

Article

How Computing Power Infrastructure Development Shapes Urban Green Innovation Performance: Evidence from China

Xiaoxiao Tian ^{1,*} , Wei Guo ² and Jingyu Liao ¹¹ School of Economics and Statistics, Guangzhou University, Guangzhou 510006, China² School of Credit Management, Guangdong University of Finance, Guangzhou 510006, China

* Correspondence: tianxiaoxiao@gzhu.edu.cn

Abstract

Computing power infrastructure (CPI), represented by data centers, cloud computing platforms, edge computing nodes, and intelligent computing clusters, has become a strategic component of new infrastructure and a direct carrier of computational resources. This study adopts a two-stage empirical framework. In the first stage, urban green innovation performance (UGIP) is measured using a global benchmark undesirable-output slacks-based measure (GB-U-SBM) model that incorporates innovation inputs, desirable patent outputs, and undesirable carbon-emission outputs. In the second stage, the estimated UGIP scores are used as the dependent variable in two-way fixed-effects panel regressions to examine the association between CPI development and UGIP. Using a city-year panel for Chinese prefecture-level cities from 2006 to 2023 and a main regression sample of 5166 observations, the results show that CPI development is positively associated with UGIP after controlling for city fixed effects, year fixed effects, and city-level covariates. The results are robust to alternative measures of UGIP and CPI, sample treatments, winsorization, lagged specifications, and instrumental-variable estimates. Channel-consistency tests indicate that CPI development is associated with human capital accumulation and informatization, two conditions that support green innovation. The study clarifies that the evidence should be interpreted primarily as conditional association rather than definitive causal mediation. The findings suggest that CPI can become a productive foundation for green innovation only when it is embedded in skilled-labor formation, data governance, high utilization, and low-carbon operation systems.

Keywords: computing power infrastructure; data centers; urban green innovation performance; two-stage DEA; undesirable-output SBM; green innovation efficiency; human capital; informatization; sustainable urban development; China



Academic Editor: Irina Georgescu

Received: 16 June 2026

Revised: 14 July 2026

Accepted: 14 July 2026

Published: 16 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The rapid expansion of computing power infrastructure (CPI) is reshaping the technological foundation of urban development. Data centers, cloud computing platforms, edge nodes, and artificial-intelligence computing clusters provide the computing, storage, transmission, and scheduling capabilities required for digital production, intelligent governance, and data-driven innovation. Unlike general digital infrastructure indicators such as broadband access or Internet penetration, CPI directly reflects the supply of computational resources and therefore constitutes a more specific infrastructure basis for advanced analytics, simulation, algorithmic optimization, and intelligent industrial applications [1–5].

CPI is also closely connected with the green transformation of cities. On the positive side, high-performance computing can reduce the cost of green technology research and development, improve the precision of environmental monitoring, support industrial energy-management systems, and accelerate knowledge spillovers across firms, universities, and research institutions. On the negative side, data centers and intelligent computing clusters require large amounts of electricity and cooling resources, and their environmental effects depend on energy structure, power usage effectiveness, equipment utilization, server life-cycle emissions, and spatial coordination with renewable-energy supply [6–8].

Existing studies have examined the effects of transport infrastructure, broadband pilots, digital-city construction, digital economy development, and environmental regulation on regional innovation, energy efficiency, and carbon emissions [9–18]. However, three issues remain insufficiently addressed. First, much of the literature treats digital infrastructure as a broad composite concept, while the independent role of CPI as a computing-resource supply system is less clearly identified. Second, many studies evaluate green development or innovation output separately, whereas UGIP should reflect the joint production of innovation outputs and environmental costs. Third, the channels through which CPI may affect UGIP require clearer empirical interpretation and should be distinguished from formal mediation claims unless indirect effects are explicitly identified.

This study responds to these gaps by constructing a city-level CPI proxy using Internet Data Center (IDC) license information and by measuring UGIP with a global benchmark undesirable-output slacks-based measure model. The research design is explicitly two-stage: the first stage produces comparable UGIP efficiency scores, and the second stage relates these scores to CPI development within a two-way fixed-effects panel framework. This clarification is important because two-stage DEA has a long methodological debate. Simar and Wilson [19] emphasize bootstrap-based inference, while Banker, Natarajan, and Zhang [20] show that second-stage OLS-type approaches can be informative under specific conditions. We therefore present the fixed-effects regressions as the baseline empirical framework and add a conservative interpretation, while recommending bootstrap and alternative second-stage specifications as robustness extensions.

The contribution of this paper is threefold. First, it clarifies how CPI differs conceptually from broadband, general digital infrastructure, cloud-service penetration, artificial-intelligence infrastructure, and data-center construction alone. CPI is defined here as a computing-resource supply system that combines physical computing nodes, storage capacity, interconnection, and scheduling capability. Second, it constructs a UGIP index that incorporates desirable innovation outputs and undesirable environmental outputs, while transparently discussing the proxy allocation of energy use and carbon emissions. Third, it examines human capital and informatization as channel-consistency variables rather than overstating them as fully proven mediators.

The analysis also has broader relevance for sustainability transitions in emerging economies. Many countries are simultaneously expanding digital infrastructure and tightening climate constraints. The Chinese case is informative because it combines rapid growth in data-center deployment, large intercity differences in industrial structure, and explicit national policies linking digitalization with low-carbon development. The evidence therefore speaks to a general governance problem: how cities can prevent computing infrastructure from becoming a new source of energy pressure while using it to improve the efficiency, quality, and diffusion of green innovation.

2. Literature Review and Hypothesis Development

The relevant literature can be organized into five streams: digital infrastructure and innovation, digital infrastructure and environmental sustainability, green innovation performance, computing and AI infrastructure, and urban innovation ecosystems.

2.1. Digital Infrastructure, Computing Infrastructure, and Innovation

The first research stream examines how infrastructure affects innovation. Earlier work emphasizes transport, communication networks, and broadband infrastructure as mechanisms that reduce transaction costs, improve market access, and facilitate knowledge diffusion [1–5]. Recent studies extend this logic to digital infrastructure and show that broadband networks, digital-city construction, and information infrastructure can support regional innovation and environmental performance [16–18,21]. However, these studies often measure digital infrastructure by communication access, digital inclusion, or composite digital-economy indices. CPI differs from these measures because it supplies computational capacity rather than merely communication connectivity. It enables simulation, model training, optimization, digital twins, and high-frequency data processing, all of which are increasingly important for green technology development.

2.2. Digital Infrastructure and Environmental Sustainability

The second stream links digital infrastructure to environmental sustainability. Digital technologies can improve energy management, environmental monitoring, green finance screening, supply-chain coordination, and carbon accounting [14–17,21–25]. At the same time, computing-intensive infrastructures generate direct and indirect environmental costs through electricity demand, cooling-water use, embodied carbon of servers and chips, land occupation, and e-waste [6–8]. Therefore, CPI is theoretically ambivalent: it can enable green innovation but may also create environmental burdens if deployed through extensive, low-utilization, or fossil-fuel-dependent expansion.

2.3. Green Innovation Performance and Efficiency Measurement

The third stream concerns the measurement of green innovation [26]. Patent counts are widely used to capture green innovation output, but they do not directly account for the input costs and undesirable environmental by-products associated with innovation activities. DEA-based efficiency models provide an alternative way to evaluate how cities transform inputs into desirable innovation outputs while controlling for undesirable outputs. The SBM approach is particularly relevant because it directly incorporates slacks in inputs, desirable outputs, and undesirable outputs [27–29]. Compared with a patent-only measure, a UGIP efficiency score is intended to capture the performance with which innovation resources are converted into green innovation outcomes under environmental constraints.

