sustainability

MDPI

Article

The Nonlinear Relationship Between AI Innovation and Carbon Emission Intensity: Evidence from Chinese Provinces

Shaoqin Shi [1] and Sanmang Wu [2,*]

[1] School of Marxism, China University of Political Science and Law, Beijing 100088, China; shishaoqin@cupl.edu.cn
[2] School of Economics and Management, China University of Geosciences (Beijing), Beijing 100083, China
* Correspondence: wusanmang@sina.com

Abstract

China's pursuit of its dual-carbon targets amid rapid artificial intelligence (AI) development raises an important question: do the environmental implications of AI innovation change as regional innovation advances? Using a balanced panel of 30 Chinese provinces from 2011 to 2024, this study measured patent-based AI innovation intensity using applications identified through a strict AI patent classification. Linear and quadratic models with province and year fixed effects were estimated, and the Lind–Mehlum test was used to assess the shape of the relationship within the observed range. The preferred specification indicates an inverted-U-shaped association: carbon emission intensity initially increases with patent-based AI innovation but declines beyond an interior turning point. The negative quadratic coefficient remains stable when the emissions data source, patent classification, sample period, treatment of outliers, and timing of the AI terms are varied. Supplementary Bartik and copula-control analyses preserve the negative curvature, although their identification limitations preclude a definitive causal interpretation. A Kaya-based exact decomposition shows that the estimated curvature is concentrated in energy intensity rather than the carbonization factor. Human capital strengthens the estimated concavity, while the clearest regional contrast is observed between central and eastern China, with the strongest curvature in the central provinces. These findings suggest that greater AI patenting does not automatically reduce emissions. Its environmental implications depend on the stage of regional innovation and its interaction with energy efficiency and absorptive capacity. Policies promoting AI innovation should therefore be coordinated with cleaner energy supply, efficiency improvements, and human capital investment. More broadly, the study provides a stage-sensitive basis for evaluating the sustainability implications of patent-based AI innovation through measurable changes in carbon emission intensity.

Keywords: patent-based AI innovation; carbon emission intensity; inverted-U-shaped relationship; energy intensity; human capital; regional heterogeneity; China

Check for updates

Academic Editors: Rafael E. Vásquez, Cesar Nieto-Londoño and Zulamita Zapata-Benabithe

Received: 7 July 2026
Revised: 10 August 2026
Accepted: 18 August 2026
Published: 20 August 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) license.

1. Introduction

Reconciling economic development with carbon mitigation remains a central challenge for China's transition toward high-quality growth. This challenge has become more complex with the rapid diffusion of artificial intelligence (AI). AI can improve production planning, energy management, transportation systems, and resource allocation, thereby creating opportunities for reducing emissions per unit of output [1]. At the same time, the

development and deployment of AI require computing facilities, data centers, equipment investment, and supporting digital infrastructure, all of which may increase electricity consumption and associated emissions [2–4]. Whether AI innovation ultimately raises or lowers the carbon emission intensity therefore depends on the relative strength of these competing effects.

Existing empirical research has not reached a uniform conclusion. Some studies associate AI and industrial intelligence with improvements in energy efficiency and carbon performance, whereas others identify additional energy demand, rebound effects, or substantial regional differences [5–13]. These mixed findings suggest that the environmental effect of AI may vary with the stage of technological development. During the initial expansion of AI-related innovation, infrastructure investment and additional computing demand may dominate. As the technology matures and becomes embedded in production and energy management systems, efficiency gains and improved resource allocation may become more important. A linear specification may obscure this transition by imposing a constant marginal effect over the entire AI development range.

Three issues remain insufficiently addressed. First, much of the existing literature measures AI development using industrial robot penetration, digital economy indices, or broad indicators of intelligent transformation. These measures primarily capture technology adoption or general digital development rather than the production of AI-related technological knowledge. Patent applications provide an observable and temporally consistent measure of regional AI innovation output, although they do not constitute a comprehensive measure of the broader innovation environment. In particular, they do not directly measure computing infrastructure, data availability, human capital, or institutional support. The present study therefore used the more precise term "patent-based AI innovation intensity" and interpreted the measure accordingly. Second, evidence based solely on the sign and significance of a quadratic coefficient is insufficient to establish an inverted-U-shaped relationship. A credible nonlinear result additionally requires opposite and statistically meaningful slopes at the two ends of the observed range and a turning point supported by actual sample observations. Existing studies have rarely examined these conditions together or assessed whether the estimated turning point remains inside the sample under alternative identification strategies. Third, the interpretation of energy intensity as a conventional mediator is complicated by the accounting relationship between carbon emissions, energy use, and economic output. Carbon emission intensity can be decomposed exactly into energy use intensity and the carbon content of energy consumption. Treating energy intensity as a statistically independent mediator would therefore mix an accounting identity with an estimated behavioral relationship. A decomposition framework provides a more transparent way of determining whether the estimated nonlinear effect operates primarily through the amount of energy used per unit of output or through changes in the carbonization of energy use.

Against this background, this study examined the relationship between patent-based AI innovation intensity and carbon emission intensity using panel data for 30 Chinese provinces from 2011 to 2024. AI innovation intensity was measured as the natural logarithm of one plus the number of patent applications identified under a strict AI classification. Carbon emission intensity was calculated using provincial carbon emissions and gross regional product expressed using constant 2010 prices. The empirical analysis combined two-way fixed-effects estimation with the Lind–Mehlum test, alternative variable and sample specifications, and two complementary approaches to potential endogeneity. The first is an internal-structure Bartik design that allows for interaction between predetermined provincial specialization across four AI technology branches with subsequent patent-family growth outside China. The second is an instrument-free nonparametric copula-control-

function approach. Randomization inference, weak-identification diagnostics, province-clustered bootstrap inference, and leave-one-province-out analysis were used to assess the limitations of these approaches.

