the American housing survey.” Transportation Research Part A: Policy and Practice, Vol. 30, Issue 5, 361–377.
8. Cervero, R. and Radisch, C. (1996). “Travel choices in pedestrian versus automobile oriented neighborhoods.” Transportation Policy, Vol. 3. Issue 3, 127–141.
9. Cervero, R. (2002). “Built Environments and Mode Choice: Toward a Normative Framework.” Transportation Research Part D, Vol. 7, 265-284.
10. Chang, H. L., and C. L. Lin. (2013). “Explore the willingness and influential factors to walk.” Ph.D. Dissertation, Department of Transportation & Logistics Management, NYCU. http://140.113.39.130/cgi-bin/gs32/tugsweb.cgi/ccd=kB8XZn/record?r1=1&h1=1#XXXX.
11. Chen, Z., Lierop, D.V., and Ettema, D. (2020). “Dockless bike-sharing systems: what are the implications?” Transport Reviews, Vol. 40, Issue 3, 333-353.
12. Cheng, Y.H., and Lin, Y.C. (2017). “Expanding the Effect of Metro Station Service Coverage by Incorporating a Public Bicycle Sharing System.” International Journal of Sustainable Transportation, Vol. 12, Issue 4, 241-252.
13. Chung, C. L. and Huang, Y. S. (2015). “Operations Assessment of the Taipei Bike Sharing System.” Master Thesis, Department of Transportation Management, TKU.
14. Cohen, J., 1988. “Statistical power analysis for the behavioral sciences, 2nd edition.” New York, Routledge.
15. Cohen, B. H., 2003. “Explaining psychological statistics, 2nd edition.” New York: Wiley.
16. Cui, Y., Mishra, S. and Welch, T.F. (2014). “Land use effects on bicycle ridership: a framework for state planning agencies.” Journal of Transport Geography, Vol. 41,

220-228.
17. Diana, M. (2010). “From Mode Choice to Model Diversion: A New Behavioural Paradigm and an Application to The Study of The Demand for Innovative Transport Services.” Technological Forecasting and Social Change, Vol. 77, 429-441.
18. dell’Olio, L., Ibeas, A., Bordagaray, M. and Ortúzar, J.D.D. (2014). “Modeling the effects of pro bicycle infrastructure and policies toward sustainable urban mobility.” Journal of Urban Planning and Development, Vol. 140, Issue 2, 1-8.
19. DeMaio, P. and Gifford, J. (2004). “Will Smart Bikes Succeed as Public Transportation in the United States?” Journal of Public Transportation, Vol. 7, No. 2, 1-15.
20. DeMaio, P. (2009). “Bike-sharing:history, impacts, models of provision, and future.” Journal of Public Transportation, Vol. 12, No.4, 41–56.
21. Department of Statistics (2016). “Kaohsiung City socioeconomic statistical data.” Taiwan: Kaohsiung City Government. http://kcgdg.kcg.gov.tw/kcgstat/page/default.aspx.
22. Department of Transportation (2020). “Introduction of YouBike system.” Taiwan: Taipei City Government. https://www.dot.gov.taipei/News_Content.aspx?n=44C34CB5F656CF90&sms=9BA74DCA64C550B3&s=FAC57225039CE3B6.
23. Department of Transportation, Taipei City Government (2016). YouBike System. Retrieved June 30, 2022.
24. Dill, J. and Voros, K. (2007). “Factors Affecting Bicycling Demand.” Transportation Research Record: Journal of the Transportation Research Board, No. 2031, 9-17.
25. El-Assi, W., Mahmound, M.S. and Habib, K.N. (2015). “Effects of built environment and weather on bike sharing demand: a station level analysis of commercial bike sharing in Toronto.” Transportation, Vol. 44, 589-613.
26. Ewing, R. and Cervero, R. (2001). “Travel and the Build Environment: A Synthesis.” Transportation Research Record: Journal of the Transportation Research Board, Vol. 1780, Issue 1, 81-114.
27. Faghih-Imani, A., Eluru, N., EI-Geneidy, A.M., Rabbat, M. and Haq, U. (2014). “How land-use and urban form impact bicycle flows: evidence from the bicycle-sharing system (BIXI) in Montreal.” Journal of Transport Geography, Vol. 41, 306-314.
28. Faghih-Imani, A. and Eluru, N. (2016a). “Determining the role of bicycle sharing system infrastructure installation decision on usage: case study of Montreal BIXI system.” Transportation Research Part A: Policy and Practice, Vol. 94, 685-698.
29. Faghih-Imani, A. and Eluru, N. (2016b). “Incorporating the impact of spatio-temporal interactions on bicycle sharing system demand: A case study of New York CitiBike system.” Journal of Transport Geography, Vol. 54, 218-227.
30. Faghih-Imani, A, Anowar, S., J. Miller, E. and Eluru, N. (2017a). “Hail a cab or ride a bike? A travel time comparison of taxi and bicycle-sharing systems in New York City.” Transportation Research Part A: Policy and Practice, Vol. 101, 11-21.
31. Faghih-Imani, A., Hampshire. R., Marla, L. and Eluru, N. (2017b). “An empirical

analysis of bike sharing usage and rebalancing: Evidence from Barcelona and Seville." Transportation Research Part A: Policy and Practice, Vol. 97, 177-191.
32. Fishman, E., Washington, S. and Haworth, N. (2012). " Barriers and facilitators to public bicycle scheme use: A qualitative approach." Transportation Research Part F: Traffic Psychology and Behaviour, Vol. 15, Issue 6, 686-698.
33. Fishman, E., Washington, S., and Haworth, N. (2013). "Bike share: A synthesis of the literature." Transport Reviews, Vol. 33, Issue 2, 148–165.
34. Fishman, E. (2016). "Bikeshare: a review of recent literature." Transportation Reviews, Vol.36, No. 1, 92-113.
35. Gao, X. and Lee, G.M. (2019). "Moment-based rental prediction for bicycle-sharing transportation systems using a hybrid genetic algorithm and machine learning." Computers and Industrial Engineering, Vol. 128, 60-69.
36. Gebhart, K. and Noland, R. (2014). "The impact of weather conditions on Bikeshare trips in Washington, DC." Transportation, Vol. 41, Issue 6, 1205–1225.
37. González, F., Melo-Riquelme, C. and Grange, L.D. (2016). "A combined destination and route choice model for a bicycle sharing system." Transportation, Vol. 43, Issue 3, 407 – 423.
38. Guidon, S., Reck, D.J. and Axhausen, K. (2020). "Expanding a(n) (electric) bicycle-sharing system to a new city: Prediction of demand with spatial regression and random forests." Journal of Transport Geography, Vo1. 84, 102692.
39. Hampshire, R. C., and Marla. L. (2012). "An analysis of bike sharing usage: Explaining trip generation and attraction from observed demand." 91st Annual Meeting of the Transportation Research Board, 12-2099.
40. Handy, S. L., Boarnet, M. G., Ewing, R., and Killingsworth, R. E. (2002). "How the built environment affects physical activity: Views from urban planning." American Journal of preventive medicine, 23(2 Suppl), 64-73.
41. Handy, S., van Wee, B., and Kroesen, M. (2014). "Promoting cycling for transport: Research needs and challenges." Transport Reviews, Vol 34, Issue 1, 4–24.
42. Hankey, S., Lindsey, G., Wang, X., Borah, J., Hoff, K., Utecht, B. and Xu, Z. (2012). "Estimating use of non-motorized infrastructure: Models of bicycle and pedestrian traffic in Minneapolis, MN." Landscape and Urban Planning, Vol. 107, Issue 3, 307-316.
43. Hensher, D. A., and Reyes, A. J. (2000). "Trip Chaining as a Barrier to The Propensity to Use Public Transport." Transportation, Vol. 27, 341-261.
44. Holmgren, J., and Ivehammar, P. (2020). "Mode Choice in Home-to-Work Travel in Mid-Size Towns: The competitiveness of Public Transport When Bicycle and Walking are Viable Options." Transportation Research Procedia, Vol. 48, 1635-1643.
45. Hu, S. R., and Liu, C. T. (2014). "An optimal location model for the bicycle sharing system: A case study of the Kaohsiung CityBike system." (in Mandarin) Transportation Planning Journal 43 (4), 367–392.
46. Hu, S.R., and Liu, C.T. (2022). "Effects of Trip Generation and Attraction Attributes on Bicycle Sharing System Ridership." forthcoming in Journal of Urban Planning and

Development.
47. Ji, Y., Fan, Y. L., Ermagun, A., Cao, X., Wang, W., and Das, K. (2017). “Public Bicycle as a Feeder Mode to Rail Transit in China: The Role of Gender, Age, Income, Trip Purpose, and Bicycle Theft Experience.” International Journal of Sustainable Transportation, Vol. 11, Issue 4, 208-317.
48. Kaohsiung City Government. (2020a). Administrative Districts. Retrieved June 9, 2020. https://www.kcg.gov.tw/EN/cp.aspx?n=C8B81163CD9BA0FD .
49. Kaohsiung City Government. (2020b). Law and Regulations Retrieving System. Retrieved June 30, 2022. https://outlaw.kcg.gov.tw/LawContent.aspx?id=GL000535#lawmenu
50. Kaohsiung Public Bike. (2020). CityBike Guidance. Retrieved June 9, 2020. https://www.c-bike.com.tw/Portal/zh-TW/Other/DownloadList..
51. Kendall, M.G. and Ord, J.K. (1990). “Time Series 3rd edition” London: Oxford University Press.
52. Kittelson and Associates (2003). “Transit capacity and quality of service manual second edition.” Transportation Research Board, Washington D.C.
53. Krizek, K.J., Barnes, G. and Thompson, K. (2009). “Analyzing the effect of bicycle facilities on commute mode share over time.” Journal of Urban Planning and Development, Vol. 135, Issue 2, 66-73.
54. Lee, B. J., Fujiwara, A., Zhang, J., and Sugie, Y. (2003). “Analysis of Mode Choice Behaviours based on Latent Class Models.” 10th International Conference on Travel Behaviour Research, Lucerne, Switzerland.
55. Lindsey, G., Wilson, J., Rubchinskaya, E., Yang, J., and Han, Y. (2007). “Estimating urban trail traffic: Methods for existing and proposed trails.” Landscape and Urban Planning, Vol. 81, Issue 4, 299-315.
56. Li, H., Huang, H., and Liu, J. (2010). “Parameter Estimation of the Mixed Logit Model and Its Application.” Journal of Transportation Systems Engineering and Information Technology, Vol. 10, Issue 5, 73-78.
57. Majumdar, B.B. and Mitra, S. (2018). “Development of level of service criteria for evaluation of bicycle suitability.” Journal of Urban Planning and Development, Vol. 144, Issue 2.
58. Marqués, R., Hernández-Herradora, V., Calvo-Salazar, M. and García-Cebrián, J.A. (2015). “How infrastructure can promote cycling in cities: Lessons from Seville.” Research in Transportation Economics, Vol. 53, 31-44.
59. Mateo-Babiano, I., Bean, R., Corcoran, J. and Pojani, D. (2016). “How does our natural and built environment affect the use of bicycle sharing?” Transportation Research Part A: Policy and Practice, Vol. 94, 295-307.
60. Mattson, J. and Godavarthy, R. (2017). “Bike share in Fargo, North Dakota: Keys to success and factors affecting ridership.” Sustainable Cities and Society, Vol. 34, 174-182.
61. Médard de Chardon, C., and Caruso, G. (2015). “Estimating bike-share trips using station level data.” Transportation Research Part B: Methodological, Vol. 78, 260-279.

62. Médard de Chardon, C., Caruso, G. and Thomas, I. (2017). “Bicycle sharing system ‘success’ determinants.” Transportation Research Part A: Policy and Practice, Vol. 100, 202-214.
63. Meddin, R. and DeMaio, P. (2022). “The bike-sharing world map.” Retrieved March 24, 2022. https://bikesharingworldmap.com/#/all/2/0/51.5/.
64. Morton, C., Kelley, S., Monsuur, F. and Hui, T. (2021). “A spatial analysis of demand patterns on a bicycle sharing scheme: Evidence from London.” Journal of Transport Geography, Vol. 94, 103125.
65. Nikita, A. (2018). “Understanding bike-sharing acceptability and expected usage patterns in the context of a small city novel to the concept: A story of ‘Greek Drama.” Transportation Research Part F: Traffic Psychology and Behaviour, Vol. 56, 306-321.
66. Noland, R. B., M. J. Smart, and Z. Guo. 2016. “Bikeshare trip generation in New York City.” Transportation Research Part A: Policy and Practice, Vol.94,164–181.
67. O’Fallon, C., Sullivan, C., and Hensher, D. A. (2004). “Constraints affecting mode choices by morning car commuters.” Transport Policy, Vol. 11, 17-29.
68. Parkin, J., Wardman, M. and Page, M. (2008). “Estimation of the determinants of bicycle mode share for the journey to work using census data.” Transportation, Vol. 35, Issue 1, 93–109.
69. Parkes, S. D., Marsden, G., Shaheen, S. A., and Cohen, A. P. (2013). “Understanding the diffusion of public bike sharing systems: Evidence from Europe and North America.” Journal of Transport Geography, Vol. 31, 94–103.
70. Rawlings, J.O., Pantula, S.G., and Dickey, D.A. (2001). “ Applied Regression Analysis: A Research Tool, Second Edition.” Berlin: Springer.
71. Rixey, R.A. (2013). “ Station-Level Forecasting of Bike sharing Ridership: Station Network effects in Three U.S Systems.” Transportation Research Record: Journal of the Transportation Research Board, Vol. 2381, Issue 1, 46-55.
72. Rowangould, G.M. and Tayarani, M. (2016). “Effect of bicycle facilities on travel mode choice decisions.” Journal of Urban Planning and Development, Vol. 142, Issue 4.
73. Saelens, B.E., Sallis, J.F. and Frank, L.D. (2003). “Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures.” Annals of Behavioral Medicine, Vol.25, Issue 2, 80–91.
74. Saltykova, K., Ma, X., Yao, L. and Kong, H. (2022). “Environmental impact assessment of bike-sharing considering the modal shift from public transit.” Transportation Research Part D: Transport and Environment, Vo1. 94, 164–181.
75. Sarwar, R., Cho, H., Cox, S.J., Mago, P.J. and Luck, R. (2017). “Field validation study of a time and temperature indexed autoregressive with exogenous (ARX) model for building thermal load prediction.” Energy, Vol. 119, 483-496.
76. Schwarz, G. (1978). “Estimating the dimension of a model.” The Annals of Statistics, Vol. 6, Number 2, 461-464.
77. Scott, D.M. and Ciuro, C. (2019). “What factors influence bike share ridership? An investigation of Hamilton, Ontario’s bike share hubs.” Travel Behaviour and

Society, Vol.16,50-58.
78. Shaheen, S., Guzman, S. and Zang, H. (2010). “Bikesharing in Europe, the Americas, and Asia.” Transportation Research Record: Journal of the Transportation Research Board, No. 2143, 159-167.
79. Shaheen, S., Cohen, A. P., and Martin, E. W. (2013). “Public bikesharing in North America: Early operator understanding and emerging trends.” Transportation Research Record: Journal of the Transportation Research Board, No. 2387, 83–92.
80. Shaheen, S., and Cohen, A. (2019). “Shared micromobility policy toolkit: Docked and dockless bike and scooter sharing. Berkeley, CA: UC Berkeley.” Transportation Sustainability Research Center, University of 814. California. https://escholarship.org/uc/item/00k897b5.
81. Shen, Y., Zhang, X. and Zhao, J. (2018) “Understanding the usage of dockless bike sharing in Singapore.” International Journal of Sustainable Transportation, Vol. 12, Issue 9, 686-700.
82. Sun, F., Chen, P. and Jiao, J. (2018) “Promoting public bike-sharing: A lesson from the unsuccessful Pronto system.” Transportation Research Part D: Transport and Environment, Vo1. 63, 533–547.
83. Taipei City Government (2020). Taipei City First Generation Bike Sharing System. Retrieved April 1, 2020. http://www.tptea.org.tw/Data/study/sdy010.pdf.
84. Transportation Bureau of Kaohsiung City Government, (2009). “Report of household interview and trip characteristics analysis in Kaohsiung metropolitan.” Taiwan: Kaohsiung City Government.
85. Wampold, B. E., and Freund, R. D. (1987). “Use of multiple regression in counseling psychology research: A flexible data-analytic strategy.” J. Couns. Psychol. 34 (4): 372–382.
86. Wang, X., Lindsey, G., Schoner, J. E. and Harrison, A. (2016). “Modeling bike share station activity: effects of nearby businesses and jobs on trips to and from stations.” Journal of Urban Planning and Development, Vol. 142, Issue 1, 1－9.
87. Wang, Y., Liu, Y., Ji, S., Hou, L., Han, S.S. and Yang, L. (2018). “Bicycle lane condition and distance: case study of public bicycle system in Xi’an, China.” Journal of Urban Planning and Development, Vol. 144, Issue 2, 1-8.
88. Wang, J., and Lindsey, G. (2019). “Do new bike share stations increase member use: A quasi-experimental study.” Transportation Research Part A: Policy and Practice, Vo1. 121, Issue 3, 1–11.
89. Washington, S.P., Karlaftis, M.G. and Mannering, F.L. (2003). “Statistical and econometric methods for transportation data analysis.” New York, NY: Chapman & Hall/CRC.
90. YouBike Corporation. 2020. CityBike Equipment Introduction. Retrieved June 9, 2020. https://taipei.youbike.com.tw/use/equipment?_id=5cb99765083e7b09bc22f2b2.
91. Yun, K., Luck, R., Mago, P.J. and Cho, H. (2012). “Building hourly thermal load prediction using an indexed ARX model.” Energy and Buildings, Vol. 54, 225-233.
92. Zhao, D., Ong, G.P., Wang, W. and Zhou, W. (2021). “Estimating Public Bicycle Trip

Characteristics with Consideration of Built Environment Data." Sustainability, Vol. 13, Issue 2, 500-512.