2.4. Computing Infrastructure, AI Infrastructure, and Urban Innovation Ecosystems

The fourth and fifth streams focus on computing infrastructure and urban innovation systems. Knowledge spillover theory suggests that spatial concentration of knowledge assets and technical workers increases innovation through localized learning [30]. Endogenous growth theory highlights knowledge accumulation and technology as engines of long-run growth [31]. Innovation-systems theory emphasizes interactions among firms, universities, governments, and intermediaries [32]. Dynamic-capability theory suggests that organizations need sensing, seizing, and reconfiguring capabilities to transform digital resources into innovation outcomes [33]. These theories jointly imply that CPI should not be treated as a stand-alone physical investment. Its effect on UGIP depends on whether com-

puting resources are integrated with human capital, data circulation, industrial application scenarios, and low-carbon governance.

Compared with the existing literature, this study is positioned at the intersection of computing infrastructure, green innovation performance, and urban sustainability. Previous studies on broadband and communication infrastructure mainly emphasize network access, Internet penetration, or telecom investment, and usually examine innovation output, productivity, or emissions. Studies on the digital economy and smart-city construction often use composite indices that combine platforms, networks, digital finance, and application scenarios, making it difficult to isolate the role of computational resource supply. Research on AI, cloud services, and data platforms provides valuable evidence on firm innovation, energy efficiency, and digital transformation, but relatively limited attention has been paid to city-level evidence on the physical deployment of computing power infrastructure.

In addition, studies of green innovation performance frequently rely on green patents or general environmental efficiency indicators, which may not simultaneously incorporate innovation inputs, desirable innovation outputs, and undesirable environmental outputs. This study addresses these gaps by measuring CPI through IDC-based city-level deployment and a province-level alternative proxy, and by linking CPI to UGIP measured through a GB-U-SBM framework under environmental constraints. In this sense, the paper does not simply replace one digital-infrastructure proxy with another; rather, it examines how computational resource supply, as a distinct infrastructure capability, is associated with the efficiency of urban green innovation under resource and carbon constraints.

2.5. Hypothesis Development

CPI can influence UGIP through factor aggregation, technological acceleration, resource allocation, and environmental governance. From the perspective of factor aggregation, data centers and cloud platforms attract digital firms, platform enterprises, software developers, and high-skilled labor. This agglomeration increases the density of knowledge exchange and expands the set of actors participating in green research and development.

From the perspective of technological acceleration, CPI provides computational support for simulation, optimization, and prediction. Green material design, energy-system dispatch, carbon-emission forecasting, pollution-source identification, and digital-twin production systems all require intensive computation and data processing. By reducing trial-and-error costs and improving the precision of R&D activities, CPI can raise the efficiency with which cities convert innovation inputs into green outputs.

From the perspective of resource allocation, CPI improves the information environment in which governments, firms, financial institutions, and research organizations interact. The integration of computing power with industrial Internet platforms and urban data centers can reduce search costs, support green finance screening, and enable more targeted environmental regulation. These effects improve the matching of green innovation projects with capital, data, talent, and policy support.

Nevertheless, the effect of CPI on UGIP is not automatic. If computing facilities are built through extensive scale expansion, low utilization, or fossil-fuel-intensive electricity use, the environmental costs of CPI may offset part of its innovation-enabling effect. Therefore, the expected relationship depends on whether CPI is embedded in efficient utilization and low-carbon operation systems. Based on the above arguments, this study proposes the following hypothesis:

Hypothesis 1 (H1). *CPI development is positively associated with urban green innovation performance.*

Human capital is an important condition for converting computing resources into green innovation capacity [34]. CPI development increases the demand for data scientists, algorithm engineers, cloud architects, environmental engineers, and interdisciplinary researchers who understand both digital technologies and low-carbon transition. Cities with stronger CPI can provide better platforms for online experimentation, cloud-based simulation, and remote collaborative R&D, thereby enhancing the productivity of high-skilled workers.

Hypothesis 2a (H2a). *CPI development is positively associated with urban human capital accumulation, which is a supportive condition for higher UGIP.*

Informatization refers to the degree to which economic and social activities are supported by information networks, digital platforms, and data circulation systems. CPI provides the hardware and platform foundation for higher informatization because computing resources enable large-scale data collection, storage, processing, and sharing. Higher informatization reduces information asymmetry among firms, research organizations, governments, and financial institutions, thus improving the matching efficiency of green innovation resources.

Hypothesis 2b (H2b). *CPI development is positively associated with urban informatization, which is a supportive condition for higher UGIP.*

3. Materials and Methods

3.1. Sample and Data Sources

This study uses a city-year panel of Chinese prefecture-level and above cities from 2006 to 2023. To avoid confusion caused by different non-missing requirements across specifications, this study distinguishes between the main regression sample and specification-specific samples. The main regression sample contains 5166 observations, corresponding to a balanced panel of 287 cities over 18 years. Some specifications use smaller samples because of additional data requirements: the municipality-exclusion robustness test uses 5094 observations, the one-period lagged specifications use 4879 observations, and the channel-consistency tests use 4892 observations after requiring non-missing channel variables. This observation accounting clarifies the sample basis for each empirical table and prevents confusion between the balanced city panel and specification-specific estimation samples.

Owing to data availability and administrative comparability, Danzhou, Sansha, Haidong, Turpan, Hami, several autonomous prefectures, Chaohu, Laiwu, Tibet, Hong Kong, Macao, and Taiwan are excluded from the core city panel. The main data sources include the China Statistical Yearbook, China City Statistical Yearbook, China Energy Statistical Yearbook, China Environment Statistical Yearbook, the National Intellectual Property Administration, the National Bureau of Statistics, the CSMAR database, CEADs, and the MIIT online administrative service platform.

The city-year panel is used to identify within-city changes in CPI and UGIP after controlling for time-invariant city characteristics and common annual shocks. All regressions include city and year fixed effects. Standard errors are clustered at the city level to account for serial correlation and heteroskedasticity within cities.

3.2. Variable Measurement

3.2.1. Dependent Variable: Urban Green Innovation Performance

UGIP is evaluated with a global benchmark undesirable-output slacks-based measure (GB-U-SBM) model within the data envelopment analysis (DEA) framework, whose foundational CCR formulation was introduced by Charnes, Cooper, and Rhodes [35]. This is the first stage of the empirical framework. The model is non-radial and non-oriented, and it simultaneously considers input excesses, desirable-output shortfalls, and undesirable-output excesses, following the undesirable-output SBM logic [27,36]. The baseline specification assumes variable returns to scale (VRS), because Chinese cities differ substantially in population size, industrial scale, fiscal capacity, and innovation endowments. The global benchmark production possibility set pools all city-year observations over 2006–2023, which makes efficiency scores comparable across years. This procedure generates efficiency scores rather than Malmquist productivity indices [37]. The complete mathematical formulation and implementation details are provided in Appendix A. Descriptive statistics are reported in Table A1.

The input, desirable-output, and undesirable-output indicators are summarized in Table 1, while Appendix A provides further details on the proxy allocation of electricity consumption and carbon emissions.

Table 1. Indicator system for evaluating urban green innovation performance.

