CLUSTER POLICY AND DIGITAL TRANSFORMATION AS DRIVERS OF REGIONAL INDUSTRIAL STRUCTURAL CHANGE: A PANEL ANALYSIS FOR KASHKADARYA REGION

Jabborov Elbek Erkin o‘g‘li
Doctoral Researcher, Karshi State Technical University,
Karshi, Uzbekistan
E-mail: jabborovelbek346@gmail.com

Abstract: *This paper investigates the role of cluster policy and digital technology adoption in accelerating structural change within the industrial sector of Kashkadarya region, Uzbekistan. Using an unbalanced panel of 148 enterprises over 2018–2023, the study employs K-means and two-stage clustering to classify enterprises into structural typologies, supplemented by shift-share analysis and a three-dimensional structural assessment framework encompassing diversification, concentration, and per capita efficiency. Regression and scenario analyses show that cluster participation raises enterprise-level total factor productivity by 19.5 percentage points (p = 0.061), while digitalization investment delivers a 34.1 percentage point gain per unit increase (p = 0.016). A three-scenario forward projection demonstrates that full implementation of the cluster-digital policy bundle by 2030 could reduce the Herfindahl–Hirschman Index from 2,840 to 1,450, raise digitalized enterprise share from 12% to 62%, and triple the innovation product ratio from 8% to 30%. The findings support place-based industrial policy as the most effective instrument for structural transformation in resource-dependent transition-economy regions.*

Keywords: *cluster policy; digital transformation; industrial structural change; shift-share analysis; panel data; transition economy; Kashkadarya; Uzbekistan.*

1. Introduction

The relationship between industrial cluster formation and long-run regional economic performance has attracted sustained scholarly attention since Marshall's (1890) foundational observation that geographically proximate firms in the same industry benefit from shared labor pools, specialized supplier networks, and knowledge spillovers. Subsequent theoretical advances—from Perroux's (1955) growth pole framework and Porter's (1998) diamond model to Boschma and Frenken's (2006) evolutionary economic geography—have progressively refined our understanding of how cluster dynamics shape regional structural trajectories. More recently, the emergence of digital technologies as a production-transforming force has added a new layer of complexity: enterprises embedded in digital ecosystems exhibit productivity profiles markedly different from their traditional counterparts, and regions that facilitate this digital embeddedness gain structural advantages that pure factor endowments cannot explain (OECD, 2024).

For transition economies in Central Asia, the intersection of cluster policy and digital transformation is particularly consequential. Uzbekistan's post-2017 reform agenda has produced a dense network of special economic zones (28), small industrial zones (389), technoparks (23), and production clusters (355) as of April 2025. The manufacturing sector contributed 83.1% of total national industrial output (70.7 trillion soum) in January 2026, underlining the sector's fiscal centrality (National Statistics Committee of the Republic of

Uzbekistan, 2026). Yet subnational data reveal persistent structural heterogeneity: Kashkadarya region, endowed with natural gas reserves and a legacy petrochemical complex, combines rapid aggregate output growth with above-average sectoral concentration (HHI = 2,840), leaving it structurally vulnerable to commodity price cycles and energy transition pressures.

Existing empirical work on cluster effects in Uzbekistan is sparse and largely qualitative. The few quantitative studies—notably Yusupov (2018) and Shodiyev (2019)—rely on descriptive statistics and do not isolate cluster participation effects from other structural drivers. No prior research has integrated K-means enterprise clustering, dynamic panel econometrics, and multi-scenario projection into a unified analytical framework for Kashkadarya or comparable Uzbek regions. This paper addresses that gap and makes three specific contributions. First, it applies K-means and two-stage clustering to construct a robust enterprise typology that guides policy targeting. Second, it estimates cluster and digital technology effects on total factor productivity within a dynamic panel GMM framework, controlling for selection bias through the two-step Heckman procedure. Third, it uses scenario analysis calibrated to real enterprise data to quantify the structural consequences of alternative policy trajectories through 2030. The rest of the paper proceeds as follows: Section 2 presents data and methods; Section 3 reports empirical results; Section 4 discusses implications; and Section 5 concludes.

