A COMPREHENSIVE MEASUREMENT MODEL FOR JOB-HOUSING BALANCE CONSIDERING SPATIAL INTERACTIONS: A CASE STUDY IN SHANGHAI

Yuting Chen [1], Changfeng Jing [2] *, Gaoran Xu [2]

[1] School of Geomatics and Urban Spatial Informatics, Beijing University of Civil Engineering and Architecture, Beijing 100044, China – (Y.C.) 2108570021112@stu.bucea.edu.cn
[2] School of Information Engineering, China University of Geosciences, Beijing 100083, China – (C.J.) jingcf@cugb.edu.cn, (G.X.) xgr010703@163.com

KEY WORDS: Job-housing Balance, Cellular Signaling Data, Structural Equation Modelling, Spatial Interaction, Measuring Method, GIS.

ABSTRACT:

Assessing jobs-housing balance (JHB) is crucial for the optimization of urban spatial pattern and transportation planning. However, due to the limitations of traditional data and the inability to integrate evaluation indicators, the evaluation results may be bias, conflicting with common sense. Considering interaction of multiple factors, a Job-housing balance measurement model (JHBM-SEM) was proposed to accurately evaluate the degree of JHB in a certain area. To address the problem of multi-indicator fusion, the structural equation model is used to incorporate the advantages, limitations and complementary relationships of each evaluation indicator into the evaluation process. A function relationship is defined between factor loadings, path coefficients, and variable scores to construct the measurement model and obtain comprehensive evaluation results. In addition, by integrating multi-source spatiotemporal big data mainly based on mobile signaling data, the paper can effectively mine the current status of job-housing balance in cities, and solve problems such as low precision and granularity of traditional data. The model was validated using Shanghai as an example, and the results show that compared with classical methods, the model's results are more consistent with the real situation of regional job-housing patterns. It can identify the pseudo-balance phenomenon in underdeveloped areas of the suburbs and make reasonable evaluations, and the intermediate process can explain the direction of the imbalance factors in the region.

1. INTRODUCTION

The places where urban residents spend the longest time in a day are their homes and workplaces, representing their primary activity areas. As a result, the Jobs-housing spatial relationship have become an important topic for urban spatial layout pattern. Specifically, achieving a balance between jobs and housing is critical to promoting a balanced city layout, shortening commute times, and alleviating traffic congestion (Wang et al., 2021). Measuring this balance can assist decision-makers in formulating appropriate urban planning strategies to achieve coordinated development between people and cities.

Job-housing balance is an important concept in Western urban design, originally stemming from Howard's "Garden City", which refers to the equitable distribution of employment and housing in a city, enabling residents to access work within walking distance of their homes (Howard, 1946). However, with the acceleration of modern urban expansion, cities have struggled to implement this design principle effectively, resulting in long-distance commuting and rush hour traffic problems (Zhang et al., 2017). To address these urban issues, numerous public institutions have introduced new policies, such as the UK's green belt regulation, which achieves its objectives from a land use perspective, and China's government has implemented multiple policies, including the provision of affordable housing and the construction of integrated new urban areas, to achieve job-housing balance (Wang et al., 2022). Despite the abundance of research on job-housing balance, there is currently no universally recognized method for quantitatively evaluating its extent, whether it be the selection of indicators or the determination of geographical scales. Additionally, many studies rely on data primarily derived from travel surveys or statistical yearbooks (Zhou et al., 2017), which may have low accuracy, small sample sizes, slow updating speeds, and other limitations that may to some extent restrict the accuracy of evaluation conclusions.

Currently, numerous researches aimed at determining the degree of balance between employment and residential factors within urban areas (Cervero, 1996). These studies typically approach the issue from three main perspectives: quantity balance, quality balance, and spatial balance (Ta et al., 2017). Several methods are available to represent these perspectives, including job-housing ratios, self-containment, and job accessibility. However, these measures offer limited interpretive effects, as evidenced by the fact that job-housing ratios, for instance, fail to capture spatial mismatches (Horner and Mefford, 2007). Similarly, self-containment, which relies on the proportion of workers residing and working within a given locality, is also affected by commuting trends (Martinus and Biermann, 2018; Zhou et al., 2012) and may, therefore, yield inaccurate results. Although some scholars have experimented with excess commuting theory to analyze the relationship between employment and residential aspects, this approach often suffers from constraints related to sample size and resolution, particularly in larger cities (Xiao et al., 2021a), and the scale itself is an important variable worth noting (Jing et al., 2022). Hence, relying on a single indicator to gauge job-housing balance may not accurately reflect reality (Qin and Wang, 2019), and it may be more appropriate to integrate multiple indicators for a more comprehensive evaluation.

* Corresponding author, Changfeng Jing jingcf@cugb.edu.cn

In order to overcome these challenges, we propose a model for measuring job-housing balance based on Structural Equation Modeling (SEM), named JHBM-SEM, which aims to provide a more accurate representation of the current state. The main contributions are as follows:

(1) To address the issue of a single indicator being inadequate to accurately reflect employment-residence relationships, and the interactions between factors affecting employment-residence balance, a comprehensive evaluation model based on SEM is designed. It integrates multiple evaluation indicators and considers the spatial interaction effects behind these indicators, aiming to derive more reasonable measurement conclusions.

(2) This presents an approach that leverages the characteristics of big data to overcome inherent limitations such as low resolution and limited sample size. It achieves this by utilizing a multi-source dataset comprising primarily mobile signaling data, complemented by various internet big data and traditional data sources.

