Proceedings of the 5th International Conference on Financial Technology and Business Analysis
DOI: 10.54254/2754-1169/2026.38926

City-Level Market Contexts and Cross-City Disparities in Chinese Firm-Survival

Yaxin Wang

Department of Mathematics, University of Pittsburgh, Pittsburgh, USA
wangyaxin62@gmail.com

Abstract. This paper investigates how city level economic and demographic conditions shape firm-survival across Chinese cities. Using 2000-2024 panel data covering 50 cities and over 5 million firms, this study applies principal-component analysis (PCA) to handle high indicator correlation and constructs three composite indices: economic opportunity, business-pressure, and population-base index. These indices are incorporated into weighted least-squares regressions. The baseline specification controls for year, industry and firm ownership fixed effects, while the preferred model further adds city fixed effects; unweighted OLS serves for robustness checks. Baseline results show the population-base index has a significantly negative association with firm-survival, whereas economic opportunity and business-pressure indices are statistically insignificant at conventional levels. Once city fixed effects are included, all three indices lose significance. This suggests observed cross-city firm-survival gaps stem largely from persistent intercity structural differences rather than short-run within-city changes. Robustness tests reveal operating pressure correlates negatively with firm-survival rates. Overall, firm-survival disparities are pronounced across Chinese cities, requiring researchers to distinguish cross sectional city differences from over time within-city variations when analysing urban environment firm-survival linkages.

Keywords: city-level market contexts, cross-city disparities, operating pressure, principal-component analysis

1. Introduction

In China's economic system, enterprises are important participants in the market economy and play an important role in promoting economic growth, creating employment, and promoting market innovation. Therefore, enterprises' survival and development are closely tied to regional economic stability and high-quality socio-economic development. With the continuous development of China's market economy, the number of market entities has increased rapidly, and market competition has become increasingly fierce. At the same time, many enterprises still face problems such as short survival cycles, high exit risk, and increasing operating pressure. These problems not only affect individual enterprises' development but may also weaken regional economic resilience and market stability. In addition, enterprises face clear differences in economic conditions and residents' demographic characteristics across cities. Therefore, the core question of this article is whether a

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2754-1169/2026.38926

systematic relationship exists between differences in the urban-level market environment and differences in enterprise survival conditions across cities in China.

Studies have analyzed the factors affecting enterprise survival from many angles. The characteristics of enterprises, such as enterprise scale, establishment years, productivity, and ownership structure, are considered to be important factors affecting the sustainable business ability of enterprises [1]. Factors such as entrepreneurial characteristics, internal financing, inventory management, and marketing strategies will also have an important impact on the survival of small and medium-sized enterprises and micro-enterprises [2]. Financing conditions also affect the survival and development of enterprises. Equity financing can improve the survival and development conditions of small and medium-sized enterprises, while insufficient intellectual property protection may limit the further growth of R&D enterprises [3]. Relevant theoretical research also points out that enterprises will gradually understand their own productivity in the process of operation. When enterprises cannot meet the market's minimum survival threshold, low-productivity enterprises are more likely to withdraw from the market [4]. The macroeconomic environment also affects enterprise survival. The decline in aggregate demand and economic contraction may increase the exit risk of low-productivity enterprises [5]. At the regional level, economic growth and residents' income may affect enterprise survival by expanding market demand and increasing factor costs [6].

Although existing literature has extensively investigated firm-survival from perspectives of firm-specific characteristics, financing, business strategies and regional economic performance, limited research explores the joint impacts of city-level economic and demographic factors on cross-city firm-survival disparities in China. Considering China's substantial inter-city differences in economy, population, labour and market size that create heterogeneous business environments, this study departs from single-factor frameworks to conduct city-level analysis. Based on 2000-2024 data covering key economic and demographic variables, it adopts PCA to reduce dimensionality and construct three composite indices: economic-opportunity, business-pressure and population-base indices. Weighted regression and city fixed-effects models are utilised for estimation. This paper makes three contributions. It connects firm-survival with PCA-constructed indices to mitigate multicollinearity and reflect comprehensive urban conditions; it supplies long-term panel-based evidence covering 50 cities, 10 industries and over 5 million firms to identify multi-dimensional heterogeneity; it differentiates cross-city gaps from within-city temporal variations via model comparison. The study theoretically expands firm-survival research and practically illuminates urban business-environment and firm-survival differences.

2. Framework

This study develops a theoretical framework linking city-level economic conditions to enterprise survival. Following the industry dynamics literature, firms are heterogeneous in productivity and remain in the market only if expected profits exceed a minimum survival threshold. There is productivity heterogeneity between enterprises. The entry, exit, and continued operation of an enterprise depend on whether its productivity and expected income are sufficient to cover the cost of continuing operation [7]. Macroeconomic conditions, household income and fiscal burdens influence enterprise survival by altering the profitability of operating within a city. The aggregate survival rate observed at the city-year level is therefore the outcome of individual firm-survival decisions. Consider firm i operating in city c during year t. Firm profit is defined as Model (1):