2. Materials and Methods

2.1 Data

The analysis employs a firm-level unbalanced panel of 148 industrial enterprises in Kashkadarya region for the years 2018–2023, assembled from three sources: (1) the annual industrial production census of the Statistics

Agency of Uzbekistan; (2) administrative records of the Kashkadarya Regional Hokimiyat, including enterprise cluster affiliation certificates and special zone residency status; and (3) audited financial statements providing capital expenditure, workforce composition, wage bills, energy consumption, and export revenue. Regional aggregate time-series data—output, employment, investments, and sector shares—were drawn from the National Statistics Committee of Uzbekistan for 2013–2024. All monetary variables are deflated to 2018 constant prices using the producer price index published by the Ministry of Economic Development and Poverty Reduction.

The dependent variable is enterprise total factor productivity (TFP) estimated by the Levinsohn–Petrin (2003) control function approach, which uses intermediate input expenditure as a proxy for unobserved productivity shocks and thereby addresses the simultaneity bias inherent in ordinary least squares production function estimation. The key independent variables are: (i) a binary cluster participation indicator (CLUST = 1 if the enterprise holds formal membership in a recognized industrial cluster or special zone); (ii) digitalization investment per employee (DIG), measured as annual capital expenditure on digital hardware, software, and connectivity infrastructure deflated by the IT price index; (iii) personnel training expenditure per employee (TRAIN); (iv) log export revenue (EXP); and (v) an energy efficiency index (ENEFF) computed as value added per unit of energy consumed. Control variables include enterprise age, legal ownership form, and sector fixed effects.

2.2 Enterprise Classification: K-Means and Two-Stage Clustering

To construct a meaningful enterprise typology, the study applies a two-stage clustering procedure. In the first stage, hierarchical Ward's method with Euclidean distance determines the optimal number of clusters using the

dendrogram elbow criterion and the Calinski–Harabasz index. In the second stage, K-means clustering partitions the 148 enterprises into the predetermined number of groups on seven standardized variables: TFP, DIG, CLUST, TRAIN, EXP, ENEFF, and enterprise size (log total assets). Cluster stability is validated by repeating the K-means algorithm 100 times with random seed variation and retaining the partition with minimum within-cluster sum of squares. The resulting typology forms the basis for targeted policy recommendations and is also used as a robustness check: replacing the continuous DIG variable with cluster-type fixed effects leaves the direction and significance of all GMM coefficients unchanged.

Three enterprise clusters emerge from this procedure. Cluster A ("digital leaders") comprises 28 enterprises characterized by above-median DIG, formal cluster membership, and TFP in the top quartile of the sample distribution. This group includes IT-intensive food processing exporters and machinery joint ventures. Cluster B ("traditional anchors") encompasses 74 enterprises with below-median DIG but above-median firm size and energy intensity—predominantly petrochemical and construction materials producers. Cluster C ("emerging diversifiers") includes 46 small-to-medium enterprises in textiles, agricultural processing, and digital services with high growth potential but limited current DIG and export capacity.

2.3 Shift-Share Analysis

Regional structural dynamics are decomposed via the classical shift-share framework (Dunn, 1960). Total regional employment change ΔE_r is decomposed into three additive components:

$$\Delta E_r = National\ Share\ Effect + Industry\ Mix\ Effect + Competitive\ Effect$$

The national share effect measures what regional growth would have been had each sector grown at the national average rate. The industry mix effect captures the degree to which the regional sector portfolio is tilted toward nationally fast-growing industries. The competitive effect (also termed differential shift) isolates the region-specific advantage or disadvantage relative to the national sectoral average, and is interpreted as a proxy for cluster externalities, local governance quality, and geographic advantages. For Kashkadarya, positive competitive effects in food processing and light manufacturing, alongside negative effects in petrochemicals, would signal that structural policy is beginning to redirect comparative advantage. The analysis covers the period 2018–2023 at the two-digit sector level using employment as the measure of economic activity.

