International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 09 Issue 05 May 2026
DOI: 10.47191/ijcsrr/V9-i5-04, Impact Factor: 8.048
IJCSRR @ 2026
www.ijcsrr.org

Explainable AI for Foreign Direct Investment Analysis: Evidence from Central Asia

Khlyostkina Xeniya[1], Diping Zhang[2]
[1]Master's Student, Applied Statistics, Zhejiang University of Science and Technology, Hangzhou, China
[2]Zhejiang University of Science and Technology, Hangzhou, China

ABSTRACT: Foreign direct investment (FDI) is an important factor in the economic development of Central Asian countries, where investment flows have traditionally been concentrated in resource-based sectors. In the context of a growing focus on diversification, the need to analyze and study the determinants of FDI is increasing.
This study examines the determinants of FDI inflows in Central Asian countries using machine learning methods (CatBoost) and explainable artificial intelligence (SHAP), and compares the results with a classical econometric approach based on a two-way fixed effects (TWFE) model. Given the limited availability of data, a transfer learning approach is applied: the model is first trained on a group of countries structurally similar to Central Asia and then fine-tuned on the regional sample.
The results show that key macroeconomic factors such as Trade (% of GDP), Current account balance (% of GDP), and several other macroeconomic variables remain significant across both methodologies. At the same time, ML identifies additional regional patterns, such as a higher importance for FDI of determinants including Adjusted savings: carbon dioxide damage (% of GNI), Urban population (% of total population), and Access to electricity (% of population), among others.
The findings indicate that XAI provides interpretable results that are consistent with classical methods and additionally allows for capturing nonlinearities and regional heterogeneity. The study extends the application of ML and XAI in data-constrained Central Asian settings and demonstrates the value of combining econometric and machine learning approaches in the analysis of FDI determinants.

KEYWORDS: Foreign Direct Investment (FDI), Investment Determinants, Machine Learning, Explainable Artificial Intelligence (XAI), SHAP, CatBoost, Central Asia

I. INTRODUCTION

Foreign direct investment (FDI) plays an important role in the economic development of Central Asian countries. Traditionally, investment flows in the region have been concentrated in sectors related to agriculture and natural resource extraction. In recent years, however, Central Asian countries have increasingly set the goal of attracting investment into more diversified and innovation-oriented sectors.

Designing effective policies in this context requires not only updating the regulatory framework, but also improving the analytical tools used to understand where investment flows are directed, why they concentrate in certain sectors, and how they can be reoriented toward new development priorities.

The use of explainable artificial intelligence (XAI) represents a promising complement to established classical econometric methods. This approach has the potential to identify patterns and relationships that are difficult to detect using traditional models, while at the same time providing interpretable information that economists, analysts, and decision-makers can use in practice.

Despite the extensive literature on FDI in Central Asian countries, applications of XAI in this area remain limited.

In this context, the present study aims to adapt existing XAI frameworks to the analysis of FDI in Central Asia and to empirically compare this approach with classical econometric methods commonly used to examine the determinants of FDI inflows.

II. LITERATURE REVIEW

Foreign direct investment (FDI) has long been and continues to be one of the key subjects of economic research. This sustained interest is largely explained by the significant role of FDI as a driver of economic growth and development. Given its importance for both national economies and global economic integration, the study of FDI and its determinants remains highly relevant in both academic and policy-oriented research (Kumari, 2014; Kehal, 2004; Udeh et al., 2025).

While GDP growth, market size, inflation, and trade openness are widely recognized as key drivers, their effects vary across countries and over time, especially under economic and political instability. This limits the explanatory power of traditional approaches.

To address this, recent research increasingly applies machine learning (ML) and explainable artificial intelligence (XAI) methods to FDI analysis, focusing on predictive performance and interpretability.

A series of studies by Devesh Singh et al. examine FDI determinants using ML and XAI. Singh and Turała (2022) show that elastic net outperforms other regression models in explaining regional FDI in Hungary. Singh (2023) demonstrates that artificial neural networks outperform linear models in forecasting FDI and firm growth. Singh (2021) applies LIME to improve interpretability, finding gradient boosting to be the most effective model. Another study by Singh (2023) combines interpretable ML with econometric techniques, highlighting differences between important and trusted determinants and confirming the superior predictive performance of ML models.

Alon et al. (2022) develop a greenfield FDI attractiveness index using factor analysis and automated ML, identifying openness and economic growth as the key drivers. Kemives et al. (2024) propose an ensemble ML model that outperforms individual algorithms and identifies key determinants of FDI.

Overall, ML-based approaches consistently outperform traditional econometric models and allow for capturing complex, non-linear relationships, while XAI methods improve interpretability and policy relevance. However, the literature remains regionally imbalanced. While developed economies, Western Europe, and parts of Asia are well studied, Central Asia remains underexplored and is typically analyzed using conventional econometric methods with limited data.

This study addresses this gap by applying ML and XAI to analyze FDI determinants in Central Asia, contributing both methodologically and empirically.

III. DATA OVERVIEW

This study employs panel data obtained from the World Development Indicators (WDI) database published by the World Bank. This data source is chosen for several reasons. First, WDI provides a broad set of harmonized macroeconomic, institutional, and socio-economic indicators across countries and over time, which is essential for panel-based analysis of foreign direct investment (FDI) determinants. Second, the database is widely used in empirical research and undergoes systematic validation, ensuring data reliability and comparability with existing studies (Prince & Fantom, 2014).