DOI: 10.54254/2754-1169/2026.38926

$$\pi_{ict} = a_i X_{ct} - Z_{ct} \tag{1}$$

Where a_i denotes firm-specific productivity, X_{ct} represents the economic opportunity available in city c at time t, and Z_{ct} captures operating pressure. A firm remains active only when its profit exceeds a minimum survival threshold $\pi_{ict} \geq \bar{\pi}$. Rearranging the expression gives as Model (2):

$$a_i \geq a_{ct}^{*} = \frac{\bar{\pi} + Z_{ct}}{X_{ct}} \tag{2}$$

Where a_{ct}^{*} denotes the minimum productivity level required for survival. This threshold provides the central mechanism of the model. Any change in local economic conditions affects enterprise survival through its impact on the productivity threshold required to remain in operation. The economic opportunity index X_{ct} reflects the ability of local markets to generate demand and revenue opportunities for firms. Formally, $X_{ct} = X(Y_{ct},\ I_{ct},\ M_{ct})$, Where Y_{ct} is the economic scale of the region, which is used to measure the overall economic volume and industrial volume of the region. I_{ct} is the income level of residents, which is used to reflect the actual purchasing power and consumption ability of the family sector. M_{ct} is the market size and population vitality, which is used to measure the regional market capacity, demographic characteristics, and dynamic potential of population development. According to theoretical inference, the economic scale, the income level of residents, and the vitality of the market population all have a positive marginal impact on the regional economic opportunity index; that is, the improvement of the three will enhance the ability of the regional market to create demand and revenue for enterprises. The corresponding theoretical expectation can be expressed as: $\frac{\partial X_{ct}}{\partial Y_{ct}} > 0,\ \frac{\partial X_{ct}}{\partial I_{ct}} > 0,\ \frac{\partial X_{ct}}{\partial M_{ct}} > 0,$ A larger total economy, stronger purchasing power of residents, and a more active population market will directly expand the scale of effective market demand, raise the revenue space and profitability of enterprises, and thus lower the productivity threshold required for enterprises to survive in the market.The operating-pressure index Z_{ct} captures the costs and constraints faced by firms: $Z_{ct} = Z(W_{ct}, T_{ct}, F_{ct})$, Where W_{ct} is the labor cost, which corresponds to all kinds of rigid human expenditures generated by the employment link of enterprises. T_{ct} is the tax and institutional burden, covering the tax burden of enterprises and the institutional transaction costs generated in the business process. F_{ct} is the availability of financial resources, reflecting the convenience of enterprises to obtain credit, financing, and other financial support. The expected relationships are: $\frac{\partial Z_{ct}}{\partial W_{ct}} > 0,\ \frac{\partial Z_{ct}}{\partial T_{ct}} > 0,\ \frac{\partial Z_{ct}}{\partial F_{ct}} < 0$. The increase in labor wages pushes up labor costs, taxes, and institutional burdens, which will directly increase the rigid operating expenses of enterprises and amplify the degree of operating pressure. The higher the availability of regional financial resources, the more relaxed the financing constraints of enterprises, and they can rely on more convenient financial support to smooth business fluctuations, thus significantly reducing operating pressure.

Because firms differ in productivity, assume productivity follows a cumulative distribution function G(a). The city-level enterprise survival rate is therefore the proportion of firms whose productivity exceeds the survival threshold. Substituting the threshold expression yields as Model (3):

DOI: 10.54254/2754-1169/2026.38926

$$SR_{ct} = 1 - G\left(\frac{\bar{\pi}+Z_{ct}}{X_{ct}}\right) = 1 - G\left(\frac{\bar{\pi}+Z(W_{ct},T_{ct},F_{ct})}{X(Y_{ct},I_{ct},M_{ct})}\right) \quad (3)$$

This equation links firm-level survival decisions to the aggregate enterprise survival rate observed in the data. The theoretical model provides a conceptual foundation for empirical analysis. Rather than directly modeling firm-level duration outcomes, the framework explains how macroeconomic conditions, demographic fundamentals, and income indicators shape aggregate enterprise survival patterns across Chinese Cities. The observed city-level survival rate can therefore be interpreted as the equilibrium outcome of heterogeneous firms operating under different local economic environments.

3. Data and facts

3.1. Data source, unit of observation, and sample structure

The dataset is primarily sourced from CnOpenData, complemented by National Bureau of Statistics and local statistical yearbooks [8]. Firm level registration, cancellation, industry and ownership records come from industrial commercial registration databases; city level macroeconomic, income and population indicators draw on official public statistics. Variable labels, statistical calibers and measurement units are standardised to guarantee comparability across different data sources. The panel runs 2000 2024 (4478 observations), covering 21 provinces, 50 cities, 51 industries and 31 ownership categories, with 5 077 854 total firms (4 603 151 surviving, 474 703 exited). Weighted firm-survival rate equals 90.65% and exit rate 9.35%. Survival and exit rates are weighted by total firm count instead of simple group averages to prevent sample size driven bias. The multi-dimensional sample supplies solid data support for empirical estimation.

3.2. Missing data and imputation strategy

Table 1 uses urban-level panel data from 2000 to 2024 to study the survival of Chinese enterprises. The final sample contains 4,478 observations, covering 50 cities, 51 industry categories, 31 types of enterprise nature, and 5,077,854 enterprises. The dependent variables, such as enterprise survival and cancellation, have no missing values. Missing data are mainly concentrated in the economic and demographic explanatory variables at the urban level, and the original missing rate for all variables is less than 2%. Among them, the resident population has the highest missing rate (1.88%), while per capita GDP has the lowest (0.04%). For missing values, this paper first uses the data of adjacent years in the same city for linear interpolation; for observations that are still missing at both ends of the time series, it uses the information of adjacent years in the city to fill in forward or backward; if there are still missing values, the median of the same province and the same year and the median of the same year in the country are used in turn. Line supplement. After processing, there are no missing values in any explanatory variable, and all panel observations are retained. For abnormal observations, this article combines variable definitions, units, and reasonable value ranges to distinguish between data input or unit problems and real differences between cities, and avoid the loss of economically significant urban heterogeneity due to mechanical deletion of extreme values.