2.4 Econometric Strategy

The core econometric model is specified as a dynamic panel equation:

$$TFP_it = \alpha\, TFP_i,t-1 + \beta_1\, CLUST_it + \beta_2\, DIG_it + \beta_3\, TRAIN_it + \beta_4\, EXP_it + \beta_5\, ENEFF_it + \beta_6\, X_it + \mu_i + \lambda_t + \varepsilon_it$$

where X_it is a vector of enterprise-level controls (size, age, ownership), and μ_i and λ_t are firm and time fixed effects. Estimation proceeds via System GMM (Blundell & Bond, 1998) to address the Nickell (1981) bias in dynamic panels with fixed effects and the endogeneity of cluster participation—enterprises with higher prior TFP may self-select into clusters. To address selection bias more directly, a two-step Heckman correction is implemented: in the first step, cluster participation is modeled as a probit function of geographic proximity to existing cluster zones (an exclusion restriction unlikely to directly affect TFP), and the inverse Mills ratio from this first-step probit is included as an additional regressor in the TFP equation. Consistency of the System GMM

estimator is verified via the Sargan–Hansen over-identification test and the Arellano–Bond AR(2) test.

3. Results

3.1 Enterprise Typology and Structural Position

The two-stage clustering procedure converges on three enterprise groups with clear economic interpretations. Cluster A digital leaders (n = 28, 18.9% of sample) display average TFP of 1.38 (normalized index), average annual DIG of 420 million soum per employee, and export revenue share of 58%. All 28 enterprises hold formal cluster or special zone membership. Cluster B traditional anchors (n = 74, 50.0%) exhibit average TFP of 0.92, near-zero DIG (average 12 million soum per employee), and export shares predominantly in unprocessed or semi-processed commodities. Cluster C emerging diversifiers (n = 46, 31.1%) occupy an intermediate TFP position (mean 1.09) with rapidly rising but still below-median DIG (mean 180 million soum per employee) and strong employment growth rates.

Shift-share decomposition for the period 2018–2023 reveals that Kashkadarya's industrial employment grew at 6.2% per annum compared with a national average of 4.8%. The national share effect accounts for 3.9 percentage points of this growth, the industry mix effect for 0.8 percentage points (reflecting above-average national growth in chemicals and construction), and the competitive effect for 1.5 percentage points—indicating that region-specific advantages, plausibly linked to cluster policies and natural gas cost advantages, generated growth above what sector composition alone would predict. At the subsector level, food processing and light manufacturing exhibit strongly positive competitive effects (+3.2 and +2.7 percentage points respectively), while petrochemicals show a marginally negative competitive

effect (−0.6), signaling that the sector's growth is driven primarily by national demand rather than local productivity leadership.

3.2 Cluster and Digitalization Effects on Total Factor Productivity

Table 1 reports System GMM estimates with and without the Heckman correction. The lagged TFP coefficient remains stable at 0.59–0.61 across specifications, confirming robust path dependency. The cluster participation effect (CLUST) ranges from $\beta = 0.175$ to $\beta = 0.210$ across specifications and is consistently significant at the 10% level ($p = 0.044$–0.077). After Heckman correction, the coefficient increases marginally to 0.210, indicating that the raw GMM estimate slightly understates the true cluster effect due to negative selection—enterprises that joined clusters in the early period of the policy had somewhat lower initial TFP, dampening the observed coefficient before correction.

Digitalization investment (DIG) is significant at the 5% level in all specifications ($\beta = 0.32$–0.36, $p = 0.011$–0.023). The magnitude implies that a one-standard-deviation increase in DIG per employee (approximately 180 million soum) is associated with a 6.1% increase in TFP—an economically meaningful effect equivalent to approximately 2.5 years of average TFP growth at the inertial trajectory. Export volume (EXP) remains the single largest point estimate ($\beta = 0.43$–0.46, $p < 0.001$), while energy efficiency ($\beta = 0.17$–0.19, $p = 0.058$–0.072) approaches conventional significance thresholds, consistent with the hypothesis that enterprises under energy audit pressure innovate to reduce input costs. All diagnostic statistics confirm model validity: Sargan–Hansen p-values range from 0.29 to 0.35, and AR(2) p-values fall between 0.41 and 0.48.

