

World Conference on Social Sciences, Law and Public Policy
Hosted Online from Toronto, Canada
Date: 26th October, 2026
Website: https://econferencia.com

ECONOMETRIC ANALYSIS OF THE IMPACT OF DIGITAL INFRASTRUCTURE ON THE DEVELOPMENT OF THE INFORMATION SERVICES MARKET

Xudayberganov Zokhidjon Djurakulovich
Independent Researcher, Samarkand
State University named after Sharof Rashidov
xudayberganov@janubiy.uztelecom.uz

Abstract

The article uses econometric methods on panel data to assess the impact of digital infrastructure on the development of the market for information and communication services. The analysis covers the fourteen regions of the Republic of Uzbekistan over 2016–2025 and comprises 140 observations. Five specifications were tested. In the model with region fixed effects only, the coefficient on the number of Internet subscribers is 0.9308 and statistically significant at the one per cent level ($t = 11.66$). After year fixed effects are added, the coefficient falls to 0.4633 and loses its significance entirely ($t = 0.71$). Variance decomposition shows that 68.4% of the variation in the dependent variable is generated by a common time factor. The results show that the strong association in the raw data reflects a common trend and that interpreting it as a causal relationship would be erroneous.

Keywords: Econometric analysis, panel data, fixed effects, digital infrastructure, common trend, spurious correlation, variance decomposition.



World Conference on Social Sciences, Law and Public Policy
Hosted Online from Toronto, Canada
Date: 26th October, 2026
Website: https://econferencia.com

INTRODUCTION

The impact of digital infrastructure on the development of the information services market is usually taken as an axiom. The logic appears simple: the higher the level of Internet connectivity, the greater the demand for digital services and the faster the industry develops. This thesis is used as a premise in regional programmes, and investment in infrastructure is planned as a tool for developing the industry.

However, this relationship is rarely tested empirically. Its apparent plausibility conceals the need for testing: both indicators grow over time, and the correlation between them is bound to be high. In statistical analysis, such a situation creates a classic risk – the correlation of two series with a common trend does not imply a causal relationship.

This study rigorously tests the hypothesised relationship using panel data. The panel structure makes it possible to separate the common trend statistically: year fixed effects capture factors common to all regions, and the coefficient is identified solely from within-region fluctuations.

LITERATURE REVIEW

A broad empirical base has been accumulated in the literature on the impact of information and communication technologies on economic growth. E. Brynjolfsson and A. McAfee showed that there is a time lag between technological renewal and productivity growth: significant results are obtained not immediately after a technology is introduced but after a certain period of accumulation [1].

Cross-country studies of the relationship between broadband Internet and economic growth have reported positive results [6; 7]. At the same time, the methodological weaknesses of these studies are widely discussed: the problem of

reverse causality may remain unresolved – economic growth makes it possible to develop infrastructure, so the relationship is likely to be bidirectional.

When working with panel data, the common trend problem is considered one of the main methodological issues in the econometric literature. The two-way fixed effects model is the standard tool for addressing this problem, as it simultaneously controls for unobserved time-invariant regional characteristics and annual shocks common to all regions [2].

UNCTAD reports pay particular attention to the relationship between digital infrastructure and value creation: infrastructure is a necessary condition, but it does not automatically translate into economic activity [3]. This thesis is important for interpreting the results of the present study.

The review of the literature concludes that the relationship is theoretically grounded, but the debate on its strength and causal direction continues. No such test based on regional panel data has been carried out for Uzbekistan.

METHODOLOGY

The empirical base of the study consists of two official open datasets of the National Statistics Committee: the growth rate of services in the field of information and communication and the number of subscribers connected to the Internet, both by region [4; 5]. The panel covers fourteen regions and 2016–2025, totalling 140 observations.

The dependent variable is a chain index of the volume of information and communication services (2015 = 100). The index was constructed for each region as the product of annual growth rates. The explanatory variable is the number of Internet subscribers. Both indicators are in logarithms, so the coefficient is interpreted as an elasticity.

The main specification is a two-way fixed effects model:

$$\ln Y_{it} = \alpha_i + \lambda_t + \beta \cdot \ln SUB_{it} + \varepsilon_{it} \qquad (1)$$

where Y(it) is the services index in region i in year t; SUB(it) is the number of Internet subscribers; α(i) is the region fixed effect; λ(t) is the year fixed effect. Standard errors are clustered by region.