2. RELATED WORK

2.1 The Analysis Methods of Job-housing Balance

Scholars proposed the concept of job-housing balance for increasingly congested urban spaces, which is widely accepted by urban planners. However, balanced designs in terms of land use cannot effectively address the market's impact on residential and employment choices (Ta et al., 2017). Cost considerations may lead residents to work outside their residential areas and employees may not be able to afford homes near their workplaces. Even with balance policies, it is still necessary to quantitatively measure work-life balance in real life.

Currently, research on analyzing the JHB mainly focuses on discussing measurement indicators or methods (Han et al., 2015), scale issues in the measurement process (Yan et al., 2019), and how the balance affects commuting (Mitra and Saphores, 2019). At the same time, some studies explore what factors influence the job-housing balance. There are many methods for measuring the degree of JHB, including the ratio method, in which the job-housing ratio and independent index are typical representatives. However, these methods have problems with the modifiable areal unit problem and cannot handle cases of job-housing separation due to boundary isolation (Zhou et al., 2018). In addition to the classic ratio method, scholars also use the cost method of commuting time to measure balance, but it does not consider the influence of resident income on commuting time. High-income individuals may choose long-distance commuting for a better quality of life, and transportation can reduce commuting time. Job accessibility is also a commonly used research method, which can directly measure the spatial mismatch between work and residence. Ong (1998) and Shen (1998) provide two different calculation methods, but its accuracy is influenced by the level of public transportation. In general, these methods have their advantages and disadvantages, and their accuracy is affected by multiple factors such as transportation facilities, population income, and job-housing costs.

2.2 Dataset for Job-housing Relationship

Scholars have shifted to new data types for addressing housing and commuting issues, due to technological advancements. Foreign scholars used government survey data (Cervero, 1996) for housing-to-work ratios and comparisons with commuting distances when Geographic Information Systems (GIS) were underdeveloped. Domestic scholars faced limitations in survey data and resorted to survey collection methods (Ta et al., 2017). However, traditional data sources have significant limitations, such as slow updates and low resolution in government census data, and limiting factors in survey questionnaires such as sample size, uniformity of sample distribution, and rational questionnaire design. Thus, this data is more suitable for analyzing factors from a personal perspective rather than providing comprehensive macro-level analysis.

The advent of sensor networks, mobile Internet, high-performance computing, and storage technologies has led to the emergence of new data types, such as trajectory big data and Internet data, which provide copious data resources to capture information about residential and employment locations. Many studies have used trajectory big data to investigate the issue of job-housing mismatch. Someone employed mobile phone signal data to evaluate the aggregated mode of job-housing relationship and commuting time cost, establishing it as a dependable substitute for travel surveys (Yan et al., 2019). Some scholars conducted a study on the job-housing relationship and commuting characteristics near urban rail transit using smart card data (Li et al., 2019). Similarly, Zhou et al. (2016) employed mobile phone signal data to investigate eight typical employment centers in Shanghai and the effect of job-housing mismatch on commuting behavior. Trajectory big data has been used in numerous studies to reveal individual activity patterns, ranges, behavior modes, and social network relationships, thus providing an extensive understanding of group activities at the macro level. This approach largely compensates for the lack of important and comprehensive data sets that are not publicly available in domestic research and eliminates the subjective interference issues involved in questionnaire design.

2.3 Related Work on Structural Equation Model

Structural Equation Modeling (SEM) is a comprehensive measurement model that integrates multiple advantages. It has been widely applied in various disciplines and fields, such as psychology, sociology, education, management, and medicine (Wang and Rhemtulla, 2021; Çakıt et al., 2020). In recent years, it has gradually expanded to the field of geoinformatics (Jaafari et al., 2020) and urban spatial science. SEM is a method for establishing, estimating, and testing causal relationships. It can replace multiple regression, path analysis, factor analysis, covariance analysis, and other methods. By clearly analyzing the effect of individual indicators on the overall situation and the interrelationships between them, SEM is a multivariate statistical modelling technique mainly used for confirmatory model analysis (Shi et al., 2018). Since 2000, hundreds of papers on SEM have been published domestically and internationally. Although it has been widely used in social and humanities sciences, the number of papers related to SEM in the field of Earth science is relatively low (Mcardle and Kadlec, 2013). Challenges still exist in using SEM in geosciences and urban geography, such as the transition from confirmatory model analysis to exploratory model analysis and from modeling of non-spatial statistical variables to modelling of spatial statistical variables.

Despite this, due to the numerous advantages of structural equation modeling (SEM), it has shown some achievements in the field of job-housing issues. Xiao et al. utilized SEM to identify the impact of people's job skills and urban structural factors on employment accessibility (Xiao et al., 2021b).

Meanwhile, Mitra et al. analyzed long-distance commuting among California households using the generalized SEM, verifying the importance of increasing affordable housing and mixed development options in reducing long-distance commuting (Mitra and Saphores, 2019). Based on previous studies, SEM can effectively analyze the impact pathways and magnitudes of different factors on job-housing balance, to some extent, distinguishing the dominant factors of regional balance, which helps to assess each region more targeted instead of using a single indicator to explain all regions. Compared with traditional regression analysis, it also has better interpretability in analyse effects.

3. METHODOLOGY

3.1 Commuter Mobility Matrix and Basic Description Indicators

To explore the job-housing relationship using mobile signaling, the first step is to extract residential and employment locations based on the original data. This is done by calculating the frequency of each user's repeated appearances in the same spatial location or its surrounding areas during the day (10:00-16:00) and night (20:00-6:00) (Yang et al., 2021). The OD points of residential and employment locations are then summarized using a 250m grid, and a commuting flow matrix is aggregated at the subdistrict level for measuring JH spatial relationships.