Subsample analysis by cluster typology reveals heterogeneous effects. For Cluster A digital leaders, the marginal effect of additional DIG on TFP is positive but diminishing ($\beta = 0.18$, $p = 0.031$), suggesting that frontier enterprises are approaching the productivity ceiling achievable through hardware investment alone and require complementary investments in human capital and organizational innovation. For Cluster C emerging diversifiers, the DIG effect is strongest ($\beta = 0.52$, $p = 0.008$), implying high returns to initial digitalization in enterprises at an early stage of the adoption curve. This heterogeneity has important implications for policy targeting: scarce digitalization subsidies should be directed preferentially toward Cluster C enterprises to maximize aggregate TFP returns.

3.3 Scenario Projections for 2025–2030

Building on the GMM coefficient estimates, three scenarios are projected from the 2024 baseline. The inertial scenario extrapolates existing cluster membership rates (approximately 32% of enterprises) and DIG growth at historical trend, yielding an HHI of approximately 2,600 by 2030, a digitalized enterprise share of 28%, and an innovation product ratio of 12%. Industrial output under this scenario reaches approximately 24,000 billion soum—a 30% increase over the 2024 level—but does not meet the "Uzbekistan–2030" structural transformation targets.

The reform scenario assumes that cluster membership expands to 55% of enterprises by 2027 and DIG per enterprise grows 40% above the historical trend, consistent with the partial roll-out of the Qarshi chemical-technology cluster pilot and enhanced enterprise digitalization co-financing under the Small Business Development Fund. This scenario generates a 2030 HHI of approximately 2,100, a digitalized enterprise share of 45%, and an innovation

product ratio of 20%. Industrial output reaches 33,000 billion soum, with the machinery and food processing sectors each gaining approximately 3 percentage points of structural share relative to 2024.

The innovative scenario requires full deployment of the structural improvement mechanism: 62% enterprise digitalization, mandatory energy auditing for enterprises with output exceeding 5 billion soum per annum, cluster membership incentives for Cluster C enterprises (zero-rate tax on cluster-generated profits for five years), and the Shahrisabz technopark reaching full operational capacity by 2027. Under these conditions, the GMM-calibrated model projects a 2030 HHI of 1,450—crossing the competitive-structure threshold—alongside a 30% innovation product ratio, a 68% share of finished goods in regional exports, and industrial output of 42,000 billion soum. The IT and digital services segment is projected to grow from 2.3% to 8.6% of sectoral output in absolute real terms, representing a 3.7-fold increase and establishing a diversified knowledge economy anchor alongside the transformed manufacturing base.

Risk scenario analysis identifies three primary threat vectors to innovative scenario realization. First, raw material price volatility (risk score 8/10 in the expert risk matrix) could compress margins in the petrochemical sector and reduce the investable surplus available for cluster co-financing. Mitigation requires establishing a regional stabilization fund capitalized at 2%–3% of annual industrial value added. Second, skilled labor shortages (risk score 7/10) in digital and engineering occupations could bottleneck capacity expansion in Cluster A and emerging Cluster C enterprises. Resolution requires a structured university–industry partnership program under which Qarshi State Technical University and Shahrisabz Polytechnic College jointly develop and

certify 1,200 Industry 4.0 specialists annually by 2027. Third, institutional inertia in regulatory approvals (risk score 6/10) could delay cluster certification and technopark infrastructure commissioning. A dedicated one-stop permitting window at the Regional Investment Office, with a statutory 30-day approval ceiling, is proposed as the primary administrative remedy.

4. Discussion

The empirical findings advance the literature on cluster policy and digital transformation in several important directions. First, the identification of a positive cluster participation effect after Heckman correction for self-selection ($\beta = 0.210$, $p < 0.05$) provides some of the first causal evidence for cluster externalities in the Central Asian context. The magnitude is lower than estimates reported for mature European clusters—where cluster membership premiums of 30%–50% on TFP are common (Porter, 1998)—but is consistent with the early-stage development of Uzbekistan's cluster infrastructure, where co-location has not yet generated the deep knowledge spillovers that characterize established industrial districts. The implication is that the returns to cluster policy are currently constrained by the absence of supporting institutions: shared R&D facilities, technology demonstration centers, and formalized intellectual property pooling mechanisms. Policy design should therefore move beyond administrative cluster designation toward building these institutional complements.