APPENDICES

Appendix A: Models M1 and M2 estimation results

A.1 Results of Yancheng, Gushan, Zuoying and Nanzih Districts-1.

Model / IV	M1 (departures) Yancheng	Gushan	Zuoying	Nanzih	M2 (departures) Yancheng	Gushan	Zuoying	Nanzih
Constant	-173513.473 (0.000***)	6808.222 (0.008**)	-261.071 (0.899)	-14297.341 (0.004**)	-55263.126 (0.000***)	-26030.130 (0.000***)	-34184.461 (0.000***)	-13116.807 (0.001***)
Lag 1					0.650 (0.000999)	0.296 (0.000***)	0.291 (0.001***)	0.228 (0.074*)
Worker trip attraction								
Worker trip generation					-0.038 (0.693)		0.531 (0.000***)	
Student trip attraction								0.405 (0.008**)
Student trip generation						0.524 (0.000***)		
Capacity of station	0.021 (0.404)	-0.014 (0.663)	0.027 (0.528)	0.020 (0.795)				
CityBike station capacity in a 1-km buffer	-0.019 (0.424)	0.027 (0.417)	-0.045 (0.400)	-0.022 (0.784)	0.004 (0.919)	-0.056 (0.229)		0.184 (0.258)
Bikeway length								
Number of streetlamps	0.770 (0.000***)			0.342 (0.000***)				0.216 (0.023*)
Length of major road		0.972 (0.000***)	0.942 (0.000***)					-0.336 (0.000***)
Number of parks	0.522 (0.000***)				0.258 (0.004**)			
Number of companies	0.409 (0.000***)				0.110 (0.004**)			
Number of factories	-0.068 (0.035**)				0.042 (0.416)	-0.017 (0.762)	-0.302 (0.000***)	
Number of hotels			-0.051 (0.480)			-0.147 (0.000***)	0.149 (0.000***)	
Number of markets		0.226 (0.001***)					0.067 (0.290)	
Number of schools						0.255 (0.000***)		
Number of public transportation yard								
Tourist	0.154 (0.000***)	0.105 (0.021**)	0.094 (0.016**)	0.282 (0.001***)	0.126 (0.002**)	0.080 (0.023*)	0.066 (0.009**)	0.249 (0.000***)
Income	0.017 (0.758)	0.014 (0.715)	0.002 (0.975)	-0.332 (0.002**)	-0.003 (0.943)	-0.106 (0.000***)	-0.177 (0.000***)	-0.106 (0.267)
Private vehicle		-0.191 (0.000***)	-0.096 (0.110)			-0.159 (0.000***)		
Land value	0.562 (0.000***)	0.073 (0.121)	0.065 (0.158)	0.117 (0.263)			-0.010 (0.819)	
Adjusted R^2	0.962	0.918	0.893	0.554	0.913	0.953	0.950	0.875
F value	225.045***	126.266***	95.157***	17.126***	125.718***	215.980***	224.982***	53.233***

DW	2.143	0.910	95.157	0.345	1.884	1.771	1.728	1.879
AIC	980.747	1063.430	978.385	783.811	1141.002	1079.481	1000.923	481.141
BIC	1003.345	1083.517	998.472	798.028	1161.517	1102.560	1021.438	494.933

Note: 1.*90% level of significance, ** 95% level of significance, *** 99% level of significance
2.*t*-statistics values are shown in parenthesis.

A.2 Results of Yancheng, Gushan, Zuoying and Nanzih Districts-2.

Model / IV	M1 (departures)				M2 (departures)			
	Yancheng	**Gushan**	**Zuoying**	**Nanzih**	**Yancheng**	**Gushan**	**Zuoying**	**Nanzih**
Constant	-32881.393	-44099.830	-50711.406	-32259.434	-55263.126	-26030.131	-34184.460	-13116.807
	(0.031)	(0.000***)	(0.000***)	(0.000***)	(0.000***)	(0.000***)	(0.000***)	(0.001**)
Lag 1					0.650	0.296	0.291	0.228
					(0.000***)	(0.000***)	(0.001***)	(0.074*)
Worker population			0.980	1.345	-0.038	0.524		0.405
			(0.000***)	(0.000***)	(0.693)	(0.000***)		(0.008**)
Student population	-0.076	1.179					0.531	
	(0.491)	(0.000***)					(0.000***)	
Capacity of station	0.029	0.022	0.020	-0.054				
	(0.549)	(0.443)	(0.700)	(0.294)				
CityBike station capacity in a 1-km buffer	-0.059	0.000	-0.053	-0.026	0.004	-0.056		0.184
	(0.193)	(0.999)	(0.407)	(0.617)	(0.919)	(0.229)		(0.258)
Length of bikeway								
Number of streetlamps				0.008				0.216
				(0.918)				(0.023*)
Length of major road								-0.336
								(0.000***)
Number of parks	0.760				0.258			
	(0.000***)				(0.004**)			
Number of companies	0.053				0.110			
	(0.258)				(0.004**)			
Number of factories	0.005			0.976	0.042	-0.017	-0.302	
	(0.940)			(0.000***)	(0.416)	(0.762)	(0.000***)	
Number of hotels			-0.362			-0.147	0.149	
			(0.001***)			(0.000***)	(0.000***)	
Number of markets		0.592					0.067	
		(0.000***)					(0.290)	
Number of schools						0.255		
						(0.000***)		
Number of public transportation yard								
Tourist	0.226	0.182	0.047	0.398	0.126	0.080	0.066	0.249
	(0.000***)	(0.000***)	(0.307)	(0.000***)	(0.002**)	(0.023*)	(0.009**)	(0.000***)
Income	0.011	-0.146	-0.002	-0.172	-0.003	-0.106	-0.177	-0.106
	(0.863)	(0.000***)	(0.969)	(0.018**)	(0.943)	(0.000***)	(0.000***)	(0.267)
Private vehicle		-0.451	-0.456			-0.159		
		(0.000***)	(0.000***)			(0.000***)		
Land value		0.198	0.141				-0.010	
		(0.000***)	(0.012**)				(0.819)	
Adjusted R^2	0.861	0.934	0.854	0.811	0.913	0.953	0.950	0.875
F value	70.677***	160.158***	67.050***	48.677***	125.718***	215.980***	224.982***	53.233***
DW	1.128	1.179	67.050	0.803	1.884	1.771	1.728	1.879
AIC	1098.196	1043.250	1006.607	717.002	1141.002	1079.481	1000.923	481.141
BIC	1118.283	1063.337	1026.694	733.589	1161.517	1102.560	1021.438	494.933

Note: 1.*90% level of significance, ** 95% level of significance, *** 99% level of significance
2.*t*-statistics values are shown in parenthesis.

A.3 Results of Yancheng, Gushan, Zuoying and Nanzih Districts-1.

Model / IV	M1 (arrival)				M2 (arrival)			
	Yancheng	**Gushan**	**Zuoying**	**Nanzih**	**Yancheng**	**Gushan**	**Zuoying**	**Nanzih**
Constant	-175016.658 (0.000***)	6584.514 (0.009**)	-505.779 (0.810)	-14895.562 (0.004**)	-56769.848 (0.000***)	-27849.274 (0.000***)	-34813.078 (0.000***)	-13280.828 (0.001***)
Lag 1					0.639 (0.000***)	0.278 (0.001***)	0.293 (0.001***)	0.234 (0.064*)
Worker trip attraction					-0.037 (0.700)	0.538 (0.000***)		
Worker trip generation								
Student trip attraction							0.541 (0.000***)	0.403 (0.009**)
Student trip generation								
Station capacity	0.020 (0.442)	-0.014 (0.648)	0.026 (0.568)	0.023 (0.775)	0.004 (0.912)			
CityBike station capacity in a 1-km buffer	-0.022 (0.386)	0.029 (0.382)	-0.048 (0.382)	-0.026 (0.748)		-0.057 (0.211)		0.185 (0.254)
Bikeway length								
Number of streetlamps	0.749 (0.000***)			0.352 (0.000***)				0.211 (0.026)
Length of major road		0.974 (0.000***)	0.949 (0.000***)					-0.332 (0.000***)
Number of parks	0.522 (0.000***)				0.264 (0.004**)			
Number of companies	0.408 (0.000***)				0.111 (0.004**)			
Number of factories	-0.066 (0.052)				0.040 (0.446)	-0.012 (0.836)	-0.308 (0.000***)	
Number of hotels			-0.054 (0.469)			-0.149 (0.000***)	0.151 (0.000***)	
Number of markets		0.225 (0.000***)					0.083 (0.200)	
Number of schools						0.269 (0.000***)		
Number of public transportation yard								
Tourist	0.171 (0.000***)	0.108 (0.016**)	0.091 (0.023**)	0.272 (0.002**)	0.135 (0.001**)	0.083 (0.017*)	0.065 (0.012*)	0.247 (0.000***)
Income	0.025 (0.669)	0.014 (0.706)	0.003 (0.951)	-0.336 (0.002**)	-0.002 (0.969)	-0.109 (0.000***)	-0.183 (0.000***)	-0.095 (0.318)
Private vehicle		-0.186 (0.000***)	-0.090 (0.146)			-0.161 (0.0009***)		
Land value	0.554 (0.000***)	0.070 (0.127)	0.073 (0.126)	0.100 (0.340)			-0.015 (0.727)	
Adjusted R^2	0.958	0.922	0.888	0.547	0.911	0.955	0.947	0.876
F value	228.685***	133.214***	89.880***	16.678***	123.146***	225.030***	213.674***	53.345***
DW	2.113	0.935	1.189	0.340	1.858	1.771	1.723	1.900
AIC	992.967	1061.425	981.859	787.393	1145.023	1078.768	1004.735	482.957
BIC	1015.564	1081.512	1001.946	801.610	1165.537	1101.847	1025.250	496.749

Note: 1.*90% level of significance, ** 95% level of significance, *** 99% level of significance
2.*t*-statistics values are shown in parenthesis.

A.4 Results of YanCheng, Gushan, Zuoying and Nanzih Districts-2.

Model / IV	M1 (arrival)				M2 (arrival)			
	Yancheng	**Gushan**	**Zuoying**	**Nanzih**	**Yancheng**	**Gushan**	**Zuoying**	**Nanzih**
Constant	-34902.327 (0.024)	-44632.418 (0.000***)	-50866.737 (0.000***)	-33229.765 (0.000***)	-50035.759 (0.002**)	-27849.274 (0.000***)	-34813.078 (0.000***)	-13280.828 (0.001**)
Lag 1					0.623 (0.000***)	0.278 (0.001***)	0.293 (0.001***)	0.234 (0.064*)
Worker population			0.985 (0.000***)		-0.122 (0.298)	0.538 (0.000***)		0.403 (0.009**)
Student population	-0.067 (0.545)	1.172 (0.000***)		1.367 (0.000***)			0.541 (0.000***)	
Station capacity	0.028 (0.558)	0.021 (0.457)	0.018 (0.730)	-0.053 (0.304)				
CityBike station capacity in a 1-km buffer	-0.061 (0.178)	0.002 (0.943)	-0.055 (0.399)	-0.030 (0.563)	-0.001 (0.985)	-0.057 (0.211)		0.185 (0.254)
Length of bikeway								
Number of streetlamps				0.004 (0.961)				0.211 (0.026)
Length of major road								-0.332 (0.000***)
Number of parks	0.757 (0.000***)				0.236 (0.009**)			
Number of companies	0.057 (0.221)				0.113 (0.003**)			
Number of factories	0.003 (0.965)			0.973 (0.000***)		-0.012 (0.836)	-0.308 (0.000***)	
Number of hotels			-0.365 (0.001***)		-0.089 (0.148)	-0.149 (0.000)	0.151 (0.000***)	
Number of markets		0.587 (0.000***)					0.083 (0.200)	
Number of schools						0.269 (0.000)		
Number of public transportation yard								
Tourist	0.242 (0.000***)	0.186 (0.000***)	0.043 (0.356)	0.389 (0.000***)	0.137 (0.001**)	0.083 (0.017*)	0.065 (0.012*)	0.247 (0.000***)
Income	0.015 (0.822)	-0.147 (0.000***)	-0.002 (0.976)	-0.156 (0.031)	0.004 (0.914)	-0.109	-0.183 (0.000***)	-0.095 (0.318)
Private vehicle		-0.444 (0.000***)	-0.451 (0.000***)			-0.161 (0.000***)		
Land value		0.193 (0.000***)	0.149 (0.010**)				-0.015 (0.727)	
Adjusted R^2	0.861	0.933	0.847	0.811	0.913	0.955	0.947	0.876
F value	70.841***	158.262***	63.228***	48.963***	125.512***	225.030***	213.674***	53.345***
DW	1.181	1.159	1.486	0.804	1.847	1.771	1.723	1.900
AIC	1100.661	1046.781	1010.023	718.978	1143.343	1078.768	1004.735	482.957
BIC	1120.748	1066.868	1030.110	735.564	1163.858	1101.847	1025.250	496.749

Note: 1.*90% level of significance, ** 95% level of significance, *** 99% level of significance
2.*t*-statistics values are shown in parenthesis.

A.5 Results of Sinsing, Cianjin, Lingya and Cianjhen District-1.

Model / IV	M1 (departures) Sinsing	Cianjin	Lingya	Cianjhen	M2 (departures) Sinsing	Cianjin	Lingya	Cianjhen
Constant	18301.060 (0.000***)	4340.095 (0.014**)	-688.468 (0.727)	6621.446 (0.000***)	30759.868 (0.000***)	21662.324 (0.079*)	26918.804 (0.000***)	27470.781 (0.000***)
Lag 1					0.608 (0.000***)	0.466 (0.000***)	0.347 (0.000***)	0.554 (0.000***)
Worker trip attraction								
Worker trip generation						-0.089 (0.416)		-0.673 (0.000***)
Student trip attraction							-0.858 (0.000***)	
Student trip generation					-0.479 (0.000***)			
Capacity of station	0.138 (0.091)	-0.038 (0.277)	0.038 (0.428)	-0.047 (0.141)		-0.001 (0.980)	-0.037 (0.554)	
CityBike station capacity in a 1-km buffer	0.237 (0.030)	0.041 (0.261)	0.097 (0.081*)	0.041 (0.197)	-0.156 (0.006**)			-0.231 (0.001***)
Bikeway length		0.969 (0.000***)						
Number of streetlamps	-0.460 (0.000***)					-0.052 (0.371)		
Length of major road	-0.129 (0.022)	-0.549 (0.000***)	0.696 (0.000***)	0.623 (0.000***)	-0.095 (0.007**)			
Number of parks								
Number of companies					-0.129 (0.009**)			
Number of factories	-0.656 (0.000***)		-0.085 (0.164)	0.161 (0.000***)	-0.162 (0.001***)		-0.217 (0.000***)	0.222 (0.000***)
Number of hotels						-0.082 (0.105)		
Number of markets							0.066 (0.331)	
Number of schools	-0.022 (0.835)			0.615 (0.000***)		-0.499 (0.000***)	0.408 (0.000***)	
Number of public transportation yard				-0.042 (0.297)				0.021 (0.585)
Tourist			0.130 (0.051*)				0.025 (0.521)	
Income	-0.647 (0.000***)	-0.019 (0.608)	-0.419 (0.000***)	-0.115 (0.007**)	-0.027 (0.621)	0.009 (0.809)	0.196 (0.001***)	0.022 (0.626)
Private vehicle			0.010 (0.913)	-0.810 (0.000***)				0.019 (0.816)
Land value		0.006 (0.902)	0.478 (0.000***)					
Adjusted R^2	0.779	0.908	0.864	0.930	0.937	0.910	0.941	0.935
F value	44.385***	135.165***	66.277***	136.662***	202.511***	138.377***	191.817***	194.751***
DW	1.010	1.618	1.255	1.090	1.461	1.754	1.912	1.820
AIC	1014.915	959.055	951.919	836.872	1013.148	1134.672	1049.816	986.843
BIC	1032.176	973.568	971.270	856.223	1031.098	1152.622	1070.331	1004.793

Note: 1.*90% level of significance, ** 95% level of significance, *** 99% level of significance
2.*t*-statistics values are shown in parenthesis.