DOI: 10.54254/2754-1169/2026.38926

Table 1. Missing values before and after panel-based imputation

Variable	Missing before	Missing before (%)	Remaining after imputation
Resident population (10,000 persons)	84	1.88%	0
Tax revenue (100 million yuan)	79	1.76%	0
Household deposits (100 million yuan)	77	1.72%	0
Birth rate (per mille)	72	1.61%	0
Rural disposable income (yuan)	27	0.60%	0
Average wage (yuan)	19	0.42%	0
Urban disposable income (yuan)	10	0.22%	0
GDP (100 million yuan)	6	0.13%	0
Registered population (10,000 persons)	3	0.07%	0
GDP per capita (yuan)	2	0.04%	0

In order to comprehensively portray the external market environment of different cities and reduce the multiple collinearity problem caused by the high correlation between multiple macroeconomic indicators, this paper first uses the PCA method to downsize the economic, population, and resident economic indicators at the urban level, and constructs the Economic Opportunity Index, Operating Pressure Index, and Demographic Fundamentals Index. Among them, the economic opportunity index is mainly composed of indicators such as GDP, per capita GDP, resident income, resident deposits, and resident population; the operating pressure index mainly reflects operating cost factors such as wage level and tax; the demographic fundamentals index is used to reflect the urban population structure and population attractiveness. Subsequently, this paper takes the enterprise survival rate as the explained variable, incorporates three comprehensive indexes into the weighted regression model, and controls for the fixed effects of year, industry, and enterprise nature. It further adds urban fixed effects to control for long-term, unchanging characteristics of different cities, distinguishing long-term structural differences between cities from changes within the same city over time. Since macroeconomic variables are measured at city-year level whereas the regression dataset follows city-year-industry panel structure, PCA is implemented based on distinct city-year observations.

Overall, Chinese enterprise survival differs markedly by time, city, and industry. Between 2000 and 2024, the overall weighted survival rate of sample enterprises was 90.65%, and the exit rate was 9.35%. From the time trend, the survival rate was relatively low at the beginning of the research and reached a low point around 2005. Then it gradually rebounded and remained high for a long period,

DOI: 10.54254/2754-1169/2026.38926

indicating that overall enterprise survival improved during the research period. At the same time, enterprise survival differs markedly across industries. Some industries such as manufacturing and finance show a high survival rate of enterprises, while the withdrawal rate of food processing, some mining, and other industries is relatively high. There are also large differences in enterprise scale and survival rates across cities, showing that the local market environment in which enterprises operate is heterogeneous. These descriptive facts show that the overall national trend alone cannot fully explain enterprise survival. It is necessary to further consider economic opportunities, operating pressures, and the population environment at the urban level, and distinguish between long-term differences between cities and changes within cities over time in the subsequent return. Descriptive statistics of enterprise outcomes and local conditions can be seen in Table 2. Survival and exit rates are row-level rates. Monetary variables are in nominal units as recorded in the source file. The table uses the imputed explanatory variables but original outcome counts.

Table 2. Descriptive statistics of key variables

Variable	N	Mean	Std. dev.	Min	Median	Max
Surviving enterprises	4,478	1,027.95	10,274.55	0.00	8.00	158,690.00
Deregistered enterprises	4,478	106.01	707.25	0.00	1.00	16,650.00
Total enterprises	4,478	1,133.96	10,858.17	1.00	11.00	167,993.00
Survival rate	4,478	0.78	0.30	0.00	0.93	1.00
Exit rate	4,478	0.22	0.30	0.00	0.07	1.00
GDP (100 million yuan)	4,478	18,882.29	15,288.75	35.80	15,420.20	53,759.50
GDP per capita (yuan)	4,478	95,761.67	62,667.93	1,212.49	80,692.00	228,167.00
Resident population (10,000 persons)	4,478	1,743.45	970.08	48.35	1,961.90	28,005.00
Registered population (10,000 persons)	4,478	57,634.85	534,265.49	44.81	1,257.80	6,785,273.00 （outlier）
Birth rate (per mille)	4,478	8.82	27.43	2.79	7.96	23.07
Average wage (yuan)	4,478	90,680.59	66,507.57	12.99	72,273.29	247,174.00
Household deposits (100 million yuan)	4,478	23,366.55	20,635.28	17.70	18,041.30	77,900.00
Tax revenue (100 million yuan)	4,478	12,106.11	87,456.62	0.97	2,251.59	1,651,452.00
Urban disposable income (yuan)	4,478	40,919.22	25,599.69	136.79	34,831.41	93,095.00
Rural disposable income (yuan)	4,478	16,420.27	10,745.30	1,098.49	13,742.00	45,644.00

DOI: 10.54254/2754-1169/2026.38926

3.3. Time-series facts

Figures 1 shows the distribution of sample enterprises among industries and regions, respectively. Sample enterprises are clearly concentrated at the industry level: manufacturing enterprises have the largest share, followed by unclassified industries, while the remaining industries have relatively few enterprises in the sample. This shows large differences in the number of enterprises across industries and reflects the sample's distinct industry structure. Similar concentration phenomena also occur at the regional level. Shanghai has significantly more enterprises than other regions in the sample. Liaoning ranks second, and Zhejiang and Beijing also have relatively large enterprise scales, while the remaining provinces account for a smaller share of the sample. Overall, the sample enterprises are highly concentrated regionally, and market size and enterprise numbers differ clearly across regions. This heterogeneity at the industry and regional levels shows that enterprises face different market environments, which also provides a basis for controlling for industry characteristics and further examining urban-level differences in subsequent empirical analysis.