The initial country sample includes Central Asian economies as well as a broader group of emerging and developing countries. The inclusion of non-Central Asian economies is motivated by the need to increase cross-country variation and to benchmark Central Asian countries against economies at comparable stages of development. Country selection follows the Emerging and Developing Economies classification adopted by the International Monetary Fund, thereby avoiding arbitrary sample construction and ensuring conceptual consistency.

Raw data were transformed into a panel structure, where each observation corresponds to a country - year pair and columns represent individual indicators. This format is standard for both fixed-effects econometric models and machine-learning algorithms applied to tabular data.

Substantial attention was devoted to data quality and missing-value treatment. At an initial stage, countries with extremely poor data coverage were excluded, based on the average proportion of missing values across numerical indicators. These countries were typically small economies or countries with irregular statistical reporting. In addition, Turkmenistan was excluded from the sample due to pervasive data limitations that prevented reliable estimation in panel settings.

For the remaining countries, missing values in explanatory variables were addressed using a combination of forward filling along the time dimension within each country and subsequent imputation using country-specific means. This approach preserves within-country dynamics while minimizing sample loss, which is particularly important for fixed-effects models that cannot accommodate missing values. The data were processed in this manner to ensure a unified and comparable input structure for both econometric and machine-learning methods.

Multicollinearity was further examined, and variables exhibiting high pairwise correlations were removed to avoid unstable coefficient estimates and redundant information. The final set of explanatory variables includes both conventional FDI determinants widely discussed in the literature and a deliberately expanded set of socio-economic and institutional indicators. This design enables subsequent evaluation of whether explainable artificial intelligence (XAI) techniques can identify economically meaningful drivers while appropriately down-weighting less relevant features.

Extreme values of FDI inflows, including observations exceeding 100% of GDP in specific years, were retained, as they reflect genuine economic events rather than measurement errors. Observations with zero values of the dependent variable were removed to ensure compatibility across estimation techniques and to maintain a consistent sample for fixed-effects and machine-learning models.

As a result of these preprocessing steps, the final dataset consists of 87 countries and 28 explanatory variables, fully cleaned and free of missing values, and is suitable for subsequent analysis using both fixed-effects econometric models and machine-learning approaches. Summary information on the processed dataset and the set of features is presented in Table 1. Additional data transformations and subsample constructions will be implemented at later stages depending on the specific requirements of each method.

Table 1. Dataset summary

Item	Description
Data source	World Bank, World Development Indicators
Country group	Emerging and Developing Economies, IMF classification
Number of countries	87
Central Asian countries	4
Time period	2000 - 2024
Number of observations	2,172
Dependent variable (Target)	Foreign direct investment, net inflows (% of GDP)
Number of explanatory variables	26

IV. METHODOLOGY

4.1 TWFE as an econometric baseline

As an econometric benchmark, this study employs a two-way fixed effects (TWFE) panel regression model. This approach is widely used in empirical research with panel data and allows for controlling unobserved, time-invariant country characteristics as well as common time shocks affecting all countries simultaneously (Woodbridge, 2002).

The estimated model can be written as:

$$FDI_{it} = \beta^{\top} X_{it} + \alpha_i + \gamma_t + \varepsilon_{it} \tag{1}$$

where FDI_{it} denotes foreign direct investment inflows to country *i* in year *t* (as a percentage of GDP), X_{it} is a vector of explanatory variables capturing macroeconomic, institutional, and structural characteristics, α_i represents country fixed effects accounting for time-invariant features (such as geography or historically determined institutions), γ_t denotes time fixed effects capturing global shocks and common trends, and ε_{it} is the error term.

The use of TWFE in this study is motivated by its ability to provide transparent and interpretable linear estimates of the determinants of FDI while rigorously controlling for both country and time heterogeneity. This makes it a natural point of reference for comparison with machine learning models and Explainable AI (XAI) methods, which are capable of capturing more complex and potentially nonlinear relationships. In this sense, TWFE serves as a classical econometric baseline reflecting a "traditional" perspective on FDI determinants.

A TWFE specification was also estimated using only the Central Asian countries. However, due to the limited number of observations and the relatively large set of explanatory variables, the resulting estimates proved to be statistically unstable and difficult to interpret. This is a well-known issue in TWFE models when the number of cross-sectional units is small relative to the number of regressors: degrees of freedom are reduced, multicollinearity becomes more severe, and the variance of coefficient estimates increases. As a result, standard errors are inflated and statistical inference becomes unreliable.

Therefore, the primary econometric benchmark in this study is based on the TWFE model estimated on the full panel of countries, including both Central Asian and other economies. In this broader sample, the model yields more stable and interpretable results. These estimates are subsequently compared with the findings obtained from Explainable AI methods in the results section of the paper.

4.2 Transfer learning and Gradient boosting

A key methodological challenge of this study is the limited amount of data available for Central Asian countries. The small number of countries and observations makes it difficult to train stable machine learning models directly on the regional subsample. To address this issue, a transfer learning approach was adopted, allowing information from a broader set of countries to improve model performance in the target region.