A.6 Results of Sinsing, Cianjin, Lingya and Cianjhen Districts-2.

Model / IV	M1 (departures) Sinsing	Cianjin	Lingya	Cianjhen	M2 (departures) Sinsing	Cianjin	Lingya	Cianjhen
Constant	54059.689 (0.000***)	9147.353 (0.076)	64956.327 (0.000***)	32045.693 (0.000***)	30759.869 (0.000***)	21662.324 (0.079*)	49221.392 (0.000***)	27470.782 (0.000***)
Lag 1					0.608 (0.000***)	0.466 (0.000***)	0.187 (0.029*)	0.554 (0.000***)
Worker population	-1.215 (0.000***)	-0.099 (0.318)	-1.171 (0.000***)	-0.861 (0.000***)				
Student population					-0.479 (0.000***)	-0.089 (0.416)	-1.027 (0.000***)	-0.673 (0.000***)
Capacity of station	0.069 (0.162)	-0.031 (0.394)	0.007 (0.861)	-0.066 (0.199)				
CityBike station capacity in a 1-km buffer	0.081 (0.222)	0.033 (0.374)	0.066 (0.179)	-0.041 (0.410)	-0.156 (0.006**)	-0.001 (0.980)	-0.022 (0.690)	-0.231 (0.001***)
Length of bikeway		0.892 (0.000***)						
Number of streetlamps	-0.356 (0.000***)					-0.052 (0.371)		
Length of major road	-0.072 (0.035)	-0.539 (0.000***)			-0.095 (0.007**)			
Number of parks								
Number of companies					-0.129 (0.009**)			
Number of factories	-0.450 (0.000***)		-0.324 (0.000***)	0.454 (0.000***)	-0.162 (0.001**)		-0.356 (0.000***)	0.222 (0.000***)
Number of hotels	-0.487 (0.000***)		-0.969 (0.000***)			-0.082 (0.105)	-0.380 (0.000***)	
Number of markets							-0.013 (0.830)	
Number of schools				-0.015 (0.824)		-0.499 (0.000***)	0.290 (0.000***)	
Number of public transportation yard								0.021 (0.585)
Tourist			0.125 (0.029)				0.009 (0.810)	
Income		-0.013 (0.721)	-0.162 (0.017*)	-0.114 (0.038)	-0.027 (0.621)	0.009 (0.809)		0.022 (0.626)
Private vehicle		0.039 (0.508)						0.019 (0.816)
Land value								
Adjusted R^2	0.920	0.908	0.890	0.818	0.937	0.910	0.951	0.935
F value	142.499***	116.017***	96.140***	62.247***	202.511***	138.388***	232.451***	194.751***
DW	1.603	1.629	1.433	0.696	1.461	1.754	1.795	1.812
AIC	926.514	959.943	933.318	914.297	1013.148	1134.672	1032.276	986.843
BIC	943.775	976.875	950.250	928.810	1031.098	1152.622	1052.790	1004.793

Note: 1.*90% level of significance, ** 95% level of significance, *** 99% level of significance
2.*t*-statistics values are shown in parenthesis.

A.7 Results of Sinsing, Cianjin, Lingya and Cianjhen Districts-1.

Model / IV	M1 (arrival)				M2 (arrival)			
	Sinsing	**Cianjin**	**Lingya**	**Cianjhen**	**Sinsing**	**Cianjin**	**Lingya**	**Cianjhen**
Constant	17948.893 (0.000***)	5330.108 (0.003**)	-1681.410 (0.386)	6693.364 (0.000***)	29840.939 (0.000***)	22934.047 (0.065*)	28933.283 (0.000***)	28078.748 (0.000***)
Lag 1					0.605 (0.000***)	0.439 (0.000***)	0.343 (0.000***)	0.558 (0.000***)
Worker trip attraction								
Worker trip generation								
Student trip attraction					-0.484 (0.000***)	-0.089 (0.422)	-0.900 (0.000***)	-0.673 (0.000***)
Student trip generation								
Station capacity	0.133 (0.103)	-0.043 (0.235)	0.027 (0.563)	-0.047 (0.142)		-0.001 (0.983)	-0.034 (0.589)	
CityBike station capacity in a 1-km buffer	0.231 (0.034)	0.045 (0.223)	0.082 (0.160)	0.037 (0.242)	-0.155 (0.006**)			-0.229 (0.001)
Bikeway length		0.956 (0.000***)						
Number of streetlamps	-0.462 (0.000***)					-0.069 (0.246)		
Length of major road	-0.134 (0.017)	-0.581 (0.000***)	0.729 (0.000***)	0.629 (0.000***)	-0.098 (0.006**)			
Number of parks								
Number of companies					-0.129 (0.009**)			
Number of factories	-0.653 (0.000***)		-0.095 (0.124)	0.163 (0.000***)	-0.162 (0.001***)		-0.230 (0.000***)	0.221 (0.000***)
Number of hotels						-0.086 (0.096*)		
Number of markets							0.025 (0.714)	
Number of schools	-0.037 (0.717)			0.612 (0.000***)		-0.535 (0.000***)	0.427 (0.000***)	
Number of public transportation yard				0.027 (0.495)				0.013 (0.719)
Tourist			0.137 (0.042)				0.028 (0.473)	
Income	-0.642 (0.000***)	-0.019 (0.608)	-0.435 (0.000***)	-0.118 (0.006)	-0.025 (0.656)	0.009 (0.805)	0.203 (0.001***)	0.022 (0.632)
Private vehicle			0.075 (0.436)	-0.803 (0.000***)				0.022 (0.781)
Land value		-0.006 (0.910)	0.481 (0.000***)					
Adjusted R^2	0.780	0.903	0.863	0.930	0.936	0.907	0.938	0.935
F value	44.480***	128.658***	64.469***	136.346***	200.933***	132.956***	180.956***	195.742***
DW	1.004	1.641	1.245	1.104	1.448	1.736	1.911	1.804
AIC	1009.089	959.505	949.410	840.243	1007.437	1135.429	1051.041	990.262
BIC	1026.350	974.018	968.761	859.594	1025.387	1153.379	1071.556	1008.212

Note: 1.*90% level of significance, ** 95% level of significance, *** 99% level of significance
2.*t*-statistics values are shown in parenthesis.

A.8 Results of Sinsing, Cianjin, Lingya and Cianjhen Districts-2.

Model / IV	M1 (arrival) Sinsing	M1 (arrival) Cianjin	M1 (arrival) Lingya	M1 (arrival) Cianjhen	M2 (arrival) Sinsing	M2 (arrival) Cianjin	M2 (arrival) Lingya	M2 (arrival) Cianjhen
Constant	51939.559	10315.955	64408.924	32769.127	29840.939	22934.047	28933.283	27583.001
	(0.000***)	(0.046)	(0.000***)	(0.000***)	(0.000***)	(0.065*)	(0.000***)	(0.000***)
Lag 1					0.605	0.439	0.343	0.556
					(0.000***)	(0.000***)	(0.000***)	(0.000***)
Worker population	-1.208	-0.105	-1.174	-0.862				
	(0.000***)	(0.301)	(0.000***)	(0.000***)				
Student population					-0.484	-0.089	-0.900	-0.653
					(0.000***)	(0.422)	(0.000***)	(0.000***)
Station capacity	0.064	-0.035	0.009	-0.065		-0.001	-0.034	
	(0.193)	(0.345)	(0.843)	(0.204)		(0.983)	(0.589)	
CityBike station capacity in a 1-km buffer	0.075	0.037	0.057	-0.045	-0.155			-0.235
	(0.257)	(0.330)	(0.251)	(0.369)	(0.006**)			(0.001***)
Length of bikeway		0.874						
		(0.000***)						
Number of streetlamps	-0.360					-0.069		
	(0.000***)					(0.246)		
Length of major road	-0.077	-0.571			-0.098			
	(0.024**)	(0.000***)			(0.006**)			
Number of parks								
Number of companies					-0.129			
					(0.009**)			
Number of factories	-0.449		-0.332	0.456	-0.162		-0.230	0.218
	(0.000***)		(0.000***)	(0.000***)	(0.001***)		(0.000***)	(0.000***)
Number of hotels	-0.468		-1.002			-0.086		
	(0.000***)		(0.000***)			(0.096*)		
Number of markets							0.025	
							(0.714)	
Number of schools						-0.535	0.427	
						(0.000***)	(0.000***)	
Number of public transportation yard				-0.030				0.015
				(0.647)				(0.679)
Tourist			0.121				0.028	
			(0.038**)				(0.473)	
Income		-0.014	-0.182	-0.149	-0.025	0.009	0.203	0.015
		(0.724)	(0.009**)	(0.033)	(0.656)	(0.805)	(0.001***)	(0.695)
Private vehicle								
Land value		0.030						
		(0.627)						
Adjusted R^2	0.919	0.903	0.887	0.816	0.936	0.907	0.938	0.936
F value	140.973***	110.553***	92.892***	61.772***	200.933***	132.956***	180.956***	230.744***
DW	1.596	1.650	1.409	0.614	1.448	1.736	180.956	1.796
AIC	921.705	960.315	932.478	918.015	1007.437	1135.429	1051.041	988.347
BIC	938.966	977.247	949.410	932.528	1025.387	1153.379	1071.556	1003.733

Note: 1.*90% level of significance, ** 95% level of significance, *** 99% level of significance
2.*t*-statistics values are shown in parenthesis.

A.9 Results of Cijin, Siaogang and Fongshan Districts-1.

Model / IV	M1 (departures)			M2 (departures)		
	Cijin	**Siaogang**	**Fongshan**	**Cijin**	**Siaogang**	**Fongshan**
Constant	-179.552 (0.823)	-21619.876 (0.004)	-9850.106 (0.001)	17859.751 (0.300)	-16874.853 (0.000***)	-15073.165 (0.000***)
Lag 1				0.698 (0.000***)	0.526 (0.000***)	0.348 (0.000***)
Worker trip attraction						
Worker trip generation					0.295 (0.000***)	
Student trip attraction						
Student trip generation				-0.155 (0.329)		0.704 (0.000***)
Capacity of station		-0.040 (0.672)	-0.025 (0.701)			
CityBike station capacity in a 1-km buffer	0.095 (0.466)	0.159 (0.132)	0.010 (0.882)	-0.079 (0.604)		
Bikeway length						
Number of streetlamps		0.340 (0.001***)				
Length of major road	-0.150 (0.361)		1.073 (0.000***)	-0.251 (0.149)		
Number of parks			0.164 (0.098*)			
Number of companies						
Number of factories			-0.052 (0.706)		-0.017 (0.791)	-0.066 (0.064)
Number of hotels					-0.047 (0.375)	
Number of markets	-0.543 (0.054)			-0.023 (0.872)	-0.035 (0.619)	0.007 (0.911)
Number of schools						0.154 (0.007**)
Number of public transportation yard						
Tourist	-0.089 (0.471)	0.033 (0.799)		0.070 (0.583)	0.086 (0.092)	
Income	0.544 (0.106)	-0.287 (0.007**)	0.045 (0.738)		-0.056 (0.163)	-0.045 (0.310)
Private vehicle			0.350 (0.028)		-0.217 (0.003)	-0.220 (0.000***)
Land value		-0.218 (0.101)				
Adjusted R^2	-0.006	0.329	0.752	0.385	0.926	0.955
F value	0.916***	7.698***	36.470***	5.383***	149.327***	286.294***
DW	0.527	1.220	1.346	1.688	2.099	1.789
AIC	744.743	839.386	825.246	418.139	920.578	899.292
BIC	756.263	853.899	842.177	428.706	941.093	917.243

Note: 1.*90% level of significance, ** 95% level of significance, *** 99% level of significance
2.*t*-statistics values are shown in parenthesis.

A.10 Results of Cijin, Siaogang and Fongshan Districts-2.

Model / IV	M1 (departures) Cijin	Siaogang	Fongshan	M2 (departures) Cijin	Siaogang	Fongshan
Constant	29257.616 (0.098*)	-54755.804 (0.008)	-17744.473 (0.000***)	17859.749 (0.300)	-16874.853 (0.000***)	-15073.165 (0.000***)
Lag 1				0.698 (0.000***)	0.526 (0.000***)	0.348 (0.000***)
Worker population	-0.278 (0.106)	0.321 (0.077)	0.962 (0.000***)		0.295 (0.000***)	
Student population				-0.155 (0.329)		0.704 (0.000***)
Capacity of station		-0.050 (0.594)	-0.026 (0.686)			
CityBike station capacity in a 1-km buffer	0.095 (0.466)	0.101 (0.352)	0.006 (0.924)	-0.079 (0.604)		
Length of bikeway						
Number of streetlamps		0.251 (0.018)				
Length of major road	-0.185 (0.299)			-0.251 (0.149)		
Number of parks			-0.414 (0.001***)			
Number of companies						
Number of factories		0.117 (0.427)	-0.105 (0.445)		-0.017 (0.791)	-0.066 (0.064)
Number of hotels					-0.047 (0.375)	
Number of markets	-0.125 (0.371)			-0.023 (0.872)	-0.035 (0.619)	0.007 (0.911)
Number of schools						0.154 (0.007**)
Number of public transportation yard		-0.108 (0.471)				
Tourist	-0.089 (0.471)			0.070 (0.583)	0.086 (0.093)	
Income		-0.252 (0.020)	-0.113 (0.367)		-0.056 (0.163)	-0.045 (0.310)
Private vehicle			-0.529 (0.000***)		-0.217 (0.003**)	-0.220 (0.000***)
Land value						
Adjusted R^2	-0.006	0.349	0.753	0.385	0.926	0.955
F value	0.916***	7.288***	36.666***	5.383***	149.327***	286.294***
DW	0.527	1.244	1.342	1.688	2.099	1.789
AIC	744.743	837.727	824.901	418.139	920.578	899.292
BIC	756.263	854.659	841.833	428.706	941.093	917.243

Note: 1.*90% level of significance, ** 95% level of significance, *** 99% level of significance
2.*t*-statistics values are shown in parenthesis.

A.11 Results of Cijin, Siaogang and Fongshan Districts-1.

Model / IV	M1 (arrival) Cijin	M1 (arrival) Siaogang	M1 (arrival) Fongshan	M2 (arrival) Cijin	M2 (arrival) Siaogang	M2 (arrival) Fongshan
Constant	-78.470	-22957.218	-9680.079	13736.237	-17599.299	-15604.153
	(0.927)	(0.004)	(0.002)	(0.455)	(0.000***)	(0.000***)
Lag 1				0.707	0.522	0.334
				(0.000***)	(0.000***)	(0.000***)
Worker trip attraction						
Worker trip generation						
Student trip attraction				-0.109	0.293	0.723
				(0.489)	(0.000***)	(0.000***)
Student trip generation						
Station capacity		-0.039	-0.017			
		(0.677)	(0.793)			
CityBike station capacity in a 1-km buffer	0.099	0.146	0.012	-0.052		
	(0.444)	(0.159)	(0.860)	(0.726)		
Bikeway length						
Number of streetlamps		0.331				
		(0.001***)				
Length of major road	0.024		1.085	-0.196		
	(0.882)		(0.000***)	(0.254)		
Number of parks						
Number of companies						
Number of factories			-0.147		-0.013	-0.067
			(0.139)		(0.839)	(0.060)
Number of hotels					-0.046	
					(0.382)	
Number of markets	-0.456			-0.080	-0.027	0.012
	(0.102)			(0.570)	(0.694)	(0.837)
Number of schools						0.161
						(0.005**)
Number of public transportation yard			0.070			
			(0.615)			
Tourist	-0.111	0.039		0.057	0.085	
	(0.363)	(0.760)		(0.655)	(0.093*)	
Income	0.257	-0.255	0.050		-0.052	-0.044
	(0.438)	(0.015**)	(0.711)		(0.183)	(0.326)
Private vehicle			0.346		-0.220	-0.224
			(0.030**)		(0.002**)	(0.000***)
Land value		-0.280				
		(0.033**)				
Adjusted R^2	0.009	0.349	0.751	0.400	0.928	0.954
F value	1.137***	8.339***	36.245***	5.674***	152.937***	283.369***
DW	0.549	1.205	1.341	1.603	2.069	1.805
AIC	754.246	849.741	825.264	423.632	927.921	899.966
BIC	765.766	864.254	842.196	434.200	948.436	917.917

Note: 1.*90% level of significance, ** 95% level of significance, *** 99% level of significance
2.*t*-statistics values are shown in parenthesis.