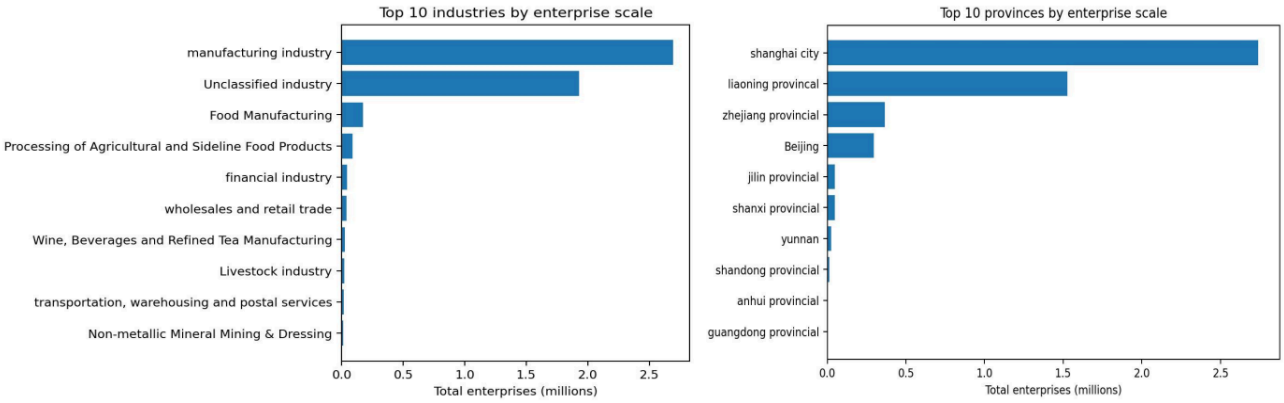


Figure 1. Top 10 industries and Top 10 provinces (picture credit: original)

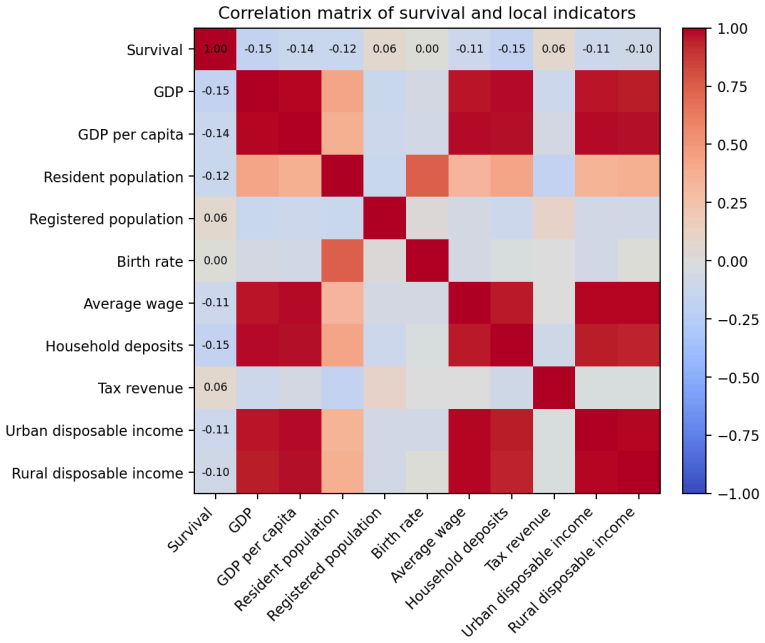


Figure 2. Correlation matrix of survival and local indicators (picture credit: original)

DOI: 10.54254/2754-1169/2026.38926

Figure 2 presents the correlation heatmap. Macroeconomic and household income variables exhibit strong mutual positive correlation, while resident population correlates moderately with economic metrics. Registered population, tax revenue and birth rate show weak links to most other variables. Firm-survival rate correlates weakly negatively with most economic income indicators, weakly positively with registered population and tax revenue, and nearly uncorrelated with birth rate. Survival's correlation with local conditions is substantially weaker than co movement among macro indicators themselves.

4. Methodology

4.1. PCA-based index construction

The explanatory variables describing macroeconomic development, household income, fiscal conditions, and demographic characteristics are highly correlated. Including these variables individually in the regression model may lead to multicollinearity and unstable coefficient estimates. To summarize the common variation while preserving the underlying economic interpretation, this study constructs three composite indices using PCA. Because macroeconomic variables are measured at the city-year level, while the regression dataset is organized at city- year-industry panel, PCA is constructed using the unique city-year observations. The resulting component scores are then merged back into the full panel so that every industry observation within the same city-year shares the same macroeconomic index. For variables measured in monetary units or economic scale (e.g., GDP, income, deposits, tax revenue, wages), a logarithmic transformation is first applied to reduce skewness. Variables that are already rates or ratios (such as the birth rate) are not log-transformed. Since the explanatory variables are measured in different units, each transformed variable is standardized before PCA as Model (4):

$$Z_{k,ct} = \frac{\widetilde{X}_{k,ct} - \bar{\widetilde{X}}_k}{s_k} \tag{4}$$

Where $X_{k,ct}$ denotes variable k for city c in year t, $\tilde{X}_{k,ct}$ is the transformed value. The first principal-component from each conceptual block is retained and interpreted as a latent index. The Economic Opportunity Index is constructed from GDP, GDP per capita, resident population, household deposits, urban disposable income, and rural disposable income. It is defined as $EOI_{ct} = \sum_{k=1}^{6} w_k^{(E)} Z_{k,ct}$. The Operating Pressure Index summarizes average wage and tax revenue. It is computed as Model (5):

$$PI_{ct} = w_1^{(O)} Z_{\text{Wage},ct} + w_2^{(O)} Z_{\text{Tax},ct} \tag{5}$$

A larger value indicates a city-year characterized by relatively higher labor costs and fiscal burden. The Demographic Fundamentals Index combines the birth rate and the resident-to-registered population ratio. The index is $DFI_{ct} = w_1^{(D)} Z_{\text{Birth},ct} + w_2^{(D)} Z_{\text{Population Ratio},ct}$. Higher values indicate stronger demographic support and greater population attractiveness. To facilitate interpretation, each retained principal-component is standardized again as Model (6):

DOI: 10.54254/2754-1169/2026.38926

$$Index_{ct} = \frac{PC_{1,ct} - \bar{PC}_1}{sd(PC_1)} \quad (6)$$

Consequently, the regression coefficients can be interpreted as the change in the enterprise survival rate associated with a one-standard-deviation increase in the corresponding latent index.