Indicator Type	Indicator	Variable Definition
Input	Human input	Full-time equivalent of R&D personnel (hundred persons/year)
	Capital input	Stock of R&D capital (100 million CNY) [38]
Desirable output	Energy input	Electricity consumption in innovation-related activities (100 million kWh)
	Patent output	Weighted number of granted domestic patents of three types (units)
Undesirable output	Carbon emissions	Carbon emissions from innovation-related activities (10,000 tons)

The desirable output is a weighted composite of patent grants. Domestic invention patents, utility model patents, and design patents are assigned weights of 50%, 30%, and 20%, respectively. The rationale is that invention patents usually involve substantive examination and higher technological novelty, utility models capture applied technical improvements, and design patents have relatively lower technical content for green production. This weighting scheme is treated as a transparent baseline assumption rather than as an externally validated technological-value scale. To address potential sensitivity to this judgment-based weighting, we further use a patent-based green innovation measure as an alternative dependent variable in the robustness checks.

3.2.2. Core Explanatory Variable: Computing Power Infrastructure Development

The core explanatory variable is the level of CPI development (dig). CPI is conceptualized as a computing-resource supply network that consists of computing nodes, storage devices, scheduling systems, and high-speed interconnection channels. Data centers are the main spatial carriers of server clusters and computing services. Following this logic, city-level CPI is measured by the cumulative stock of approved data centers based on IDC-related telecommunications business licenses issued by MIIT [39].

This measure should be interpreted as a proxy for locally embedded computing-infrastructure deployment rather than as a direct measure of computing capacity. It captures the spatial establishment of IDC facilities, which is highly relevant for city-level access to data storage, cloud services, and computing nodes. However, it does not directly observe installed server racks, GPU resources, utilization rate, AI-computing capability, power usage effectiveness, or data-center electricity consumption. This study therefore avoids claiming that the proxy fully measures real-time computing capacity. Instead, it treats

IDC license stock as an observable deployment measure and uses province-level CPI as an alternative proxy to account for broader computing-resource sharing within provincial computing networks.

The construction procedure includes four steps. First, historical records of telecommunications business operation licenses are collected from MIIT's online administrative service platform. Second, the raw records are cleaned and deduplicated, and only licenses containing IDC-related business items are retained. Third, server-room location and approval date are extracted. Fourth, newly approved data centers are aggregated to the city-year level and then cumulatively summed from the year in which a city first receives an IDC license.

3.2.3. Channel Variables and Control Variables

Two channel variables are used. Human capital (*hum*) is measured by the number of students enrolled in regular institutions of higher education per 10,000 residents. This indicator captures the local stock of potential high-skilled labor and the educational foundation for green and digital innovation. Informatization (*info*) is measured by Internet broadband access users per 100 residents. This indicator reflects the degree of information-network penetration and the basic digital environment required for data circulation and platform-based innovation.

The control variables include foreign direct investment (*fdi*), measured by the ratio of actually utilized foreign direct investment to GDP; industrial structure (*ind*), measured by the share of secondary-industry value added in GDP; service-sector development (*ser*), measured by the share of tertiary-industry value added in GDP; consumption level (*con*), measured by retail sales of consumer goods relative to GDP; urbanization rate (*urb*), measured by the proportion of urban residents in total residents; and government science-and-technology support (*gov*), measured by science-and-technology fiscal expenditure relative to GDP.

3.3. Econometric Specification

The second stage of the empirical framework uses the UGIP scores obtained from the first-stage GB-U-SBM model as the dependent variable. The baseline specification is a two-way fixed-effects model:

$$UGIP_{it} = \alpha_1 + \beta dig_{it} + \gamma'_1 X_{it} + \mu_i + \lambda_t + \varepsilon_{it} \quad (1)$$

where i and t denote city and year, respectively. $UGIP_{it}$ is urban green innovation performance, dig_{it} is CPI development, X_{it} is the vector of control variables, μ_i denotes city fixed effects, λ_t denotes year fixed effects, and ε_{it} is the error term. The coefficient β is the parameter of interest. Following the two-stage DEA literature, the regression is interpreted as a second-stage contextual-variable analysis rather than as a direct production-frontier estimation.

To examine whether CPI is associated with the proposed channel conditions, the following model is estimated:

$$M_{it} = \alpha_2 + \rho dig_{it} + \gamma'_2 X_{it} + \mu_i + \lambda_t + v_{it} \quad (2)$$

where M_{it} represents human capital or informatization. A positive and significant ρ indicates that CPI development is associated with the improvement of the corresponding channel condition. Because this specification does not estimate the full indirect effect from CPI to UGIP through the channel variable, the results are interpreted as channel-consistency evidence rather than as causal mediation evidence.

Heterogeneity is examined through subsample regressions and permutation-based group-difference tests. The subgroup estimates are interpreted together with the reported group-difference statistics. A significant coefficient in one subgroup and an insignificant coefficient in another subgroup are not treated as sufficient evidence of statistically different effects.

3.4. Robustness, Bootstrap Inference, and Endogeneity Strategy

The empirical analysis uses several robustness and endogeneity checks: replacing the dependent variable with a super-efficiency CCR measure, replacing the efficiency-based UGIP measure with a patent-based green innovation measure, replacing the city-level CPI variable with a province-level CPI indicator, excluding municipalities where applicable, winsorizing the key variables at the 1% level, using one-period lagged regressors, and reporting two instrumental-variable specifications. These checks test whether the baseline association is sensitive to model choice, dependent-variable construction, CPI proxy, sample composition, extreme values, simultaneity concerns, and potential endogeneity.

In response to the two-stage DEA concern, cluster bootstrap inference is reported for the baseline second-stage fixed-effects regression in Appendix B. The bootstrap procedure resamples city clusters and re-estimates the baseline specification over 999 replications, thereby providing an additional check on whether the statistical significance of the CPI coefficient is sensitive to the generated-dependent-variable setting and within-city error dependence. The baseline fixed-effects model is retained because the research question focuses on within-city changes over a long panel rather than cross-sectional efficiency determinants alone. This choice is consistent with the panel nature of the data, while the bootstrap results provide supplementary inference for the main coefficient.

The one-period lag is not treated as a full causal identification strategy. It can reduce contemporaneous simultaneity because CPI planning, approval, construction, and commissioning generally take time, but it cannot fully eliminate omitted time-varying confounders. The two instrumental-variable specifications provide additional evidence on potential endogeneity, but their interpretation still depends on the validity of the relevance condition and exclusion restriction. Therefore, the estimates are interpreted as robust conditional associations supported by lagged, bootstrap, and IV evidence rather than as definitive causal effects.

4. Results

4.1. Baseline Results

Table 2 reports the baseline regression results. Column (1) includes city and year fixed effects without controls. Column (2) reports a pooled specification without fixed effects. Columns (3) to (5) progressively include control variables and fixed effects. In the full specification in Column (5), the coefficient of *dig* is 0.367 and statistically significant at the 5% level, indicating that CPI development is positively associated with UGIP after accounting for city fixed effects, year fixed effects, and city-level covariates. This result supports H1 in the sense of a robust conditional association rather than a fully identified causal effect.

In terms of economic significance, the coefficient of *dig* in Column (5) indicates that a one-standard-deviation increase in CPI is associated with an estimated increase of about 0.00276 in UGIP, which is approximately 0.86% of the sample mean of UGIP. This suggests that the CPI-UGIP association is not only statistically significant but also economically meaningful, although the magnitude should be interpreted conservatively.

Table 2. Baseline estimation results.