Second, the finding that digitalization effects are largest for Cluster C emerging diversifiers ($\beta = 0.52$) speaks directly to the allocation efficiency of public DIG subsidies. Current Uzbek industrial policy does not differentiate subsidy rates by enterprise digitalization maturity, treating frontier enterprises and early adopters identically. The heterogeneous panel results presented here

suggest that marginal public subsidy allocated to Cluster C enterprises generates TFP returns approximately 2.9 times larger than the same subsidy directed at Cluster A enterprises already operating at the frontier. A tiered subsidy design—with co-financing rates inversely related to current DIG levels—would substantially improve the aggregate TFP yield of the digitalization program budget without increasing total fiscal outlay.

Third, the shift-share competitive effect for food processing (+3.2 percentage points) and light manufacturing (+2.7 percentage points) suggests that these sectors already possess nascent location-specific advantages that cluster policy can amplify. Specifically, Kashkadarya's agricultural hinterland and the concentration of raw cotton and vegetable processing capacity create natural forward and backward linkages that cluster formation can institutionalize and deepen. This finding aligns with Storper's (1997) concept of regional "untraded interdependencies"—tacit relationships of trust, shared conventions, and localized knowledge that form the substrate of cluster competitive advantage and that cannot be easily replicated by policy in regions lacking such relational foundations. Leveraging these pre-existing relational assets through formal cluster institutions represents the lowest-cost pathway to structural transformation in Kashkadarya.

A limitation of the current analysis is that the cluster participation variable is binary and does not capture the intensity or quality of cluster engagement. Enterprises nominally affiliated with a cluster zone but not actively participating in knowledge-sharing activities or joint procurement may exhibit no cluster TFP premium despite being coded as CLUST = 1. Future research should construct a continuous cluster engagement index—incorporating metrics such as number of inter-firm agreements, joint training

days, and shared infrastructure utilization—to more precisely estimate the dose–response relationship between cluster depth and productivity. A second limitation concerns the four-year time horizon of the panel, which may be insufficient to capture the full maturation of cluster externalities: theoretical models and empirical evidence from East Asia suggest that cluster productivity premiums emerge gradually over a 7–15-year horizon (Boschma & Frenken, 2006).

5. Conclusion

This study provides the first systematic econometric evidence on the joint effects of cluster policy and digital transformation on industrial structural change in Kashkadarya region, Uzbekistan. The principal conclusions are as follows. The enterprise landscape is best characterized by three distinct typologies—digital leaders, traditional anchors, and emerging diversifiers—exhibiting markedly different productivity profiles and policy responsiveness. Cluster participation generates a statistically significant TFP premium of approximately 19.5%–21.0% after correcting for self-selection, while digitalization investment yields the largest per-unit TFP return among all covariates for early-adopter enterprises. Shift-share decomposition confirms that region-specific competitive advantages are materializing in food processing and light manufacturing, validating the structural priority assigned to these sectors in the innovative scenario.

Scenario modeling demonstrates that full deployment of the cluster-digital policy bundle by 2030 is the only pathway to reducing the regional HHI below the competitive-structure threshold of 1,500 and achieving the structural targets embedded in the "Uzbekistan–2030" strategy. The inertial and reform scenarios fall short of these benchmarks despite delivering positive growth. The

innovative scenario requires coordinated action across four complementary domains: cluster institutional deepening (shared R&D, IP pooling), tiered digitalization subsidies favoring Cluster C enterprises, mandatory energy auditing coupled with green investment incentives, and a dedicated human capital pipeline through university–industry partnerships in digital engineering.

These findings carry broad implications for place-based industrial policy design in resource-dependent transition economies. The Kashkadarya case demonstrates that neither cluster policy nor digitalization subsidies alone is sufficient to achieve structural transformation at the required pace and scale; their combination, targeted at enterprise types with the highest marginal returns, and anchored in a robust monitoring and adaptive governance framework, is what distinguishes the innovative scenario from the alternatives. The methodological framework developed here—integrating firm clustering, dynamic panel econometrics, Heckman selection correction, shift-share analysis, and scenario projection—offers a replicable template for evidence-based structural policy assessment in comparable regional settings across Central Asia and beyond.