4.2. Econometric specifications

The empirical model follows the idea from the theoretical framework. Two weighted least-squares specifications are estimated. The baseline model is as Model (7):

$$SR_{cijt} = \alpha + \beta_1 EOI_{ct} + \beta_2 OPI_{ct} + \beta_3 DFI_{ct} + \lambda_t + \gamma_j + \delta_i + \varepsilon_{cijt} \quad (7)$$

The preferred mode is as Model (8):

$$SR_{cijt} = \alpha + \beta_1 EOI_{ct} + \beta_2 OPI_{ct} + \beta_3 DFI_{ct} + \mu_c + \lambda_t + \gamma_j + \delta_i + \varepsilon_{cijt} \quad (8)$$

The baseline model includes year, industry, and enterprise fixed effects but omits city fixed effects. It combines cross-city differences with within-city changes over time. The preferred model adds city fixed effects and identifies the PCA coefficients from within-city changes after ruling out time-invariant local characteristics. Year effects λ_t absorb nationwide shocks and common trends; industry effects absorb persistent sector-specific differences; enterprise proper effects absorb differences across organizational forms. WLS models weight w= observations by total enterprise count. Due to the large difference in the number of enterprises included in different city-industry observations, this paper uses the total number of enterprises as the observation weight in the weighted least-squares model. When there are differences in the estimation accuracy or the overall scale represented by different observations, the weighting method can improve the applicability of estimation [9]. Standard errors are clustered by city to allow arbitrary heteroskedasticity and serial correlation within cities. Considering that there may be error correlations in the observations of different years in the same city, this paper clusters the standard errors at the city level to obtain a more robust statistical inference of the correlation in the group [10]. An unweighted OLS model with the full set of fixed effects is reported as a robustness check. Because the dependent variable is bounded, the linear specification is interpreted as a conditional mean approximation. The distinction between the baseline and preferred specification is central. In the baseline model, a coefficient may reflect persistent differences between high- and low-development cities. After city fixed effects are introduced, the coefficient captures whether a city's survival rate changes when its own macroeconomic index changes relative to its city-level average.

5. Empirical results

Table 3 shows the factor loadings of each observed variable on the first principal-component and the proportion of the interpretation of the variance of the principal-component. The results of the principal-component analysis effectively support the theoretical connotation of the three potential

DOI: 10.54254/2754-1169/2026.38926

variables proposed in this paper and verify the good construct validity of the scale. For the economic opportunity index group, six variables, such as GDP, per capita GDP, and disposable income of rural residents, all have positive loadings on PC1. The first principal-component explains 64.1% of the group's overall variation, and the loadings of each variable can be weighted to build a comprehensive indicator of economic opportunities. The business-pressure dimension includes two observed variables, average wage and tax income. Both factor loadings are 0.707, and the variance interpretation rate of PC1 is 60.5%. The two indicators exhibit highly co-directional changes, which can stably measure the level of regional business-pressure. The ratio load of the birth rate and the resident household registration population under the basic dimension of the population is also 0.707. The interpretation rate of PC1 variance in this dimension reaches 76.5%, the highest among the three groups of indicators. The two population indicators have strong inherent consistency, which can effectively characterize the fundamental situation of the regional population. The interpretation of the variance of the first principal-component of the three groups of indicators is higher than 60%, and the load coefficients of all variables are positive, which confirms that the three potential variables of economic opportunities, business-pressure, and population-base are statistically reasonable, and the load of each variable can be used as the calculation weight of the synthetic composite index.

Table 3. PCA loadings and variance explained

Index	Variable	Loading	PC1 variance explained
Economic opportunity	GDP	0.482	64.1%
Economic opportunity	GDP per capita	0.459	64.1%
Economic opportunity	Rural disposable income	0.444	64.1%
Economic opportunity	Household deposits	0.423	64.1%
Economic opportunity	Urban disposable income	0.398	64.1%
Economic opportunity	Resident population	0.220	64.1%
Operating pressure	Average wage	0.707	60.5%
Operating pressure	Tax revenue	0.707	60.5%
Demographic fundamentals	Birth rate	0.707	76.5%
Demographic fundamentals	Resident-to-registered population ratio	0.707	76.5%

Table 4 reports how the three indices affect enterprise survival rates across three model specifications. The baseline weighted least-squares (WLS) controls for year, industry and enterprise-ownership fixed effects. The preferred WLS further adds city fixed-effects to capture time-invariant city-specific characteristics, and unweighted OLS acts as a robustness check for weighting-driven estimation differences. Baseline WLS returns negative associations between all three indices and firm-survival. The Demographic Fundamentals index produces the most robust result (coefficient −0.0392, SE=0.0109, $p<0.001$): a one-standard-deviation increase correlates with a 3.9-percentage-point survival-rate reduction, suggesting larger demographic bases intensify competition, firm turnover and industrial restructuring. The economic-opportunity (−0.1225, $p=0.066$) and operating-pressure (−0.0232, $p=0.155$) indices show no conventional statistical significance. With city fixed-effects incorporated into preferred WLS, all three indices lose significance (−0.0180, $p=0.661$; −0.1465, $p=0.097$; −0.1560, $p=0.196$), as city fixed-effects absorb persistent local heterogeneity and widen confidence intervals. In the unweighted-OLS robustness test, only operating pressure remains significantly negative (−0.1447, $p=0.013$). Negative

DOI: 10.54254/2754-1169/2026.38926

demographic-fundamentals coefficients persist across models yet turn insignificant under city fixed-effects, indicating baseline correlations stem from cross-city structural disparities rather than within-city shifts. Result divergence underscores that weighting schemes substantially influence empirical inference, especially for operating-pressure estimates.

