Proceedings of the 4th International Conference on Management Research and Economic Development
DOI: 10.54254/2754-1169/2026.33772

The Clustering and Persistence of Foreign Direct Investment Inflows: The Acceleration and Cumulative Effects of Initial Entry in the Same City and Industry

Suyang Duan

Department of Social Science, The University of Manchester, Manchester, UK
3105900094@qq.com

Abstract. This paper examines the dynamics of foreign direct investment (FDI) at the city-industry level. We ask two questions. Does past investment make future investment in the same place more likely? And does it arrive faster as more projects accumulate? The first is a persistence effect—more investment attracts more investment. The second is an acceleration effect — the gaps between projects get shorter over time. To test these, we use panel count models for persistence. For acceleration, we turn to survival models — log-interval regression and Cox/AFT models. The results hold up across both sets of tests, even after we account for macroeconomic conditions, city-industry fixed effects, and clustered standard errors. The findings point to clear micro-level mechanisms behind FDI clustering. Together, they give local governments a clearer picture of when and provide evidence to judge whether in designing investment promotion strategies or deciding how to allocate resources across industrial parks.

Keywords: Foreign Direct Investment, Industrial Cluster, Path Dependence, Count Model, Survival Analysis

1. Introduction

Foreign Direct Investment (FDI) drives regional industrial upgrading and global value chain integration. It typically clusters spatially and temporally, showing two key dynamics: persistence (more investment attracts more investment) and acceleration (faster investment brings faster investment).

This study takes city-industry units to examine whether FDI entry triggers faster and larger subsequent investment. It uses project-level date data to build panel and interval data and aligns samples by relative time. Employing panel count models and survival analysis (Cox/AFT), it uses IRR and HR to measure effects.

This paper aims to address two specific research questions:

(1) After controlling for year fixed effects and introducing city × industry fixed effects, when necessary, can the number of projects in the previous year and the cumulative number of historical projects significantly predict the number of new projects in the current year? If their incidence rate

DOI: 10.54254/2754-1169/2026.33772

ratio (IRR) is significantly greater than 1, can this be used to confirm the persistence of "more investment attracts more investment?"

(2) In the project-level arrival sequence, as the cumulative number or value of projects rises, does the time interval between adjacent projects systematically shorten? Equivalently, does the hazard of the next project's arrival increase significantly, thereby demonstrating the acceleration of "faster investment brings faster investment"?

2. Literature review and theoretical framework

2.1. Literature review

Krugman revealed the self-reinforcing mechanism of industrial agglomeration, based on increasing returns and economic geography models [1]. From the perspective of locational advantages, Dunning future incorporated the factor endowments and institutional environment of host countries into the location choices of multinational enterprises [2]. They theoretically established the foundation of location-specific agglomeration advantages. Li summarized three major mechanisms: external economies of scale, information accumulation, and industrial supporting facilities, directly supporting the hypothesis of persistent FDI agglomeration [3]. Wang and Yin further used the system GMM method to prove that service sector FDI has significant foreign capital agglomeration effects and industrial agglomeration effects, directly responding to the dynamic feature of "more investment attracts more investment" [4]. Also, Wang and Yin systematically demonstrated the relationship between the entry duration of foreign-funded enterprises and the productivity of domestic enterprises, revealing the nonlinear change of FDI effects over time, which is consistent with the topics of "acceleration" and "entry interval" [5].