The baseline estimates indicate a concave relationship between patent-based AI innovation intensity and carbon emission intensity. The Lind–Mehlum test supports an inverted-U-shaped pattern over the observed sample range, with observations located on both sides of the estimated turning point. Alternative measurements, sample treatments, and lag structures retain the negative quadratic term. The internal-structure Bartik and copula-based estimates also produce negative curvature and an interior point estimate, although the precision of the turning point differs between identification strategies and the Bartik estimates remain subject to weak-identification concerns. The accounting decomposition attributes most of the estimated curvature to the energy-use-intensity component rather than to the carbonization component. Human capital makes the conditional curve more pronounced, while the regional estimates indicate that the nonlinear relationship is most clearly identified in central China.

This study contributes to the literature in three respects. First, it distinguishes patent-based AI innovation from broader measures of digitalization and technology adoption. The strict patent classification improves the technological specificity of the explanatory variable, while broader patent definitions are retained as sensitivity checks. This approach does not claim to measure the complete regional AI innovation environment; instead, it provides a transparent and reproducible measure of one observable dimension of that environment, namely, AI-related knowledge production. Taken together, this study provides a quantitative framework for evaluating the sustainability implications of patent-based AI innovation through measurable changes in carbon emission intensity, thereby linking innovation measurement to sustainability assessment. Second, the study evaluates nonlinearity through curvature and sample support. It supplements the quadratic fixed-effects model with endpoint-slope tests, an explicitly estimated turning point, observations on both sides of that point, and complementary identification strategies. The resulting evidence is therefore not based solely on the statistical significance of the squared term. Third, the analysis replaces conventional energy-intensity mediation with an exact accounting decomposition of carbon emission intensity. This separates the energy-use and carbonization components and avoids interpreting a mechanical accounting relationship as an independent causal pathway. The subsequent analysis of human capital and regional heterogeneity examines the conditions under which the estimated nonlinear relationship is more pronounced.

The remainder of the paper is organized as follows: Section 2 reviews the relevant literature. Section 3 develops the theoretical framework and research hypotheses. Section 4 describes the variables, data, and empirical specifications. Section 5 presents the baseline estimates, robustness checks, and identification analyses. Section 6 reports the decomposition, moderation, and regional heterogeneity results. Section 7 discusses the findings and their limitations. Section 8 concludes and presents the policy implications.

2. Literature Review

2.1. AI Measurement and Carbon Outcomes

A growing body of research has examined the environmental consequences of artificial intelligence and related digital technologies. Studies based on industrial and provincial data have generally found that AI or industrial intelligence can reduce carbon intensity by improving production efficiency, facilitating technological upgrading, and reallocating production factors [5,6]. Related evidence suggests that industrial intelligence may accelerate the transition toward peak emissions [7]. At the provincial level, AI has also been

associated with reductions in local and neighboring carbon emissions [8], while system-GMM estimates indicate that technological progress and industrial restructuring constitute important mitigation channels [9].

The empirical literature nevertheless differs substantially in how AI is measured. Some studies use industrial robots as an indicator of intelligent production [10,14], whereas others rely on city-level AI applications [15], composite indicators incorporating several dimensions of AI development [8], or broader measures of the digital economy [16]. These measures primarily capture downstream adoption, industrial application, or the wider process of digital transformation. They do not necessarily measure the production of AI-related knowledge. This distinction matters because applied AI and innovative AI may have different environmental consequences. Dong et al. [11], for example, report that applied AI is associated with lower emissions, whereas innovative AI displays a nonlinear relationship over its development cycle.

Patent data provide an observable measure of codified inventive output and are widely used to compare technological activity between regions and over time [17,18]. They are therefore appropriate for examining the upstream knowledge-production dimension of AI innovation. However, patent applications do not constitute a complete measure of the regional innovation environment. In particular, they do not directly capture computing infrastructure, data availability, skilled labor, financing conditions, or institutional support. Accordingly, the explanatory variable used in this study is interpreted more narrowly as patent-based AI innovation intensity, rather than as a comprehensive index of the AI innovation environment.

The study further adopted a strict patent classification that focuses on inventions whose technological content is directly associated with core AI functions. This restriction reduces the risk of including conventional software, general digital systems, and commercial data-processing patents that are only loosely connected with AI. A broader patent definition is retained for robustness analysis. The distinction between strict and broad measures allows the analysis to prioritize technological specificity while evaluating whether the main results depend on the patent classification boundary [19].

2.2. Nonlinear Effects and the Decomposition of Carbon Emission Intensity

The environmental effect of AI need not be constant across stages of technological development. The construction of computing infrastructure, the expansion of data-processing capacity, and the digital retrofitting of production systems may initially increase electricity use and embodied energy demand. At later stages, technological diffusion, process optimization, predictive maintenance, and improved resource allocation may strengthen the efficiency-enhancing effects of AI. Existing evidence is consistent with the coexistence of these forces. Production-scale expansion can offset part of the emission reductions associated with technological progress and industrial restructuring [9]. Nonlinear effects have also been documented for industrial robots [10], innovative AI [11], and the digital economy [16].

These findings provide empirical motivation for a stage-dependent relationship, but they do not establish that every AI indicator must follow an inverted-U-shaped path. A negative coefficient on a squared term is not sufficient on its own to demonstrate such a relationship. The slopes at both ends of the observed explanatory-variable range and the location of the turning point must also be considered [20,21]. This distinction is important because previous studies frequently infer nonlinearity from coefficient signs without establishing whether the turning point is supported by observations within the sample.

The channels connecting AI innovation with carbon emission intensity also require careful definition. AI may affect energy demand and energy efficiency through

green technological progress, production scheduling, and energy management applications [12,13,22]. Meanwhile, decomposition studies show that changes in China's carbon emission intensity reflect the amount of energy used per unit of output and the carbon content of the energy mix [23,24].