Table 4. PCA indices and enterprise survival

Variable	Specification	Coef. (SE)	95% CI	p-value
Economic opportunity index	Baseline WLS	-0.1225 (0.0666)	[-0.2531, 0.0081]	0.066
Operating pressure index	Baseline WLS	-0.0232 (0.0164)	[-0.0553, 0.0088]	0.155
Demographic fundamentals index	Baseline WLS	-0.0392 (0.0109)	[-0.0604, -0.0179]	0.000
Economic opportunity index	Preferred WLS	-0.1465 (0.0883)	[-0.3196, 0.0267]	0.097
Operating pressure index	Preferred WLS	-0.1560 (0.1206)	[-0.3925, 0.0804]	0.196
Demographic fundamentals index	Preferred WLS	-0.0180 (0.0411)	[-0.0985, 0.0625]	0.661
Economic opportunity index	Unweighted OLS robustness	0.0429 (0.0293)	[-0.0145, 0.1003]	0.143
Operating pressure index	Unweighted OLS robustness	-0.1447 (0.0581)	[-0.2585, -0.0309]	0.013
Demographic fundamentals index	Unweighted OLS robustness	-0.0160 (0.0141)	[-0.0436, 0.0117]	0.258

Table 5 summarizes the estimation specifications and overall model fit for the three regression models. For baseline WLS and Preferred WLS, the R- squared and adjusted R-squared are from 0.80081 to 0.8182, suggesting that these models fit the data well, which captures much of the systematic variation in enterprise survival. The unweighted OLS model tends to underfit the data, with a lower R-squared and adjusted R-squared. The large difference in R-squared between WLS and unweighted OLS partly reflects the weighting scheme and should not be interpreted as a direct model-quality comparison. The preferred WLS increases R-squared only modestly relative to the baseline WLS, from 0.8126 to 0.8182.

Table 5. Model specification and goodness of fit

Statistic	Baseline WLS	Preferred WLS	Unweighted OLS
City fixed effects	No	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes
Enterprise-property fixed effects	Yes	Yes	Yes
Observations	4478	4478	4478
R-squared	0.8126	0.8182	0.1609
Adjusted R-squared	0.8081	0.8129	0.1363

Figure 3 compares coefficient estimates for the three PCA indices from baseline WLS and city-fixed-effects preferred WLS, demonstrating shifts in estimated relationships as identification moves from cross-sectional to within-city temporal variation. The Demographic Fundamentals index exhibits the most distinct pattern. Its coefficient is negative and statistically significant in the baseline model, with a confidence interval entirely below zero. After introducing city fixed effects, the coefficient shrinks and the confidence interval broadens to include zero, losing significance. This indicates the baseline significant negative effect stems mainly from cross-city disparities instead of within-city changes: enterprise survival differs between cities of contrasting demographic

DOI: 10.54254/2754-1169/2026.38926

foundations, yet survival does not necessarily respond when a single city's demographic conditions evolve. The other two indices follow a similar pattern. Already insignificant in baseline WLS with confidence intervals crossing zero, the economic-opportunity and operating-pressure indices see slightly larger absolute coefficient values alongside widened confidence intervals under city fixed effects and remain non-significant. Overall, city fixed-effect inclusion widens confidence intervals and lowers estimation precision. This highlights the critical role of unobserved city-level attributes and the limited power of identification relying purely on within-city variation.