Variable	(1)	(2)	(3)	(4)	(5)
dig	0.355 *	0.811 ***	0.599 ***	1.265 ***	0.367 **
	(0.198)	(0.210)	(0.171)	(0.255)	(0.154)
fdi		0.145 ***	0.171 ***	0.140 ***	0.123 ***
		(0.007)	(0.007)	(0.008)	(0.008)
ind		−0.208 ***	0.240 ***	−0.216 ***	0.126 ***
		(0.028)	(0.045)	(0.029)	(0.046)
ser		−0.077 **	0.337 ***	−0.056	0.166 ***
		(0.035)	(0.051)	(0.035)	(0.056)
con		0.007 ***	0.001	0.007 ***	−0.000
		(0.001)	(0.001)	(0.001)	(0.001)
urb		0.028 **	−0.085 ***	0.025 *	−0.098 ***
		(0.013)	(0.012)	(0.013)	(0.012)
gov		−0.213 **	−0.076	−0.268 ***	−0.172 **
		(0.087)	(0.078)	(0.091)	(0.085)
_cons	0.318 ***	0.353 ***	0.033	0.348 ***	0.192 ***
	(0.002)	(0.023)	(0.039)	(0.023)	(0.042)
City FE	Yes	No	Yes	No	Yes
Year FE	Yes	No	No	Yes	Yes
N	5166	5166	5166	5166	5166
R ²	0.693	0.215	0.272	0.132	0.709

Notes: City-clustered standard errors are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

4.2. Robustness and Endogeneity Analyses

Table 3 reports the robustness checks, lagged specifications, and IV estimates. Column (1) replaces the GB-U-SBM dependent variable with a super-efficiency CCR measure. Column (2) uses a patent-based green innovation measure as an alternative dependent variable to examine whether the baseline finding is sensitive to the efficiency-based construction of UGIP. This specification directly addresses the concern that UGIP may be closer to environmental efficiency than to conventional green innovation output. Column (3) replaces city-level CPI with a province-level CPI indicator to reflect broader computing-resource sharing. Column (4) reports the municipality-exclusion robustness where applicable. Column (5) winsorizes the key variables at the 1% level. Columns (6) and (7) use one-period lagged CPI, with and without control variables, respectively. Columns (8) and (9) further report IV estimates based on two alternative instruments for CPI development. Across these specifications, the coefficient of CPI remains positive and statistically significant, indicating that the baseline association is not driven by a single model specification, dependent-variable construction, CPI proxy, sample treatment, outlier treatment, lag structure, or IV specification.

The positive and significant coefficient in Column (2) further shows that CPI development is also associated with patent-based green innovation output. This result strengthens confidence in the main finding because it indicates that the CPI-UGIP relationship is not solely driven by the DEA-based efficiency framework. The larger sample size in Column (3) arises because the province-level CPI proxy can be assigned to the full 287-city balanced panel. The lagged specifications in Columns (6) and (7) mitigate contemporaneous simultaneity concerns by using one-period lagged CPI. The IV estimates in Columns (8) and (9) provide supplementary evidence that the positive association remains after using alternative instruments to alleviate potential endogeneity concerns. Appendix B further reports cluster-bootstrap inference for the baseline two-way fixed-effects specification, and the CPI coefficient remains positive and significant. Nevertheless, the interpretation remains conservative because neither lagged regressors nor IV estimates can fully eliminate all identification concerns.

Table 3. Robustness checks, lagged specifications, and IV estimates.

Variable	(1)	(2)	(3)	(4)	(5)
dig	0.265 ** (0.118)	0.466 *** (0.123)		0.539 *** (0.195)	0.370 ** (0.188)
Dig			0.482 *** (0.029)		
fdi	0.073 *** (0.006)	0.121 *** (0.005)	0.090 *** (0.008)	0.108 *** (0.008)	0.121 *** (0.008)
ind	0.016 (0.033)	0.089 *** (0.029)	0.136 *** (0.045)	0.131 *** (0.046)	0.125 *** (0.045)
ser	0.018 (0.040)	0.135 *** (0.035)	0.155 *** (0.054)	0.163 *** (0.056)	0.155 *** (0.054)
con	0.002 * (0.001)	0.001 * (0.001)	−0.002 * (0.001)	−0.000 (0.001)	−0.000 (0.001)
urb	−0.061 *** (0.009)	−0.054 *** (0.008)	−0.080 *** (0.012)	−0.088 *** (0.013)	−0.098 *** (0.012)
gov	−0.065 (0.060)	0.153 *** (0.054)	−0.112 (0.082)	−0.089 (0.085)	−0.161 ** (0.082)
_cons	0.412 *** (0.030)	0.010 (0.041)	0.198 *** (0.041)	0.188 *** (0.042)	0.197 *** (0.041)
City FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
N	5166	5166	5166	5094	5166
R ²	0.779	0.822	0.724	0.709	0.718
Variable	(6)	(7)	(8)	(9)	
L.dig		0.352 * (0.206)	0.407 ** (0.202)		
dig				0.838 ** (0.394)	0.229 *** (0.038)
fdi			0.121 *** (0.009)	0.072 (0.058)	0.109 *** (0.016)
ind			0.146 *** (0.050)	0.025 (0.295)	0.099 (0.090)
ser			0.178 *** (0.060)	−0.166 (0.384)	0.076 (0.110)
con			−0.000 (0.001)	−0.019 * (0.011)	−0.005 ** (0.002)
urb			−0.095 *** (0.013)	−0.161 * (0.084)	−0.115 *** (0.024)
gov			−0.223 ** (0.089)	−3.162 ** (1.508)	−0.979 *** (0.214)
_cons		0.324 *** (0.002)	0.183 *** (0.045)		
City FE		Yes	Yes	Yes	Yes
Year FE		Yes	Yes	Yes	Yes
N		4879	4879	5166	5166
R ²		0.694	0.709		
Anderson canon. corr. LM statistic				4.898 **	50.378 ***
Cragg–Donald Wald F statistic				4.608	47.811

Notes: Column (1) uses an alternative DEA-based dependent variable; Column (2) uses a patent-based green innovation measure as an alternative dependent variable; Column (3) uses province-level CPI; Column (4) reports the municipality-exclusion robustness where applicable; Column (5) winsorizes key variables; Columns (6) and (7) report one-period lagged CPI specifications without and with controls, respectively; Columns (8) and (9) report IV estimates using historical telecommunications base $\times$ national CPI expansion and terrain ruggedness $\times$ national CPI expansion as instruments for CPI development, respectively. City-clustered standard errors are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

For the IV specifications, we construct two instruments following the logic of shift-share, or Bartik-type, identification. This approach commonly interacts a pre-determined local exposure variable with an aggregate time-varying shock, so that the identifying variation comes from differential local exposure to a common national trend [40,41]. The first instrument is constructed as the interaction between a city's pre-sample telecommunications infrastructure base and national CPI expansion. Cities with stronger historical telecommunications conditions are more likely to host subsequent IDC facilities, cloud-service nodes, and computing-related infrastructure because they generally have better network access, lower connection costs, and stronger digital-service foundations. Since the telecommunications base is measured before the main sample period and the national CPI expansion is common across cities, this instrument captures heterogeneous exposure to national computing-infrastructure growth rather than contemporaneous local innovation shocks.

The second instrument is constructed as the interaction between terrain ruggedness and national CPI expansion. Terrain conditions are predetermined geographical characteristics that affect the construction cost, site selection, power-grid connection, and network deployment feasibility of data centers and IDC facilities. The use of terrain-based variables as predetermined geographical constraints is common in applied empirical research, because such variables are plausibly exogenous to short-run economic and innovation shocks [42]. In this study, the direct level effect of terrain ruggedness is absorbed by city fixed effects. Its interaction with national CPI expansion captures the fact that cities with different geographical construction conditions may be differentially exposed to the same national wave of computing-infrastructure expansion. The validity of the IV strategy rests on the relevance condition that the instruments predict CPI development and the exclusion restriction that, conditional on city fixed effects, year fixed effects, and control variables, they affect UGIP mainly through CPI development. The diagnostic statistics in Table 3 provide mixed but informative evidence. For Column (8), the Anderson canonical correlation LM statistic is 4.898 and significant at the 5% level, indicating that the instrument passes the underidentification test. However, the Cragg–Donald Wald F statistic is 4.608, which is below conventional Stock–Yogo critical values, suggesting potential weak-instrument concerns. For Column (9), the Anderson LM statistic is 50.378 and significant at the 1% level, while the Cragg–Donald Wald F statistic is 47.811, indicating stronger first-stage relevance. Therefore, the IV estimates are interpreted as supplementary evidence rather than as definitive causal proof [43].