The essence of the identification strategy is as follows. Region effects capture time-invariant characteristics – population density, the presence of educational institutions, proximity to the capital. Year effects account for factors common to all regions – macroeconomic conditions, changes in national policy, technological trends. As a result, the coefficient β is identified only from within-region deviations from the common trend.

To check the robustness of the results, five specifications were tested: with region effects only; with two-way effects; with a lagged explanatory variable; with both current and lagged variables; and a first-differences model. In addition, a variance decomposition was performed.

RESULTS AND ANALYSIS

The key test of the hypothesis is whether the estimated effect of Internet subscribers survives once the common time trend is controlled for. To this end, the five specifications described above were estimated; the results are presented in Table 1.



Table 1. Estimation results of the panel regression models

Specification	Fixed effects	Variable	Coefficient	Standard error	t-statistic	Within R^2
Model 1	Region FE	ln SUB(t)	0.9308	0.0798	11.66	0.818
Model 2	Region and year FE	ln SUB(t)	0.4633	0.6565	0.71	0.083
Model 3	Region and year FE	ln SUB(t−1)	0.3520	0.6360	0.55	0.057
Model 4	Region and year FE	ln SUB(t)	0.4195	0.3111	1.35	0.083
		ln SUB(t−1)	0.0507	0.4059	0.12	
Model 5	Region and year FE	Δln SUB (first differences)	0.1183	0.1721	0.69	0.012

The analysis of the table provides the central result of the study. In Model 1, i.e. the specification with region fixed effects only, the coefficient is 0.9308 and is highly significant at the one per cent level (t = 11.66). The within R^2 is 0.818 – the model explains 81.8% of the variation in the dependent variable.

At first glance, this result convincingly confirms the hypothesis: the elasticity is close to one, i.e. a ten per cent increase in the number of subscribers raises the volume of services by about nine per cent. Had the analysis stopped at this stage, the conclusion would have been clear and strong.

In Model 2, however, the picture changes completely. After year fixed effects are added, the coefficient falls to 0.4633, i.e. by more than half, while the standard error increases from 0.0798 to 0.6565, i.e. more than eightfold. As a result, the t-statistic drops from 11.66 to 0.71 and the coefficient loses its statistical significance entirely.

The within R^2 also falls from 0.818 to 0.083. This means that almost all of the variation explained in Model 1 is attributable not to the number of subscribers but to a time factor common to all regions.

The remaining three specifications confirm the result. The model with a lagged variable (Model 3) also yields an insignificant result; when current and lagged variables are included together (Model 4), the coefficients do not differ significantly from zero either. The first-differences model (Model 5), the most rigorous specification relating growth rates to growth rates, gives a coefficient of 0.1183, which is also insignificant.

The difference between the specifications is easier to grasp when the point estimates are shown together with their confidence intervals, as in Figure 1.