Although transfer learning is often associated with neural networks (Pan & Yang, 2010), modern gradient boosting decision tree algorithms also support staged learning: a model can first be pretrained on a large dataset and then fine-tuned on a smaller target sample. Neural networks were not used as the main modeling tool, as they require substantially larger datasets for stable training and are less interpretable in small-sample macroeconomic settings. In the preliminary stage, three algorithms from the gradient boosting family were evaluated - XGBoost, LightGBM, and CatBoost. All three implement the same core idea of sequentially adding tree (Friedman, 2001):

$$F_m(x) = F_{m-1}(x) + \eta h_m(x) \tag{2}$$

where $F_m(x)$ is the model after m boosting iterations, $h_m(x)$ is a new tree fitted to the gradient of the loss function, and η is the learning rate.

Preliminary experiments indicated that CatBoost provided the most stable and reliable performance in the transfer learning setup, and therefore subsequent analysis focused on CatBoostRegressor. The choice of gradient boosting models is motivated by several considerations.

First, such models effectively capture nonlinear relationships and feature interactions, which are typical for FDI determinants. Second, they are relatively robust to multicollinearity among macroeconomic indicators, since predictions are based on sequential feature splits rather than a single linear combination. Third, they are fully compatible with theoretically grounded and computationally stable explainability methods such as SHAP, making them particularly suitable for Explainable AI applications.

Macroeconomic and institutional indicators were used as features, while the variable year was deliberately excluded. Year is treated as a temporal index capturing common trends and global shocks rather than a substantive economic determinant of FDI. Including it as a feature would lead XAI methods to interpret it as an explanatory driver, which would be methodologically misleading. To control for persistent cross-country heterogeneity, country dummy variables were included as an analogue of country fixed effects.

Model evaluation for Central Asian countries followed a Leave-One-Country-Out (LOCO) protocol, a group-based variant of cross-validation (Shao & Er, 2016). In each fold, one country was completely excluded from training and used as the test country, ensuring that performance reflects the model's ability to generalize to unseen national contexts. Transfer learning was implemented in two stages. First, the model was pretrained on countries outside Central Asia (source domain). Second, the pretrained model was fine-tuned on Central Asian countries (target domain), excluding the held-out country in each LOCO fold. Early stopping was applied in both stages to prevent overfitting. Performance was evaluated using *RMSE, MAE,* and R^2 based on out-of-fold predictions across all Central Asian observations. A control experiment without pretraining (training only on Central Asian countries) yielded substantially worse performance, confirming the benefit of transferring knowledge from a broader international panel.

The comparison between models with and without transfer learning reveals substantial differences in predictive performance for Central Asian countries. The transfer learning model achieved a positive coefficient of determination ($R^2 \approx 0.20$) along with lower

RMSE and *MAE* values. In contrast, the CA-only model produced a negative R^2, indicating that its predictions were, on average, worse than simply using the sample mean.
Country-level analysis under the LOCO scheme further shows that the CA-only model exhibits strong instability, with highly negative R^2 values for some countries, reflecting overfitting and poor generalization in a small-sample setting. The transfer learning model, in contrast, delivers more stable and balanced performance across all countries in the region.
Hyperparameter tuning was also found to improve the performance of the pretrained source model, enhancing its ability to capture generalizable patterns before the fine-tuning stage. However, similar tuning did not yield comparable improvements for the CA-only model. This suggests that, under extremely limited data, performance constraints are driven primarily by data scarcity rather than model configuration. Overall, these findings demonstrate that knowledge transfer from a broader international panel is crucial for improving model performance and enables more reliable analysis of FDI determinants in data-scarce regional contexts.
Predictive results (Table 2) demonstrate that transfer learning substantially improves model performance for Central Asian economies. The CatBoost model pretrained on non-Central Asian countries and subsequently fine-tuned on the regional sample achieves RMSE = 3.50 and a positive coefficient of determination ($R^2 = 0.20$), whereas a model trained solely on Central Asian data performs considerably worse (RMSE = 4.24, $R^2 = -0.17$).
Given the extremely limited regional sample (four countries), the negative R^2 obtained by the CA-only model indicates insufficient data for stable standalone learning. Transfer learning therefore serves not only as a performance enhancement strategy but also as a mechanism for stabilizing inference under small-sample conditions.

Table 2. Predictive performance: Transfer learning vs CA-only training

Model	Training Strategy	RMSE	MAE	R²
CatBoost + Transfer Learning	Pretrain (non-CA) → Fine-tune (CA, LOCO)	3.50	2.52	0.20
CatBoost (CA-only)	Train only on CA (LOCO)	4.24	3.29	− 0.17

4.3 SHAP
Explainable Artificial Intelligence (XAI) has emerged as an important methodological framework aimed at improving the interpretability of complex machine learning models while preserving their predictive performance. Although modern algorithms such as gradient boosting machines demonstrate strong accuracy in economic prediction tasks, including the analysis of foreign direct investment (FDI), their nonlinear structure makes direct interpretation difficult compared to traditional econometric models. This limitation is particularly relevant in empirical economic research, where understanding the contribution of individual determinants is as important as prediction accuracy. To address this issue, this study employs SHapley Additive Explanations (SHAP), a theoretically grounded XAI method based on cooperative game theory (Shapley, 1988; Lundberg & Lee, 2017).
Consider a prediction model defined as $f(x) = f(x_1, x_2, \dots, x_m)$, where x denotes an observation vector and x_i represents the i-th feature among MMM total explanatory variables. SHAP interprets the prediction as the outcome of a cooperative game in which each feature acts as a player contributing to the final prediction. The contribution of feature $i$is quantified using the Shapley value