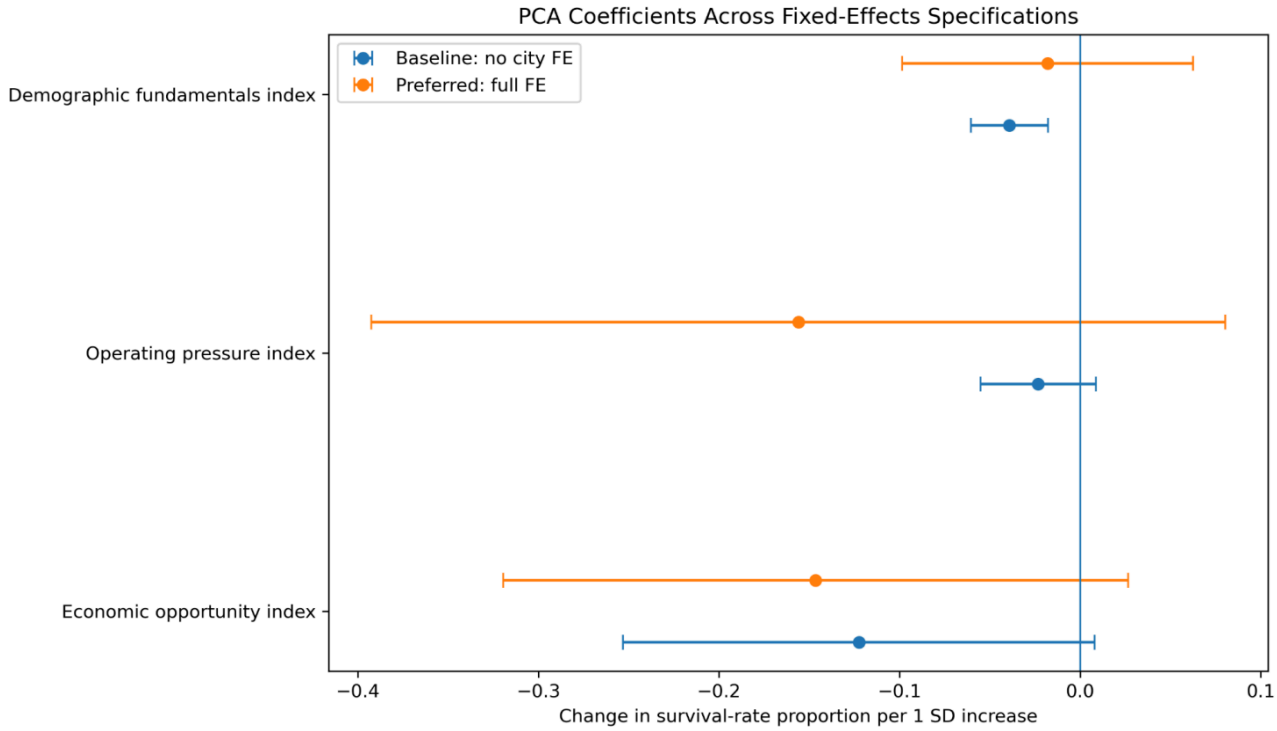


Figure 3. Comparison of PCA coefficients (picture credit: original)

Figure 4 shows the estimated city fixed effects in the preferred WLS model. The city fixed effects reveal persistent heterogeneity after observed macroeconomic conditions. Among the positive fixed-effect cities with relatively accurate estimates and confidence intervals that do not exceed zero, the urban fixed effects of Shanghai, Shenyang, Guangzhou, Zhuhai, and Liuzhou are all positive, indicating that the enterprise survival rate of these cities is higher than that of benchmark cities after controlling for the PCA index and other fixed effects. On the contrary, the urban fixed effects of Yiyang, Baishan, Meishan, Loudi, Anyang, Baoshan, and Leshan are negative, indicating that the survival probability of enterprises in these cities is lower than that of benchmark cities. These fixed effects summarize all time-invariant or slowly changing city characteristics omitted from the observed indices, like industrial history, local business culture, market access, and administrative capacity. Effects are relative to the omitted reference city and conditional on the PCA indices and the other fixed effects.

DOI: 10.54254/2754-1169/2026.38926

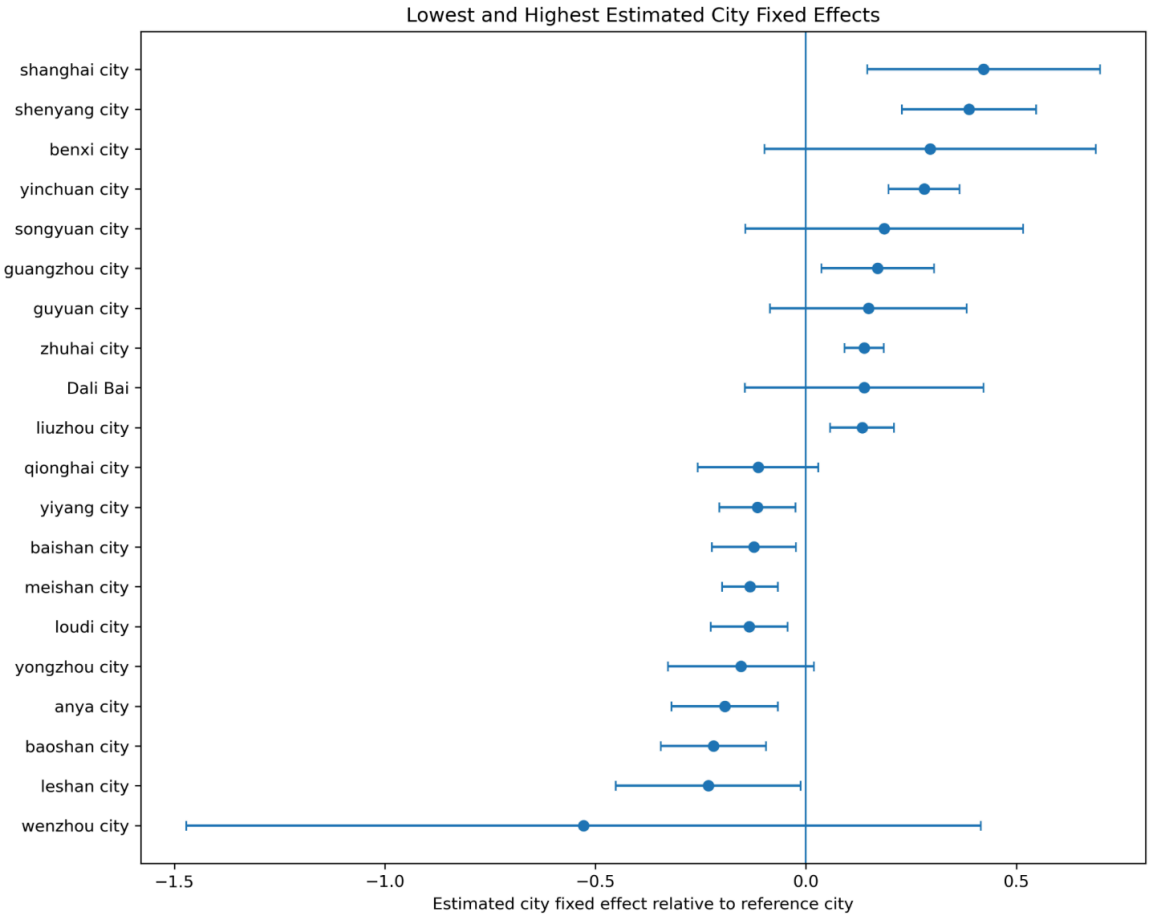


Figure 4. Lowest and highest estimated city fixed effects (picture credit: original)