Besides, existing studies have methodological divergences. On the one hand, annual panel studies mostly focus on the cross-sectional or long-term differences in entry levels, with a relatively thin description of the arrival process. scholars systematically demonstrated the relationship between the entry duration of foreign-funded enterprises and the productivity of domestic enterprises at the micro level, revealed the nonlinear change of FDI effects over time, and explored the moderating role of rhythmic and irregular entry modes [6]. On the other hand, literature using event history or point process methods can describe the "arrival-after-arrival" mechanism but often lacks comparability and mutual verification with the count regression framework. Salamaga used event history analysis and the Cox proportional hazards model to study the persistence and exit patterns of FDI [7]. In addition, in specific empirical operations, by studying the ex-ante and ex-post effects of Chinese enterprises' outward foreign direct investment, involving the handling of the "starting point" effect, Baiardi, Gattai and Natale provide an empirical reference paradigm for the "first-year bias" problem [8]. Xie et al. extended this perspective to the spatial dimension of Chinese multinational enterprises. Collectively, these studies provide both theoretical and empirical support for the cumulative effects and acceleration dynamics discussed in this paper [9].

2.2. Mechanism overview: why does "once started" lead to more and faster investment

To shed light on why FDI tends to cluster — and why the clustering tends to accelerate over time — we identify four mechanisms.

(1) Less uncertainty, more information

Early movers matter. The first projects into a location create a demonstration effect. Later investors can look at what worked, adjust their expectations about the market, the regulatory

DOI: 10.54254/2754-1169/2026.33772

environment, and the supply chain. Entry barriers decline, more enterprises enter the market, and the gaps among new entrants gradually narrow.

(2) Improves supportive infrastructure and share fixed costs

As the local supply chain matures, transaction costs fall, coming with shorter cycle of site selection and construction. In addition, one-off investments by industrial parks and local governments are spread across more projects, thereby achieving economies of scale.

(3) Networks externalities and knowledge spillovers

The flow of tacit knowledge and talent among enterprises improves their operational performance after market entry. This raises the expected return of market entry and strengthens the sustainability of investment.

(4) Dynamic Allocation of Policies and Resources

Preferential policies and priority factor allocation for projects, that have been launched or are under negotiation, enable existing cumulative investment to better attract new projects.

2.3. Theoretical framework and testable hypotheses

To realise the above mechanisms, we define the arrival intensity of a city × industry unit (c, i) as:

$$\lambda_{ci}(t) = \mu_{ci}(t) + g\left(cum_{ci}(t^{-})\right) \tag{1}$$

where $\mu_{ci}(t)$ denotes the time-varying baseline intensity (macro/industry/year fixed effects, FE), and $cum_{ci}(t^{-})$ represents the cumulative number or value of projects up to the previous moment. Based on this, we propose the following hypotheses:

H1 Persistence

After controlling for year fixed effects and necessary city × industry fixed effects, the number of projects in the previous year or the cumulative projects up to the previous year has an incidence rate ratio (IRR) significantly greater than 1 for the number of new projects in the current year.

H2 Acceleration

A higher cumulative project level is associated with a shorter survival time for the next project (coefficient < 0 in log-interval regression), or a hazard ratio (HR) greater than 1 in the Cox model.

H3 Nonlinearity

The function $g(\cdot)$ may exhibit diminishing marginal returns or an inverted-U shape. Congestion inhibition emerges in the high cumulative interval.

H4 Seasonality

Month/quarter fixed effects show more concentrated arrivals at the end of the year or in Q4, aligned with the policy implementation and budget cycle.

H5 Heterogeneity

Technology/capital-intensive industries and source countries with smaller institutional distance are more likely to exhibit stronger persistence and acceleration.

3. Data and variables

This study uses data from the National Enterprise Information Publicity System, covering foreign-invested enterprises established in mainland China by foreign investors in accordance with Chinese law, as of August 2024. The raw dataset includes key fields such as enterprise name, establishment date, enterprise type, registered capital, registered province, and industry type.

DOI: 10.54254/2754-1169/2026.33772

3.1. Analysis unit and variable construction

We adopt the "city × industry × year" as the analysis unit. For each "city × industry" pair, we sort observations by year and calculate the number of projects in the current year, the number of projects in the previous year, and the cumulative number of projects up to the previous year. Each observation unit represents the foreign direct investment (FDI) activity in a specific city and industry during a given year.