These components are linked through the accounting identity $\frac{C_{it}}{Y_{it}} = \frac{E_{it}}{Y_{it}} \times \frac{C_{it}}{E_{it}}$, where $\frac{C_{it}}{Y_{it}}$ is the carbon emission intensity, $\frac{E_{it}}{Y_{it}}$ is the energy use intensity, and $\frac{C_{it}}{E_{it}}$ is the carbonization factor. Consequently, energy intensity is not statistically independent of the carbon emission intensity: the two indicators share the same output denominator, and carbon emissions are themselves calculated from energy consumption. Treating energy intensity as an ordinary mediator may therefore confound an accounting relationship with a behavioral transmission mechanism. Following the decomposition literature [25], the present study instead separated the estimated relationship into an energy use component and a carbonization component. This approach identifies whether the nonlinear pattern arises primarily from changes in the quantity of energy used or from changes in the carbon content of that energy.

2.3. Human Capital, Regional Heterogeneity, and Research Gaps

The environmental consequences of AI innovation may also depend on regional absorptive capacity. The absorptive capacity literature emphasizes that the benefits of new knowledge depend on the ability to recognize, assimilate, and apply it [26]. Human capital facilitates technological diffusion and enables regions to adapt general technologies to local production conditions [27,28]. It may therefore affect not only the average environmental effect of AI innovation but also the shape of the relationship. Regions with stronger human capital foundations may initially expand AI-related investment and application more rapidly, while also possessing a greater capacity to convert mature AI technologies into energy efficiency improvements.

Related studies connect AI and digital transformation with green economic growth and green economic efficiency [29,30]. However, these findings do not establish that human capital produces a uniformly emission-reducing effect. In a nonlinear setting, human capital may alter the slope and curvature of the relationship. It should therefore be treated as a boundary condition rather than simply as an additional transmission variable.

Regional differences provide a further boundary condition. Existing studies report substantial geographical variation in the carbon effects of AI, although the estimated regional patterns are not uniform [8,9,11]. Differences in industrial composition, energy structure, technological capability, and development stage may cause otherwise similar increases in AI activity to produce different environmental outcomes. The broader literature on regional development and carbon emissions likewise shows that China's provinces follow heterogeneous transition paths [31].

Three gaps follow from this literature: First, much of the existing research measures industrial intelligence, robot adoption, AI applications, or the broader digital economy. Less attention has been paid to the environmental consequences of upstream AI knowledge production measured using a technologically strict patent classification. Second, although nonlinear effects have been reported, the empirical evidence often relies primarily on the sign of the quadratic coefficient and does not formally verify the endpoint slopes and the sample location of the turning point. Third, energy intensity has frequently been interpreted as a conventional mediation variable despite its accounting relationship with carbon emission intensity.

The present study addressed these gaps by examining patent-based AI innovation intensity, formally testing the nonlinear relationship over the observed sample range,

decomposing carbon emission intensity into energy-use and carbonization components, and evaluating human capital and regional location as boundary conditions. This design does not treat patent counts as a complete measure of the AI innovation ecosystem; rather, it isolates one observable and comparable dimension, namely, the intensity of AI-related inventive activity, and investigates its association with provincial carbon emission intensity.

3. Theoretical Framework and Research Hypotheses

The proposed relationship is organized around two competing forces: (1) the expansion effect arises from the additional computing capacity, experimentation, capital investment, and organizational adjustment associated with the development of AI; (2) the efficiency effect arises when accumulated knowledge is translated into process optimization, improved energy management, cleaner technological change, and more efficient resource allocation. The observed relationship with carbon emission intensity depends on which of these forces dominates at a given stage of AI development.

3.1. Expansion and Efficiency Effects

AI has the characteristics of a general purpose technology: it can be applied across a wide range of sectors, but its economic benefits depend on complementary investments and organizational adaptation [32]. Such investments often precede measurable productivity gains. The productivity J-curve provides a temporal interpretation of this process: implementation costs and intangible investment are concentrated in the early stage, whereas the benefits emerge only after complementary capabilities have been accumulated [33].

At relatively low levels of AI innovation, the expansion effect is likely to dominate. AI-related research and experimentation require computing equipment, data storage, model training, and supporting digital infrastructure. These activities can increase electricity consumption before corresponding improvements in output or energy efficiency are realized [2–4]. Existing production systems and new digital systems may also operate simultaneously during the transition, adding to energy demand. Moreover, efficiency improvements at the task level need not reduce the aggregate energy consumption if lower operating costs stimulate greater production or wider technology use. This rebound mechanism can partially offset initial efficiency gains [34].

The balance may change as AI-related knowledge accumulates and diffuses into production. Mature applications can improve production scheduling, predictive maintenance, energy system management, transport coordination, and the allocation of material inputs. These uses reduce avoidable energy consumption and raise output per unit of energy. Innovation may also be redirected toward cleaner inputs and low-carbon production processes when environmental incentives and technological capabilities become sufficiently strong [35]. Empirical studies accordingly find that AI applications can reduce energy intensity through technological progress [36] and improve environmental performance through energy efficiency and green technological development [37].

The expansion and efficiency effects therefore operate with different timings. At an early stage, complementary investment and additional energy demand may outweigh efficiency gains, producing a positive marginal association with carbon emission intensity. At a more advanced stage, technological diffusion and operational efficiency may become dominant, causing the marginal association to turn negative. This argument leads to the following hypothesis:

Hypothesis 1 (H1). *Patent-based AI innovation intensity has an inverted-U-shaped relationship with regional carbon emission intensity: carbon emission intensity initially increases with AI innovation intensity but declines after AI innovation reaches a sufficiently advanced level.*

3.2. Energy-Use and Carbonization Margins

The relationship between AI innovation and carbon emission intensity can be examined through the accounting structure of emissions. Carbon emission intensity can be written as $\frac{C_{it}}{Y_{it}} = EI_{it} \times CF_{it}$, where EI_{it} denotes energy intensity and CF_{it} denotes the carbonization factor, measured as carbon emissions per unit of aggregate energy consumption. In logarithmic form, changes in carbon emission intensity are exactly divided between these two components [25].