A.12 Results of Cijin, Siaogang and Fongshan Districts-2.

Model / IV	M1 (arrival)			M2 (arrival)		
	Cijin	**Siaogang**	**Fongshan**	**Cijin**	**Siaogang**	**Fongshan**
Constant	14880.886	-59154.925	-17638.282	13736.234	-17599.299	-15604.153
	(0.426)	(0.007)	(0.000***)	(0.455)	(0.000***)	(0.000***)
Lag 1				0.707	0.522	0.334
				(0.000***)	(0.000***)	(0.000***)
Worker population	-0.131	0.325			0.293	
	(0.438)	(0.069*)			(0.000***)	
Student population			0.972	-0.109		0.723
			(0.000***)	(0.487)		(0.000***)
Station capacity		-0.049	-0.018			
		(0.596)	(0.778)			
CityBike station capacity in a 1-km buffer	0.099	0.087	0.008	-0.052		
	(0.444)	(0.415)	(0.901)	(0.726)		
Length of bikeway						
Number of streetlamps		0.237				
		(0.023**)				
Length of major road	0.008			-0.196		
	(0.966)			(0.254)		
Number of parks			-0.437			
			(0.001***)			
Number of companies						
Number of factories		0.179	-0.124		-0.013	--0.067
		(0.217)	(0.371)		(0.839)	(0.060*)
Number of hotels					-0.046	
					(0.382)	
Number of markets	-0.258			-0.080	-0.027	0.012
	(0.065*)			(0.570)	(0.694)	(0.837)
Number of schools						0.161
						(0.005**)
Number of public transportation yard						
Tourist	-0.111	-0.105		0.057	0.085	
	(0.363)	(0.476)		(0.655)	(0.093*)	
Income		-0.223	-0.130		-0.052	-0.044
		(0.035**)	(0.380)		(0.183)	(0.326)
Private vehicle			-0.542		-0.220	-0.224
			(0.000***)		(0.002**)	(0.000***)
Land value						
Adjusted R^2	0.009	0.371	0.752	0.400	0.928	0.954
F value	1.137***	7.917***	36.436***	5.674***	152.937***	283.369***
DW	0.549	1.229	1.336	1.603	2.069	1.805
AIC	754.246	847.802	824.927	423.632	927.921	899.966
BIC	765.766	864.734	841.859	434.200	948.436	917.917

Note: 1.*90% level of significance, ** 95% level of significance, *** 99% level of significance
2.*t*-statistics values are shown in parenthesis.

A.13 Descriptive statistics of the collected data (urban district)

Variables	Min	Max	Mean	Std. deviation	Data resource
Station-level data					
CityBike departures (per station/ per month)	46	4,745	1,378	811	
CityBike arrivals (per station/ per month)	46	4,747	1380		KRTC
Capacity of station (units)	12	32	28.34	6	
CityBike station capacity in a 1-km buffer (units)	0	254	87	57	
District-level data					
Population (people/ per square kilometer)	3,428	28,205	14,415		
Worker population (worker/ per square kilometer)	2,414	18,601	9,749	3,933	
Student population (student/ per square kilometer)	517	3,466	1,991	688	DOBAS, TBKCG, DOSKCG
Tourist (people/ per square kilometer)	0	711,902	18,510	52,999	
Income (NT$/ year)	446,578	826,201	544,905	54,730	
Private vehicle (units / per square kilometer)	3,552	30,923	15,258	6334	
Bikeway length (meter/ per square kilometer)	270	58,482	4,937	9,706	PBWKCG
Number of streetlamps (units / per square kilometer)	293	3,095	1113	687	
Length of major road (meter/ per square kilometer)	6,986	35,116	17,976	6,451	LABKCG
Land value (NT$/ per square kilometer)	824	245,834	41,854	43,526	
Number of parks (units / per square kilometer)	2	12	4	2	
Number of companies (units / per square kilometer)	65	2,348	705	582	
Number of factories (units / per square kilometer)	2	29	10	9	DOSKC, MOTC
Number of hotels (units / per square kilometer)	0	85	12	22	
Number of markets (units / per square kilometer)	0	6	1	1	
Number of schools (units / per square kilometer)	1	4	2	1	
Number of public transportation yard (units / per square kilometer)	1	25	7	7	TBKCG
Trip attraction (trips/ per square kilometer)	8,612	84,034	36,754	21,012	
Trip generation (trips/ per square kilometer)	2,371	58,913	28,826	16,010	DOSKC, MOTC (Census data)
Worker trip attraction (trips/ per square kilometer)					
Worker trip generation (trips/ per square kilometer)					

Student trip attraction (trips/ per square kilometer)	1,387	13,782	6,083	3,484
Student trip generation (trips/ per square kilometer)	417	9,662	4,691	2,590
Simple size			8459	

Note:

1. KRTC: Kaohsiung Rapid Transit Corporation
2. DOBAS: Department of Budget, Accounting and Statistic
3. TBKCG: Transportation Bureau of Kaohsiung City Government
4. DOSKCG: Department of Statistics of Kaohsiung City Government
5. PBWKCG: Public Work Bureau of Kaohsiung City Government
6. LABKCG: Land Administration Bureau of Kaohsiung City Government
7. MOTC: Ministry of Transportation and Communications

A.14 Descriptive statistics of the collected data (suburban district)

Variables	Min	Max	Mean	Std. deviation	Data resource
Station-level data					
CityBike departures (per station/ per month)	0	2,225	604	300	KRTC
CityBike arrivals (per station/ per month)	0	2,222	612	297	
Capacity of station (units)	16	32	25	7	
CityBike station capacity in a 1-km buffer (units)	0	124	40	40	
District-level data					
Population (people/ per square kilometer)	459	2,325	1,729	352	DOBAS, TBKCG, DOSKCG
Worker population (worker/ per square kilometer)	311	1,666	1,211	240	
Student population (student/ per square kilometer)	59	334	230	69	
Tourist (people/ per square kilometer)	0	8,108	1,845	2,602	
Income (NT$/ year)	399,531	647,060	547,873	58,567	
Private vehicle (units / per square kilometer)	475	2,405	1,793	366	
Bikeway length (meter/ per square kilometer)	0	2,167	858	821	PBWKCG
Number of streetlamps (units / per square kilometer)	124	313	221	64	
Length of major road (meter/ per square kilometer)	6,280	12,536	10,156	1,217	LABKCG
Land value (NT$/ per square kilometer)	470	11,750	4,680	3,485	
Number of parks (units / per square kilometer)	0	1	0.5	0.5	DOSKC, MOTC
Number of companies (units / per square kilometer)	8	73	39	13	
Number of factories (units / per square kilometer)	2	22	12	7	
Number of hotels (units / per square kilometer)	0	0	0	0	
Number of markets (units / per square kilometer)	0	0	0	0	
Number of schools (units / per square kilometer)	0	0	0	0	
Number of public transportation yard (units / per square kilometer)	0	2	1	1	TBKCG
Trip attraction (trips/ per square kilometer)	1,152	6,000	4,492	1,389	DOSKC, MOTC (Census data)
Trip generation (trips/ per square kilometer)	1,275	7,547	4,781	1,220	
Worker trip attraction (trips/ per square kilometer)	399	1,856	1,428	344	
Worker trip generation (trips/ per square	441	2,649	1,555	438	

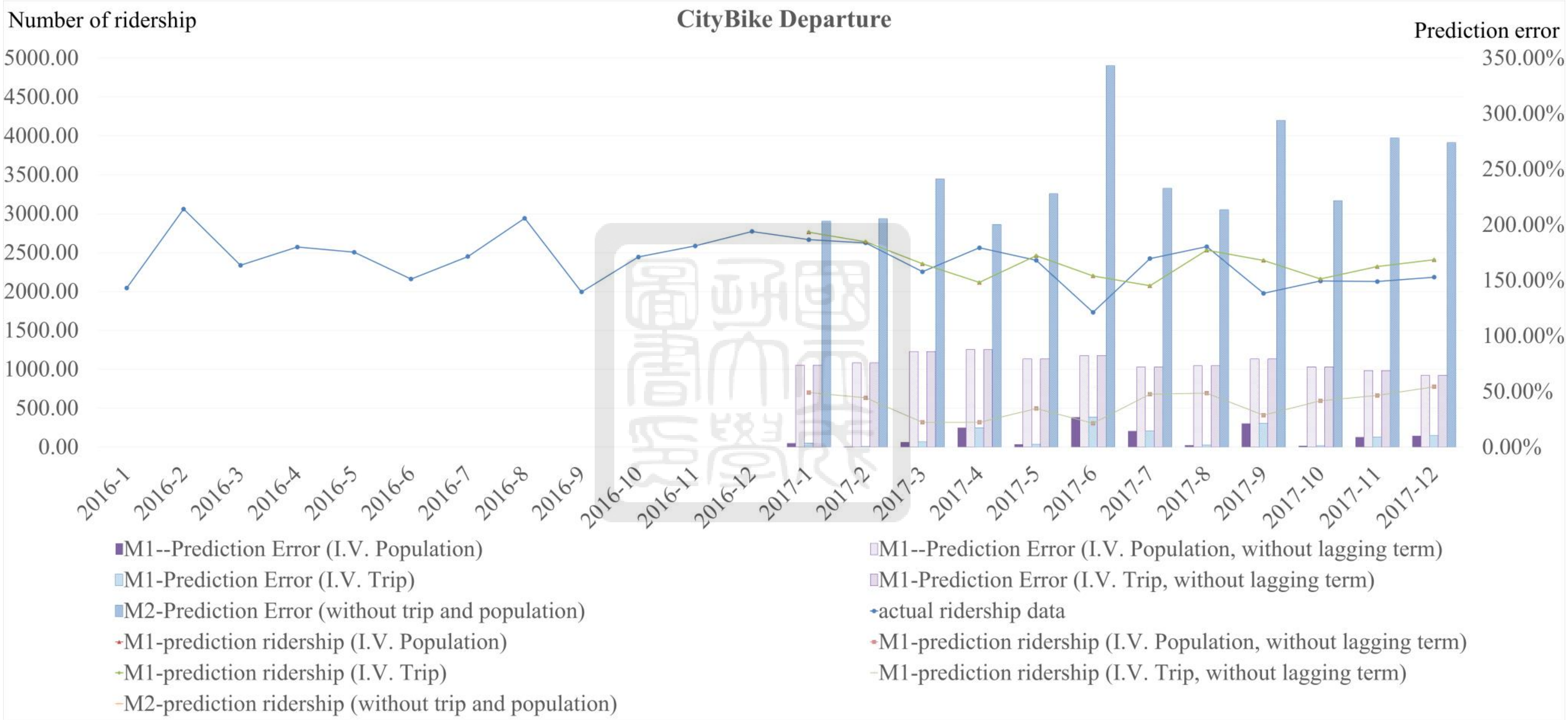
kilometer)				
Student trip attraction (trips/ per square kilometer)	147	830	600	151
Student trip generation (trips/ per square kilometer)	163	1,185	645	161
Simple size			795	

Note:

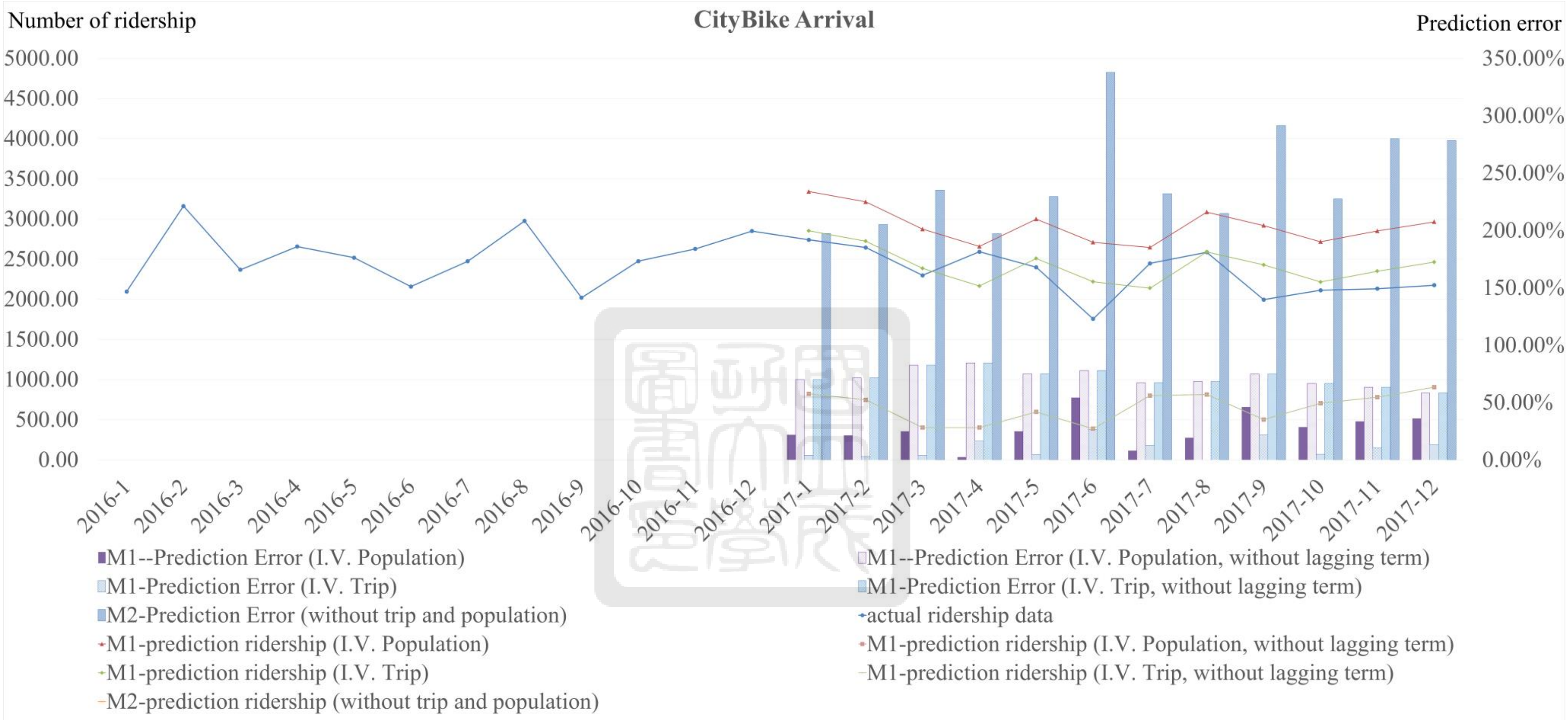
1. KRTC: Kaohsiung Rapid Transit Corporation
2. DOBAS: Department of Budget, Accounting and Statistic
3. TBKCG: Transportation Bureau of Kaohsiung City Government
4. DOSKCG: Department of Statistics of Kaohsiung City Government
5. PBWKCG: Public Work Bureau of Kaohsiung City Government
6. LABKCG: Land Administration Bureau of Kaohsiung City Government
7. MOTC: Ministry of Transportation and Communications

Appendix B: Comparison of predicted and actual CityBike departure in individual districts