Assuming there are X subdistricts to be summarized, analyze the spatial relationship between the boundary of each subdistrict and the grid, determine which home and job grids are included in each subdistrict, and add up the commuting population within each OD information according to the corresponding subdistricts to form an X-by-X commuting flow matrix, where x_{ij} represents the commuting population who live in subdistrict *j* and work in subdistrict *i*. Extract basic information from this matrix, as shown in Table 1.

Variable1	Variable2	Description
Commuting Signaling	Total Commuting	Total number of commuting OD flows
	Commuting distance	Linear Geo-distance between pairs of commuting OD points
Commuting Flow	Job signaling	Commuter signaling in the *ith* workplace
	Home signaling	Commuter signaling in the *ith* residence
	Job-housing co-location	Commuter signaling for both work and residence on the *ith* subdistrict
	Commuting outflow	Commuter Signals that reside but do not work at *ith* subdistrict
	Commuting inflow	Commuter Signals that work but do not reside at *ith* subdistrict

Table 1. Basic information in the commuter matrix.

Based on the basic information obtained from the commuting matrix, calculated the fundamental indicators describing job-housing balance. This study employs two indices, namely job-housing ratio and job-housing separation ratio, to represent the main situations of quantity balance and inter-district commuting. Cervero defined job-housing ratio (JHR) as "the ratio of the number of jobs to the number of residents in a designated area", and considers the job-housing relationship to be balanced when the ratio is between 0.8 and 1.2 (Cervero, 1996). For mobile signaling data, we use the number of employed individuals, i.e., job signaling, to replace the number of job positions, and the calculation formula is as follows:

$$JHR_i = \frac{J_i}{H_i} \tag{1}$$

Where JHR_i is the jobs–housing ratio of the *ith* Subdistrict, J_i is the number of workers for the *ith* Subdistrict, and H_i is the number of residents for the *ith* Subdistrict.

Due to interactions between different spatial units in the real world, job-housing ratio cannot reveal actual imbalances caused by cross-district commuting (Yi et al., 2022). Job-housing separation ratio is therefore introduced to depict cross-district commuting activity within a region. This index comprises two parts: the outflow rate of commuting (residential separation rate), and the inflow rate of commuting (employment separation rate). In this study, these indicators can be explained by the following formulas:

$$JHSRO_i = \frac{outflow_i}{H_i} \tag{2}$$

$$JHSRI_i = \frac{inflow_i}{J_i} \tag{3}$$

where $JHSRO_i$ is the commuting outflow rate of the *ith* Subdistrict, $JHSRI_i$ is the commuting inflow rate of the *ith* Subdistrict, $inflow_i$ is the number of commuting inflow for the *ith* Subdistrict, $outflow_i$ is the number of commuting outflow for the *ith* Subdistrict, and H_i or J_i is the number of Home signaling or job signaling in the *ith* Subdistrict.

3.2 Path Analysis Based on SEM

It is evident that a single indicator can only reflect the job-housing relationship from a certain perspective. By analyzing multiple indicators, we can discover that these indicators are influenced by certain factors that have some degree of overlap or interaction. Instead of simply adding up multiple indicators, this study uses SEM to conduct path analysis on multiple indicators and their influencing factors, making it easier to determine the strength and path of each factor's influence on the overall result. The output results obtained from SEM's path analysis also serve as input data for our JHB measurement model.

The structural equation model is a comprehensive technique that integrates various multivariate analysis methods, including factor analysis, latent variable estimation, and path analysis (Shi et al., 2018). This modeling approach combines the knowledge-driven and data-driven modeling features to enable the direct, indirect, and overall effects of a variable on another to be analyzed. Based on previous research, we have designed latent variables that reflect the balance levels of quantity, quality, and space at three different levels, starting from the influencing factors and measurement indicators of job-housing balance. We selected observable variables that reflect the principles of latent variable interactions and made assumptions about the relationships between these variables. The basic structure of the structural equation model is shown in Figure 1. The quantity balance aims to measure the aggregation of job-housing choices in a region through employment density and residential density, providing a shallow but intuitive level of measurement of job-housing relationships. the quality balance is measured by housing prices, property management fees, service facility density, and job density, indicating whether a region has the ability to attract people to live there for a long time from the perspective of the market; and finally, the space balance is measured by the density of bus stops, metro stations, and road networks, indicating the level of commuting convenience in a region. We posit that convenient commuting leads to lower costs for workers, enabling them to reach their workplace within a reasonable timeframe even if local employment is limited. This job-housing balance is achieved through transportation. We hypothesize that there are three fundamental paths in SEM, H1, H2, and H3, where job-housing aggregation, job-housing attraction, and commuting convenience all have a direct positive impact on job-housing balance in a city. The first three are exogenous latent variables, and job-housing balance is an endogenous latent variable with measured variables corresponding to job-housing ratio, commuting distance, and job uniformity.

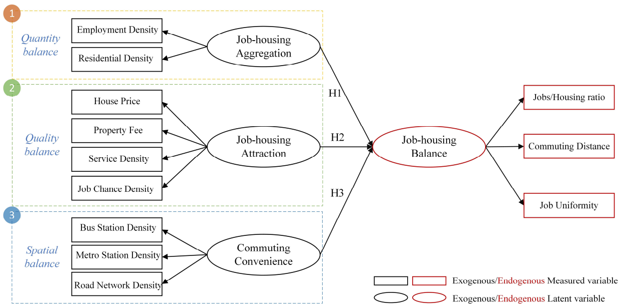

Figure 1. The Structure of path diagram.