The expansion and efficiency effects are expected to appear most directly in energy use intensity. During the expansion stage, computing infrastructure, digital retrofitting, and the parallel operation of old and new production systems can increase energy use relative to output. At a later stage, intelligent scheduling, process optimization, predictive maintenance, and improved resource allocation can reduce energy requirements per unit of output [36,37].

The carbonization factor may also respond to AI innovation, but the connection is less direct. AI can support renewable energy integration, grid management, and the development of cleaner technologies, thereby reducing the carbon content of energy use. However, these changes also depend on the regional energy mix, infrastructure, regulation, and access to clean energy. The carbonization component may therefore be weaker or less stable than the energy use component.

This decomposition is not interpreted as a conventional mediation model. Energy and carbon emission intensities share the same output denominator, while carbon emissions are themselves calculated from energy consumption. Treating energy intensity as an independent mediator would therefore mix an accounting identity with a behavioral mechanism. The empirical analysis instead estimates the two components separately and examines how their coefficients combine to produce the total relationship.

Hypothesis 2 (H2). *The nonlinear relationship between patent-based AI innovation intensity and carbon emission intensity is expected to be reflected primarily in the energy use intensity component, while the contribution of the carbonization factor component is expected to be comparatively limited and more dependent on regional energy conditions.*

3.3. Human Capital as a Boundary Condition

The environmental consequences of AI innovation depend not only on the stock of technological knowledge but also on the capacity to absorb and apply it. Absorptive capacity theory emphasizes that the ability to recognize, assimilate, and use external knowledge rests on the recipient's existing knowledge base [26]. Human capital also facilitates technological diffusion and reduces the distance between new knowledge and productive application [27,28,38,39].

Human capital is examined not because it is the only possible moderator, but because it is most directly related to the conversion of codified technological knowledge—the dimension captured by patent data—into operational changes in production and energy management. Marketization, environmental regulation, and financial development may also condition this process, but they represent complementary institutional channels rather than the specific absorptive capacity mechanism examined here.

Higher human capital can strengthen both stages of the proposed relationship. At relatively low levels of AI innovation, regions with stronger skills and technical capabilities may undertake complementary investment, experimentation, and digital transformation more rapidly. This can amplify the initial expansion effect. Once AI innovation becomes sufficiently developed, the same capabilities allow technologies to be integrated more effectively into production, energy management, and environmental governance. Human

capital can therefore accelerate the realization of efficiency gains and strengthen the decline in carbon emission intensity at higher AI innovation levels.

In the quadratic specification, this argument implies a positive interaction between human capital and the linear AI term and a negative interaction between human capital and the squared AI term. It does not impose a predetermined direction on the movement of the turning point. The location of the turning point depends jointly on both interaction coefficients and should therefore be determined from the estimates rather than assumed in advance.

Hypothesis 3 (H3). *Human capital strengthens the stage-dependent relationship between patent-based AI innovation intensity and carbon emission intensity: it amplifies the initial positive slope while making the subsequent downward curvature more pronounced.*

3.4. Regional Heterogeneity

The balance between the expansion and efficiency effects is unlikely to be uniform across China. Provinces differ in industrial composition, technological maturity, energy systems, and their capacity to absorb and apply AI-related knowledge [40,41]. Existing evidence also shows that the environmental consequences of AI and digital technologies vary substantially between regions [8,9,11,31].

Eastern provinces generally possess more developed digital infrastructure, stronger innovation systems, and greater experience in applying advanced technologies. Their AI-related activities may therefore be closer to the stage in which expansion and efficiency effects partially offset one another, producing a comparatively flatter relationship. Central provinces combine a substantial manufacturing base with continuing digital upgrading. The costs of technological expansion and the gains from process improvement are likely to be visible, making the stage transition more pronounced. In western provinces, smaller AI innovation bases and more limited absorptive capacity may weaken the precision and magnitude of the estimated relationship.

These differences do not imply that every province within a region follows the same trajectory; rather, they indicate that the relative strength of expansion and efficiency effects is conditioned by regional development stages and production structures.

Hypothesis 4 (H4). *The relationship between patent-based AI innovation intensity and carbon emission intensity exhibits regional heterogeneity. The inverted-U-shaped pattern is expected to be more pronounced in central China, comparatively flatter in eastern China, and less distinct in western China.*

4. Variable Selection and Model Construction

4.1. Variable Selection

4.1.1. Carbon Emission Intensity

The dependent variable was carbon emission intensity, defined as carbon dioxide emissions per unit of real output. Its natural logarithm was used in the regressions.