Table 1. Main variables and descriptions

Variable Type	Variable Name	Variable Symbol	Variable Definition
Dependent Variables	Number of New Projects in Current Year	$New\ projects_{i,t}$	The number of new FDI projects in city-industry unit i in year t.
	Project Interval Time	$gap\ \_\ days$	The difference in days between two adjacent projects.
	Logarithm of Interval Time	$Log(gap_{days})$	The natural logarithm of the time interval, used for conventional linear interval regression.
Core Explanatory Variables	Number of Projects in Previous Year	$Last\ Yeat\ Projects_{i,t-1}$	The number of projects in city-industry unit i in year t-1.
	Cumulative Historical Projects	$\mathrm{Cum\ \_\ at\ \_\ prev}$	The total number of projects in the city-industry unit before the current project occurs.
	Logarithm of Cumulative Projects	$\mathrm{Log\ \_\ cum}$	Used to capture the diminishing marginal effect.
Robustness Explanatory Variables	Average Projects in Past 2 Years	$Avg\ 2Year\ Projects_{i,t}$	The average number of projects in city-industry unit i in years t-1 and t-2.
	Total Projects in Past 3 Years	$Sum\ 3Year\ Projects_{i,t}$	The total number of projects in city-industry unit i in years t-1, t-2 and t-3.
Control Variable	Cumulative Projects up to $t-2$	$\mathrm{cumProjectsToT2}_{i,t-2}$	The cumulative number of projects in city-industry unit i up to year t-2.
Fixed Effects	Year Fixed Effects	γ_t	Annual dummy variables, absorbing common shocks at the time level.
	City × Industry Fixed Effects	α_i	City-industry interactive fixed effects, absorbing time-invariant heterogeneity.

We adopt the city × industry × year as the observation unit and estimate a fixed-effects Poisson regression:

$$NewProjects_{i,t} = exp(\beta_1 * LastYearProjects_{i,t-1} + \beta_2 \tag{2}$$

DOI: 10.54254/2754-1169/2026.33772

$$*CumProjectsToT2_{i,t-2} + \gamma_t + \alpha_i + \varepsilon_{i,t})$$

It is critical to note the interpretation of the incidence rate ratio (IRR) in Poisson regression: the coefficient β is exponentiated to obtain the IRR.

$$IRR\,(LastYearProjects) = \exp(\beta_1) \tag{3}$$

$$IRR\,(CumProjectsToT2) = \exp(\beta_2) \tag{4}$$

If IRR > 1 and statistically significant: The variable has a positive impact on the number of new projects in the current year. That is, "the more active or accumulated in the past, the more active this year", indicating persistence and self-reinforcing effects.

If IRR < 1 and statistically significant: The variable has a negative impact. That is, "the more active or accumulated in the past, the less active this year".

To examine whether "past activity" and "historical accumulation" can significantly predict "current activity", we further adopt two models: interval regression and survival analysis (Cox and AFT models), as detailed below.

Interval regression

$$\ln(T) = \alpha + \beta \bullet cum\_at\_prev + e \tag{5}$$

Cox Proportional Hazards Model

$$h\,(tX) = h_0(t)exp(\beta \bullet ln(1 + cum\_at\_prev)) \tag{6}$$

C. Accelerated Failure Time model (AFT)

$$\ln(T) = \mu + \beta \bullet \ln(1 + cum_{at_{Prev}}) + \sigma\omega \tag{7}$$

4. Empirical methods and identification strategy

4.1. Panel count regression

To examine whether "past activity" and "historical accumulation" can significantly predict "current activity", we use fixed-effects Poisson regression to investigate the dynamic evolution of projects at the city-industry level.

In the model specification, we adopt fixed-effects controls. We strictly include dual fixed effects of Year FE and City-Industry FE to eliminate the interference of time-varying macro shocks and time-invariant regional-industry heterogeneity on the estimation results.

Meanwhile, to address the autocorrelation of observations within individual units, we adjust the standard errors by clustering at the city × industry level (Clustered Standard Errors), ensuring the robustness of statistical inference.