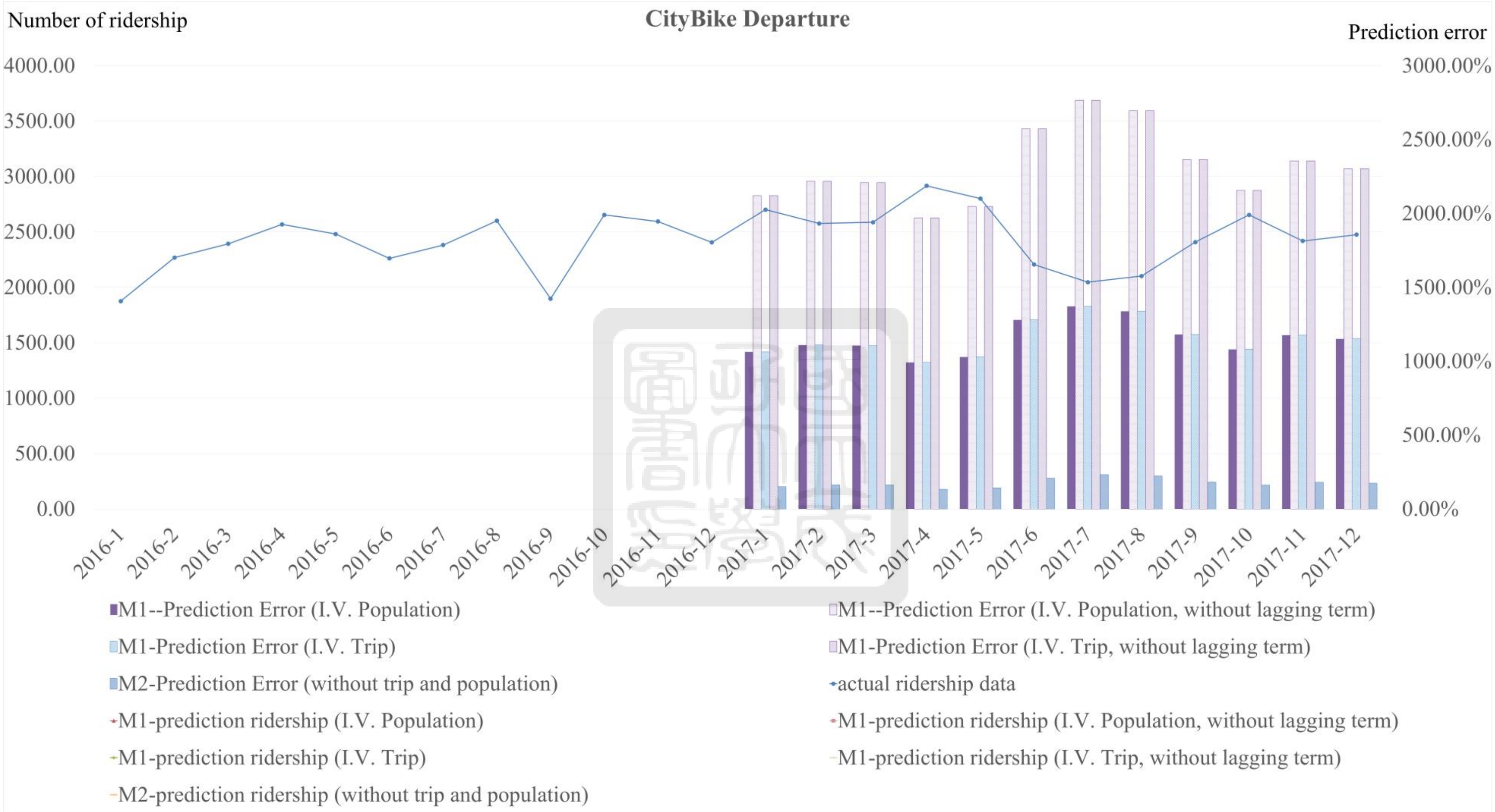
B.1 Prediction result in Yancheng District (departure).



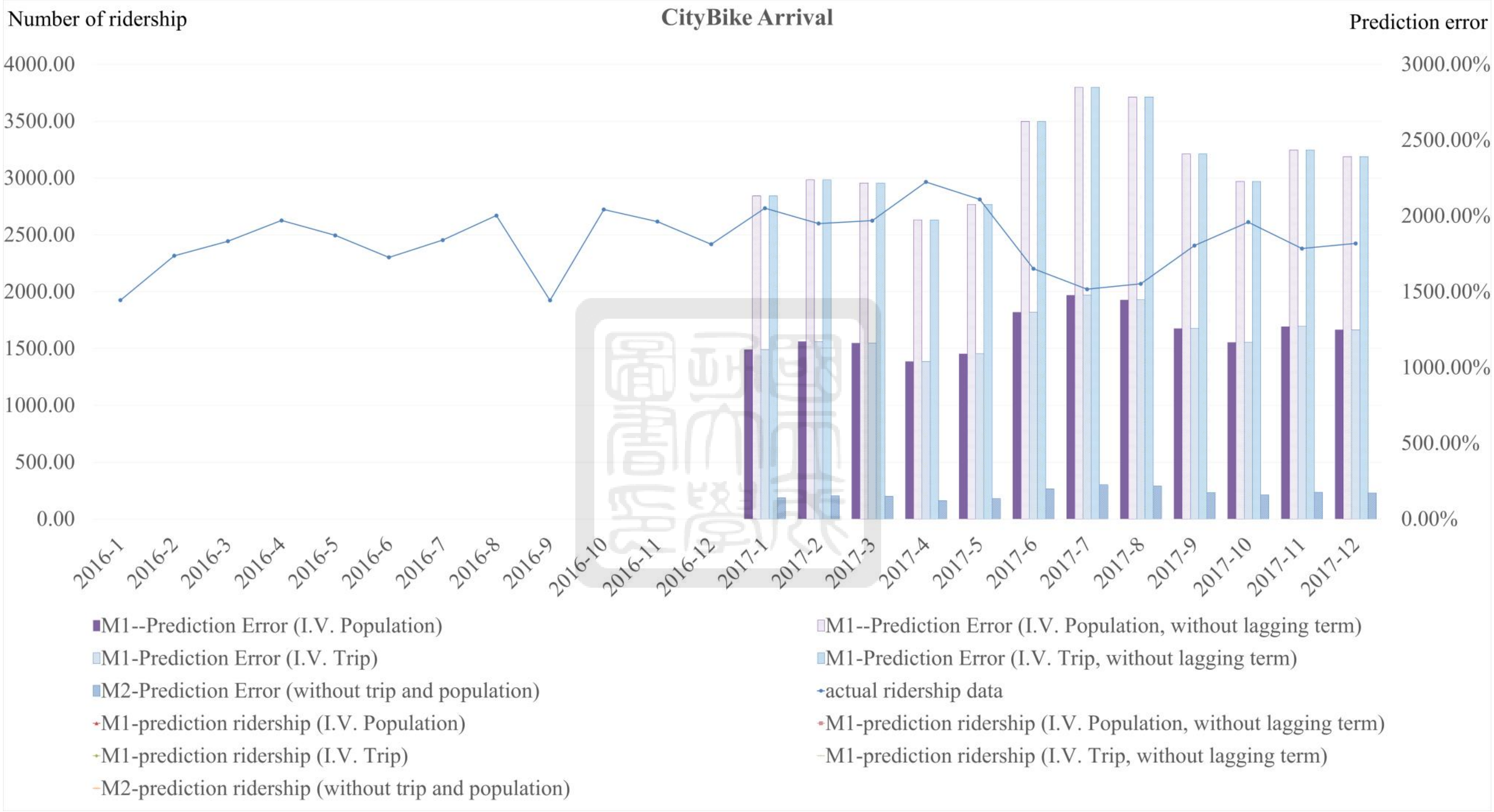
B.2 Prediction result in Yancheng District (arrival).



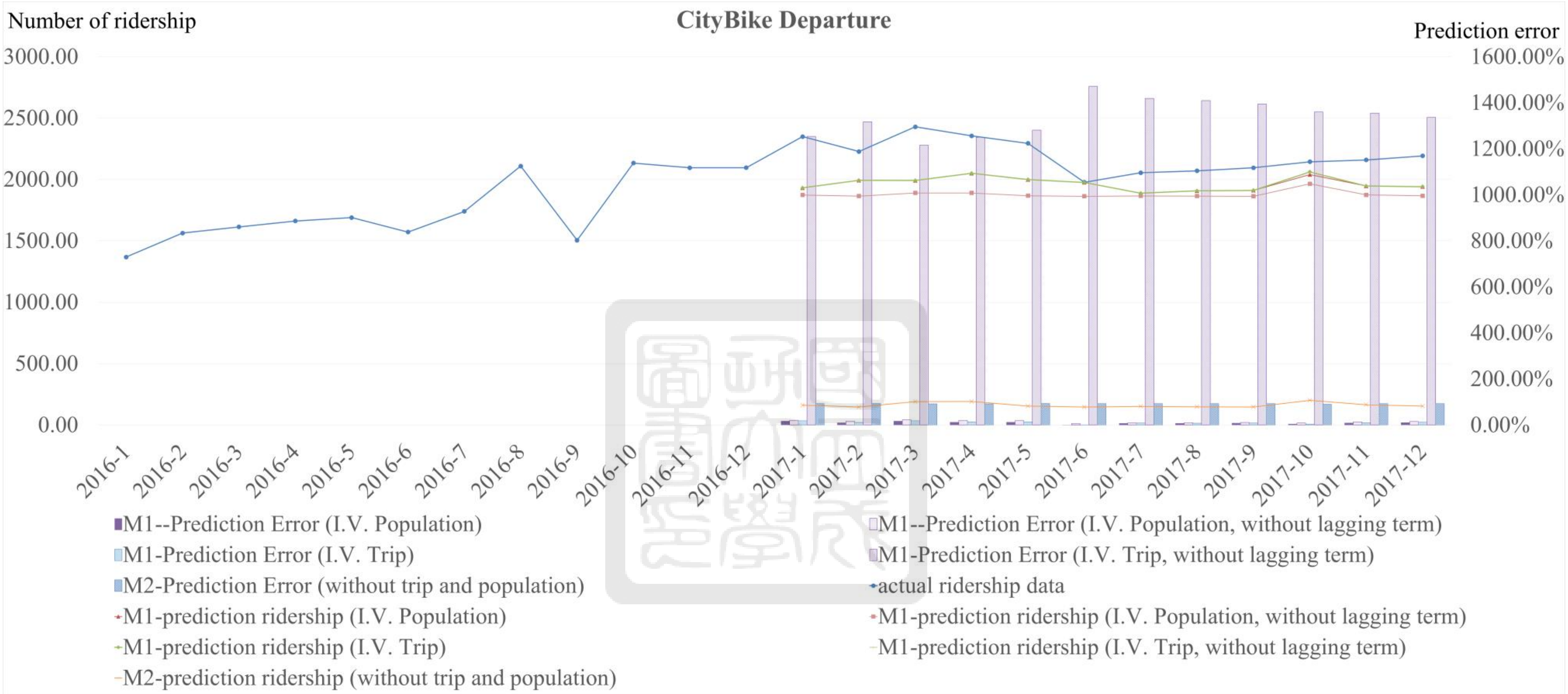
B.3 Prediction result in Gushan District (departure).



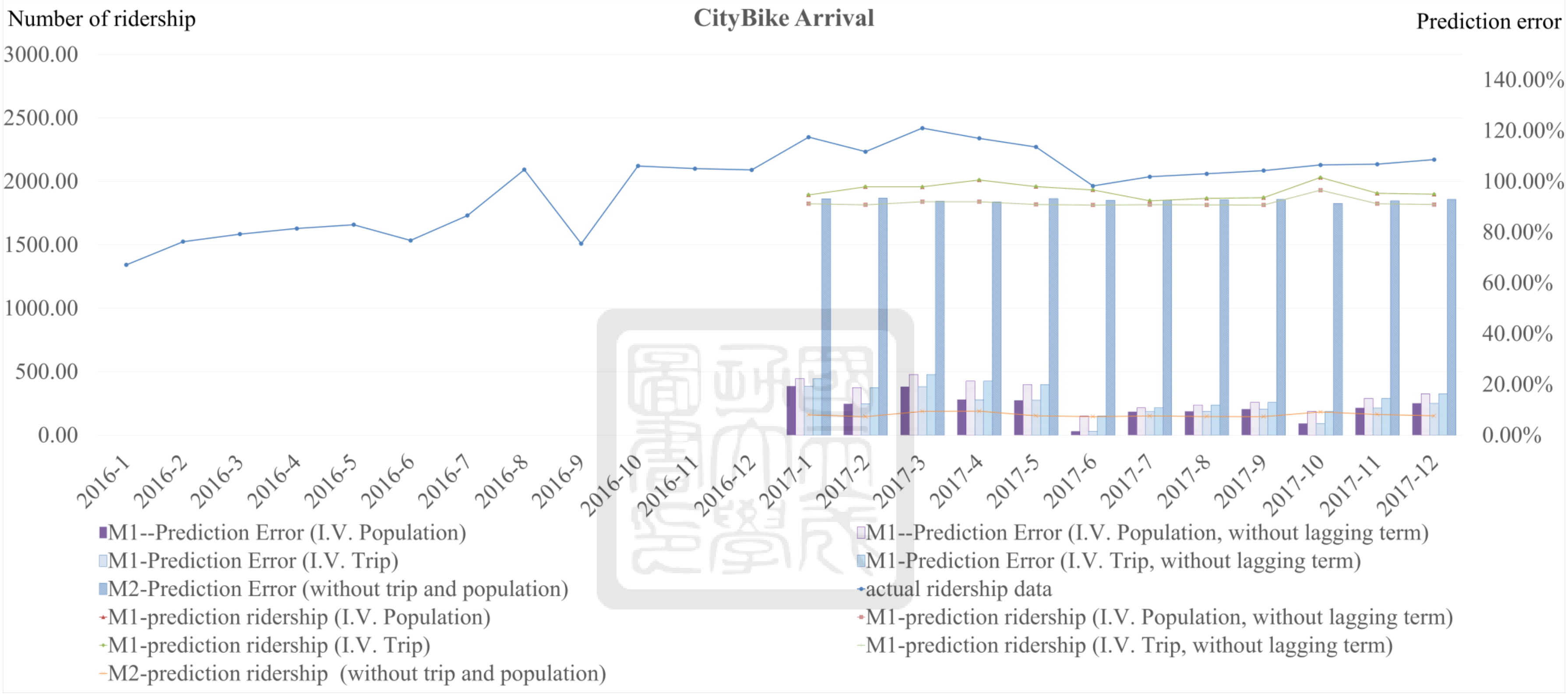
B.4 Prediction result in Gushan District (arrival).



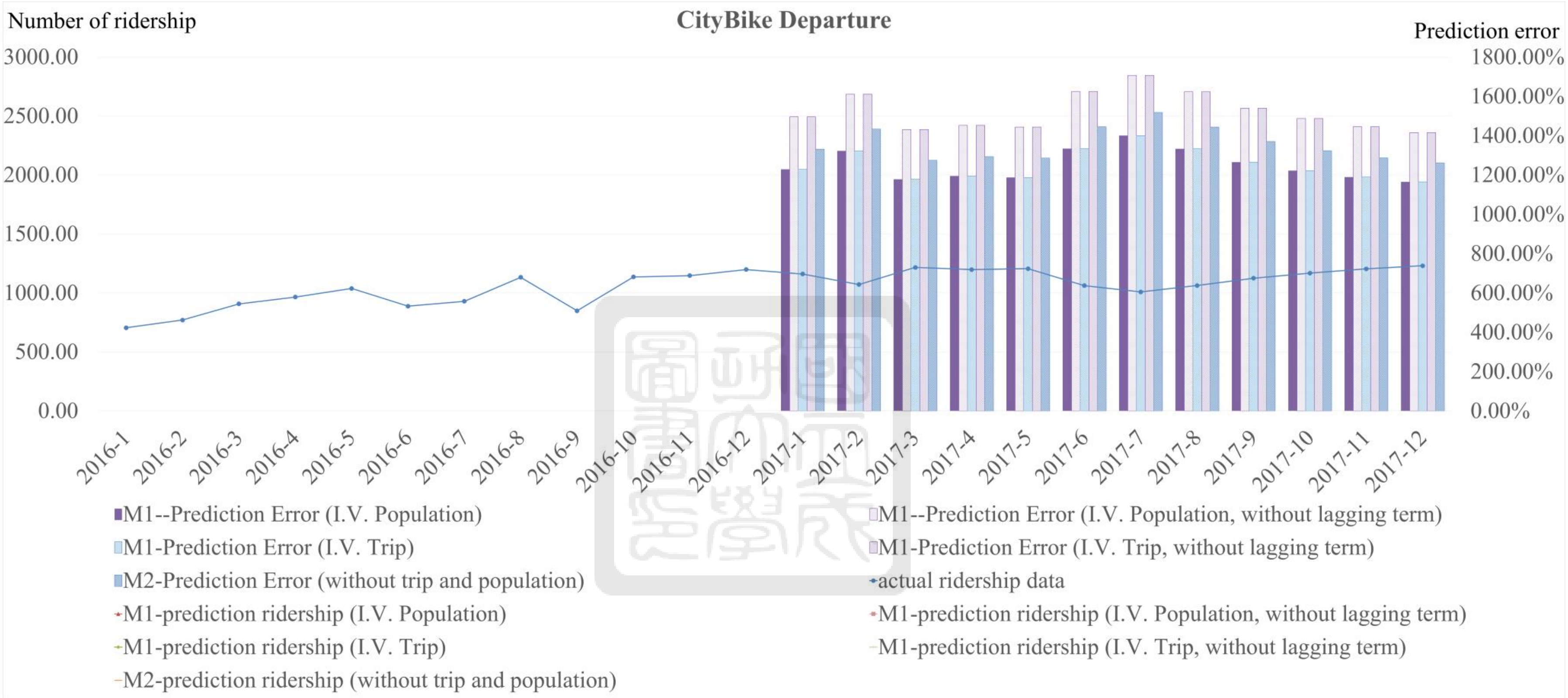
B.5 Prediction result in Zuoying District (departure).