Check variable set reliability before SEM modeling. Proceed if Cronbach's $\alpha > 0.7$ and model fit parameters meet standards. Record factor loading and path coefficients to explore influence strengths and verify model hypotheses.

3.3 JHB Measurement Model

The JHB measurement model proposed in this article is a measurement method constructed using the idea of multi-index fusion and calculated using a structural equation model. The core idea is to use the factor loadings and path coefficients of SEM to measure the weight of different factors in measuring the work-life balance. By integrating the variable scores of each factor with weights, considering the interactions between factors in the fusion process, it is possible to allocate the contribution of each factor to the final result in a more reasonable way. This is the main innovation of this article. We define the SEM-based Job-housing balance measurement model (referred to as JHBM-SEM) as follows:

$$F = \sum_{i=1}^{3} A_i \gamma_i \tag{4}$$

$$A_i = \sum_{j=1}^{k} x_{ij} \lambda_{ij} \tag{5}$$

where A_i is the balance potential score of the *ith* exogenous latent variable with k observed variables, x_{ij} denotes the actual score of the *jth* measured variable, which is an index normalized to the raw data of the relevant indicator; λ_{ij} is the effect factor loadings between the *jth* observed variables of the *ith* latent variable, and γ_i is the path coefficient of the intensity of the effect of the *ith* potential variable on the job-occupancy balance.

It is important to note that the latent variable path coefficients and the measured variable factor loadings in the structural equations are first normalized before constructing the model, i.e.:

$$\sum \gamma_i = 1 , \sum \lambda_{ij} = 1 \tag{6}$$

The units of indicators calculated from multiple data sources are not unified. To eliminate the dimensional impact on the indicator values, it is necessary to standardize the actual scores x_{ij} of the measured variables. In this paper, the zero-mean normalization method is used, which serves to eliminate the dimensional influence between the data and to remove the impact of different scales of different indicators on the analysis results. The transformation function is:

$$x^* = \frac{x - \mu}{\delta} \tag{7}$$

Where μ represents the mean value of all sample data, δ represents the standard deviation of all sample data.

4. STUDY AREA AND DATASET

4.1 Study Area

In this study, Shanghai was selected as a case-study, specifically 15 districts within it, except for Chongming District (shown in Figure 2). Chongming District, as a separate island, was excluded to ensure geographic continuity. Wei et al.'s (2016) research categorizes these districts into four groups: traditional city proper area (Huangpu, Jing'an), extended central city area (Xuhui, Changning, Putuo, Hongkou, Yangpu), inner suburban area (Pudong, Minhang, Baoshan), and outer suburban area (Jiading, Qingpu, Songjiang, Jinshan, Fengxian), roughly indicating the development level of each district. The central urban area of Shanghai is within the Outer Ring Road and covers around 660 km² with a population of over 11 million people (Zhou et al., 2016). Commuting behaviors are highly concentrated in this area, which is the focus of this study. However, due to the city's polycentric urban spatial form and the large-scale expansion of key development areas, several new towns have emerged (e.g., Jiading New Town, Qingpu New Town, and Songjiang New Town), which will significantly impact workers' residential and job choices. Therefore, this study gives secondary focus to suburban and exurban areas beyond the central urban area.

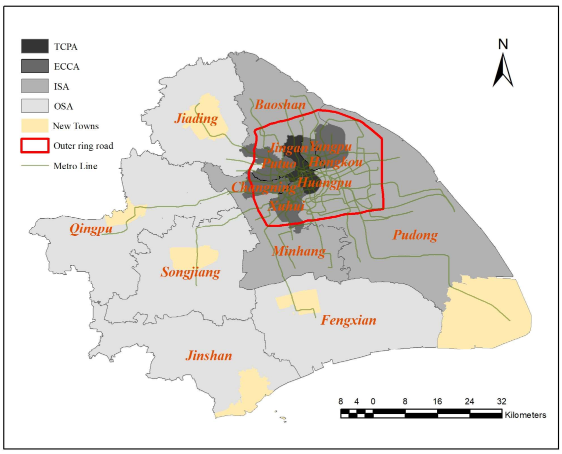

Figure 2. Districts and ring road in Shanghai.

4.2 Dataset

Our dataset is characterized by its diverse sources, primarily comprised of mobile phone signaling data. The mobile signaling data is obtained from the Smart Steps Core Insight Platform and comprises of a one-month data collection period from June 2019. The data, in the form of a 250m×250m grid network, summarizes the commuting patterns of residents who live in grid A and work in grid B, as presented in Table 2. Additionally, we employ 2020 Shanghai POI data, December 2022 secondary market housing price and property data, and the 2019 population census and fourth economic census data. The Amap open platform API provided the POI data for us to analyze the distribution of commercial and transportation facilities in a given region. The housing data is obtained from *Fang.com*, an online real estate platform, and represents the average secondary housing prices and property fees of residential communities, characterizing the accessibility of housing in a specific area and the quality of living environment. Lastly, statistical data from the Shanghai Municipal Government and the Shanghai Statistics Bureau, specifically the population census and economic census, aided in validating employment position distribution, as the economic census data was collected at a similar time to the mobile phone signaling data.

	home_id	work_id	pop_total
0	100105	90175	2.62
1	100105	92322	2.86
2	100105	93197	3.33
……	……	……	……

Table 2. Example of the job-housing data.