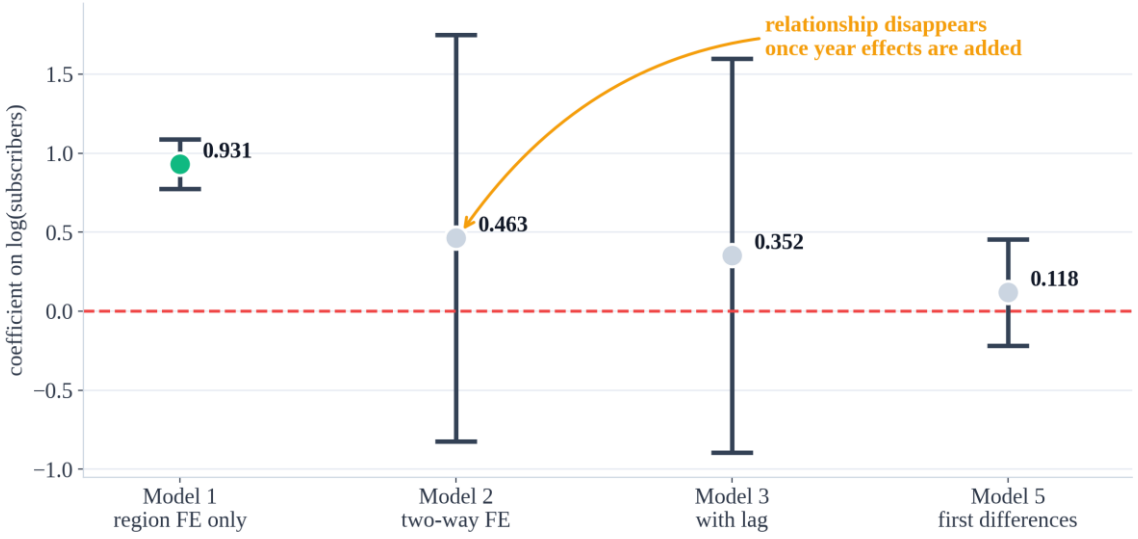


Figure 1. Coefficient estimates and 95% confidence intervals by specification

The chart visualises the essence of the results. In the first model the confidence interval is very narrow and lies well away from zero. Once year effects are included, the interval widens sharply and covers the zero line – the graphical expression of statistical insignificance.

The width of the confidence intervals is also informative. In Model 2 it stretches from about −0.82 to +1.75. Such a wide interval does not indicate the absence of a relationship but that the available data are insufficient to detect it. This distinction is scientifically important and is discussed separately below.

In the first-differences model (Model 5) the confidence interval is considerably narrower – from about −0.22 to +0.46. This specification is the most rigorous, and it rejects the existence of a strong relationship much more convincingly: if the elasticity were indeed close to one, it would lie outside this interval.

It is also useful to quantify the difference between the models. The decline of the coefficient from 0.9308 to 0.4633 amounts to 0.4675 in absolute terms – i.e. more than half of the estimate in the first model is driven by the common trend. The eightfold increase in the standard error shows that the remaining half cannot be estimated reliably either.

In the econometric literature such a situation is called omitted variable bias. In this case, time plays the role of the omitted variable: it affects both series simultaneously and creates an artificial relationship between them. Including year fixed effects is the standard tool for eliminating this bias.

To identify the source of this result, the raw data were compared with the data after removing fixed effects (Figure 2).

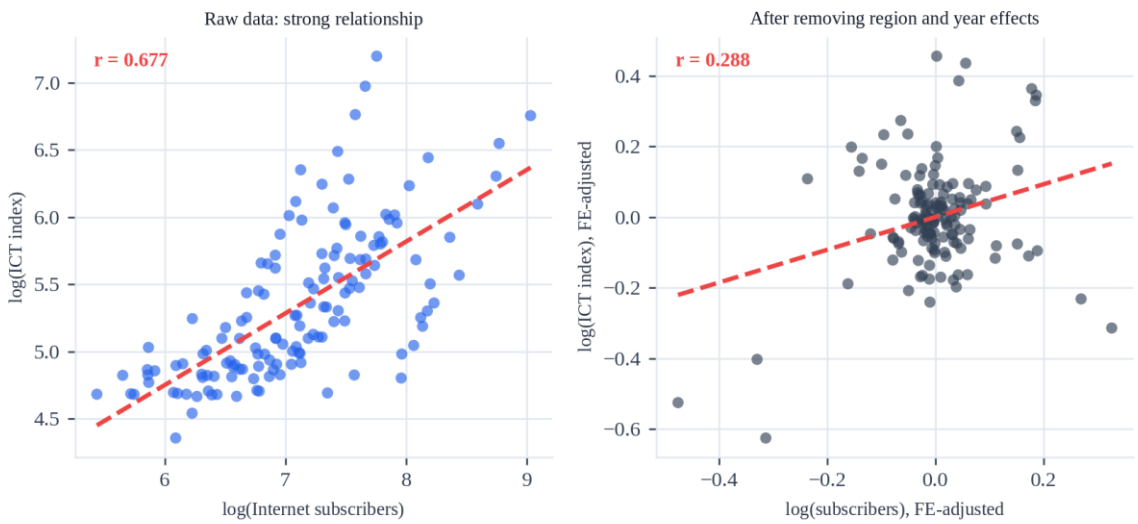


Figure 2. Relationship in the raw data and in the data adjusted for fixed effects

The chart in the left panel shows a strong positive relationship in the raw data: the correlation coefficient is 0.677, and the points are relatively tightly clustered around the regression line. It is precisely such a picture that is usually cited as evidence of a relationship.

The chart in the right panel shows the situation after region and year effects have been removed. The points are scattered around the centre, the correlation coefficient drops to 0.288, and the slope of the regression line decreases considerably. Comparing the two panels clearly shows the contribution of the common trend.