Provincial carbon-emission data for 2011–2022 were obtained from the China Emission Accounts and Datasets (CEADs) [42,43]. CEADs compiles provincial emission inventories primarily from fossil fuel consumption reported in provincial energy balance sheets. It follows the IPCC sectoral accounting framework and employs emission factors calibrated to the characteristics of fuels used in China [44–46]. Its energy classifications and accounting boundaries are closely aligned with China's energy statistical system and remain consistent over the sample period. The database has consequently been widely used in empirical studies of provincial carbon emissions in China [44]. Subsequent work has updated the

provincial carbon accounts for more recent years [47]. When the dataset for this study was assembled, the publicly available CEADs provincial inventory extended through 2022. Provincial emissions for 2023 and 2024 were therefore extrapolated using province-specific emissions-to-energy conversion coefficients anchored to the 2022 observations. The complete emission series is defined as

$$CE_{it}^{*} = \begin{cases} CE_{it}^{CEADs}, \ t = 2011, \ldots, 2022, \\ \widehat{CE}_{it} = \varphi_i E_{it}, \ t = 2023, \ 2024, \end{cases} \tag{1}$$

where

$$\varphi_i = \frac{CE_{i,\,2022}^{CEADs}}{E_{i,2022}} \tag{2}$$

Here, E_{it} denotes total energy consumption in province i and year t, and φ_i is the province-specific emissions-to-energy conversion coefficient calculated for 2022. Energy consumption data for 2023 and 2024 were obtained from the relevant statistical yearbooks and were expressed using the same units and accounting conventions as the energy consumption data used to construct the 2022 coefficients. No cross-province mean imputation was employed.

Because the conversion coefficients were calculated separately for each province, this approach preserves cross-provincial heterogeneity in energy composition and effective emissions intensity. By construction, applying φ_i to the corresponding 2022 energy consumption observation reproduced the CEADs value for that province, thereby avoiding a mechanical level discontinuity at the point where the data source changes. The central assumption underlying the extrapolation was that each province's emissions-to-energy conversion coefficient remained approximately stable over the short 2023–2024 extrapolation period. Changes in fuel composition, fuel quality, or accounting boundaries may introduce prediction errors. The direction of such errors was not imposed in advance. Instead, sensitivity to this assumption was assessed through a backcasting exercise around the anchor year and a regression based on a truncated sample.

For the backcasting exercise, the province-specific coefficients constructed from the 2022 data were applied to 2021 energy consumption, and the resulting estimates were compared with the CEADs values reported for 2021 across all 30 sample provinces. The 2021 CEADs emission observations were not used when constructing the coefficients. The mean absolute percentage error is 5.32%, and the median absolute percentage error is 2.99%. The correlation between the reconstructed and CEADs series is 0.997. At the aggregate level, the reconstructed national total is 0.85% lower than the corresponding CEADs total, while 27 of the 30 provinces have absolute percentage errors below 10%. These results indicate that the proportional conversion method performs well in a year adjacent to the anchor year, although prediction errors remain for individual provinces. The exercise therefore provides evidence of short-run stability in the conversion coefficients but does not eliminate the possibility of measurement error in the extrapolated values for 2023 and 2024.

Carbon emission intensity is defined as

$$CI_{it} = \frac{CE_{it}^{*}}{RGDP_{it}^{2010}} \tag{3}$$

where $RGDP_{it}^{2010}$ denotes the provincial gross regional product measured at constant 2010 prices. The variable used in the regression analysis was

$$lnCI_{it} = ln\left(\frac{CE_{it}^{*}}{RGDP_{it}^{2010}}\right) \tag{4}$$

Using real output in the denominator removes changes in nominal output attributable solely to price movements and improves the comparability of the carbon emission intensity between years.

Two additional checks addressed potential sensitivity to the construction of the dependent variable. First, the baseline model was re-estimated using an alternative carbon emission intensity measure constructed from the EDGAR database [48], thereby examining sensitivity to the emissions data source and accounting methodology. Second, a more conservative truncated sample covering 2011–2021 was used. This specification excluded the 2022 anchor year and the extrapolated observations for 2023 and 2024, which allowed us to determine whether the main result was already present before the data-extension procedure was introduced. EDGAR is designed to ensure methodological consistency and comparability between countries, whereas the provincial accounting structure of CEADs is more closely aligned with China's energy balance system. CEADs was therefore retained as the primary data source, with EDGAR used as an alternative measure in the robustness analysis.

4.1.2. Core Explanatory Variable

The core explanatory variable was patent-based AI innovation intensity. Patent-level records were first screened using the strict AI classification described above, and the resulting applications were aggregated by province and application year. The variable was defined as

$$\ln AI_{it} = \ln\left(1 + AI\ Patent_{it}^{strict}\right) \tag{5}$$

where $AI\ Patent_{it}^{strict}$ denotes the number of strict-definition AI patent applications in province i and year t. The addition of one retains province–year observations with no qualifying applications.

Patent applications provide a consistently recorded and technologically disaggregated measure of regional AI-related inventive activity. Application–year counts were preferred to grants, citations, and patent-value indicators because the latter are affected by examination delays, citation windows, and delayed value realization [17,18]. These timing differences were particularly important in the final years of the sample.

The baseline analysis applied a strict technological classification to reduce the inclusion of conventional software, general digital systems, semiconductor technologies, and data-processing patents only loosely related to AI [19]. A broader classification was used in the robustness analysis. A composite innovation-environment index was not adopted because consistent province–year data for its principal components are unavailable over the full sample period, while indicator selection and weighting would introduce additional discretion.

The resulting variable was therefore interpreted as patent-based AI innovation intensity rather than as a complete measure of the regional AI innovation environment. It captures one observable dimension of that environment—AI-related knowledge production—and provides the branch-level technological information required for the internal-structure Bartik design.

4.1.3. Control Variables

To limit omitted-variable bias without overparameterizing the baseline model, the analysis included six controls: urbanization (K_1), the share of secondary industry (K_2), fiscal support (K_3), population density (K_4), trade openness (K_5), and digital infrastructure (K_6). The urbanization rate was the share of urban residents in the year-end resident population, reflecting the level of regional urbanization. Fiscal support was measured as the ratio of local general budgetary expenditure to GDP, capturing the extent of government

fiscal input into economic and social development. Trade openness was the ratio of total goods imported and exported to GDP, indicating a region's participation in international trade. Digital infrastructure was constructed from the texts of provincial government work reports: keywords related to "new digital infrastructure" were first identified from existing reports; the reports were then segmented using the jieba package (version 0.42.1) in Python 3.12, and the ratio of new-digital-infrastructure keyword counts to total word counts was used to measure regional digital-infrastructure development. In addition, given the direct bearing of industrial structure and population agglomeration on emission intensity, the model further controlled for the share of secondary industry (K_2, secondary-industry value added relative to GDP) and population density (K_4, year-end resident population divided by administrative land area).