5. RESULTS AND DISCUSION

5.1 The Classical Description of Job-housing Relationships

This study examines the spatial distribution of residential and employment areas in Shanghai at the subdistrict level using mobile phone signaling data (shown in Figure 3). The research area comprises 15 districts and 204 subdistricts, which were classified into 7 levels using Jenks Natural Breaks Classification. Overall, Shanghai's residential and employment areas exhibit a single-point centre pattern, gradually decreasing in density along the TCPA-ECCA-ISA-OSA chain. In the far suburbs, the density for both employment and residence is relatively uniform, perhaps due to less developed transportation and connectivity with other regions, leading to self-sufficiency. The study also shows medium-density areas near new cities, highlighting the effectiveness of Shanghai's multi-centre development planning. In high-density areas, regions with residential densities over 11,000 people/km² account for 8.3% of the city, concentrated in central urban areas such as Changning, Yangpu, and Putuo. The areas with employment densities over 30,000 people/km² account for 2.5% of the city and are mainly distributed in Xuhui, Jing'an, and Putuo, indicating that highly busy and active commuting areas are still concentrated in these districts.

After roughly assessing the spatial distribution of residential and employment areas, this article calculated the job-housing ratio for each subdistrict based on residential and employment signaling data. A ratio of 1 indicates complete balance, while research suggests that ratios between 0.8 and 1.2 indicate balance. The job-housing ratios were further processed to determine their relative distance from 1, represented as $|JHR_i - 1|$ in Figure 4. This approach enables us to prioritize the balance level within a region, temporarily setting aside the region's inclination towards Ji or Hi. As a result, the graph's information becomes more easily discernible. The dark blue area (0-0.2) represents the balanced region, with warmer colors indicating greater imbalance. The overall spatial distribution pattern is quite blurred, with balanced areas scattered throughout each district of the city rather than concentrated in specific clusters. However, most balanced regions are found in the outer suburban area, while a minority are located in the south-eastern and southern parts of the central urban area, and highly imbalanced regions are almost entirely concentrated in the central urban area.

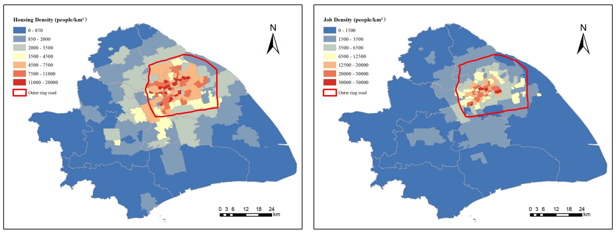

Figure 3. Spatial distributions of resident (left) and employment (right) density.

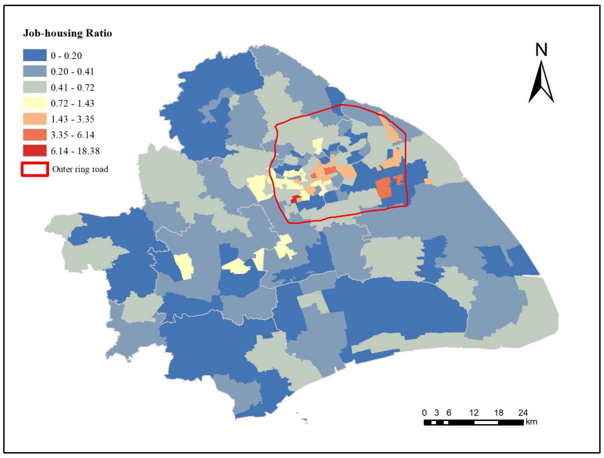

Figure 4. Job–housing ratio in Shanghai (after processing).

Owing to cross-district commuting, the job-housing ratio falls short of capturing scenarios where people live and work in separate areas. Therefore, we have introduced the job-housing separation rate to compensate for this deficiency. While the job-housing ratio captures the disparity in magnitude between job and housing quantities, the job-housing separation rate gauges the actual correspondence between them, considering spatial separation in residents' job-housing behaviors. Figure 5 displays the computation of two aspects of the job-housing separation rate: the outflow rate and the inflow rate. Of all subdistricts, 55 subdistricts (27% of the total) have both inflow and outflow rates above 95%, while 89 subdistricts (43.6% of the total) have both rates above 90%. This indicates a severe job-housing separation and a high proportion of two-way commuting, leading to traffic congestion and disorder in urban spatial order. In certain areas with high inflow of commuting to employment centers, the outflow is also high, suggesting that local residents are unable to match the employment function of the area. Even if there are sufficient job opportunities, cross-district employment cannot be avoided. Comparison between Figures 4 and 5 shows that in certain areas with good job-housing ratios, both commuting inflow and outflow rates are high, with the average commuting time in the corresponding area not being less than 10km. This implies that the job-housing situation in the area is imbalanced, indicating certain limitations in the interpretation of the JHR.

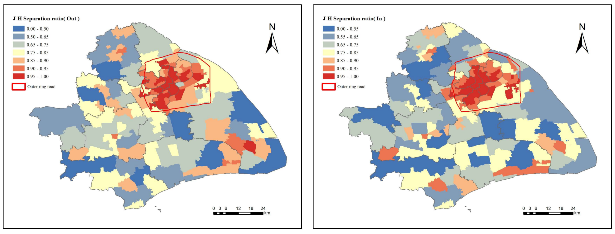

Figure 5. Commuter outflow (left) and commuter inflow (right) in Shanghai.

5.2 Results of the JHB Measurement Model

To overcome the limitations of traditional measurement methods, this paper has developed a comprehensive measurement method based on structural equation modeling. During the SEM construction process, we used nine exogenous observed variables to construct three exogenous latent variables and one endogenous latent variable. All observed data were sorted and cleaned, and records with zero signal accumulation and too small area were deleted. As shown in Table 3, we conducted reliability analysis on the three groups of observed variables. Their Cronbach's α values were all greater than 0.7, indicating good reliability of the data, which passed the reliability test.