$$\phi_i = \sum_{s \subseteq N \setminus \{i\}} \frac{|S|!(M-|S|-1)!}{M!} [f(S \cup \{i\}) - f(S)], \tag{3}$$

where N denotes the set of all features, S is any subset of features excluding feature i, $|S|$ is the number of elements in subset S, and $f(S)$ represents the model prediction obtained when only features in subset S are included. The weighting term $\frac{|S|!(M-|S|-1)!}{M!}$ ensures that the marginal contribution of each feature is averaged across all possible feature orderings, thereby providing a fair allocation of predictive importance. SHAP belongs to the class of additive feature attribution methods, in which each prediction can be decomposed as

$$f(x) = \phi_0 + \sum_{i=1}^{M} \phi_i, \tag{4}$$

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 09 Issue 05 May 2026
DOI: 10.47191/ijcsrr/V9-i5-04, Impact Factor: 8.048
IJCSRR @ 2026
www.ijcsrr.org

where $\phi_i = E[f(x)]$ denotes the expected model prediction (baseline value) and $\phi_i$represents the contribution of feature i to the deviation of the prediction from this baseline. This additive representation allows SHAP values to be interpreted analogously to regression coefficients while retaining nonlinear interactions captured by machine learning models. In empirical applications, global feature importance is commonly summarized using the mean absolute SHAP value,

$$Importance_i = \frac{1}{N}\sum_{j=1}^{N} |\phi_{ij}|, \quad (5)$$

where N denotes the number of observations and ϕ_{ij} represents the SHAP value of feature i for observation j.

In this study, SHAP is computed for (i) a global CatBoost model trained on the full panel for comparison with TWFE estimates and (ii) a transfer-learning CatBoost model fine-tuned on Central Asian economies to identify region-specific determinants.

Although SHAP values and coefficients obtained from the two-way fixed effects (TWFE) model measure different quantities, their comparison is theoretically justified. In a linear model

$$f(x) = \beta_0 + \sum_i \beta_i x_i \quad (6)$$

SHAP values reduce to

$$\phi_i(x) = \beta_i(x_i - E[X_i]) \quad (7)$$

SHAP explanations represent a local decomposition of prediction effects analogous to regression contributions (Lundberg and Lee, 2017).

Consequently, regression coefficients capture average marginal effects under linear assumptions, whereas SHAP values quantify feature contributions while allowing nonlinearities and interactions. Aggregating SHAP values across observations using the mean absolute contribution therefore provides a model-agnostic measure of determinant importance that can be meaningfully compared with econometric estimates.

In this study, such comparison enables evaluation of whether machine learning and traditional TWFE models identify similar economic drivers of FDI inflows despite differing functional assumptions.

V. RESULTS

5.1 Baseline econometric results (TWFE)

As an econometric benchmark, a Two-Way Fixed Effects (TWFE) panel model with country-clustered standard errors is estimated to account for time-invariant country characteristics and common time shocks. The model is estimated on the full panel comprising 2,172 observations across 87 countries over 25 years (Table 3). Overall, the model demonstrates moderate explanatory power for within-country variation and jointly significant regressors.

When the estimation is restricted to Central Asian economies (100 observations across 4 countries), the TWFE results become unstable and diverge from the full-sample findings, including a sign reversal for gross fixed capital formation and the loss of significance for several previously significant variables. This reflects the limited statistical support of a small regional subsample with few cross-sectional units and low residual degrees of freedom.

Table 3. Summary diagnostics of the TWFE benchmark models

Statistic	Full Sample	Central Asia
Observations	2,172	100
Countries	87	4
Time period (years)	25	25
R-squared	0.1068	0.7383
R-squared (Within)	0.1151	− 3.5601

F-statistic (robust)	27.450	3.327e+15
P-value	0.000	0.000
Distribution	F(26,2035)	F(27,45)
Fixed effects	Country & Year	Country & Year
Standard errors	Clustered by country	Clustered by country

In the full sample, statistically significant associations with FDI inflows are mainly observed for indicators of external openness and investment activity, including trade openness, trade in services, gross fixed capital formation, and the current account balance, while most other variables are not robustly significant. More details about top determinants are provided in Table 4.

Table 4. Top determinants of FDI inflows (TWFE full sample)

Variable	Estimated effect on FDI inflows (% of GDP)	SE	p-value
Trade (% of GDP)	0.030	0.010	0.002
Trade in services (% of GDP)	0.098	0.035	0.005
Gross fixed capital formation (% of GDP)	0.186	0.067	0.005
Current account balance (% of GDP)	− 0.083	0.039	0.034
Life expectancy at birth (years)	− 0.096	0.050	0.057
Inflation, consumer prices (annual %)	0.009	0.006	0.169
Individuals using the Internet (% of population)	− 0.025	0.021	0.235
CO_2 emissions per capita	0.322	0.272	0.236
GDP per capita (current US$)	− 0.00007	0.00006	0.258
Gross domestic savings (% of GDP)	− 0.019	0.017	0.268

Although Central Asia represents the primary regional focus of this study, TWFE estimation based solely on this subsample does not yield sufficiently reliable or interpretable results. Therefore, the identification of economically meaningful FDI determinants within the econometric framework relies mainly on the broader comparative panel of structurally comparable emerging economies. While these countries were selected to approximate the macroeconomic and institutional characteristics of Central Asian economies, the TWFE estimates should be interpreted as capturing general patterns in the extended sample rather than purely region-specific effects. This limitation is taken into account in the subsequent analysis.