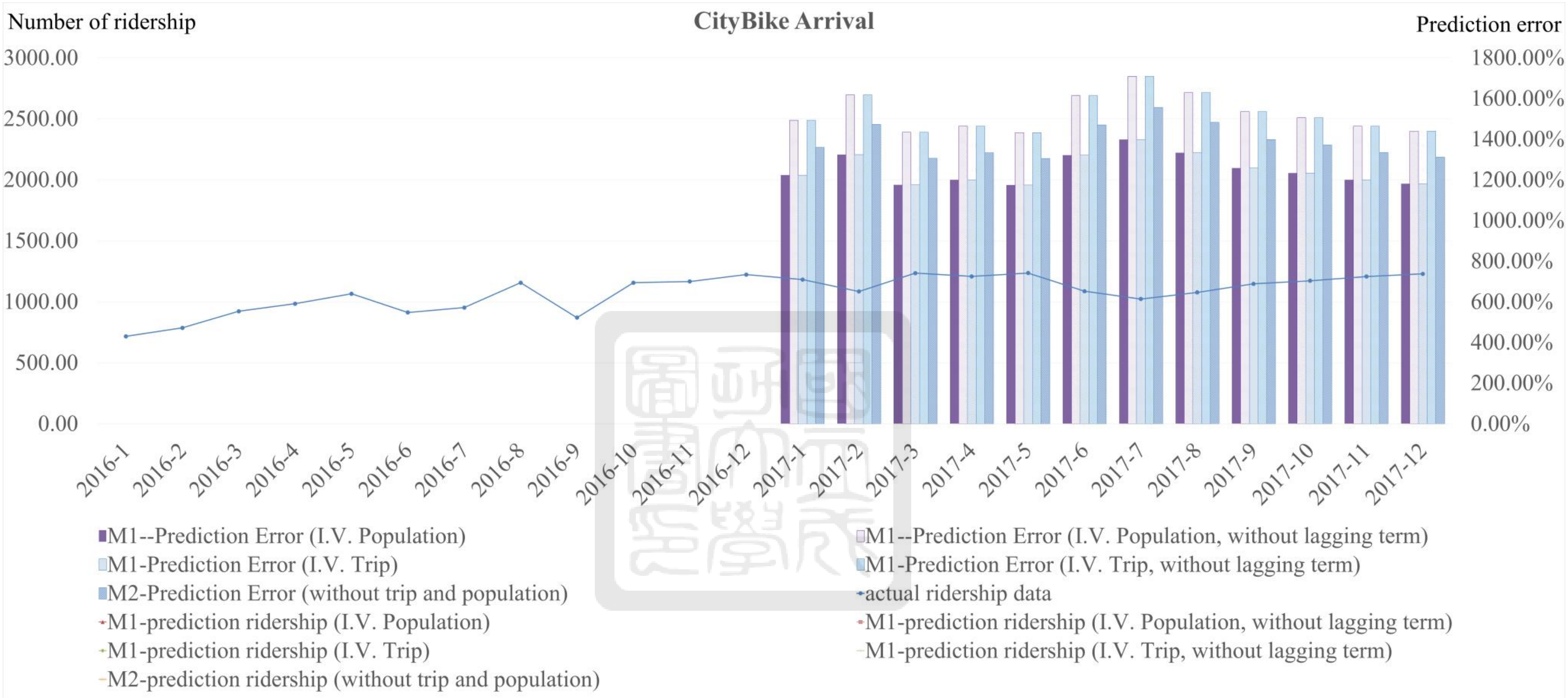
B.6 Prediction result in Zuoying District (arrival).



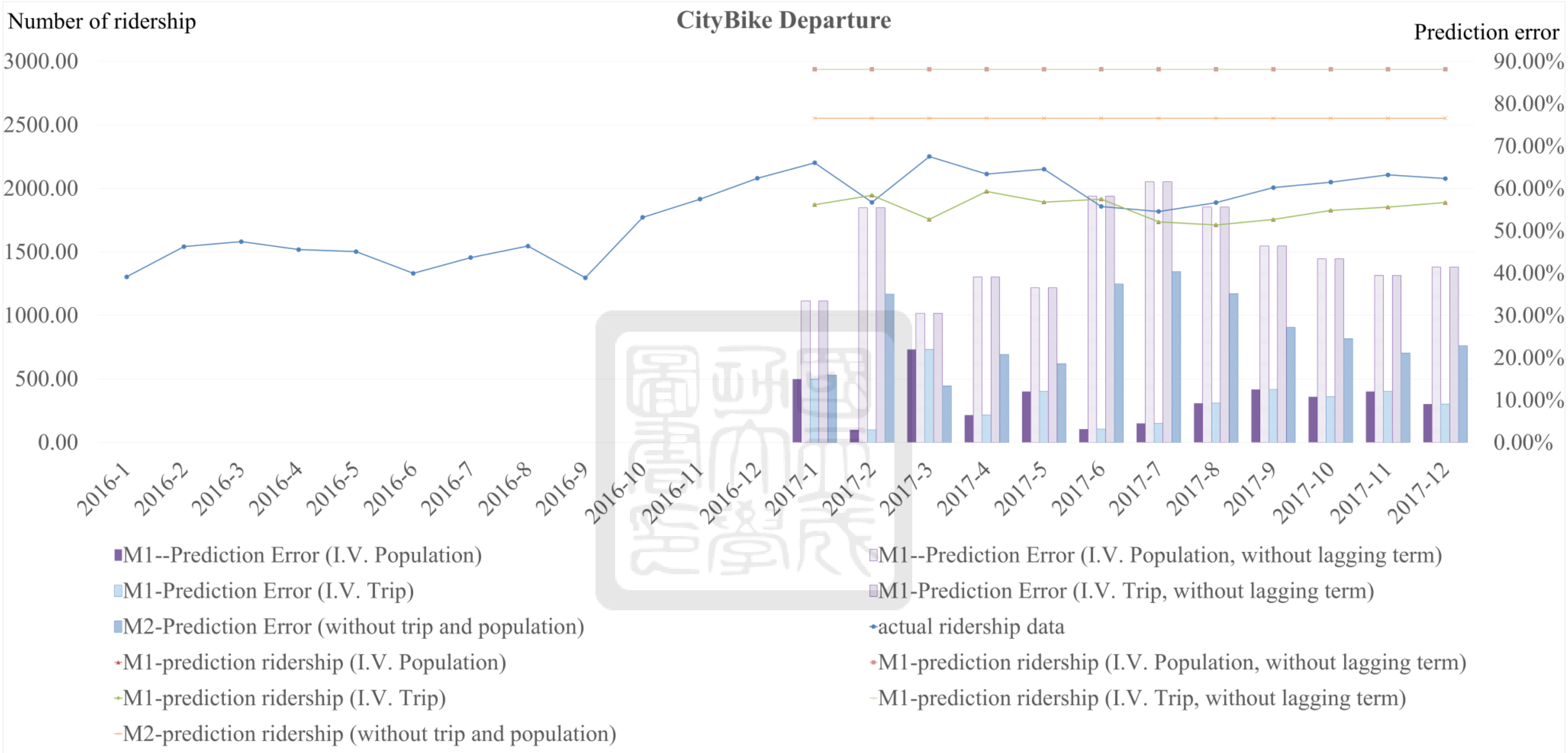
B.7 Prediction result in Nanzih District (departure).



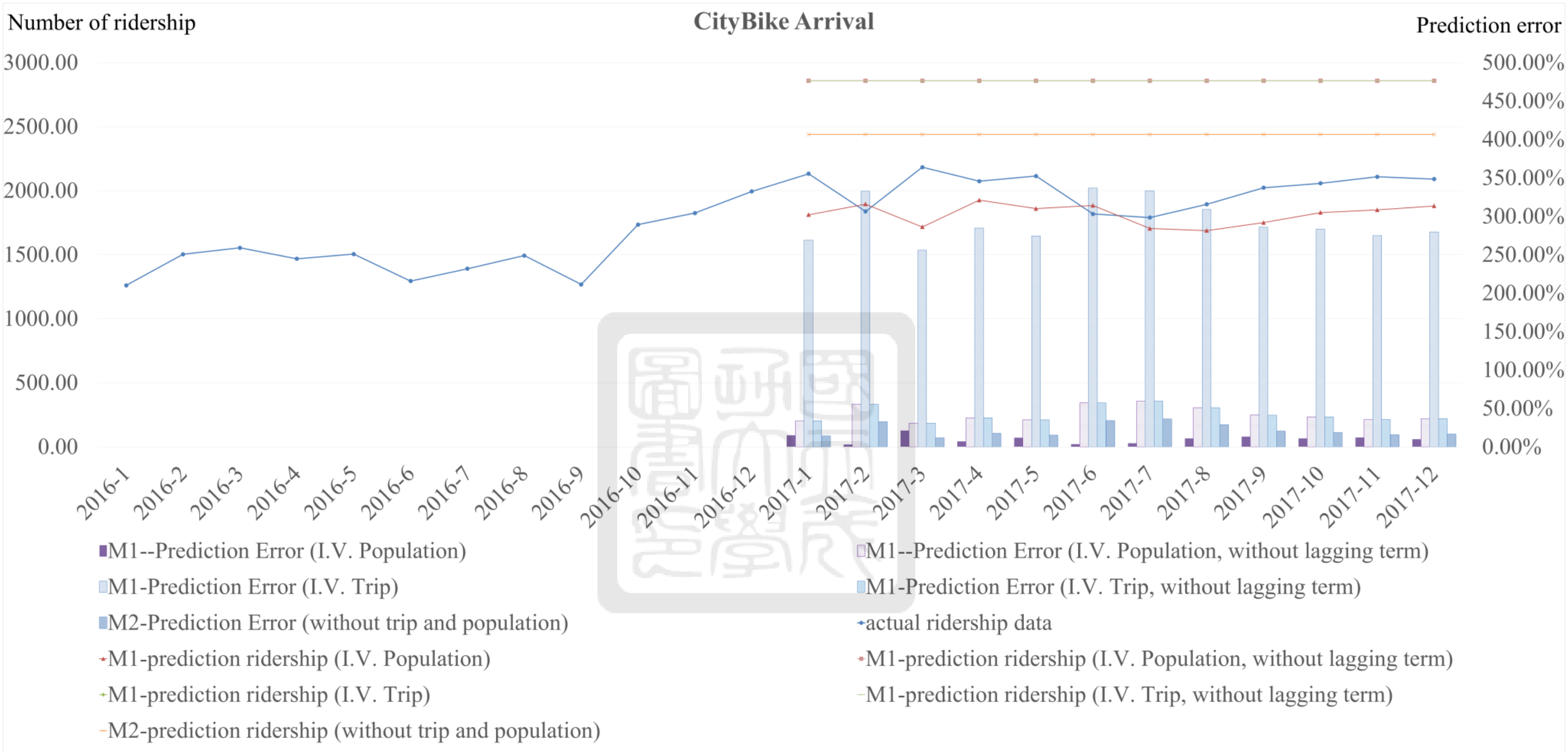
B.8 Prediction result in Nanzih District (arrival).



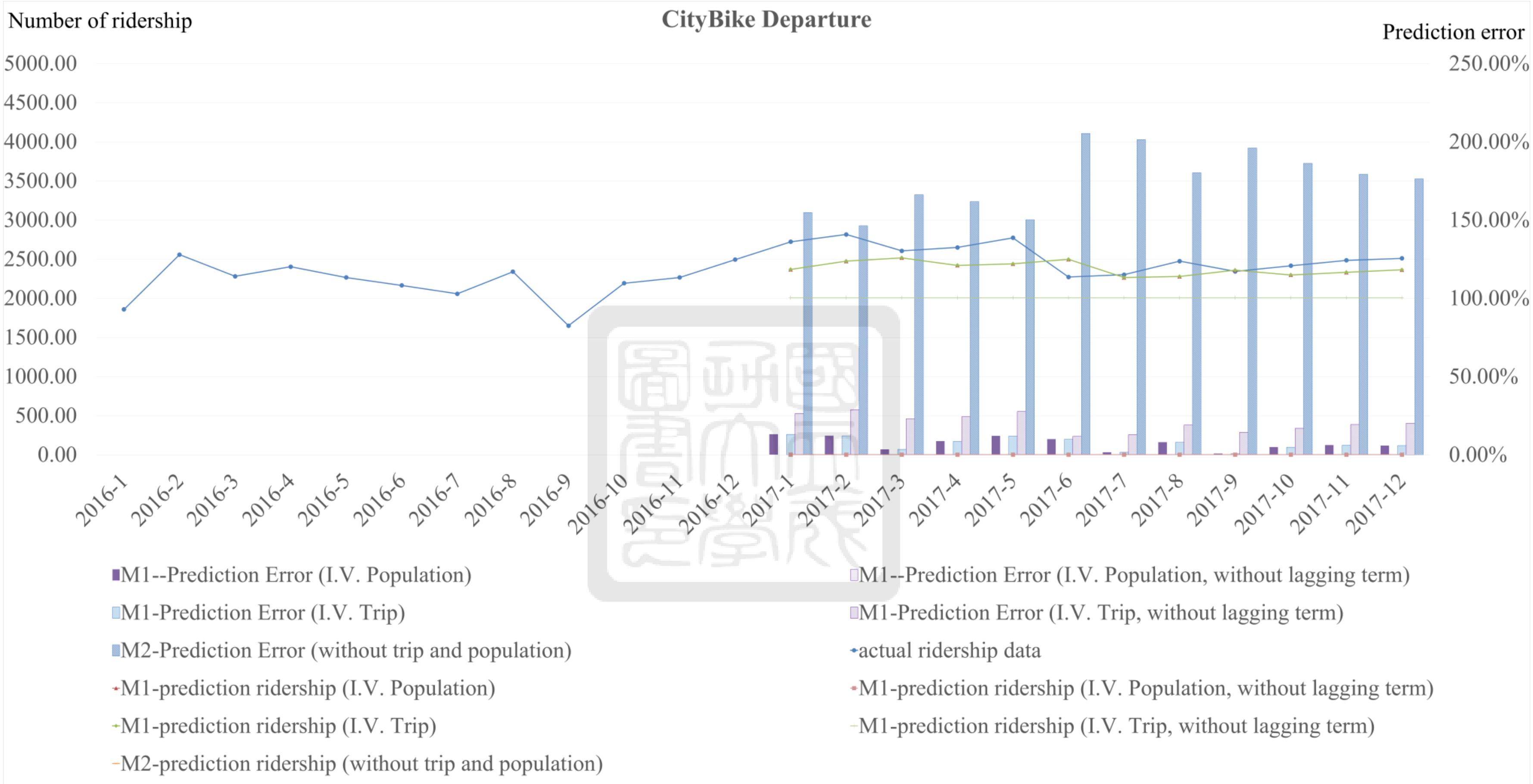
B.9 Prediction result in Sinsing District (departure).