4.3. Channel-Consistency Tests

Table 4 examines whether CPI development is associated with the proposed channel variables. Columns (1) and (2) use human capital as the dependent variable, and Columns (3) and (4) use informatization as the dependent variable. Columns (1) and (3) include city and year fixed effects without controls, while Columns (2) and (4) further add the full set of city-level covariates. The coefficients of dig are positive and significant in all columns. These findings indicate that CPI development is associated with stronger human-capital accumulation and higher informatization.

These estimates are consistent with the theoretical expectation that CPI improves the local conditions required for green innovation. Because this subsection focuses on whether CPI is associated with the proposed intermediate conditions, the results are not presented as a fully decomposed mediation effect. This conservative interpretation avoids overstating the statistical evidence while still clarifying how human capital and informatization enter the CPI-UGIP linkage.

Table 4. Channel-consistency tests: CPI, human capital, and informatization.

Variable	(1) Human Capital	(2) Human Capital	(3) Informatization	(4) Informatization
dig	0.055 *** (0.011)	0.034 *** (0.010)	0.334 *** (0.087)	0.273 *** (0.086)
fdi		0.008 *** (0.000)		0.048 *** (0.004)
ind		−0.014 *** (0.002)		−0.020 (0.021)
ser		−0.012 *** (0.003)		0.007 (0.025)
con		0.000 *** (0.000)		0.003 *** (0.001)
urb		−0.004 *** (0.001)		−0.012 ** (0.006)
gov		0.077 *** (0.004)		0.038 (0.037)
_cons	0.002 *** (0.000)	0.010 *** (0.002)	0.006 *** (0.001)	−0.009 (0.019)
City FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	4892	4892	4892	4892
R ²	0.634	0.686	0.389	0.420

Notes: Columns (1) and (2) use human capital as the dependent variable; Columns (3) and (4) use informatization as the dependent variable. Columns (1) and (3) include city and year fixed effects without controls, while Columns (2) and (4) further include the full set of controls. City-clustered standard errors are reported in parentheses. **, and *** denote significance at the 5% and 1% levels, respectively.

4.4. Heterogeneity Heterogeneous Effects Across City Groups

Table 5 reports the heterogeneity estimates and group-difference tests. Columns (1) and (2) show that the coefficient of dig is positive and significant in high-CPI regions but insignificant in low-CPI regions. The group-difference test indicates that the coefficient difference between the two groups is statistically significant, suggesting that CPI development is more strongly associated with UGIP where computing infrastructure and related digital applications have reached a relatively stronger foundation. Columns (3) and (4) further show that the CPI coefficient is positive and significant for coastal cities but insignificant for inland cities. The group-difference test provides marginal evidence of a coastal–inland difference, implying that coastal cities may be better able to convert computing resources into green innovation performance because of stronger industrial bases, denser innovation networks, and better access to skilled labor.

Table 5. Heterogeneity analysis and group-difference tests.

Variable	(1) High CPI	(2) Low CPI	(3) Coastal	(4) Inland
dig	0.935 *** (0.291)	−0.207 (0.268)	1.282 ** (0.575)	0.273 (0.214)
fdi	0.101 *** (0.011)	0.128 *** (0.013)	0.099 *** (0.016)	0.149 *** (0.011)
ind	0.284 *** (0.066)	0.121 * (0.069)	0.070 (0.086)	0.195 *** (0.063)
ser	0.296 *** (0.079)	0.194 ** (0.082)	−0.089 (0.105)	0.272 *** (0.074)
con	−0.001 (0.002)	−0.001 (0.002)	−0.009 *** (0.003)	0.001 (0.001)
urb	−0.073 *** (0.018)	−0.087 *** (0.017)	−0.062 ** (0.026)	−0.098 *** (0.015)
gov	−0.195 (0.131)	−0.319 *** (0.120)	−0.405 * (0.214)	−0.053 (0.097)
_cons	0.043 (0.060)	0.201 *** (0.062)	0.296 *** (0.075)	0.104 * (0.058)
City FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	2514	2652	1433	3733
R ²	0.760	0.747	0.803	0.707

Notes: Columns (1)–(4) correspond to high-CPI, low-CPI, coastal, and inland groups, respectively. City-clustered standard errors are reported in parentheses. Group-difference tests are based on permutation tests with 1000 replications. The estimated coefficient difference for the high-CPI versus low-CPI comparison is 1.142, with a *p*-value of 0.010. The estimated coefficient difference for the coastal versus inland comparison is 1.009, with a *p*-value of 0.050. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

5. Discussion

5.1. Main Findings and Theoretical Implications

The results indicate that CPI development is positively associated with UGIP. This finding extends the literature on digital infrastructure by focusing on computing resources rather than general network access. CPI differs from broadband infrastructure because it is not only a communication channel but also a computing, storage, and algorithmic-processing system. This distinction is important for green innovation because many green technologies depend on simulation, optimization, monitoring, and data-intensive coordination.

The findings also contribute to innovation-systems and dynamic-capability perspectives. Computing resources do not automatically generate green innovation; they must be combined with skilled labor, interoperable data systems, industrial application scenarios, and low-carbon electricity supply. This interpretation explains why the same physical stock of computing facilities can produce different green-innovation outcomes across cities.

The results also suggest that CPI should be understood as a conditional enabling infrastructure rather than as a direct innovation input. Computing resources may reduce the cost of simulation, data processing, and environmental monitoring, but these advantages can be transformed into UGIP only when cities possess complementary absorptive conditions, including skilled labor, digital platforms, industrial application scenarios, and low-carbon energy governance. This interpretation is consistent with innovation-systems and dynamic-capability perspectives: infrastructure expands the technological opportunity set, while local capabilities determine whether such opportunities are converted into measurable green innovation performance.

The channel-consistency tests show that CPI is associated with human capital and informatization. These results support the theoretical expectation that computing infrastructure strengthens the local conditions for green innovation. The discussion therefore separates empirical evidence from theoretical interpretation. The regressions show association with the two channel variables; they do not by themselves prove causal mediation.

The heterogeneity results suggest that regional context matters. Regions with advanced CPI and coastal cities appear to have stronger capacity to transform computing resources into green innovation performance. The group-difference tests further indicate that the estimated effects differ statistically across the reported subgroup comparisons, although these results should still be interpreted as heterogeneous associations rather than causal heterogeneity.

5.2. Policy Implications

First, CPI construction should shift from scale expansion to green and application-oriented deployment. This recommendation follows directly from the positive CPI-UGIP association and from the environmental-cost discussion. Cities should avoid duplicative data-center construction and instead coordinate computing nodes with industrial parks, university research platforms, green technology incubators, and renewable-energy bases.

Second, low-CPI and inland regions should prioritize shared computing platforms and application-oriented nodes rather than large-scale stand-alone data-center projects. This recommendation is aligned with the heterogeneity evidence and group-difference tests, but it should still be implemented cautiously because subgroup differences may reflect broader regional development conditions as well as CPI capacity.