The economic interpretation is as follows. Both indicators grew steadily in 2016–2025: the number of Internet subscribers increased more than fourfold nationwide, and the volume of services more than sixfold. The correlation between two growing series is mathematically bound to be high, but it does not

imply causality – both may be the outcome of a third factor, namely the overall process of digital development.

The weight of this common trend can be measured directly by decomposing the variance of the dependent variable into its components. The results are presented in Figure 3 and Table 2.

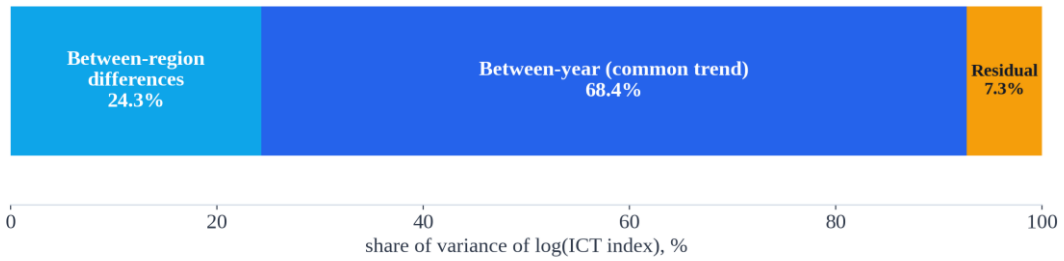


Figure 3. Decomposition of the variance of the services index

The figure shows that the common trend accounts for more than two-thirds of the variance, while the residual within-region component, from which the effect of infrastructure must be identified, is the smallest of the three. The content of each component and its econometric implications are detailed in Table 2.

Table 2. Variance decomposition and methodological interpretation of the results

Variance component	Share, %	What it captures	Econometric implication
Between-region differences	24.3	Time-invariant regional characteristics: population, education base, location	Absorbed by region FE
Between-year (common trend)	68.4	Macroeconomic conditions, national policy, technological trend	Absorbed by year FE; the main component
Residual	7.3	Time-varying regional factors	The only space left for identification

The analysis of the table provides the deepest explanation of the result. 68.4% of the variation in the dependent variable is generated by the common time factor. This is natural: the industry is developing rapidly across the whole country at the same time, and this development is determined by national policy, technological opportunities and global demand.

Persistent between-region differences account for 24.3% and are absorbed by region fixed effects. As a result, the space left for the explanatory variable is only 7.3%. It is within this narrow space that the effect of infrastructure has to be identified.

This raises the issue of statistical power. Even if infrastructure had a genuine effect within this space, it would be difficult to detect it reliably with data on fourteen regions over a ten-year period. The result obtained therefore does not mean the absolute absence of a relationship, but that the available data do not provide sufficient evidence to confirm it.

This distinction is of fundamental scientific importance. Statistical insignificance does not prove the absence of an effect – it only shows that the available evidence is insufficient. In this case, however, there is additional evidence: the narrow confidence interval in Model 5 convincingly rejects the hypothesis of a strong effect, while a weak effect cannot be ruled out.

For the results to inform decision-making, each econometric finding needs to be translated into a policy conclusion. The practical interpretation of the results is summarised in Table 3.

Table 3. Practical interpretation of the econometric results and implications for policy

Result	Possible explanation	Implication for policy
Coefficient is high and significant with region FE only	The common time trend is not controlled for	Justifications based on simple correlation should be reconsidered



Result	Possible explanation	Implication for policy
Significance disappears once year FE are added	The relationship is mainly driven by the common trend	Infrastructure on its own does not generate growth of the industry
Residual variance of 7.3%	Very little room for time-varying regional factors	Short-term regional measures will not produce significant results
Between-year component of 68.4%	Development is determined by national-level factors	National policy instruments take priority over regional ones
Also insignificant in the first-differences model	The result is robust and does not depend on the specification	The conclusion is reliable

The analysis of Table 3 makes it possible to separate the methodological and practical significance of the results. Methodologically, the main conclusion is that drawing causal conclusions from the correlation between two growing indicators is a mistake. Although this is a simple rule, it is regularly violated in practice.

In practical terms, the most important conclusion follows from the second row: digital infrastructure on its own does not generate growth of the information services market. This result is fully consistent with the situation identified in a previous study for Surkhandarya region – although the region's infrastructure block index was 0.491, its market block remained at 0.324.