Table 2. Main poisson regression (IRR)

	Poisson (Main)
Last-year projects (lag1)	1.001***
	(<0.001)

DOI: 10.54254/2754-1169/2026.33772

Table 2. (continued)

Cumulative projects to t-2 (cum_to_t2)	1.000***
	(0.002)
Num.Obs.	56856
*p < 0.1, **p < 0.05, *** p < 0.01	

the incidence rate ratios (IRRs) of both the lagged 1-period last-year projects (lag1) and cumulative projects up to t-2 (cum_to_t2) are significantly greater than 1 at the 1% statistical level. This indicates that both short-term historical investment activity and long-term cumulative project stock have a significant positive predictive effect on the number of current new FDI projects.

4.2. Analysis of "interval time" for arrival

This section attempts to address the second research question: whether historical accumulation shortens the time interval between adjacent projects, leading to "faster investment brings faster investment". We verify this using two methods: interval regression and survival analysis (AFT and Cox models).

Table 3. Log-interval regression (exp coefficients)

	Log-interval (OLS on ln gap)
Cumulative projects (at previous event)	1.000***
	(<0.001)
Month: Feb	1.001
	(0.966)
Month: Mar	0.895***
	(<0.001)
Month: Apr	0.929***
	(<0.001)
Month: May	0.888***
	(<0.001)
Month: Jun	0.872***
	(<0.001)
Month: Jul	0.852***

* p < 0.1, ** p < 0.05, *** p < 0.01

Notes: Dependent variable is ln(inter-arrival days).

Coefficients are exponentiated: values < 1 imply shorter gaps; values > 1 imply longer gaps.

Robust SE clustered by cell; covariates measured at the previous event time.

DOI: 10.54254/2754-1169/2026.33772

Table 3. (continued)

	(<0.001)
Month: Aug	0.894***
	(<0.001)
Month: Sep	1.000
	(0.997)
Month: Oct	0.852***
	(<0.001)
Month: Nov	0.855***
	(<0.001)
Month: Dec	0.975
	(0.265)
Num.Obs.	185090

Based on the interpretation logic of the interval regression in Table 3, where "Coefficients are exponentiated: values<1 imply shorter gaps", the table shows strong seasonal fluctuations, with January as the baseline. The interval shortened notably from March to August, October and November, which witnessed the most rapid growth acceleration. From March to August, as well as in October and November, the time interval shortened significantly, representing the peak periods with the fastest growth acceleration. Compared with January, the speed of enterprise settlement showed no obvious difference in February, September and December. This phenomenon may be related to factors such as the Spring Festival holiday and the fiscal year cycle.

Table 4. Cox proportional hazards model (hazard ratios)

	Cox PH (hazard ratios)
Cumulative projects (at previous event)	1.000***
	(<0.001)
Month: Feb	1.028**
	(0.039)
Month: Mar	1.075***
	(<0.001)
Month: Apr	1.046***
	(<0.001)
Month: May	1.072***
	(<0.001)

DOI: 10.54254/2754-1169/2026.33772

Table 4. (continued)

Month: Jun	1.083***
	(<0.001)
Month: Jul	1.101***
	(<0.001)
Month: Aug	1.071***
	(<0.001)
Month: Sep	1.011
	(0.359)
Month: Oct	1.102***
	(<0.001)
Month: Nov	1.102***
	(<0.001)
Month: Dec	1.013
	(0.290)
Num.Obs.	185090

It should be noted that the hazard ratio (HR) from the Cox model is converted into a percentage reduction in waiting time, effectively translating "higher event probability" into "shorter intervals between project entries", as expressed in Equation (8):

$$Percentage\ Reduction\ in\ Waiting\ Time \approx (1 - 1/HR) \times 100 \tag{8}$$

The project entry process exhibits pronounced seasonality. Relative to January, hazard ratios for the second quarter (March–June) and early fourth quarter (October–November) are significantly greater than 1 at the 1% level. This confirms that cities and industries attract new projects more efficiently during these periods, supporting the hypothesis that "accumulation drives both growth and acceleration." also, the findings align with those from the interval regression analysis.