After the confirmatory factor analysis, the structural model was established using the path diagram method with the assistance of *Amos* software, and the factor loadings, path coefficients, and error terms in the model were calculated. SEM provides various fit indices that reflect the overall goodness of fit between the model and the data. The recommended and actual values of each index are shown in Table 4. It is generally believed that the model fits the data well only when all the indices meet the recommended values. Our actual indices basically meet the recommended values, but the RMSEA slightly exceeds the threshold value.

Latent Variable	Measured Variable	Cronbach's α
Job-housing Aggregation	Employment Density	0.814
	Residential Density	
Job-housing Attraction	House Price	0.749
	Property Fee	
	Service Density	
	Job Chance Density	
Commuting Convenience	Bus Station Density	0.851
	Metro Station Density	
	Road Network Density	

Table 3. Reliability test of variables (n=204).

	X^2/df	RMSEA	GFI
Recommended	<5	<0.08	>0.9
Actual	4.788	0.086	0.963

Table 4. Evaluations of model fitting effect.

Utilizing the factor loadings and path coefficients calculated from SEM, we applied them to our proposed model and obtained the results displayed in Figure 6. As the weights calculated from the path coefficients and factor loadings may assume either positive or negative values, and the variable scores have been standardized, larger output values are not necessarily indicative of better results. When the outcome is 0, it is deemed that the impact of various variables on the final result is relatively balanced, or that the complementary strengths and weaknesses of diverse indicators have attained an ideal state. Our initial range for balanced outcomes was set between -0.2 to 0.2, as indicated by the yellow-colored patches in the diagram. Analysis of the results demonstrated that the outcomes for the outer suburban areas were primarily less than -0.5, while those of the central city areas, especially the inner-city areas, were largely greater than 0.5. These two types of regions are not categorized as balanced. The inner suburban regions between the two, particularly those near the outer ring road, possess a relatively good job-housing balance. The overall result shows a distribution pattern of worse on both ends and better in the middle. From the data analysis, it can be seen that about half of the city has a medium level of JH balance in Shanghai. The number of areas with extremely good and poor balances is comparable. The data is classified into 9 levels, we selected the lowest and highest two levels, and tallied a total of 52 subdistricts within these two extreme intervals, accounting for 25.5% of all. We also selected the three intermediate levels and tallied a total of 58 subdistricts within this relatively balanced range, accounting for 28.4% of all subdistricts.

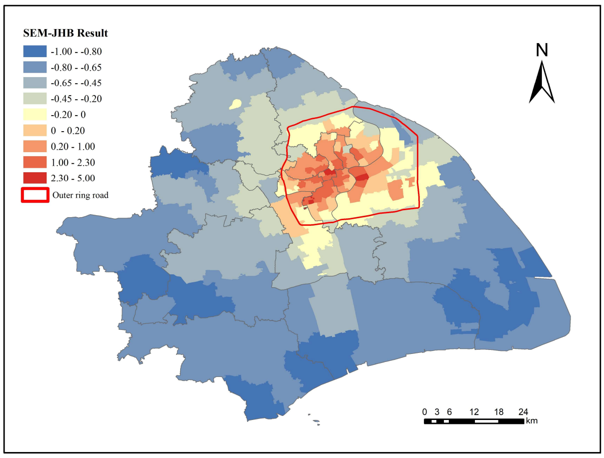

Figure 6. Results of the JHB measurement model.

5.3 Calculate the fundamental Comparing the performance of both

The main goal of this article is to achieve multi-index integration by building models. The reason for integration is that a single index cannot explain the balance between work and residential areas well. When interpreting traditional indices, it is found that phenomena that the job-housing ratio cannot explain can be supplemented by the job-housing separation rate, as shown in Figure 7. Put the commute outflow rate, commute inflow rate, and processed job-housing ratio in the same coordinate system, the definition of balance is satisfied when JHRto1 < 0.2. From the figure, it can be seen that areas with a relative job-housing ratio less than 0.2 have high inflow and outflow rates, while areas with a high job-housing ratio may have low inflow and outflow rates. We selected several typical areas from the set of job-housing ratios and job-housing separation rates that do not match and compared their performance with JHBM-SEM, as shown in Table 5. These areas are located in the developed districts of Putuo, Jing'an, Yangpu, and Pudong in the main city area. They are all in the 0.8~1.2 job-housing ratio interval and close to JHR=1, suggesting excellent balance. However, their corresponding job-housing separation rates are almost all above 95%, implying that it is difficult for residents in these areas to find work locally. Combining the corresponding average commuting distance shows that these workers need to travel more than 10 km, taking more than an hour on rail transit, indicating

Subdistrict Name	District Name	JHR	Inflow	Outflow	Travel Distance	JHBM-SEM
Caoyang	Putuo	1.06	0.97	0.97	10.19	1.51
BaoshanRoad	Jing'an	0.89	0.98	0.99	8.5	1.33
GuangzhongRoad	Jing'an	1.00	0.96	0.96	9.30	1.7
Seaport EDZ	Pudong	1.15	0.94	0.93	18.34	-0.70
SipingRoad	Yangpu	1.15	0.95	0.94	12.70	1.4

Table 5. Part of the parameters for evaluating abnormal regions.

long-distance commuting. Our model measured these areas, and the results were far from 0, demonstrating that the degree of job-housing balance in the research target is poor. This conclusion is more reasonable after our comprehensive discussion, illustrating that the model's results are more valid.