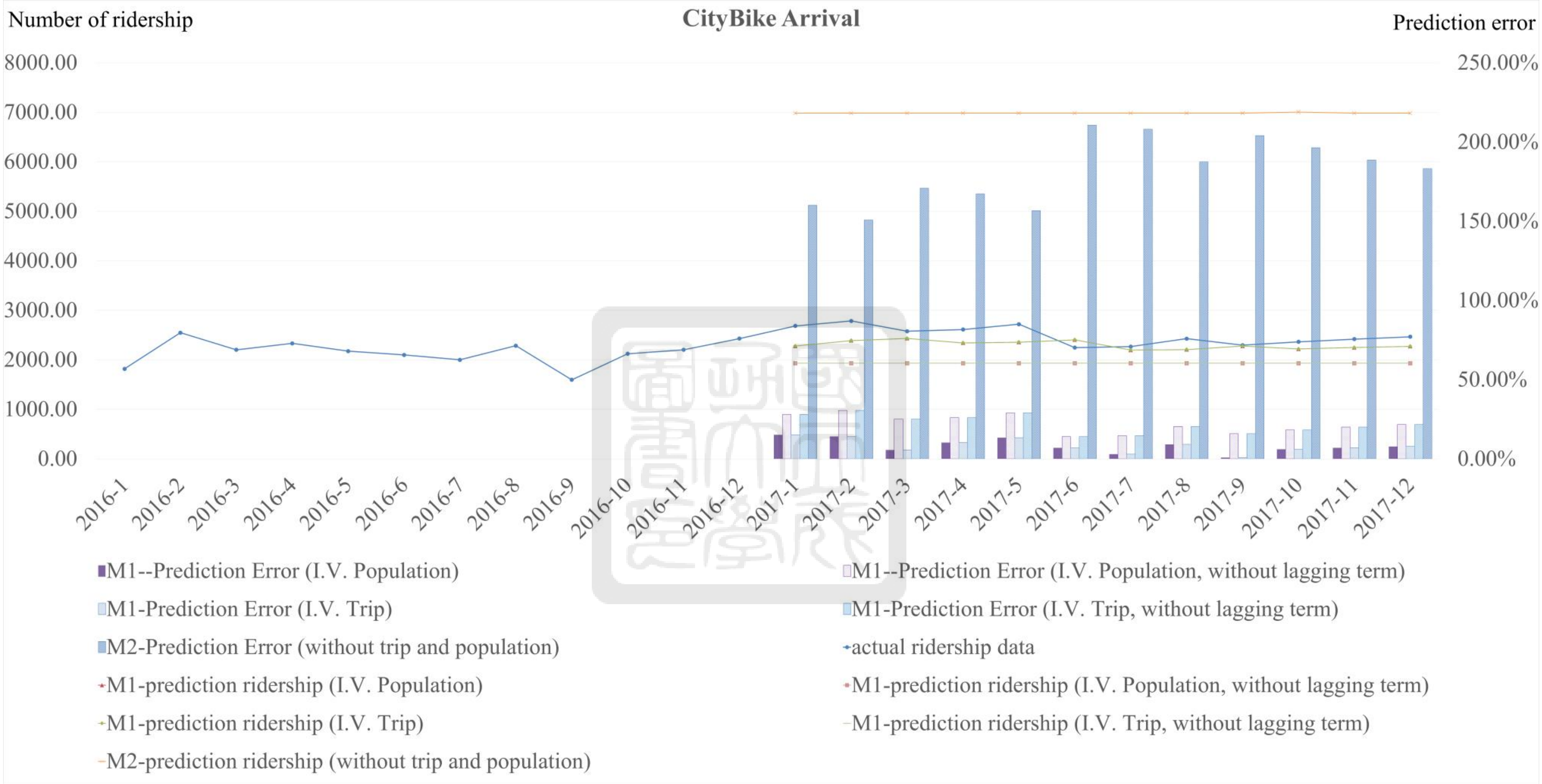
B.10 Prediction result in Sinsing District (arrival).



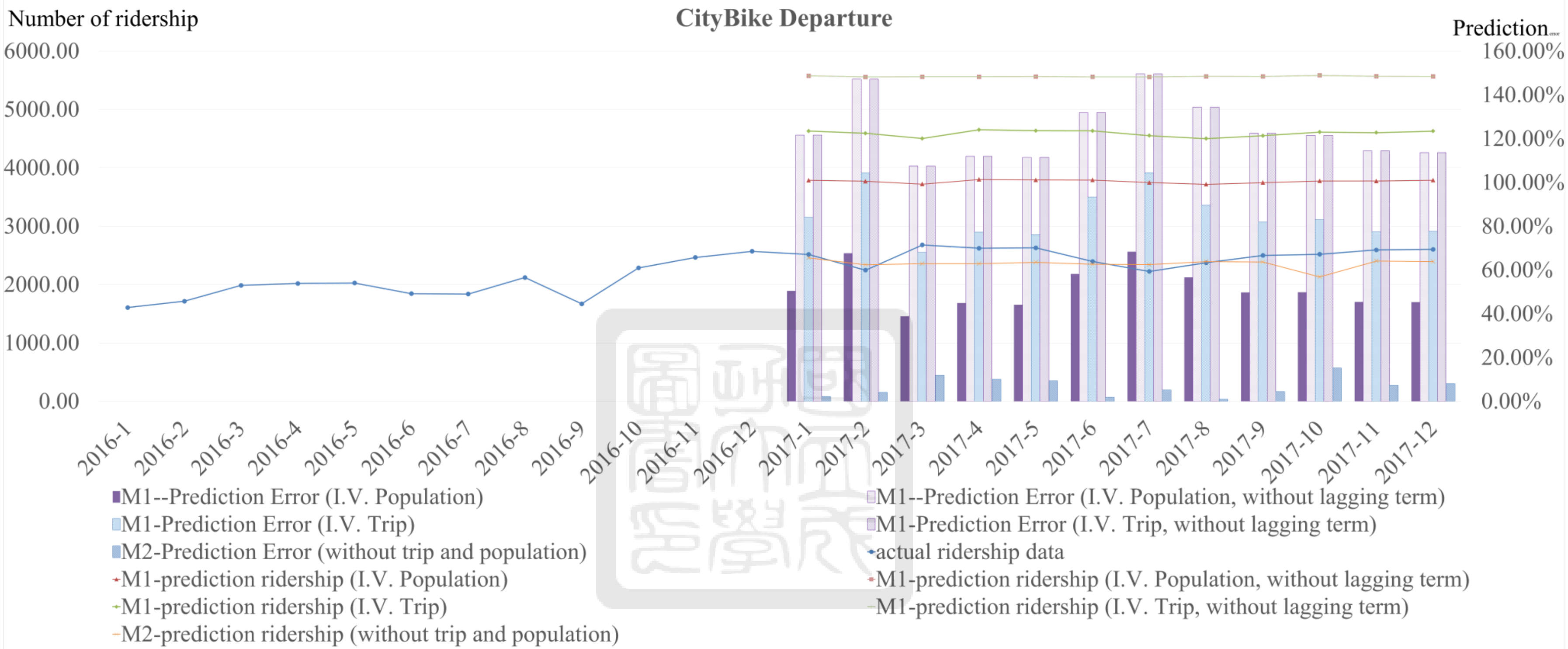
B.11 Prediction result in Cianjin District (departure).



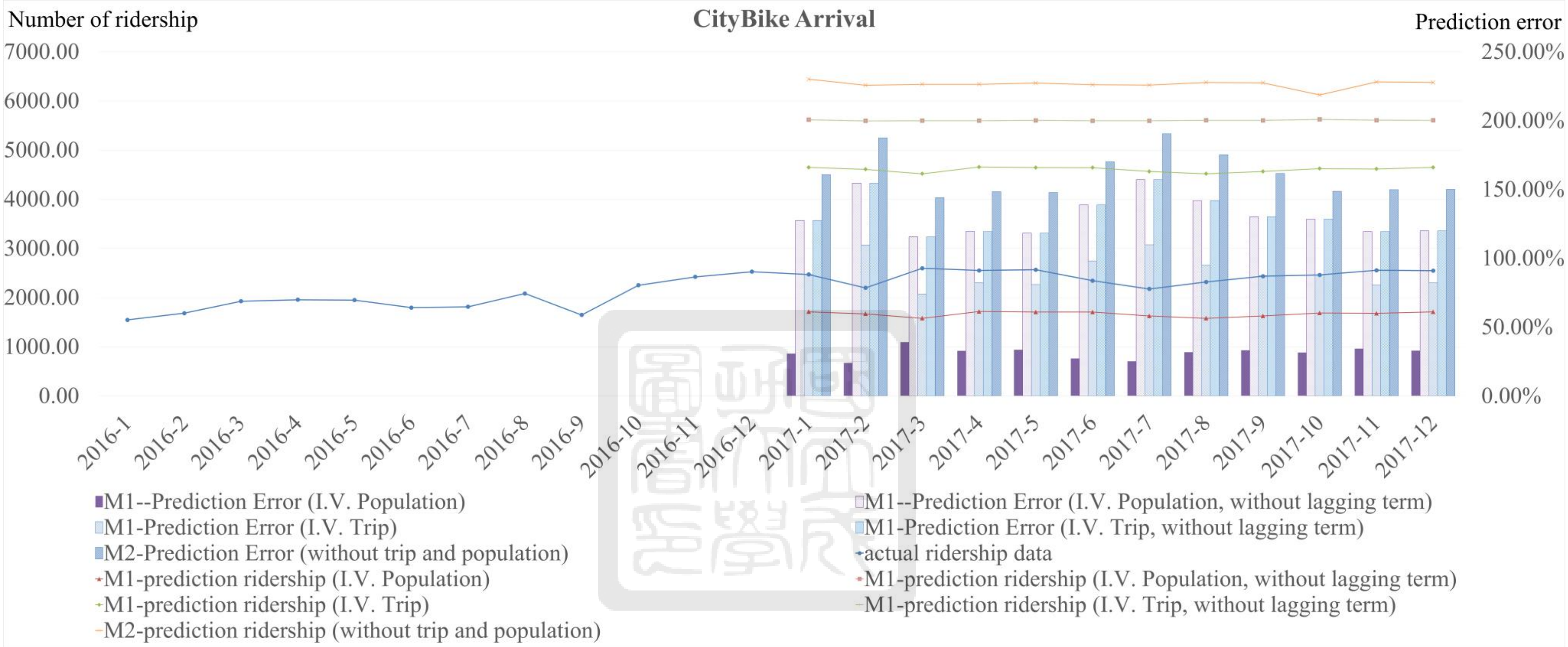
B.12 Prediction result in Cianjin District (arrival).



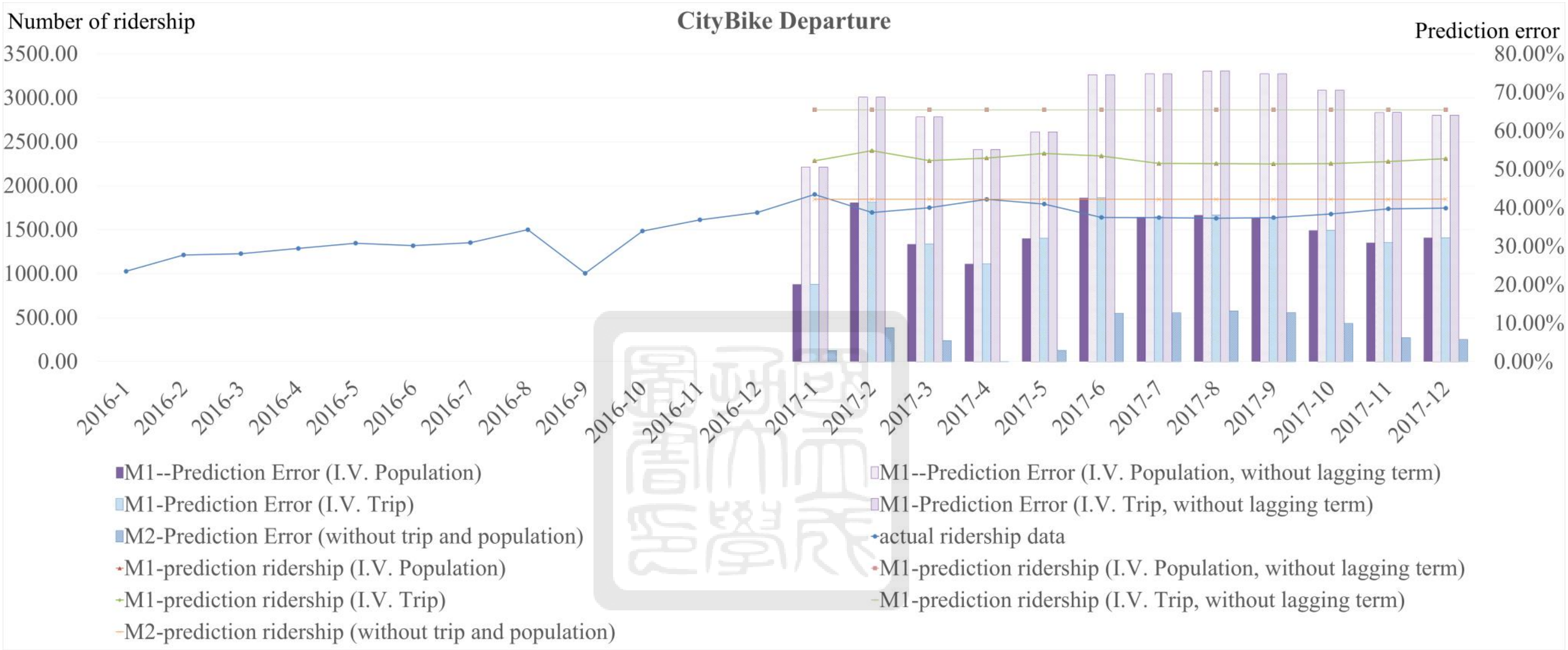
B.13 Prediction result in Lingya District (departure).



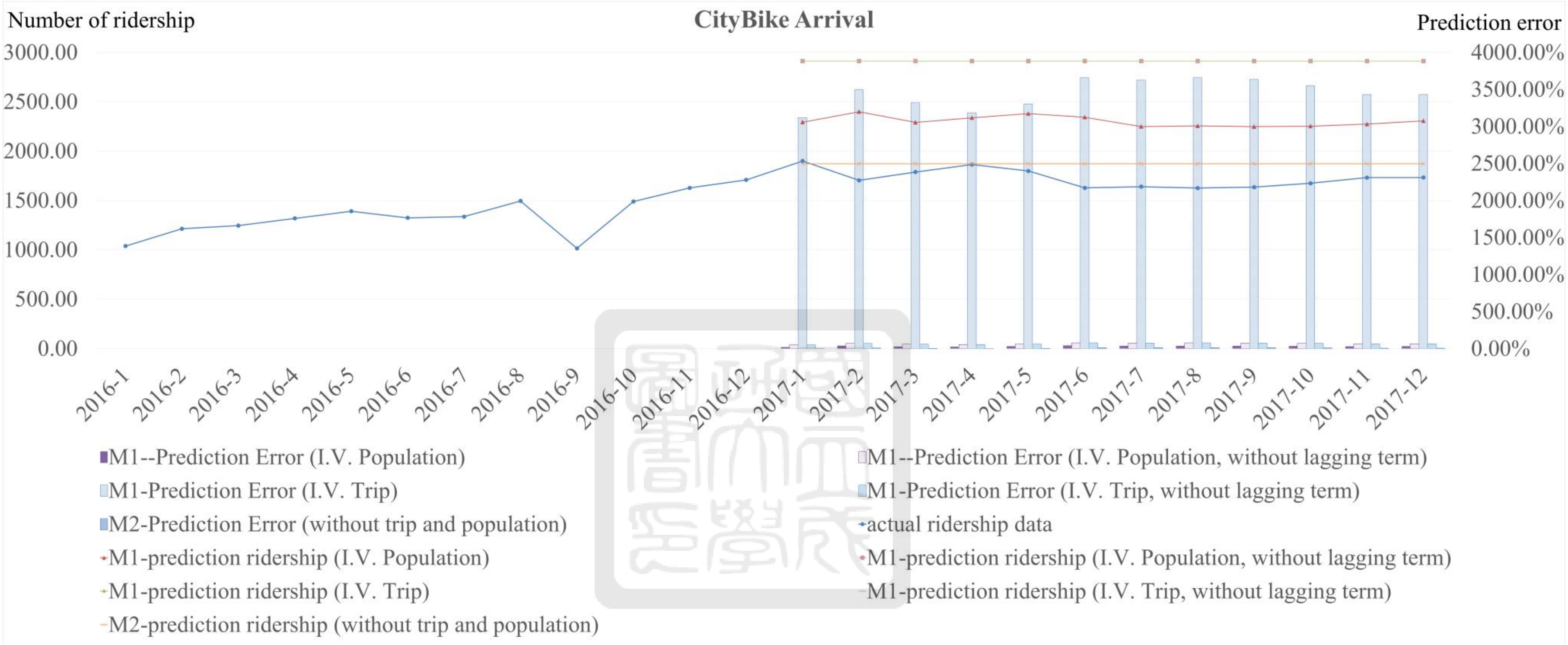
B.14 Prediction result in Lingya District (arrival).