The agreement of results obtained by two independent methods increases the reliability of the conclusion. Infrastructure is a necessary condition – without it the industry does not develop at all. However, it is not a sufficient condition: economic activity does not emerge by itself wherever connectivity is available; it requires personnel, entrepreneurship and channels of access to markets.

An important conclusion for regional policy follows from the fourth row. Since 68.4% of development is determined by national-level factors, the scope of influence of regional measures is limited. This does not mean that regional policy is unnecessary, but that it must be coordinated with national policy.

The limitations of the study should be stated openly. First, the number of Internet subscribers is not a complete measure of digital infrastructure – it measures the number of connections, not speed or quality. Second, the panel is relatively short and the number of regions is limited, which reduces statistical power. Third, the model tests only a linear relationship; threshold effects, i.e. the possibility that infrastructure has an effect only above a certain level, cannot be ruled out.

The fourth limitation concerns the nature of the dependent variable. The services index measures volume growth, not competitiveness. As substantiated in previous studies, volume growth does not imply an increase in value added. To fully assess the effect on competitiveness, it would be advisable to use an integral index as the dependent variable, but no database exists for calculating it as a long time series.

At the same time, the negative result itself has scientific and practical value. It has tested a widespread but empirically unverified assumption. Such results are usually not published, although they are no less important than positive results for preventing resources from being spent in ineffective directions.

CONCLUSIONS AND RECOMMENDATIONS

First, the analysis carried out in five specifications on panel data covering fourteen regions and 2016–2025 showed that there is no statistically reliable relationship between digital infrastructure and the development of the information services market.

Second, in the model with region fixed effects only, the coefficient is 0.9308 and significant at the one per cent level ($t = 11.66$). After year effects are added, it falls to 0.4633 and loses its significance entirely ($t = 0.71$). This is a classic example of a common trend effect.

Third, variance decomposition showed that 68.4% of the variation in the dependent variable is generated by the common time factor and 24.3% by persistent between-region differences, while the residual accounts for only 7.3%.

Fourth, the result independently agrees with the conclusion obtained through integral assessment in a previous study: infrastructure is a necessary but not a sufficient condition. Economic activity does not emerge by itself wherever connectivity is available.

Fifth, the following recommendations are made: to abandon setting infrastructure indicators as stand-alone targets in regional programmes; to plan them mandatorily together with measures for training personnel and supporting entrepreneurship; and to reconsider justifications based on simple correlation.

Sixth, the limitations of the study should be taken into account: the number of subscribers is not a complete measure of infrastructure, the panel is short and the model is linear. Threshold effects and quality indicators will be examined in further research.

In the next study, an organisational and economic mechanism for increasing competitiveness in the information services industry will be developed.

REFERENCES

1. Brynjolfsson E., McAfee A. The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. – New York: W.W. Norton & Company, 2014. – 306 p.
1. Wooldridge J.M. Econometric Analysis of Cross Section and Panel Data. 2nd ed. – Cambridge, MA: MIT Press, 2010. – 1064 p.
2. Digital Economy Report 2019: Value Creation and Capture – Implications for Developing Countries. – Geneva: UNCTAD, 2019. – 194 p.

3. National Statistics Committee of the Republic of Uzbekistan. Growth rate of services in the field of information and communication (1.05.01.0020) (in Uzbek). – URL: https://stat.uz.
4. National Statistics Committee of the Republic of Uzbekistan. Number of subscribers connected to the Internet, total (in Uzbek). – URL: https://stat.uz.
5. Measuring the Information Society Report. – Geneva: International Telecommunication Union (ITU), 2018.
6. OECD Digital Economy Outlook 2020. – Paris: OECD Publishing, 2020. – 314 p.
7. Krugman P. Increasing returns and economic geography // Journal of Political Economy. – 1991. – Vol. 99, No. 3. – P. 483–499.
8. Baldwin R. The Great Convergence: Information Technology and the New Globalization. – Cambridge, MA: Harvard University Press, 2016. – 344 p.
9. Arora A., Gambardella A. The globalization of the software industry: Perspectives and opportunities for developed and developing countries // Innovation Policy and the Economy. – 2005. – Vol. 5. – P. 1–32.
10. Decree of the President of the Republic of Uzbekistan No. PF-6079 of 5 October 2020 “On approval of the Digital Uzbekistan – 2030 Strategy and measures for its effective implementation” (in Uzbek) // Lex.uz.
11. IT Park Uzbekistan. Official data. – URL: https://it-park.uz.