5.2 Global comparison of TWFE and SHAP results

In this section, we report intermediate findings obtained from a comparative assessment of econometric and machine-learning evidence based on the full panel dataset. Although such comparison does not constitute the primary objective of the study, it represents an important methodological step that helps evaluate whether explainable machine learning provides results broadly consistent with the baseline econometric framework. As demonstrated in the previous section, the TWFE specification applied exclusively to the Central Asian subsample produces unstable and economically difficult-to-interpret estimates due to the limited number of observations and restricted cross-sectional variation. For this reason, both approaches are first examined on the extended global panel to ensure comparability of empirical evidence.

The SHAP results for the global model identify current account balance, trade openness, trade in services, and gross fixed capital formation as the most influential predictors of FDI inflows. These variables largely coincide with those found statistically significant in the TWFE estimates, indicating a meaningful degree of convergence between structural econometric interpretation and predictive feature importance. In particular, trade openness and gross fixed capital formation exhibit strong positive contributions in both approaches, while current account balance consistently demonstrates a negative association with FDI inflows.

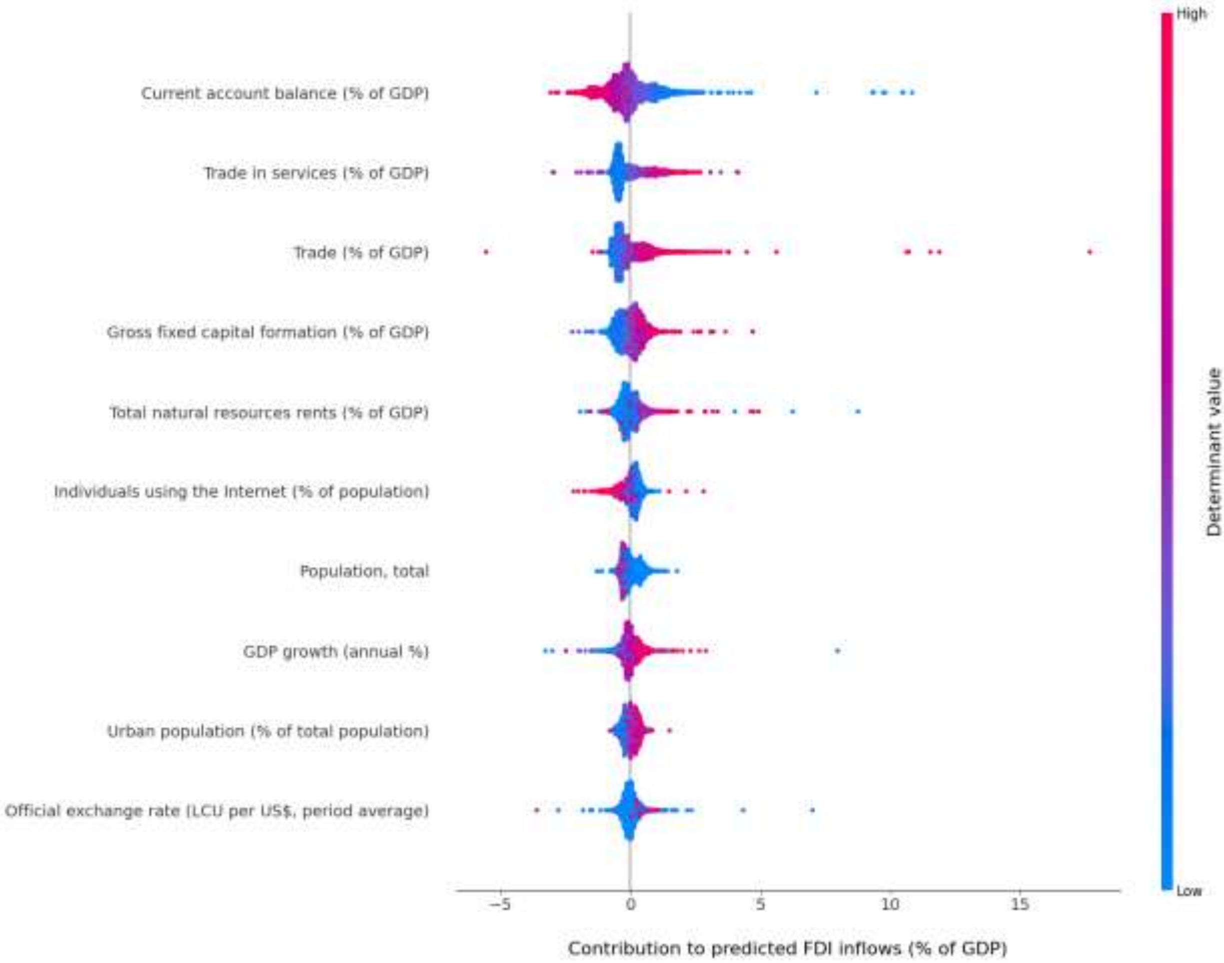

Figure 1. SHAP summary plot, full sample

Figure 1 presents the SHAP summary plot for the global model. The vertical ordering reflects the relative importance of determinants based on the average absolute contribution of each variable to model predictions. The horizontal dispersion of observations represents the magnitude and direction of individual contributions to predicted FDI inflows. Points located to the right indicate positive contributions to predicted FDI, whereas points on the left indicate negative contributions. The colour gradient reflects the value of the determinant in each observation, with higher values shown in red and lower values in blue. This visualisation highlights both the overall strength of key determinants and the heterogeneity of their effects across countries and time.