4.1.4. Decomposition and Moderating Variables

Energy intensity and the carbonization factor were used in the accounting decomposition rather than as conventional mediators. Energy intensity (*EI*) was defined as total energy consumption per unit of real GDP, while the carbonization factor (*CF*) was defined as carbon emissions per unit of energy consumption. The logarithmic sum of these two components is exactly equal to the logarithm of the carbon emission intensity.

Human capital (*HR*) was measured as the number of students enrolled in higher education relative to the year-end resident population, and its natural logarithm was used in the moderation analysis. It was introduced as a moderator of the linear and quadratic AI terms. This specification allows human capital to alter the slope and curvature of the relationship without assuming that it has a uniformly emission-reducing effect. Table 1 summarizes the definitions and data sources of the principal variables used in the empirical analysis.

Table 1. Variable definitions and data sources.

Variable	Symbol	Definition	Data Source
Carbon emission intensity (dependent variable)	$\ln CI_{it}$	Natural logarithm of provincial CO_2 emissions divided by real GDP at constant 2010 prices.	CEADs; China Statistical Yearbook; 2023–2024 estimated by the authors (see Section 4.1.1)
Carbon emission intensity (dependent variable, robustness)	$\ln CI_{it}^{EDGAR}$	Ratio of regional CO_2 emissions to GDP, as natural logarithms, reflecting emissions per unit of economic output.	EDGAR; China Statistical Yearbook
Patent-based AI innovation intensity (core explanatory variable)	$\ln AI_{it}$	Natural logarithm of one plus the annual number of strict-definition AI patent applications.	IncoPat patent database
Urbanization rate (control)	$\ln K_1$	Urban resident population/year-end resident population.	Provincial statistical yearbooks
Share of secondary industry (control)	$\ln K_2$	Secondary-industry value added/GDP.	Provincial statistical yearbooks
Fiscal support (control)	$\ln K_3$	Local general budgetary expenditure/GDP.	Provincial statistical yearbooks
Population density (control)	$\ln K_4$	Year-end resident population/administrative land area.	Provincial statistical yearbooks
Trade openness (control)	$\ln K_5$	Total goods imported and exported/GDP.	Provincial statistical yearbooks

Table 1. *Cont.*

Variable	Symbol	Definition	Data Source
Digital infrastructure (control)	$\ln K_6$	Constructed from provincial government work reports: texts were segmented in Python, and the frequency share of "new digital infrastructure" keywords measures regional digital-infrastructure development.	Provincial government work reports
Energy intensity (decomposition component)	EI	Energy consumption per unit of real GDP at constant 2010 prices; higher values indicate greater energy intensity.	Provincial statistical yearbooks
Human capital (moderator)	$\ln HR_{it}$	Students enrolled in higher education/year-end resident population.	Provincial statistical yearbooks

To examine potential multicollinearity in the nonlinear specification, variance inflation factors (VIFs) were calculated using the within-transformed regressors after accounting for province and year fixed effects. For this diagnostic, the AI variable was mean-centered before its squared term was constructed, thereby reducing the mechanical correlation between the linear and quadratic terms. The results are reported in Table 2.

Table 2. Variance inflation factor diagnostics.

Variable	VIF	Tolerance
Centered AI innovation intensity (c $\ln AI_{it}$)	1.49	0.671
Squared centered AI term ((c $\ln AI_{it})^2$)	2.04	0.490
Urbanization ($\ln K_1$)	1.55	0.646
Secondary industry share ($\ln K_2$)	1.27	0.786
Fiscal support ($\ln K_3$)	1.15	0.868
Population density ($\ln K_4$)	1.28	0.780
Trade Openness ($\ln K_5$)	1.25	0.798
Digital infrastructure ($\ln K_6$)	1.16	0.864
Mean VIF	1.40	-

Notes: VIFs were calculated using regressors residualized with respect to province and year fixed effects. The AI variable was centered on its sample mean before constructing the squared term. Tolerance is the reciprocal of the VIF.

The individual VIF values range from 1.15 to 2.04, with a mean of 1.40, while the tolerance values range from 0.490 to 0.868. All VIF values are below the conventional threshold of 5, and all tolerance values are well above 0.10. These results indicate that multicollinearity is unlikely to materially affect the coefficient estimates. Mean-centering was used only for this diagnostic and does not alter the fitted quadratic relationship.

4.2. Model Construction

4.2.1. Baseline Regression Model

To estimate the conditional relationship, we first specified a linear benchmark:

$$\ln CI_{it} = \beta_1 \ln AI_{it} + \gamma' \ln X_{it} + \lambda_i + \delta_t + \mu_{it} \quad (6)$$

where X_{it} is the vector of time-varying provincial controls, λ_i denotes province fixed effects, δ_t denotes year fixed effects, and μ_{it} is the error term. The quadratic specification is

$$\ln CI_{it} = \beta_1 \ln AI_{it} + \beta_2 \left(\ln AI_{it} \right)^2 + \gamma' \ln X_{it} + \lambda_i + \delta_t + \mu_{it} \quad (7)$$

Province fixed effects absorb time-invariant provincial characteristics, while year fixed effects account for nationwide shocks common to all provinces. Standard errors were clustered at the provincial level to allow for heteroskedasticity and within-province serial correlation [49]. The logarithmic transformations reduce skewness and facilitate proportional interpretation. In the quadratic specification, however, the marginal effect of AI innovation was $\beta_1 + 2\beta_2 \ln AI_{it}$, and therefore, varied with the level of the explanatory variable.