Based on Figure 6, our model differs significantly from the job-housing ratio, particularly in outer suburban areas. While these areas show balanced job-housing ratios, they do not fall within the equilibrium range based on JHBM-SEM. Most outer suburban areas have JHBM-SEM values less than -0.5, indicating a "shallow balance" state that only maintains a self-sufficient balance in terms of residents and jobs but lacks real development. In urban planning, it is believed that the ability of residents to walk to work is a result of the city having well-equipped service facilities, and reasonable configuration of jobs and housing spaces, rather than a backward area that achieves a "balanced" state passively under market forces with underdeveloped living and working conditions. Our model identifies such pseudo-equilibrium states and provides a method for distinguishing them. The model's high-value areas can explain common imbalances caused by spatial mismatches and long-distance commuting. As shown in Figures 4 and 6, central urban areas, especially the inner city, received unbalanced evaluations under both measurement methods. The model identifies Shanghai's job-housing balanced areas mainly in the vicinity of the outer ring road (yellow patches), which belong to the inner suburban area. Most of them are located in Baoshan, Minhang, and Pudong districts, where the residential density, employment density, and job-housing separation rate are at a moderate level. These areas are classified as balanced by our model and have an average commuting distance of 2-6km due to good transportation coverage and close connection with the central urban area. Our model's classification of these areas is more in line with actual residents' travel situations, as reflected by multiple sources of data.

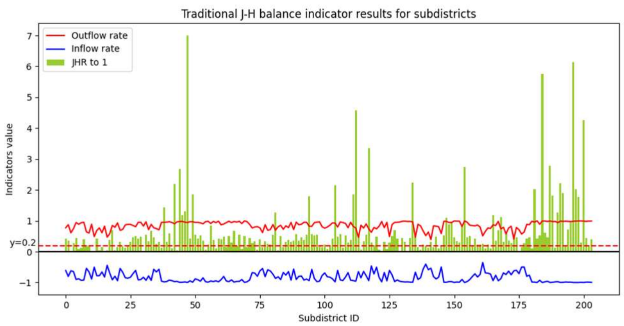

Figure 7. Traditional JHB indicator results for subdistricts.

6. CONCLUSIONS

In this paper, we propose a job-housing balance measurement model that accounts for multi-factor interactions. It integrates various internet big data and mobile signaling to comprehensively evaluate the job-housing balance level of a region. Compared with traditional measurement methods such as the job-housing ratio and the job-housing separation rate, our model can identify the pseudo-balance phenomenon in underdeveloped areas in the suburbs and make reasonable evaluations. It can also produce similar results to classical methods to measure job-housing imbalance. We examine the job-housing balance status at the subdistrict level in Shanghai and discover that balanced areas are mainly near the outer ring road, covering around half the study area, and located in the middle of the central city and suburbs. The imbalance in the unbalanced areas is influenced by the spatial layout of public transportation and urban service facilities, as well as job positions. However, our model's expressive power is somewhat limited, and it is dependent on the selection of measurement variables in SEM. Our attempt only focuses on three aspects, and future studies can select more indicators to improve the model's suitability.

As cities rapidly expand, achieving regional job-housing balance within a jurisdiction is no longer enough to meet residents' needs. Commuting across districts is common in mega-cities, making it essential to use public transportation planning to keep commuting time reasonable and achieve job-housing balance. Additionally, a polycentric urban structure can alleviate population pressure and improve urban carrying capacity, significantly aiding job-housing balance. The JHBM-SEM can modify the model's measurement variables to produce job-housing balance results suitable for different cities in different states. This helps interpret job-housing patterns in cities and provides data and empirical support for decision-making departments to develop urban planning strategies.

ACKNOWLEDGMENTS

This research was funded by the Beijing Natural Science Foundation (Grant #8222009), Training Programme for the Talents by Xicheng, Beijing (Grant #202137) and China Scholarship Council High-level Talent Training Program (Grant #20221007).

REFERENCES

Çakıt, E., Olak, A. J., Karwowski, W., Marek, T., Hejduk, I.,Taiar, R., 2020: Assessing safety at work using an adaptive neuro-fuzzy inference system (ANFIS) approach aided by partial least squares structural equation modeling (PLS-SEM). *International Journal of Industrial Ergonomics*, 76, 102925, doi.org/10.1016/j.ergon.2020.102925.

Cervero, R., 1996: Jobs-Housing Balance Revisited: Trends and Impacts in the San Francisco Bay Area. *Journal of the American Planning Association*, 62, 492-511, doi.org/10.1080/01944369608975714.

Han, H., Yang, C., Wang, E., Song, J.,Zhang, M., 2015: Evolution of jobs-housing spatial relationship in Beijing

Metropolitan Area: A job accessibility perspective. *Chinese Geographical Science*, 25, 375-388, doi.org/10.1007/s11769-015-0738-1.

Horner, M. W.,Mefford, J. N., 2007: Investigating Urban Spatial Mismatch Using Job–Housing Indicators to Model Home–Work Separation. *Environment and Planning A: Economy and Space*, 39, 1420-1440, doi.org/10.1068/a37443.

Howard, E.: Garden cities of tomorrow, Faber London1946.

Jaafari, S., Shabani, A. A., Moeinaddini, M., Danehkar, A.,Sakieh, Y., 2020: Applying landscape metrics and structural equation modeling to predict the effect of urban green space on air pollution and respiratory mortality in Tehran. *Environmental Monitoring and Assessment*, 192, 412, doi.org/10.1007/s10661-020-08377-0.

Jing, C., Zhang, H., Xu, S., Wang, M., Zhuo, F.,Liu, S., 2022: A hierarchical spatial unit partitioning approach for fine-grained urban functional region identification. *Transactions in GIS*, 26, 2691-2715, doi.org/10.1111/tgis.12979.