At the same time, the SHAP ranking assigns relatively high predictive importance to several variables that are not statistically significant in the econometric specification, including natural resource rents, population size, GDP growth, urbanisation, exchange

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 09 Issue 05 May 2026
DOI: 10.47191/ijcsrr/V9-i5-04, Impact Factor: 8.048
IJCSRR @ 2026
www.ijcsrr.org

rate dynamics, and Internet penetration. This pattern suggests that the machine learning model captures additional predictive structure in the data, potentially reflecting non-linear relationships or interaction effects that are not fully represented within the linear fixed-effects framework. Furthermore, certain discrepancies in directional effects - most notably for trade in services - indicate that average linear coefficients and model-based contribution measures may reflect different dimensions of the underlying economic relationships.

A more detailed comparison between TWFE estimates and SHAP-based importance rankings is summarised in Table 5, which presents the key determinants jointly identified or differently evaluated by the two approaches.

Overall, the comparison shows that explainable machine learning tends to reinforce rather than overturn the main econometric findings, while simultaneously providing complementary insights into the relative predictive relevance of additional macroeconomic factors. These intermediate results support the methodological validity of combining econometric and machine-learning approaches in the analysis of FDI determinants. A more focused investigation of model behaviour and determinant importance within the specific context of Central Asian economies is presented in the following section.

Table 5. Comparison of key FDI determinants identified by TWFE and SHAP (full sample)

Determinant	TWFE coefficient	TWFE p-value	SHAP importance rank	SHAP direction
Panel A. Determinants significant in TWFE				
Trade (% of GDP)	0.030	0.002	3	Positive
Trade in services (% of GDP)	0.098	0.005	2	Negative
Gross fixed capital formation (% of GDP)	0.186	0.005	4	Positive
Current account balance (% of GDP)	− 0.083	0.034	1	Negative
Life expectancy at birth, total (years)	− 0.096	0.057	-	-
Panel B. Additional predictors highlighted by SHAP				
Total natural resources rents (% of GDP)	0.002	0.968	5	Negative
Individuals using the Internet (% of population)	− 0.025	0.235	6	Positive
Population, total	0.000	0.879	7	Negative
GDP growth (annual %)	0.027	0.556	8	Negative
Urban population (% of total population)	0.013	0.735	9	Positive
Official exchange rate (LCU per US$, period average)	0.000	0.837	10	Negative

Panel A reports determinants that are statistically significant in the TWFE model at conventional levels.

Panel B reports additional variables ranked highly by SHAP but not statistically significant in TWFE. SHAP importance rank is based on mean absolute SHAP values from the global machine learning model.

5.3 XAI findings for region of Central Asia

The central objective of this study is to identify economically meaningful determinants of FDI inflows in Central Asian economies using ML and XAI. Unlike the baseline econometric specification, TWFE, which proved unstable when applied to the small regional subsample, the transfer learning framework allows the model to exploit information from a broader comparative panel while still

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 09 Issue 05 May 2026
DOI: 10.47191/ijcsrr/V9-i5-04, Impact Factor: 8.048
IJCSRR @ 2026
www.ijcsrr.org

producing region-specific explanatory insights. The overall ranking of determinants derived from the SHAP analysis is summarised in Table 6, while graphical representations of the main patterns are provided in Figure 2.

The SHAP results for Central Asia reveal a determinant structure that partially overlaps with the global findings but also exhibits several distinct regional features. Consistent with both the global SHAP analysis and the TWFE estimates, indicators of external openness and investment activity remain among the most influential predictors. In particular, current account balance, trade openness, and gross fixed capital formation continue to rank as key drivers of FDI inflows. This stability across modelling approaches suggests that these factors represent robust structural channels shaping investment dynamics in the region.

At the same time, the regional SHAP ranking reveals several determinants whose relative importance increases when the model is fine-tuned specifically for Central Asian economies. In particular, access to electricity and carbon-related environmental indicators emerge as substantially more influential predictors compared to their position in the global machine learning model and their statistical insignificance in the TWFE specification. This pattern likely reflects infrastructure constraints and resource-dependent investment structures characteristic of land-locked transition economies.

Similarly, urban population and GDP growth display stronger contributions in the regional model. While these variables were present in the global SHAP ranking, their elevated importance in the Central Asian specification suggests that demographic concentration and short-term macroeconomic performance may play a more pronounced signalling role for foreign investors in relatively small and volatile markets. Such shifts in determinant ranking indicate that transfer learning helps uncover region-specific investment dynamics that are only partially visible in broader cross-country patterns.

Overall, these findings demonstrate the practical value of the transfer learning approach adopted in this study. By combining cross-country information with region-specific fine-tuning, the proposed ML/XAI framework is able to uncover determinant patterns that would remain hidden in a purely local econometric analysis. The results therefore suggest that explainable machine learning can serve as a meaningful complementary analytical tool for FDI research in data-scarce regional settings, providing both confirmation of established relationships and new evidence on context-dependent investment drivers.