A negative estimate of β_2 indicates concavity but is not sufficient to establish an inverted-U-shaped relationship within the observed range. The Lind–Mehlum test is therefore used to evaluate the slopes at the lower and upper bounds and the sample location of the estimated turning point [20,21].

4.2.2. Kaya-Based Exact Decomposition

Energy intensity is not treated as a conventional mediator in this study. Carbon emission intensity and energy intensity share the same GDP denominator, while carbon emissions are themselves calculated from fossil energy consumption. A conventional mediation model would therefore mix an accounting relationship with an estimated behavioral relationship and could overstate the interpretation of the energy intensity channel. To address this issue, this study adopted a Kaya-type accounting decomposition that separated carbon emission intensity into an energy use component and a carbonization component [25].

Carbon emission intensity can be written as

$$\frac{C_{it}}{RGDP_{it}} = \frac{E_{it}}{RGDP_{it}} \times \frac{C_{it}}{E_{it}} \tag{8}$$

where C_{it} denotes the total carbon emissions, E_{it} denotes the aggregate energy consumption, and $RGDP_{it}$ is the real gross regional product measured at constant 2010 prices. The two components are defined as

$$EI_{it} = \frac{E_{it}}{RGDP_{it}}, CF_{it} = \frac{C_{it}}{E_{it}} \tag{9}$$

where EI_{it} is the energy intensity and CF_{it} is the carbonization factor, reflecting the average carbon content of aggregate energy use. Taking logarithms yields the exact identity

$$\ln CI_{it} = \ln EI_{it} + \ln CF_{it} \tag{10}$$

The decomposition distinguishes between two margins through which patent-based AI innovation may be associated with carbon emission intensity. The energy use component captures changes in the amount of energy consumed per unit of real output, whereas the carbonization component captures changes in the carbon content of energy consumption, principally those arising from the composition of the energy mix.

To quantify the contribution of each component, the same quadratic two-way fixed-effects specification was estimated separately for carbon emission intensity, energy intensity, and the carbonization factor:

$$\ln Y_{it}^{m} = \beta_1^{m} \ln AI_{it} + \beta_2^{m} (\ln AI_{it})^2 + \gamma^{m\prime} X_{it} + \lambda_i + \delta_t + \varepsilon_{it}^{m}, m \in \{CI, EI, CF\} \tag{11}$$

where all three equations use an identical estimation sample, control variables, and province and year fixed effects. Standard errors were clustered at the provincial level. Because

Equation (10) is an exact logarithmic identity and the three equations have the same regressor matrix, the estimated coefficients satisfy

$$\widehat{\beta_k^{CI}} = \widehat{\beta_k^{EI}} + \widehat{\beta_k^{CF}}, k = 1, 2 \tag{12}$$

Accordingly, the linear and quadratic associations estimated for carbon emission intensity can be allocated exactly between the energy-use and carbonization components. This procedure is an accounting decomposition rather than a causal mediation analysis. It does not interpret energy intensity as a statistically independent mediator or estimate an indirect effect through a separate behavioral pathway.

As an additional check on the common-denominator concern, supplementary specifications use total carbon emissions and total energy consumption as dependent variables and include real GDP directly among the regressors. These level-based specifications remove the shared GDP denominator and help determine whether the estimated nonlinear pattern is driven mechanically by the construction of the intensity measures.

4.2.3. Moderation Model

Building on the nonlinear analysis above, human capital was introduced as a moderating variable to examine how regional human capital conditions the effect of AI innovation on carbon emission intensity:

$$\begin{aligned} lnCI_{it} = \alpha_1 \ln AI_{it} + \alpha_2 (\ln AI_{it})^2 + \alpha_{11} (\ln AI_{it} \times \ln HR_{it}) + \alpha_{12} \left[(\ln AI_{it})^2 \times \ln HR_{it} \right] + \alpha_{13} \ln HR_{it} \\ + \gamma \ln X_{it} + \lambda_i + \delta_t + \mu_{it} \end{aligned} \tag{13}$$

where AI_{it} denotes the AI innovation intensity, HR_{it} denotes the human capital, λ_i and δ_t denote the province and year fixed effects, and μ_{it} denotes the error term.

In Equation (13), the interactions with the linear and squared terms capture the joint moderating role of human capital in AI innovation's mitigation effect and its nonlinearity. The significance of α_{11} or α_{12} would indicate that human capital alters the direction, strength, or turning point of the AI innovation effect on emission intensity.

To further characterize the moderation mechanism, differentiating with respect to patent-based AI innovation intensity yields the marginal effect:

$$\frac{\partial lnCI_{it}}{\partial \ln AI_{it}} = \alpha_1 + 2\alpha_2 \ln AI_{it} + \alpha_{11} \ln HR_{it} + 2\alpha_{12} \ln AI_{it} \times \ln HR_{it} \tag{14}$$

This expression describes the marginal impact of AI innovation on emission intensity at different levels of human capital: a positive derivative indicates that AI innovation raises the emission intensity, whereas a negative one indicates that mitigation has set in.