Li, T., Chen, Y., Wang, Z., Liu, Z., Ding, R.,Xue, S., 2019: Analysis of Jobs-Housing Relationship and Commuting Characteristics Around Urban Rail Transit Stations. *IEEE Access*, 7, 175083-175092, doi.org/10.1109/ACCESS.2019.2957416.

Martinus, K.,Biermann, S., 2018: Strategic planning for employment self-containment in metropolitan sub-regions. *Urban Policy and research*, 36, 35–47, doi.org/10.1080/08111146.2017.1294539.

McArdle, J. J.,Kadlec, K. M., 2013: Structural equation models. *The Oxford handbook of quantitative methods*, 2, 295–337.

Mitra, S. K.,Saphores, J.-D. M., 2019: Why do they live so far from work? Determinants of long-distance commuting in California. *Journal of Transport Geography*, 80, 102489, doi.org/10/gh2sjd.

Ong, P.,Blumenberg, E., 1998: Job access, commute and travel burden among welfare recipients. Urban Studies, 35, 77–93.

Qin, P.,Wang, L., 2019: Job opportunities, institutions, and the jobs-housing spatial relationship: Case study of Beijing. *Transport Policy*, 81, 331-339, doi.org/10.1016/j.tranpol.2017.08.003.

Shen, Q., 1998: Location Characteristics of Inner-City Neighborhoods and Employment Accessibility of Low-Wage Workers. Environment and Planning B: Planning and Design, 25, 345-365, doi.org/10.1068/b250345.

Shi, D., Lee, T.,Terry, R. A., 2018: Revisiting the Model Size Effect in Structural Equation Modeling. *Structural Equation Modeling-a Multidisciplinary Journal*, 25, 21-40, doi.org/10.1080/10705511.2017.1369088.

Ta, N., Chai, Y., Zhang, Y.,Sun, D., 2017: Understanding job-housing relationship and commuting pattern in Chinese cities: Past, present and future. *Transportation Research Part D: Transport and Environment*, 52, 562-573, doi.org/10.1016/j.trd.2016.11.011.

Wang, H., Zeng, W.,Cao, R., 2021: Simulation of the Urban Jobs-Housing Location Selection and Spatial Relationship Using a Multi-Agent Approach. *Isprs International Journal of Geo-Information*, 10, 16, doi.org/10.3390/ijgi10010016.

Wang, J., Zhou, C., Rong, J., Liu, S.,Wang, Y., 2022: Community-detection-based spatial range identification for assessing bilateral jobs-housing balance: The case of Beijing. *Sustainable Cities and Society*, 87, 104179, doi.org/10/grxd97.

Wang, Y. A.,Rhemtulla, M., 2021: Power analysis for parameter estimation in structural equation modeling: A discussion and tutorial. *Advances in Methods and Practices in Psychological Science*, 4, 2515245920918253.

Wei, Y. D., Bi, X., Wang, M., & Ning, Y., 2016: Globalization, economic restructuring, and locational trajectories of software firms in Shanghai. The Professional Geographer, 68(2), 211–226.

Xiao, L., Lo, S., Liu, J., Zhou, J.,Li, Q., 2021a: Nonlinear and synergistic effects of TOD on urban vibrancy: Applying local explanations for gradient boosting decision tree. *Sustainable Cities and Society*, 72, 103063, doi.org/10.1016/j.scs.2021.103063.

Xiao, W., Wei, Y. D.,Li, H., 2021b: Spatial inequality of job accessibility in Shanghai: A geographical skills mismatch perspective. *Habitat International*, 115, 102401.

Yan, L., Wang, D., Zhang, S.,Xie, D., 2019: Evaluating the multi-scale patterns of jobs-residence balance and commuting time–cost using cellular signaling data: a case study in Shanghai. *Transportation*, 46, 777-792, doi.org/10.1007/s11116-018-9894-3.

Yang, Y., Xiong, C., Zhuo, J.,Cai, M., 2021: Detecting Home and Work Locations from Mobile Phone Cellular Signaling Data. *Mobile Information Systems*, 2021, 5546329, doi.org/10.1155/2021/5546329.

Yi, H., Wang, L., Li, Q.,Li, X., 2022: Investigate Jobs-Housing Spatial Relationship with Individual-Based Mobility Big Data of Public Housing Tenants: A Case Study in Chongqing, China. *Sustainability*, 14, 3211, doi.org/10.3390/su14063211.

Zhang, P., Zhou, J.,Zhang, T., 2017: Quantifying and visualizing jobs-housing balance with big data: A case study of Shanghai. *Cities*, 66, 10-22, doi.org/10.1016/j.cities.2017.03.004.

Zhou, J., Wang, Y.,Schweitzer, L., 2012: Jobs/housing balance and employer-based travel demand management program returns to scale: Evidence from Los Angeles. *Transport Policy*, 20, 22-35, doi.org/10.1016/j.tranpol.2011.11.003.

Zhou, J., Wang, Y., Cao, G.,Wang, S., 2017: Jobs-housing balance and development zones in China: a case study of Suzhou Industry Park. *Urban Geography*, 38, 363-380, doi.org/10.1080/02723638.2016.1152845.

Zhou, X., Chen, X.,Zhang, T., 2016: Impact of megacity jobs-housing spatial mismatch on commuting behaviors: A case study on central districts of Shanghai, China. *Sustainability*, 8, 122.

Zhou, X., Yeh, A. G. O.,Yue, Y., 2018: Spatial variation of self-containment and jobs-housing balance in Shenzhen using cellphone big data. *Journal of Transport Geography*, 68, 102-108, doi.org/10.1016/j.jtrangeo.2017.12.006.