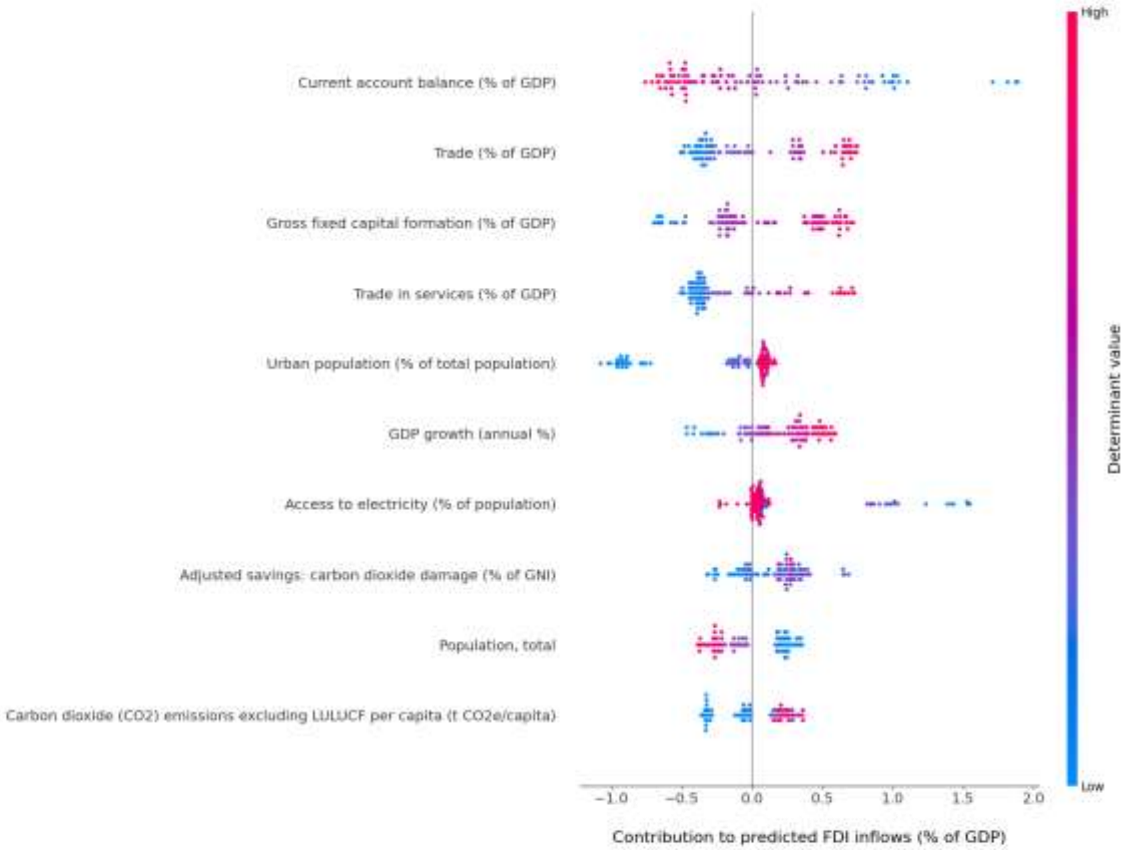

Figure 2. Determinant importance in Central Asia: SHAP summary

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 09 Issue 05 May 2026
DOI: 10.47191/ijcsrr/V9-i5-04, Impact Factor: 8.048
IJCSRR @ 2026
www.ijcsrr.org

Table 6. Determinant importance in Central Asia, SHAP evidence

Rank	Determinant	Mean absolute SHAP value	Direction of average effect
1	Current account balance (% of GDP)	0.518	Positive
2	Trade (% of GDP)	0.383	Positive
3	Gross fixed capital formation (% of GDP)	0.375	Positive
4	Trade in services (% of GDP)	0.364	Negative
5	Urban population (% of total population)	0.294	Negative
6	GDP growth (annual %)	0.285	Positive
7	Access to electricity (% of population)	0.248	Positive
8	Adjusted savings: carbon dioxide damage (% of GNI)	0.226	Positive
9	Population, total	0.224	Positive
10	Carbon dioxide emissions per capita (t CO_2e)	0.205	Positive
11	Tax revenue (% of GDP)	0.120	Positive
12	Total reserves (includes gold, current US$)	0.116	Positive
13	Individuals using the Internet (% of population)	0.094	Positive
14	Official exchange rate (LCU per US$, period average)	0.086	Positive
15	Rule of Law (estimate)	0.075	Negative

VI. CONCLUSION

The main objective of this study is to apply machine learning and explainable artificial intelligence methods to the analysis of foreign direct investment inflows in Central Asian countries. This research contributes to the literature by addressing two important gaps. First, empirical studies on FDI determinants in Central Asia remain relatively limited. Second, the application of machine learning and, in particular, explainable AI to this regional context has so far received little attention.

During the course of the analysis, several data-related constraints became evident. The number of Central Asian countries is inherently limited, and data availability is further restricted by missing values and inconsistencies across time. In practice, this required excluding Turkmenistan from the analysis due to severe data limitations. As a result, both the cross-sectional and temporal dimensions of the dataset are constrained, which poses challenges for traditional econometric as well as machine learning approaches.