Figure 1, based on the Cox model, plots the dynamic probability of project entry under different levels of historical accumulation. It illustrates how the probability of the next project's launch evolves over time, across varying cumulative levels and time horizons (30, 50, 100, 150, and 200 days).

According to Figure 1, The analysis yields three key conclusions:

DOI: 10.54254/2754-1169/2026.33772

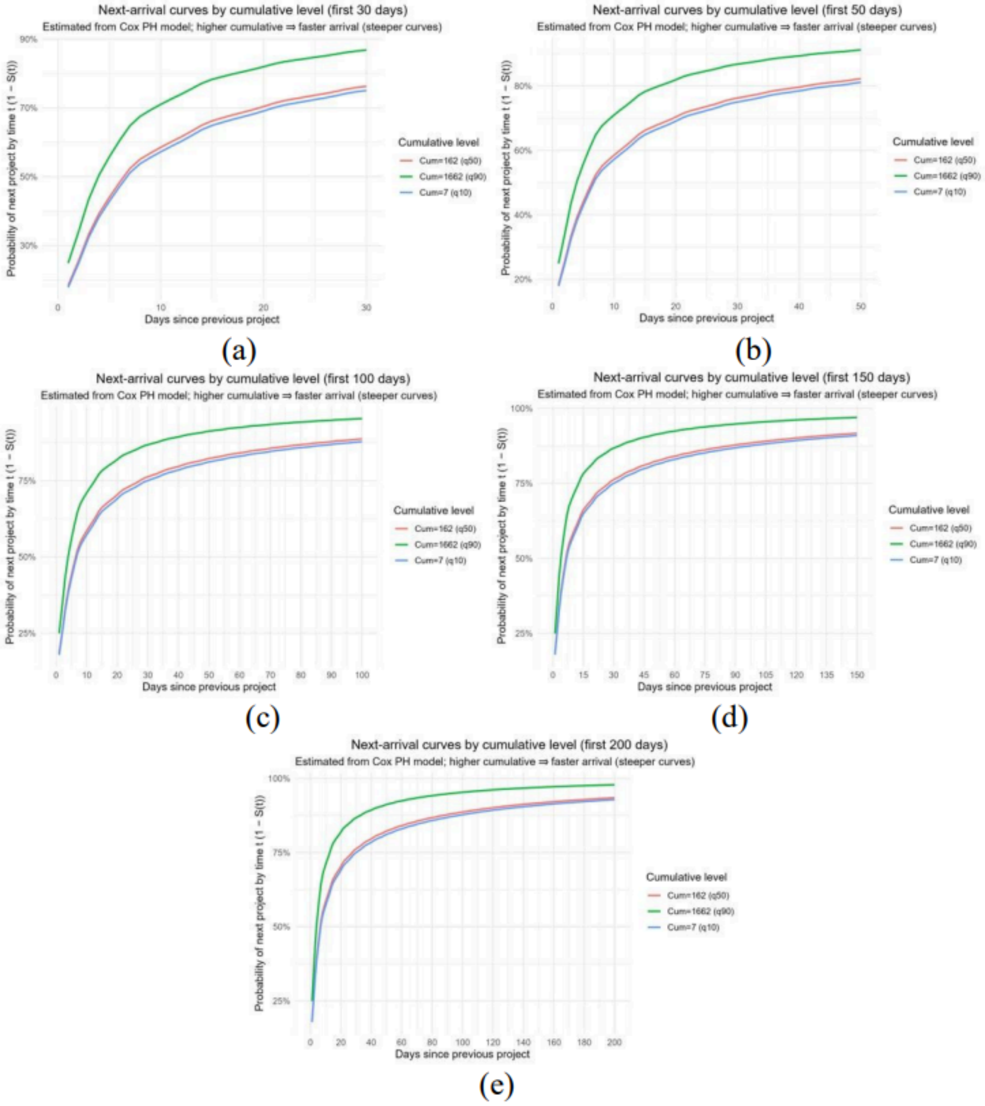

Figure 1. Next-arrival curves

First, cumulative investment exerts a significant acceleration effect. The green curve, representing the highest accumulation group, exhibits a steeper slope than the red and blue curves across all observation windows, indicating a higher likelihood of the next project being initiated.