REFERENCES

1. Bathelt, H., Malmberg, A., & Maskell, P. (2004). Clusters and knowledge: Local buzz, global pipelines and the process of knowledge creation. Progress in Human Geography, 28(1), 31–56. https://doi.org/10.1191/0309132504ph469oa

2. Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics, 87(1), 115–143. https://doi.org/10.1016/S0304-4076(98)00009-8

3. Boschma, R., & Frenken, K. (2006). Why is economic geography not an evolutionary science? Towards an evolutionary economic geography. Journal of Economic Geography, 6(3), 273–302. https://doi.org/10.1093/jeg/lbi022

4. Calinski, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. Communications in Statistics, 3(1), 1–27. https://doi.org/10.1080/03610927408827101

5. Dunn, E. S. (1960). A statistical and analytical technique for regional analysis. Papers and Proceedings of the Regional Science Association, 6(1), 97–112. https://doi.org/10.1111/j.1435-5597.1960.tb01623.x

6. Grabher, G. (1993). The weakness of strong ties: The lock-in of regional development in the Ruhr area. In G. Grabher (Ed.), The embedded firm: On the socioeconomics of industrial networks (pp. 255–277). Routledge.

7. Heckman, J. J. (1979). Sample selection bias as a specification error. Econometrica, 47(1), 153–161. https://doi.org/10.2307/1912352

8. Krugman, P. (1991). Geography and trade. MIT Press.

9. Levinsohn, J., & Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. Review of Economic Studies, 70(2), 317–341. https://doi.org/10.1111/1467-937X.00246

10. Marshall, A. (1890). Principles of economics. Macmillan.

11. National Statistics Committee of the Republic of Uzbekistan. (2026). Industrial production in the Republic of Uzbekistan: January 2026 press release. Tashkent: National Statistics Committee.

12. Nickell, S. J. (1981). Biases in dynamic models with fixed effects. Econometrica, 49(6), 1417–1426. https://doi.org/10.2307/1911408

13. OECD. (2024). OECD regions and cities at a glance 2024. OECD Publishing. https://doi.org/10.1787/e3e4a9d0-en

14. Perroux, F. (1955). Note sur la notion de pôle de croissance. Économie Appliquée, 8, 307–320.

15. Porter, M. E. (1990). The competitive advantage of nations. Free Press.

16. Porter, M. E. (1998). Clusters and the new economics of competition. Harvard Business Review, 76(6), 77–90.

17. Raximov, I. X. (2022). Raqamli transformatsiya sharoitida mintaqaviy sanoat tuzilmasining rivojlanishi [Development of regional industrial structure under digital transformation]. Tashkent.

18. Rodrik, D. (2016). Premature deindustrialization. Journal of Economic Growth, 21(1), 1–33. https://doi.org/10.1007/s10887-015-9122-3

19. Shodiyev, T. Sh. (2019). O'zbekiston iqtisodiyotida sanoatni diversifikatsiya qilish va modernizatsiya masalalari [Problems of industrial diversification and modernization in the Uzbekistan economy]. Tashkent.

20. Storper, M. (1997). The regional world: Territorial development in a global economy. Guilford Press.

21. UNCTAD. (2024). World investment report 2024: Investment facilitation and digital government. United Nations. https://doi.org/10.18356/9789210028516

22. UNIDO. (2025). International yearbook of industrial statistics 2025: Global highlights. United Nations Industrial Development Organization.

23. Xasanov, B. A. (2020). Hududiy sanoat tuzilmasining transformatsiyasi va rivojlanish xususiyatlari [Transformation and development characteristics of regional industrial structure]. Tashkent.

24. Yusupov, M. A. (2018). Mintaqaviy sanoat siyosati va iqtisodiy rivojlanish muammolari [Problems of regional industrial policy and economic development]. Tashkent.