These limitations motivated the use of a transfer learning framework as an integral part of the methodology. The model was initially trained on a broader set of countries selected based on their structural and macroeconomic similarity to Central Asia, as described in the Data Overview section. It was then fine-tuned on the Central Asian sample. This approach allows the model to leverage global information while adapting to region-specific patterns. Empirically, transfer learning improves both predictive performance and the interpretability of results, leading to a more stable and economically meaningful identification of FDI determinants in the region.

The results show that core macroeconomic determinants-such as trade openness, gross fixed capital formation, and current account balance-remain consistently important across econometric and machine learning frameworks. At the same time, the regional specification reveals several additional patterns. In particular, variables such as urban population, access to electricity, and carbon-related indicators become more prominent in the Central Asian context. While these variables are present in the global model, their increased importance in the regional specification suggests that they play a more significant role in shaping investment dynamics under local structural constraints.

An important aspect of this study is that it does not aim to position machine learning as a replacement for classical econometric methods. Instead, the comparison between TWFE and ML-XAI models is used as a consistency check. The overall alignment of results indicates that the proposed machine learning framework is economically meaningful and produces interpretations consistent with established approaches. At the same time, it provides additional insights by capturing nonlinearities and regional heterogeneity that are not fully reflected in linear models. The use of explainable AI is essential in this context, as it allows the model to move beyond a “black box” and produce interpretable, economically grounded results.

It is also important to emphasise that the study focuses on identifying determinants rather than forecasting FDI inflows. Given the complex and partly unobservable nature of investment decisions, predictive performance is not the primary objective. Instead, the emphasis is placed on understanding the structure of relationships between macroeconomic factors and FDI.

Several limitations should be noted. The analysis is based on a restricted set of macroeconomic indicators and a relatively small regional sample. In addition, while time fixed effects are included in the econometric model, dynamic temporal dependencies are not explicitly modelled. These constraints limit the scope of inference and point to directions for further research.

Future work may extend this framework by incorporating richer datasets, including institutional and geopolitical variables, as well as dynamic features such as lagged effects. With sufficient data availability, more complex modelling approaches, including neural networks and transformer-based architectures, could be explored. Given the observed effectiveness of transfer learning in the current setting, such approaches may further improve both predictive performance and the ability to capture complex economic relationships.

Finally, the proposed framework is not limited to the analysis of FDI. Similar approaches can be applied to other economic contexts characterised by limited data and structural heterogeneity, where combining econometric methods, machine learning, and explainable AI may provide additional analytical value.

REFERENCES

1. Kumari, J. (2014). Foreign Direct Investment and Economic Growth: A Literature Survey.
2. Udeh, E., Ibrahim, D., & Sesay, B. (2025). The Relationship between Foreign Direct Investment and Economic Growth: A Review of the Literature (1994-2023). *International Journal of Science and Management Studies (IJSMS)*.
3. Kehal, H. (2004). Foreign investment in developing countries.
4. Singh, D., & Turała, M. (2022). Machine Learning and Regularization Technique to Determine Foreign Direct Investment in Hungarian Counties. *DANUBE, 13*, 269 - 291.
5. Singh, D. (2023). Comparison between artificial neural network and linear model prediction performance for FDI disparity and the growth rate of companies in Hungarian counties. *Int. J. Bus. Inf. Syst., 43*, 542-552.
6. Singh, D. (2021). Interpretable Machine-Learning Approach in Estimating FDI Inflow: Visualization of ML Models with LIME and H2O. *TalTech Journal of European Studies, 11*, 133 - 152.
7. Singh, D. (2023). Foreign direct investment and local interpretable model-agnostic explanations: a rational framework for FDI decision making. *Journal of Economics, Finance and Administrative Science*.
8. Alon, I., Bretas, V.P., Sclip, A., & Paltrinieri, A. (2022). Greenfield FDI attractiveness index: a machine learning approach. *Competitiveness Review: An International Business Journal*.
9. Kemives, A., Randelovic, M., Barjaktarović, L., Dikanovic, P., Čabarkapa, M.D., & Randelovic, D.M. (2024). Identifying Key Indicators for Successful Foreign Direct Investment through Asymmetric Optimization Using Machine Learning. *Symmetry, 16*, 1346.

10. Prince, W., & Fantom, N.J. (2014). World development indicators 2014 : highlights.
11. Woodbridge, J.M. (2002). Econometric Analysis of Cross Section and Panel Data.
12. Pan, S.J., & Yang, Q. (2010). A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering, 22*, 1345-1359.
13. Friedman, J.H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics, 29*, 1189-1232.
14. Shao, Z., & Er, M.J. (2016). Efficient Leave-One-Out Cross-Validation-based Regularized Extreme Learning Machine. *Neurocomputing, 194*, 260-270.
15. Shapley, L.S. (1988). A Value for n-person Games.
16. Lundberg, S.M., & Lee, S. (2017). A Unified Approach to Interpreting Model Predictions. *Neural Information Processing Systems*.

Cite this Article: Xeniya, K., Zhang, D. (2026). Explainable AI for Foreign Direct Investment Analysis: Evidence from Central Asia. International Journal of Current Science Research and Review, 9(5), pp. 2254-2266. DOI: https://doi.org/10.47191/ijcsrr/V9-i5-04