Third, human capital policies should be integrated with CPI investment. The channel-consistency evidence supports the view that computing resources are more productive when matched with skilled workers capable of applying cloud computing, AI, environmental engineering, energy management, and carbon accounting tools.

Fourth, cities should strengthen data governance and informatization systems. CPI can support green innovation only when data are collected, shared, protected, and applied effectively. Urban big-data centers and industrial Internet platforms should integrate R&D, energy-consumption, emission, and technology-transfer information while respecting data security and privacy constraints.

Fifth, the environmental costs of CPI itself must be addressed. The policy discussion avoids unsupported engineering prescriptions and focuses on principles supported by the study: higher utilization, low-carbon electricity procurement, transparent energy-carbon accounting, and coordination between data-center deployment and local renewable-energy capacity.

5.3. Limitations and Future Research

First, although the fixed-effects, lagged, robustness, and IV specifications control for many observable and time-invariant factors and provide supplementary evidence against simple endogeneity concerns, they do not fully solve endogeneity caused by unobserved time-varying policies, industrial dynamics, or innovation shocks. Future research should use stronger quasi-natural experiments, improved instrumental variables, dynamic panel models, or spatial econometric models to strengthen causal identification.

Second, the measurement of UGIP relies on proxy allocation of electricity consumption and carbon emissions to innovation-related activities using R&D expenditure intensity. This approach is necessary under current data constraints but may not capture heterogeneous energy intensity across sectors and innovation activities. Future studies should use direct data on R&D energy use, laboratory electricity consumption, data-center electricity consumption, or firm-level green R&D inputs where available.

Third, CPI is measured by the cumulative stock of approved data centers. This indicator captures spatial deployment but not actual computing capacity, utilization, server type, AI-computing capability, or energy efficiency. Future work can construct a multidimensional CPI index by combining data-center stock, installed computing capacity, cloud-service penetration, intelligent-computing capacity, network latency, and power usage effectiveness.

Fourth, this paper provides channel-consistency evidence rather than a complete mediation decomposition. Future work should report b-path estimates, indirect effects, bootstrap confidence intervals, and additional channels such as industrial upgrading, energy-efficiency improvement, green finance, and knowledge spillovers.

6. Conclusions

This study examines the relationship between CPI development and UGIP using city-level panel data for China from 2006 to 2023. The empirical framework is explicitly two-stage: UGIP is first measured using a GB-U-SBM model, and the estimated efficiency scores are then used in two-way fixed-effects regressions. The results show that CPI development is positively associated with UGIP, and the association remains stable across alternative dependent variables, alternative CPI proxies, sample treatments, winsorization, lagged specifications, bootstrap inference, and two IV specifications.

CPI development is also positively associated with human capital and informatization, suggesting that computing infrastructure contributes to green innovation when it is embedded in talent accumulation and digital information systems. These results are interpreted as channel-consistency evidence rather than causal mediation evidence.

Overall, the evidence supports a cautious but policy-relevant conclusion: CPI can contribute to sustainable urban innovation when its deployment is coordinated with green-energy supply, talent cultivation, data governance, and industrial upgrading. The policy

objective should therefore be to improve the green productivity of computing power rather than to maximize the number of data centers alone.

The heterogeneity results indicate that regions with stronger CPI foundations and coastal cities display more evident positive associations between CPI and UGIP. The group-difference tests support statistically distinguishable subgroup patterns. These findings imply that computing-power policies should be adapted to regional industrial bases, talent conditions, and energy-resource endowments rather than implemented through uniform expansion targets, while stronger causal heterogeneity claims should be left for future research with more demanding identification designs.

Author Contributions: Conceptualization, X.T.; methodology, X.T. and W.G.; formal analysis, X.T.; data curation, W.G. and J.L.; writing—original draft preparation, X.T.; writing—review and editing, X.T., W.G. and J.L.; visualization, J.L.; supervision, X.T.; project administration, X.T.; funding acquisition, X.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the General Program of the National Social Science Fund of China, grant number 24BJY088.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Restrictions apply to the availability of these data. Certain data used in this study were obtained from the China Stock Market & Accounting Research (CSMAR) database and are available to authorized users through the CSMAR database website (<https://data.csmar.com/>), subject to the database provider's applicable access and licensing conditions. The authors are not authorized to publicly redistribute the licensed CSMAR data or the complete compiled dataset containing these data. Other source data used in this study are available from the relevant official statistical yearbooks, the Ministry of Industry and Information Technology, CEADs, the China National Intellectual Property Administration, and the National Bureau of Statistics.

Acknowledgments: The authors thank the data providers and colleagues who offered helpful comments on earlier drafts of this manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Measurement of Urban Green Innovation Performance Using the GB-U-SBM Model

Appendix A.1. Purpose and Model Rationale

This appendix provides the detailed first-stage measurement procedure for urban green innovation performance (UGIP). It clarifies the two-stage DEA design, the reproducibility of the DEA model, the model orientation, the returns-to-scale assumption, the distinction between the global benchmark undesirable-output SBM and a standard period-specific SBM, and the difference between efficiency scores and Malmquist-type productivity indices.

The dependent variable is an efficiency-based measure of green innovation performance rather than a simple count of green innovation output. The model evaluates whether each city-year decision-making unit can generate more innovation output with fewer innovation inputs and lower associated carbon emissions. Traditional CCR and BCC DEA models are radial and may ignore non-proportional input excesses, desirable-output shortfalls, and undesirable-output excesses. The slacks-based measure (SBM) is therefore more suitable for this study because it explicitly incorporates these non-radial slacks. In addition, the global benchmark framework pools all city-year observations into a common production frontier, which makes UGIP scores comparable across years and avoids the problem that period-specific frontiers may generate non-comparable efficiency values.

Appendix A.2. Global Benchmark Production Possibility Set with Alternative Notation

To avoid repeating the notation used in the main-text equations, this appendix adopts a separate notation system. Let k denote a city-year decision-making unit, where $k = 1, \dots, K$ after pooling all cities and years. For each unit k , let p_k be the vector of innovation inputs, q_k be the vector of desirable innovation outputs, and r_k be the vector of undesirable environmental outputs. Let P , Q , and R denote the corresponding pooled matrices constructed from all city-year observations. Let ω be the intensity vector used to form convex combinations of observed city-year units. Under the variable-returns-to-scale assumption, the global benchmark technology is defined as follows:

$$G = \{(p, q, r) | p \geq P\omega, q \leq Q\omega, r \geq R\omega, 1'\omega = 1, \omega \geq 0\}. \quad (A1)$$

In this formulation, the inequalities for inputs and undesirable outputs allow input redundancy and excessive emissions to be identified, whereas the inequality for desirable outputs captures the possibility of innovation-output shortfalls. The constraint $1'\omega = 1$ imposes variable returns to scale, which is appropriate because Chinese cities differ substantially in population size, industrial scale, fiscal capacity, and innovation endowments.