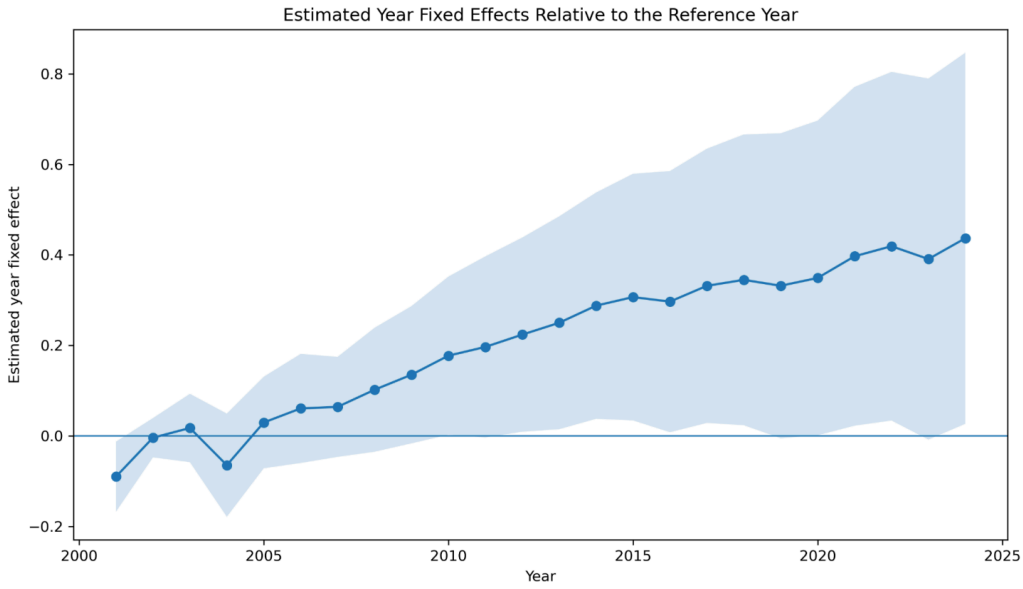


Figure 5. Estimated year fixed effects relative (picture credit: original)

Figure 5 shows the estimated year fixed effects relative to the reference year. The year effects show an upward trajectory after the mid-2000s. Relative to 2000, the effects of 2001 and 2004 are

DOI: 10.54254/2754-1169/2026.38926

negative, while most effects of 2009 are positive. This pattern indicates a broad nationwide increase in the conditional survival rate that is not explained by the city-level PCA indices, industry composition, or time-invariant city characteristics. This trend may reflect nationwide improvements in market conditions, digital infrastructure, business environment, and financial access.

6. Conclusion

This article investigates the correlation between urban external environments and enterprise survival based on China's urban-level enterprise data from 2000 to 2024, covering three core dimensions: economic opportunity, business-pressure, and population-base. The results reveal prominent urban disparities in enterprise survival across the country. The benchmark weighted regression confirms a significant negative correlation between the population-base index and enterprise survival rate, whereas the economic opportunity and business-pressure indices show no statistically significant influence. After incorporating urban fixed effects, all three comprehensive indices lose statistical significance, suggesting that the benchmark model's observed correlations largely stem from inherent urban structural differences, while intra-city temporal changes in economic and population conditions exert unstable impacts on enterprise survival. Robustness tests further verify that unweighted regression by enterprise size yields a notable negative link between business-pressure and enterprise survival, proving that sample weights and urban structural differences affect estimation outcomes. Overall, enterprise survival is shaped by local economic conditions, market competition, operating costs and urban development traits, requiring differentiated urban business environment optimization strategies instead of blind economic and population scale expansion.

This study's core contribution lies in constructing a multi-dimensional urban-level analytical framework integrating economic opportunity, business-pressure and population-base. It adopts principal-component analysis to synthesize highly correlated economic and demographic indicators, enabling a systematic examination of the urban external environment's impact on enterprise survival. The long-term, multi-city and multi-industry dataset effectively identifies regional heterogeneity in enterprise development, and comparative fixed effect models further distinguish cross-city long-term structural differences from intra-city temporal changes. The findings provide new empirical evidence for explaining China's urban disparities in enterprise survival and offer valuable references for research on urban business environments, enterprise exit behaviors and regional economic resilience. Nevertheless, this study has limitations. It relies on aggregated urban data, failing to observe individual enterprises' entry, exit and operation dynamics, and the sample presents regional and industrial concentration. Thus, the results only reflect conditional variable correlations rather than causal relationships. Future research can expand urban and industrial coverage, adopt long-term enterprise-level tracking data, and apply rigorous causal identification methods to clarify the underlying influencing mechanisms.

References

[1] Rocha, V., Carneiro, A., & Varum, C. A. (2015). Entry and exit dynamics of nascent business owners. *Small Business Economics, 45*(1), 63-84.

[2] Woldehanna, T., Amha, W., & Yonis, M. B. (2018). Correlates of business survival: Empirical evidence on youth-owned micro and small enterprises in urban Ethiopia. *IZA Journal of Development and Migration, 8*, 14.

[3] Liu, J. (2012). The dynamics of listed SMEs in China. *International Journal of the Economics of Business, 19*(3), 421-450.

[4] Jovanovic, B. (1982). Selection and the evolution of industry. *Econometrica, 50*(3), 649-670.

DOI: 10.54254/2754-1169/2026.38926

[5] Sedláček, P. (2020). Lost generations of firms and aggregate labor market dynamics. *Journal of Monetary Economics, 111*, 16-31.
[6] Audretsch, D. B., & Fritsch, M. (2002). Growth regimes over time and space. *Regional Studies, 36*(2), 113-124.
[7] Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long-run equilibrium. *Econometrica, 60*(5), 1127-1150.
[8] CnOpenData. (2026). Chinese enterprise entry and exit statistics. [DB/OL]
[9] Solon, G., Haider, S. J., & Wooldridge, J. M. (2015). What are we weighting for? *Journal of Human Resources, 50*(2), 301-316.
[10] Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources, 50*(2), 317-372.