Second, this effect is most pronounced in the short run. As shown in Panel (a), just 10 days after a project's completion, the predicted probability for the green group surges to 70%, while the low-accumulation blue group remains at approximately 55%. This "experience gap" is particularly stark within the first 30 days, reflecting the robust capacity of mature industrial ecosystems to generate new projects continuously in the short term.

Third, the figure reveals a nonlinear pattern of marginal acceleration. Observing the gaps between the curves across time points, we find that while all three groups show an upward trend, the probability increase from the medium (Cum=162) to the high accumulation group (Cum=1662) remains substantial. This suggests that even at very large scales, the acceleration effect still makes significant marginal contributions with no signs of stagnation.

4.3. Sensitivity and robustness checks

To ensure the reliability of the conclusions regarding the "acceleration effect" and "persistence," this study conducts a series of robustness checks. By controlling for city-industry interactive fixed effects, we eliminate static differences driven by region advantages and inherent industrial characteristics. The inclusion of year and month fixed effects further control disturbances from macroeconomic cycles and year-end statistical fluctuations.

Table 5. Robustness: NegBin, alternative lags, and subsample (IRR)

	NegBin (Main)	Poisson (2-year average)	Poisson (3-year sum)	Poisson (Subsample: first-seen year >= 2015)
Last-year projects (lag1)	1.002***			1.012
	(0.003)			(0.152)
2-year average (avg2)		1.001***		
		(<0.001)		
3-year sum (sum3)			1.000***	
			(0.009)	
Cumulative projects to t-2 (cum_to_t2)	1.000***	1.000***	1.000***	0.983***
	(0.004)	(<0.001)	(<0.001)	(0.002)
Num. Obs.	56856	56856	56856	24890

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Notes: Coefficients are reported as incidence rate ratios (IRR); values > 1 imply a positive association.
Fixed effects: city × industry (cell) and year. Standard errors clustered by cell.
Subsample column: only cells with first-seen year >= 2015.
Dependent variable: new projects in the current year.

The robustness regression results in Table 5 show that, regardless of adjusting the length of the time window (i.e., using the average number of projects over the past two years or the total over the past three years) or switching the model estimation method (i.e., using Quasi-Poisson instead of the negative binomial distribution to address overdispersion under high-dimensional fixed effects), the IRRs of the core explanatory variables remain stably above 1 and statistically significant at the 0.1% level.

5. Empirical results

5.1. Persistence of FDI entry

The findings show that past investment activity has strong predictive power for current new projects: IRR of the number of projects in the previous year (lag1) is significantly greater than 1 (1.001***). This indicates that after removing the inherent endowments of cities and industries as well as macro time trends, each successful investment in the previous year significantly raises the entry probability in the next year.

In addition, the cumulative stock of projects up to the previous year (cum_to_t2) also exhibits a significant positive pulling effect. This demonstrates that FDI agglomeration in specific regions and industries is not randomly distributed but shows clear path dependence.

DOI: 10.54254/2754-1169/2026.33772

5.2. Acceleration of FDI entry

Log-linear regression results show that the cumulative number of projects has a significant negative impact on the number of days between projects, directly proving that as the scale of experience expands, the time span between adjacent investments is continuously compressed.