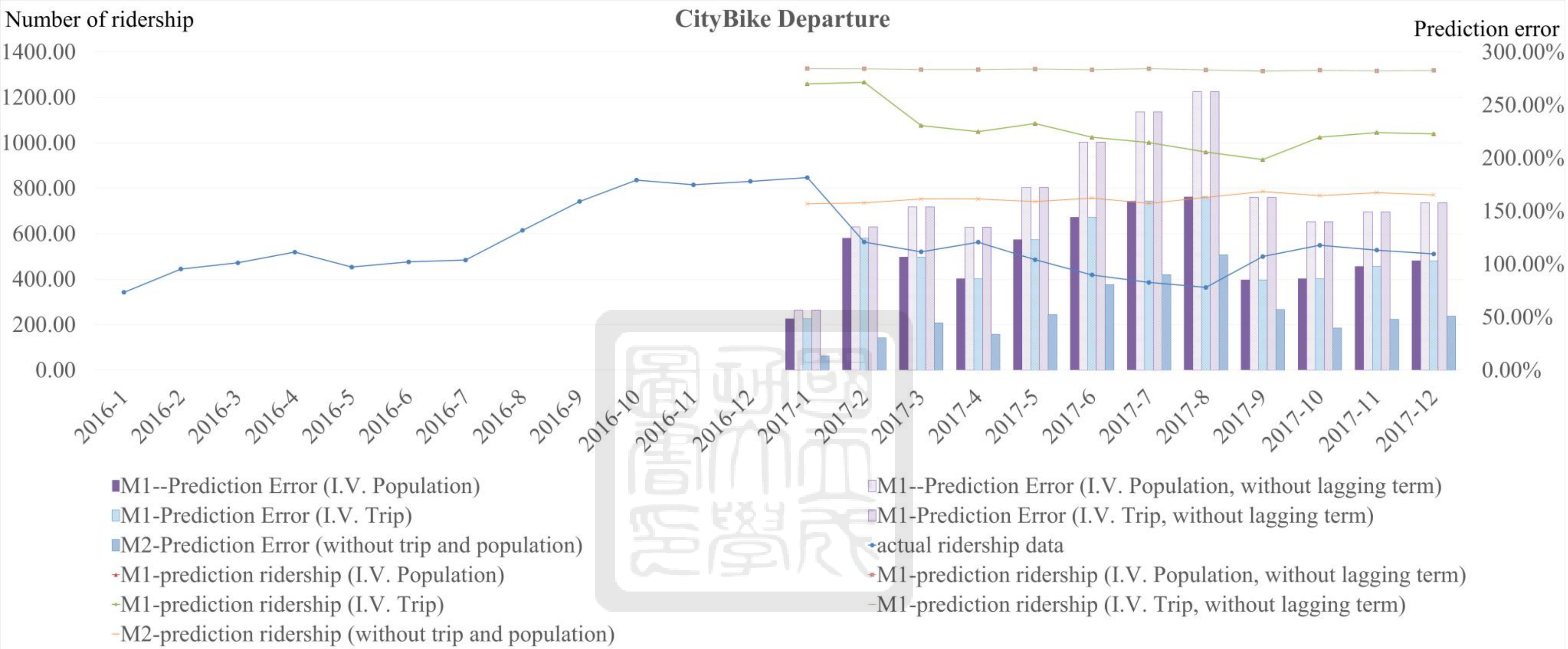
B.15 Prediction result in Cianjhen District (departure).



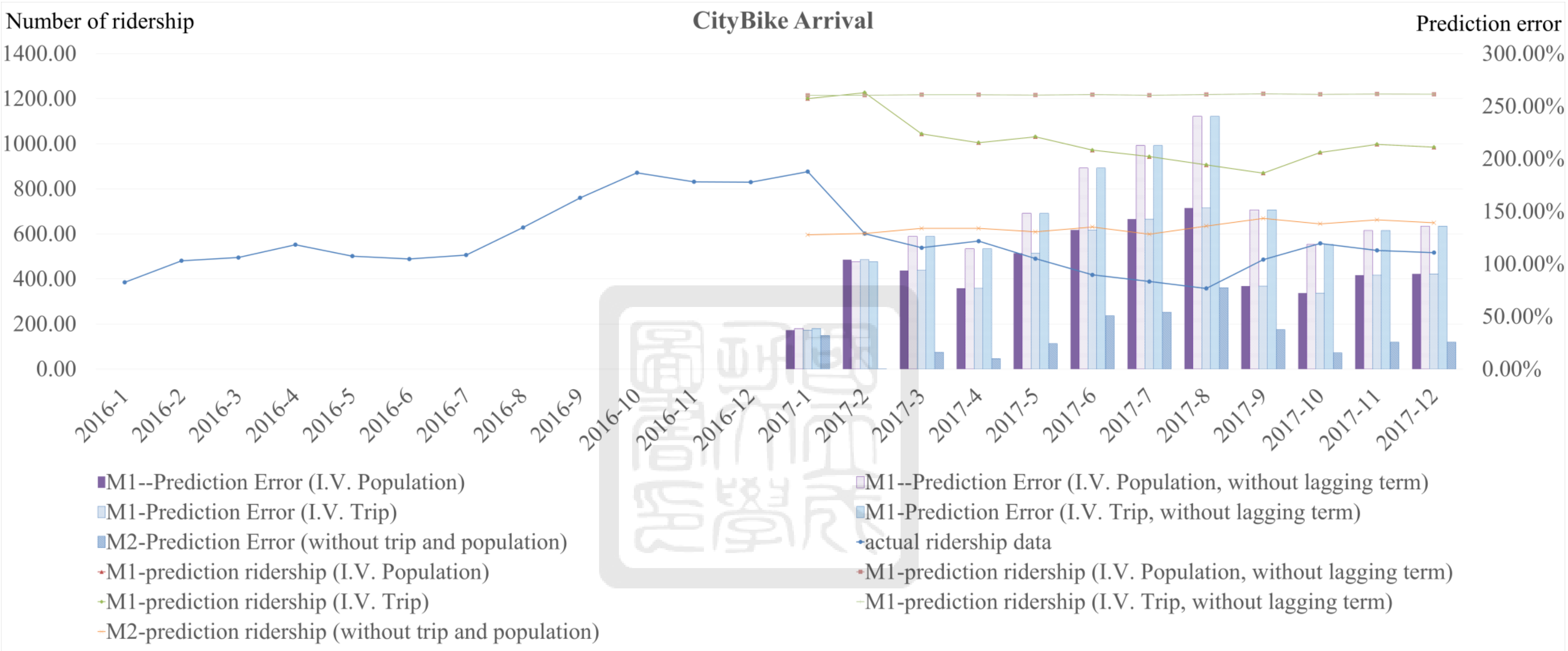
B.16 Prediction result in Cianjhen District (arrival).



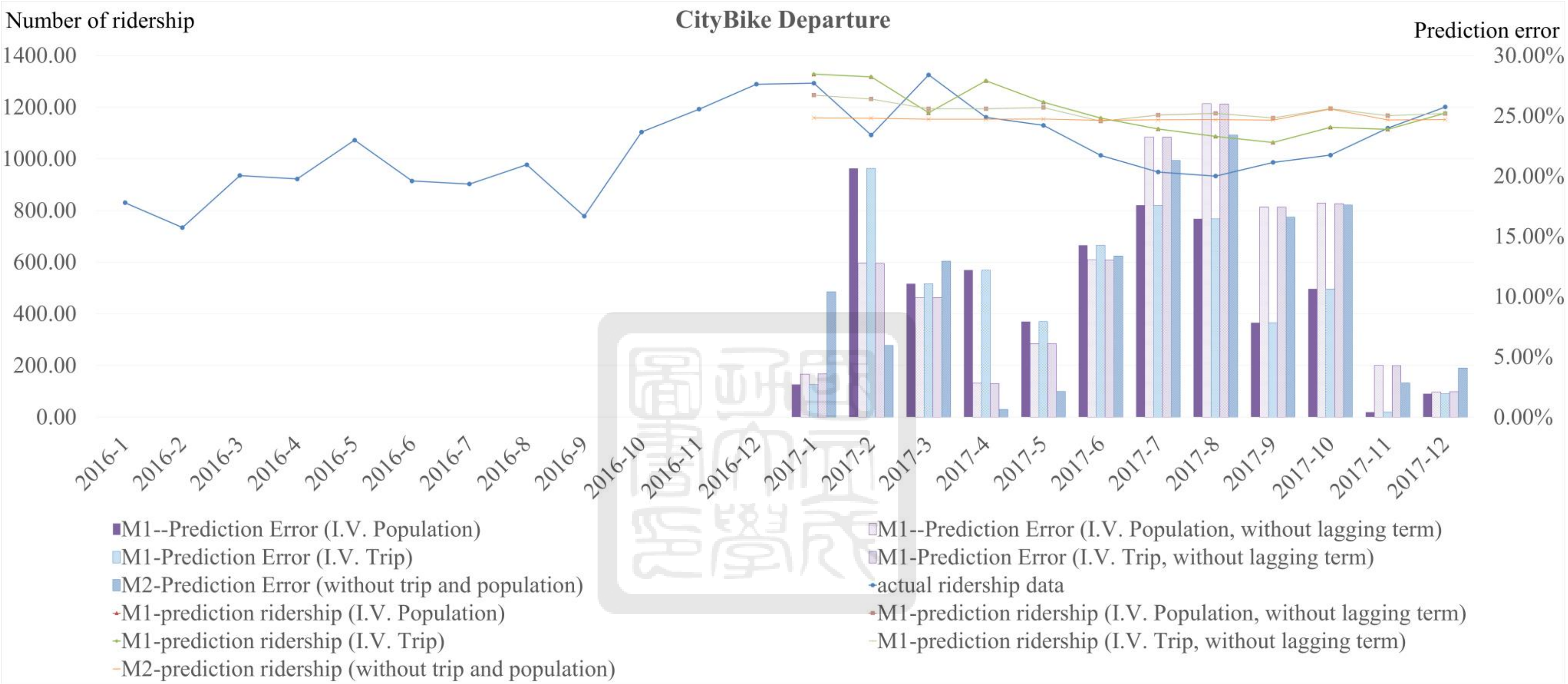
B.17 Prediction result in Cijin District (departure).



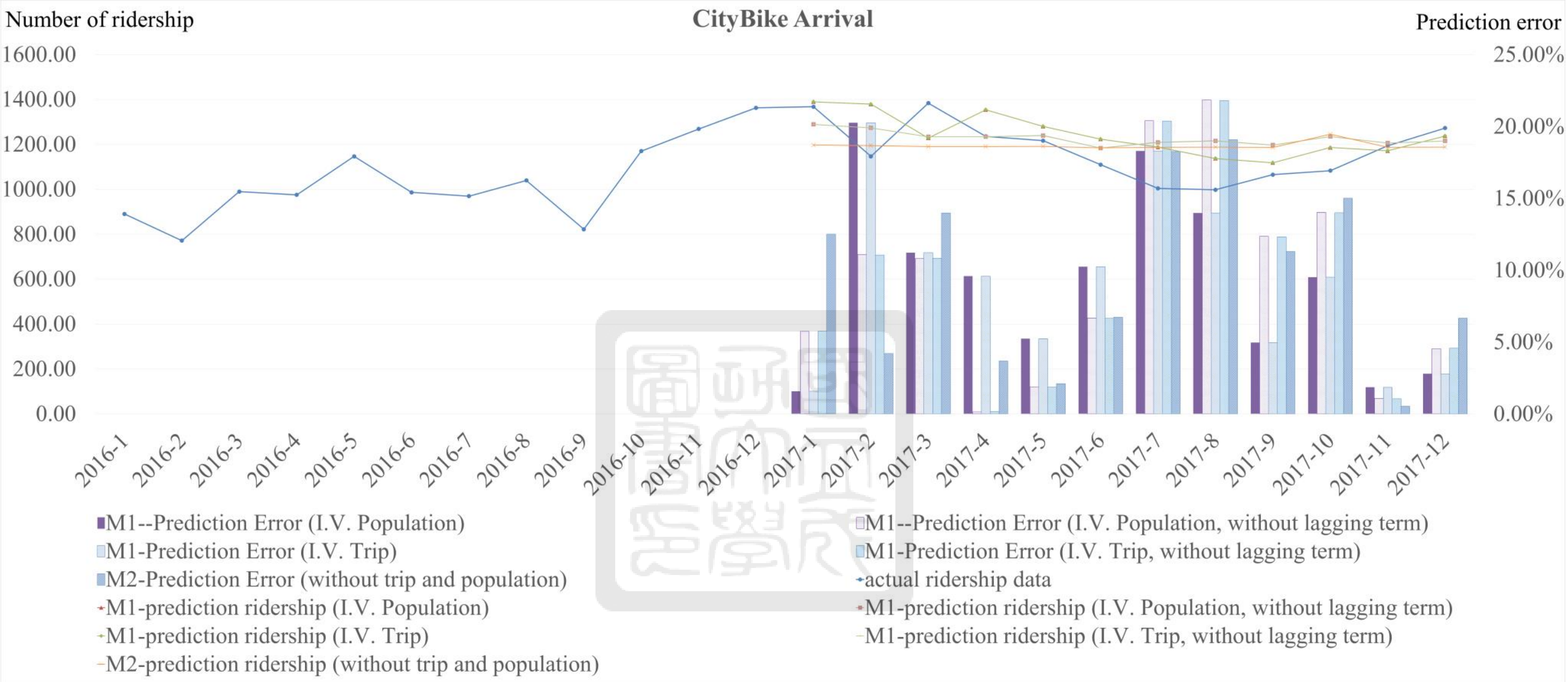
B.18 Prediction result in Cijin District (arrival).



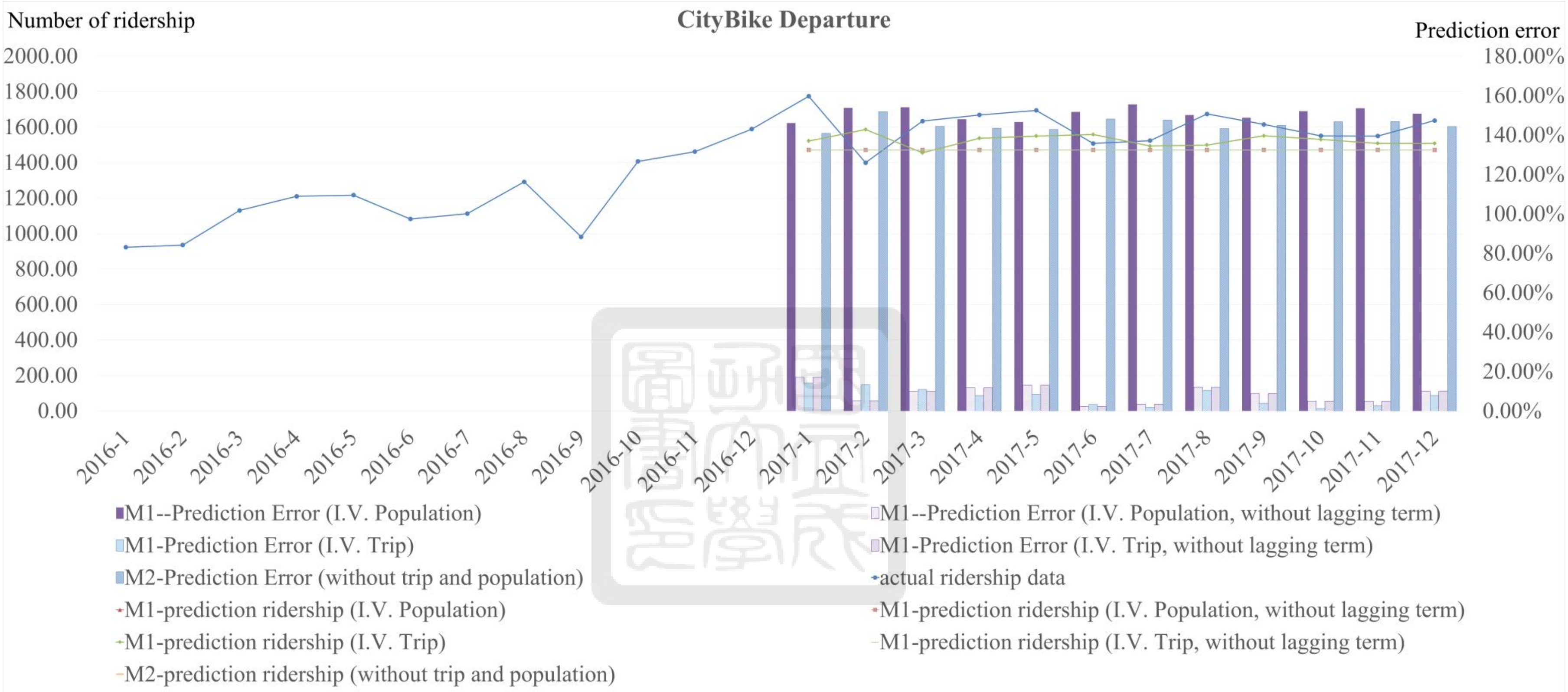
B.19 Prediction result in Siaogang District (departure).



B.20 Prediction result in Siaogang District (arrival).



B.21 Prediction result in Fongshan District (departure).



B.22 Prediction result in Fongshan District (arrival).