Appendix A.3. Undesirable-Output SBM Efficiency Score

For a target unit e , define Δp as the input-excess slack vector, Δq as the desirable-output shortfall vector, and Δr as the undesirable-output excess vector. The GB-U-SBM score φ_e is obtained by solving the following non-oriented fractional programming problem:

$$\varphi_e = \min \left\{ \frac{1 - \frac{1}{I} \sum_{i=1}^I \left(\frac{\Delta p_i}{p_{ie}} \right)}{1 + \left(\frac{1}{J+L} \right) \left[\left(\sum_{j=1}^J \left(\frac{\Delta q_j}{q_{je}} \right) + \sum_{l=1}^L \left(\frac{\Delta r_l}{r_{le}} \right) \right) \right]} \right\}. \quad (A2)$$

$$s.t. : p_e = P\omega + \Delta p, q_e = Q\omega - \Delta q, r_e = R\omega + \Delta r, 1'\omega = 1, \omega \geq 0, \Delta p \geq 0, \Delta q \geq 0, \Delta r \geq 0. \quad (A3)$$

The score increases when a city-year unit reduces redundant innovation inputs, increases desirable innovation outputs, or reduces undesirable carbon-emission outputs. Under the baseline SBM efficiency setup, a value closer to one indicates that the unit is closer to the global green innovation frontier. If a super-efficiency extension is used in a robustness check, scores may exceed one; this is conceptually distinct from the baseline GB-U-SBM score used as the main dependent variable.

Appendix A.4. Difference from Standard Undesirable-Output SBM and Malmquist Indices

The GB-U-SBM model differs from a standard period-specific undesirable-output SBM in the construction of the reference technology. A standard SBM usually compares a unit with other observations in the same period, so the frontier can move from year to year and efficiency values may not be directly comparable across time. By contrast, the global benchmark approach constructs a single reference frontier from all city-year observations in the sample period. This design is used to obtain intertemporally comparable UGIP scores.

The model also differs from a Malmquist or global Malmquist productivity index [37]. Malmquist-type indices measure productivity changes between adjacent periods and typically decompose productivity growth into efficiency change and technical change. This study does not calculate such productivity-change indices. It calculates city-year efficiency scores under a common global benchmark and then uses those scores as the dependent variable in the second-stage fixed-effects regressions.

Appendix A.5. Indicator Construction

The model treats green innovation as a production process jointly driven by human input, capital input, and energy input. The desirable output is innovation output, and the undesirable output is the carbon-emission burden associated with innovation-related activities. The indicator system used in the GB-U-SBM model is reported in Table 1 in the main text. This appendix further explains the rationale for indicator construction, proxy allocation, and robustness interpretation.

Appendix A.6. Weighting, Measurement Error, and Robustness Interpretation

The 50–30–20% patent-weighting scheme is used as a baseline assumption rather than as an externally validated technological-value scale. Invention patents are given the highest weight because they generally require substantive examination and involve higher technical novelty; utility models capture applied technical improvements; design patents usually contain weaker technological content for green production. To address potential sensitivity to this judgment-based weighting, the robustness checks report a patent-based green innovation measure as an alternative dependent variable.

UGIP should therefore be interpreted as an efficiency-based measure of green innovation performance under environmental constraints rather than as a simple patent-count measure. The desirable output captures innovation output, while the undesirable output captures the carbon-emission burden associated with innovation-related activities. This construction is consistent with the idea that green innovation performance depends not only on how many innovation outputs a city produces, but also on how efficiently it produces them under resource and environmental constraints. To address the concern that this indicator may differ from patent-based green innovation, the robustness checks further use a patent-based green innovation measure as an alternative dependent variable.

The allocation of electricity consumption and carbon emissions by R&D expenditure intensity is also a proxy-based strategy. It is adopted because direct city-level data on R&D electricity use and R&D-related carbon emissions are not systematically available. If the allocation error is classical measurement error in the dependent variable, it mainly increases noise and weakens precision. If the error is systematically related to industrial structure or energy intensity, the estimates should be interpreted with greater caution. For this reason, the manuscript describes UGIP as an efficiency-based green innovation performance measure and interprets the second-stage estimates as robust conditional associations rather than definitive causal effects.

Appendix A.7. Implementation Steps and Link to the Second-Stage Regression

The implementation proceeds in five steps. First, all input, desirable-output, and undesirable-output indicators are assembled at the city-year level. Second, the pooled matrices P , Q , and R are constructed using all valid city-year observations from 2006 to 2023. Third, the global VRS technology in Equation (A1) is imposed. Fourth, Equations (A2) and (A3) are solved for each city-year unit to obtain the UGIP efficiency score. Fifth, the resulting UGIP scores are merged with CPI and control variables and used in the second-stage two-way fixed-effects regressions.

For reproducibility, the DEA model was implemented by solving the global VRS undesirable-output SBM program after pooling all valid city-year observations. Observations with missing input, desirable-output, or undesirable-output variables were excluded before constructing the pooled production matrix. Zero or extremely small denominator values were treated consistently with the actual data-processing procedure before solving the fractional program so that the SBM ratios were well defined. The first-stage UGIP scores were then merged with the CPI and control-variable dataset for the second-stage

two-way fixed-effects regressions. This procedure ensures that the first-stage efficiency scores are constructed under a common global benchmark and are comparable across city-year observations.

Table A1. Descriptive Statistics of the First-Stage GB-U-SBM Efficiency Scores.

Statistic	UGIP
N	5166
Mean	0.321
SD	0.144
Min	0
Max	1
p50	0.297

Appendix B. Bootstrap Inference for the Second-Stage Fixed-Effects Regression

Cluster Bootstrap Robustness Check

Table A2. Cluster bootstrap inference for the baseline two-way fixed-effects regression.

Variable	Bootstrap
dig	0.367 *** (0.078)
fdi	0.123 *** (0.010)
ind	0.126 (0.090)
ser	0.166 *** (0.055)
con	−0.000 (0.002)
urb	−0.098 * (0.050)
gov	−0.172 (0.171)
City FE	Yes
Year FE	Yes
N	5166
Bootstrap reps	999

Notes: This table reports cluster bootstrap-based inference for the baseline two-way fixed-effects specification. Bootstrap standard errors, clustered at the city level, are reported in parentheses. The bootstrap procedure uses 999 replications. City and year fixed effects are included. * and *** denote significance at the 10% and 1% levels, respectively.

References

1. Wang, X.; Fan, G.; Liu, P. Transformation of China's economic growth pattern and growth sustainability. *Econ. Res. J.* **2009**, *44*, 4–16.
2. Liu, S.; Hu, A. Transport infrastructure and China's regional economic integration. *Econ. Res. J.* **2011**, *46*, 72–82.
3. Zhang, J.; Fu, K. Can information network infrastructure construction drive the improvement of urban innovation? A quasi-natural experiment based on the Broadband China pilot. *Ind. Econ. Res.* **2021**, *5*, 1–14.
4. Fan, S.; Zhang, X. Infrastructure and regional economic development in rural China. *China Econ. Rev.* **2004**, *15*, 203–214. [[CrossRef](#)]
5. Fan, X.; Song, D.; Zhao, X. Has infrastructure construction broken domestic market segmentation? *Econ. Res. J.* **2017**, *52*, 20–34.
6. Granoff, I. Nested barriers to low-carbon infrastructure investment. *Nat. Clim. Change* **2016**, *6*, 1065–1071. [[CrossRef](#)]
7. Muller, N.Z. Carbon emissions of infrastructure development. *Environ. Sci. Technol.* **2013**, *47*, 11739–11746. [[CrossRef](#)] [[PubMed](#)]
8. Xiao, T.; Nerini, F.F.; Matthews, H.D.; Tavoni, M.; You, F. Environmental impact and net-zero pathways for sustainable artificial intelligence servers in the USA. *Nat. Sustain.* **2025**, *8*, 1541–1553. [[CrossRef](#)]