The results of the Cox proportional hazards model confirm this finding: the hazard ratio (HR) corresponding to the logged cumulative projects is significantly greater than 1, meaning that for each additional unit of accumulated experience, the instantaneous probability that the city-industry unit will attract the next investment increases accordingly. The survival function curves in Figure 1 clearly illustrates this accelerating trend.

6. Identification, limitations and extrapolation

6.1. Identification boundaries and potential biases

We adopt three strategies to address the "first-year self-reinforcement bias": a "first-year +1" rule for sample inclusion, "relative-to-first year" alignment in descriptive and heterogeneity analyses, and first-year threshold-based subsampling to test robustness. While city-industry fixed effects absorb time-invariant.

In addition, while the fixed effects can absorb time-invariant endowments, unobservable factors such as firm-level strategic adjustments, industrial park governance quality, and knowledge spillovers may affect entry timing. Accordingly, robustness checks may include city- or industry-specific linear trends and report sensitivity analysis results under different model specifications.

6.2. Generalizability and applicable scenarios

From the industry perspective, technology-intensive or capital-intensive industries tend to exhibit a more pronounced self-reinforcing effect. In contrast, labor-intensive and pollution-sensitive industries are more strictly constrained by environmental regulations, energy consumption, and land prices, leading to stronger nonlinear and threshold effects. Therefore, spline functions or heterogeneity analysis should be introduced in econometric analysis to identify marginal changes across different intervals.

From the temporal perspective, global financial and exchange rate cycles shape cross-border capital allocation. During contractionary periods, the self-reinforcing effect may persist, but its marginal effect and realization speed will decline significantly. Conversely, during expansionary periods, if public infrastructure and factor supply expand accordingly, both the persistence and acceleration effects are more likely to be amplified.

7. Conclusion

This paper centers on the core proposition: once FDI enters the same city and industry, does it lead to faster and more investment? It proposes a unified mutual verification framework centered on count regression and interval/survival analysis, which verifies that FDI exhibits significant dynamic characteristics of persistence and acceleration.

DOI: 10.54254/2754-1169/2026.33772

References

[1] Krugman, P. (1991). Increasing returns and economic geography. Journal of Political Economy, 99(3), 483–499. https: //doi.org/10.1086/261763
[2] Dunning, J. H. (1998). Location and the multinational enterprise: A neglected factor? Journal of International Business Studies, 29(1), 45–66. https: //doi.org/10.1057/palgrave.jibs.8490024
[3] Li, Y. (2010). Research on China's competitive advantage enhancement strategy: On the advantages of industrial clusters in attracting foreign direct investment. Academic Circle, pp. 192–198+288.
[4] Wang, S., & Yin, F. (2021). The impact of agglomeration effects on FDI location choice in service industry: Re-decomposition and measurement from the industrial dimension. World Economy Studies, (12), 103–115+134. https: //doi.org/10.13516/j.cnki.wes.2021.12.008
[5] Zhang, Y., Li, Y., & Li, H. (2014). FDI spillovers over time in an emerging market: The roles of entry tenure and barriers to imitation. Academy of Management Journal, 57(3), 698–722. https: //doi.org/10.5465/amj.2011.0351
[6] Li, S. H., & Li, Z. (2025). Information gap synergy and FDI location choice. Journal of Finance and Economics, 46(09), 124–140. https: //doi.org/10.19795/j.cnki.cn11-1166/f.20250915.002
[7] Salamaga, M. (2023). Study on the survival rate of foreign direct investments in the countries of the Visegrad Group using event history analysis. Argumenta Oeconomica, 2023(2), 83–99. https: //doi.org/10.15611/aoe.2023.2.04
[8] Baiardi, D., Gattai, V., & Natale, P. (2021). Estimating the ex-ante and the ex-post effects of Chinese outward FDI. The World Economy, 44(12), 3641–3673. https: //doi.org/10.1111/twec.13161
[9] Xie, Y., et al. (2024). Spatial outward FDI: Evidence from China's multinational firms. Review of International Economics, 32(2), 574–603. https: //doi.org/10.1111/roie.12698