9. Ma, J.; Yang, C.; Wang, Y. Impact of carbon emissions trading policy on urban green innovation in China and its heterogeneity analysis. *Econ. Issues* **2025**, *1*, 27–36.
10. Xu, W.; Zhao, Y. Coordinated high-quality development under the dual-carbon targets and digital-intelligent empowerment. *Stat. Res.* **2025**, *42*, 60–75.
11. Yang, Q.; Liu, X.; Sun, S. Spatio-temporal pattern and convergence test of China's scientific and technological innovation efficiency. *J. Quant. Tech. Econ.* **2021**, *38*, 105–123.
12. Chen, J.; Gao, M.; Cheng, S.; Hou, W.; Song, M.; Liu, X.; Liu, Y. Global 1 km $\times$ 1 km gridded revised real gross domestic product and electricity consumption during 1992–2019 based on calibrated nighttime light data. *Sci. Data* **2022**, *9*, 202. [[CrossRef](#)] [[PubMed](#)]
13. Li, H.; Su, Y.; Ding, C.J.; Tian, G.G.; Wu, Z. Unveiling the green innovation paradox: Exploring the impact of carbon emission reduction on corporate green technology innovation. *Technol. Forecast. Soc. Change* **2024**, *207*, 123562. [[CrossRef](#)]
14. Bai, C.; Du, K.; Yu, Y.; Feng, C. Understanding the trend of total factor carbon productivity in the world: Insights from convergence analysis. *Energy Econ.* **2019**, *81*, 698–708. [[CrossRef](#)]
15. Ma, H.; Zou, J.; Qin, C. Environmental regulations implemented by local officials in China: Is there a hometown effect? *Energy Econ.* **2024**, *130*, 107340. [[CrossRef](#)]
16. Tang, K.; Yang, G. Does digital infrastructure cut carbon emissions in Chinese cities? *Sustain. Prod. Consum.* **2023**, *35*, 431–443. [[CrossRef](#)]
17. Yang, Z.; Gao, W.; Han, Q.; Qi, L.; Cui, Y.; Chen, Y. Digitalization and carbon emissions: How does digital city construction affect China's carbon emission reduction? *Sustain. Cities Soc.* **2022**, *87*, 104201. [[CrossRef](#)]
18. Li, C.; Wen, M.; Jiang, S.; Wang, H. Assessing the effect of urban digital infrastructure on green innovation: Mechanism identification and spatial-temporal characteristics. *Humanit. Soc. Sci. Commun.* **2024**, *11*, 320. [[CrossRef](#)]
19. Simar, L.; Wilson, P.W. Estimation and inference in two-stage, semi-parametric models of production processes. *J. Econom.* **2007**, *136*, 31–64. [[CrossRef](#)]
20. Banker, R.; Natarajan, R.; Zhang, D. Two-stage estimation of the impact of contextual variables in stochastic frontier production function models using data envelopment analysis: Second-stage OLS versus bootstrap approaches. *Eur. J. Oper. Res.* **2019**, *278*, 368–384. [[CrossRef](#)]
21. Dong, F.; Li, Y.; Qin, C.; Zhang, X.; Chen, Y.; Zhao, X.; Wang, C. Information infrastructure and greenhouse gas emission performance in urban China: A difference-in-differences analysis. *J. Environ. Manag.* **2022**, *316*, 115252. [[CrossRef](#)] [[PubMed](#)]
22. Luo, S.; Wang, D. The effect of carbon trading policy on provincial total factor energy efficiency. *Econ. Geogr.* **2022**, *42*, 53–61. [[CrossRef](#)]
23. Zhang, W.; Liu, X.; Wang, D.; Zhou, J. Digital economy and carbon emission performance: Evidence at China's city level. *Energy Policy* **2022**, *165*, 112927. [[CrossRef](#)]
24. Guo, X. Intelligent manufacturing, human-machine collaboration and energy use efficiency: Evidence from the application of industrial robots. *Environ. Econ. Res.* **2024**, *9*, 43–65.
25. Guo, J.; Wang, F.; Liu, S.; Wang, D.Y. AI policies empowering the improvement of urban energy efficiency. *China Popul. Resour. Environ.* **2025**, *35*, 24–37.
26. Liu, X.; Lu, X. Developing new-quality productive forces in accordance with local conditions: Growth potential and improvement path of green innovation performance in China's old industrial bases and resource-based cities. *Stat. Res.* **2025**, *42*, 45–60.
27. Tone, K. A slacks-based measure of efficiency in data envelopment analysis. *Eur. J. Oper. Res.* **2001**, *130*, 498–509. [[CrossRef](#)]
28. Johnson, A.L.; Kuosmanen, T. One-stage and two-stage DEA estimation of the effects of contextual variables. *Eur. J. Oper. Res.* **2012**, *220*, 559–570. [[CrossRef](#)]
29. Emrouznejad, A.; Brzezicki, L.; Lu, C. Development and evolution of slacks-based measure models in data envelopment analysis: A comprehensive review of the literature. *J. Econ. Surv.* **2025**, *39*, 2106–2133. [[CrossRef](#)]
30. Audretsch, D.B.; Feldman, M.P. R&D spillovers and the geography of innovation and production. *Am. Econ. Rev.* **1996**, *86*, 630–640.
31. Romer, P.M. Endogenous technological change. *J. Political Econ.* **1990**, *98*, S71–S102. [[CrossRef](#)] [[PubMed](#)]
32. Lundvall, B.-Å. *National Systems of Innovation: Toward a Theory of Innovation and Interactive Learning*; Pinter: London, UK, 1992.
33. Teece, D.J.; Pisano, G.; Shuen, A. Dynamic capabilities and strategic management. *Strateg. Manag. J.* **1997**, *18*, 509–533. [[CrossRef](#)]
34. Montero Guerra, J.M.; Danvila-del-Valle, I.; Mendez-Suarez, M. The impact of digital transformation on talent management. *Technol. Forecast. Soc. Change* **2023**, *188*, 122291. [[CrossRef](#)]
35. Charnes, A.; Cooper, W.W.; Rhodes, E. Measuring the efficiency of decision-making units. *Eur. J. Oper. Res.* **1978**, *2*, 429–444. [[CrossRef](#)]
36. Tone, K. *Dealing with Undesirable Outputs in DEA: A Slacks-Based Measure Approach*; GRIPS Research Report Series; I-2003-0005; National Graduate Institute for Policy Studies: Tokyo, Japan, 2003. [[CrossRef](#)]
37. Pastor, J.T.; Lovell, C.A.K. A global Malmquist productivity index. *Econ. Lett.* **2005**, *88*, 266–271. [[CrossRef](#)]

38. Wu, Y. R&D stock, knowledge function and production efficiency. *China Econ. Q.* **2006**, *5*, 1129–1156.
39. Sun, B.; Yu, B.; Hu, Y. Does computing power infrastructure construction inhibit enterprises' shift from the real to the virtual economy? Evidence from data center construction in Chinese prefecture-level and above cities. *J. Financ. Res.* **2025**, *7*, 95–112.
40. Goldsmith-Pinkham, P.; Sorkin, I.; Swift, H. Bartik instruments: What, when, why, and how. *Am. Econ. Rev.* **2020**, *110*, 2586–2624. [[CrossRef](#)]
41. Borusyak, K.; Hull, P.; Jaravel, X. Quasi-experimental shift-share research designs. *Rev. Econ. Stud.* **2022**, *89*, 181–213. [[CrossRef](#)] [[PubMed](#)]
42. Nunn, N.; Puga, D. Ruggedness: The blessing of bad geography in Africa. *Rev. Econ. Stat.* **2012**, *94*, 20–36. [[CrossRef](#)]
43. Stock, J.H.; Yogo, M. Testing for weak instruments in linear IV regression. In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*; Andrews, D.W.K., Stock, J.H., Eds.; Cambridge University Press: Cambridge, UK, 2005; pp. 80–108.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.