The turning point of the AI–emission intensity relationship follows as

$$\ln AI_{it}{}^* = \frac{-\alpha_1 - \alpha_{11} \ln HR_{it}}{2(\alpha_2 + \alpha_{12} \ln HR_{it})} \tag{15}$$

Differentiating Equation (15) with respect to human capital gives the response of the turning point to the moderator:

$$\frac{\partial \ln AI_{it}{}^*}{\partial \ln HR_{it}} = \frac{2(\alpha_1 \alpha_{12} - \alpha_2 \alpha_{11})}{(2\alpha_2 + 2\alpha_{12} \ln HR_{it})^2} \tag{16}$$

Equation (16) shows that when $\alpha_1 \alpha_{12} - \alpha_2 \alpha_{11} > 0$, rising human capital shifts the turning point of the AI–emission intensity relationship rightward—the marginal mitigation curve is pushed back, meaning AI innovation must reach a higher intensity

before its mitigation effect fully materializes. Conversely, when $\alpha_1\alpha_{12} - \alpha_2\alpha_{11} < 0$, stronger human capital shifts the turning point leftward, implying that AI innovation begins to reduce emissions at a lower intensity. Human capital thus not only shapes the diffusion efficiency and green conversion capacity of AI innovation but also plays a substantive moderating role in the nonlinear process through which AI innovation affects carbon emission intensity.

5. Empirical Results and Robustness Checks

Figures 1 and 2 present the spatial distributions of carbon emission intensity and strict-definition AI patent applications in 2011, 2015, 2020, and 2024. Carbon emission intensity declines across most provinces over the sample period but remains high in several energy-producing provinces, particularly Shanxi, Inner Mongolia, and Ningxia. AI patenting increases substantially and gradually diffuses inland, although activity remains concentrated in the eastern coastal provinces. The spatial distribution of AI patenting therefore differs from that of carbon emission intensity. This contrast motivates the inclusion of province and year fixed effects and the subsequent analysis of regional heterogeneity.

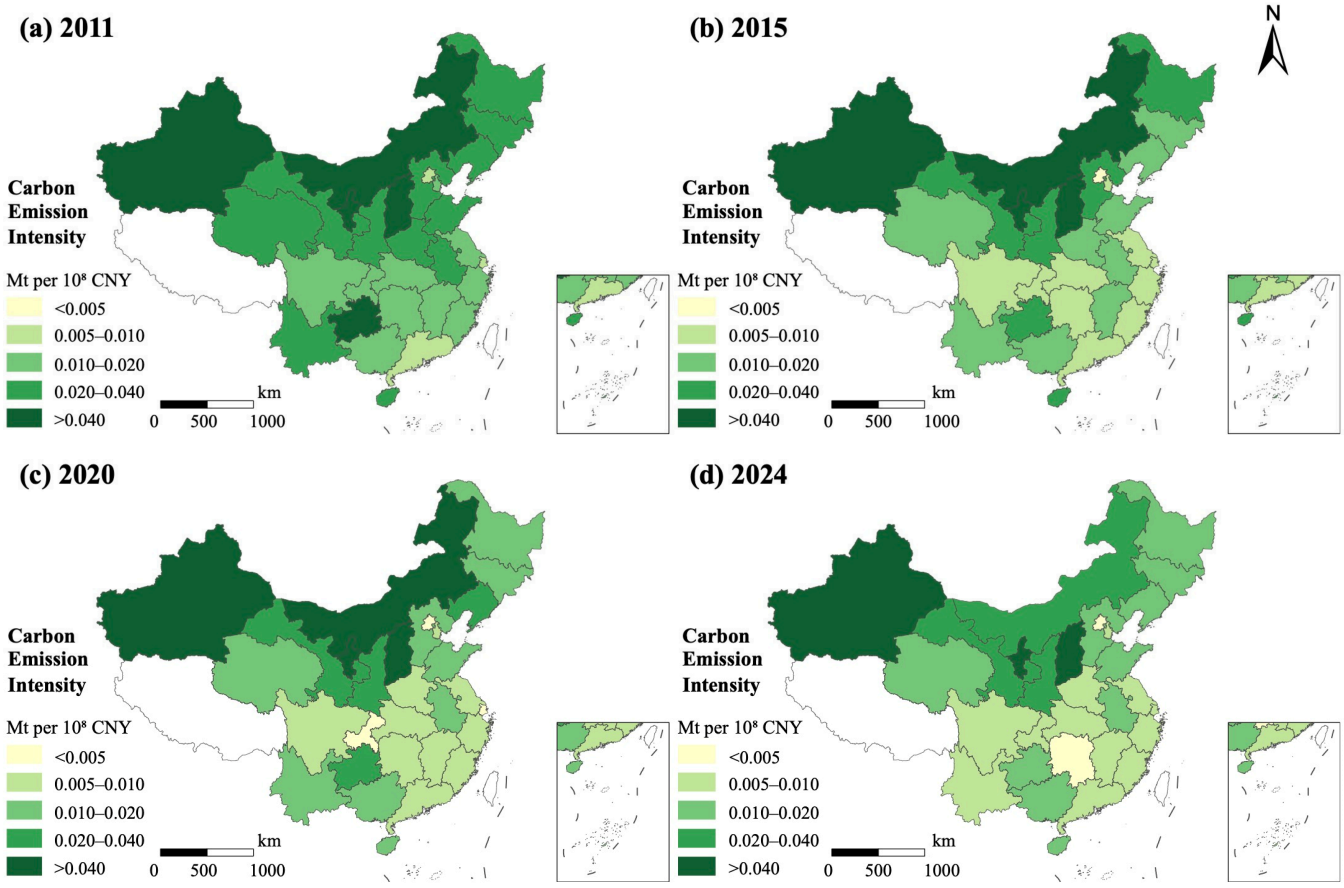


Figure 1. Spatial distribution of carbon emission intensity across provincial-level regions in China (2011, 2015, 2020, and 2024). Notes: Carbon emission intensity was provincial CO_2 emissions divided by real GDP at constant 2010 prices, in Mt per 10^8 CNY. Emission data for 2011–2022 were obtained from CEADs; values for 2023–2024 were extrapolated as described in Section 4.1.1. The five class intervals were identical across all panels. Tibet, Hong Kong, Macao, and Taiwan are outside the estimation sample and are shown unshaded. The base map follows the standard map of the Ministry of Natural Resources of China, with no modifications to the boundaries.

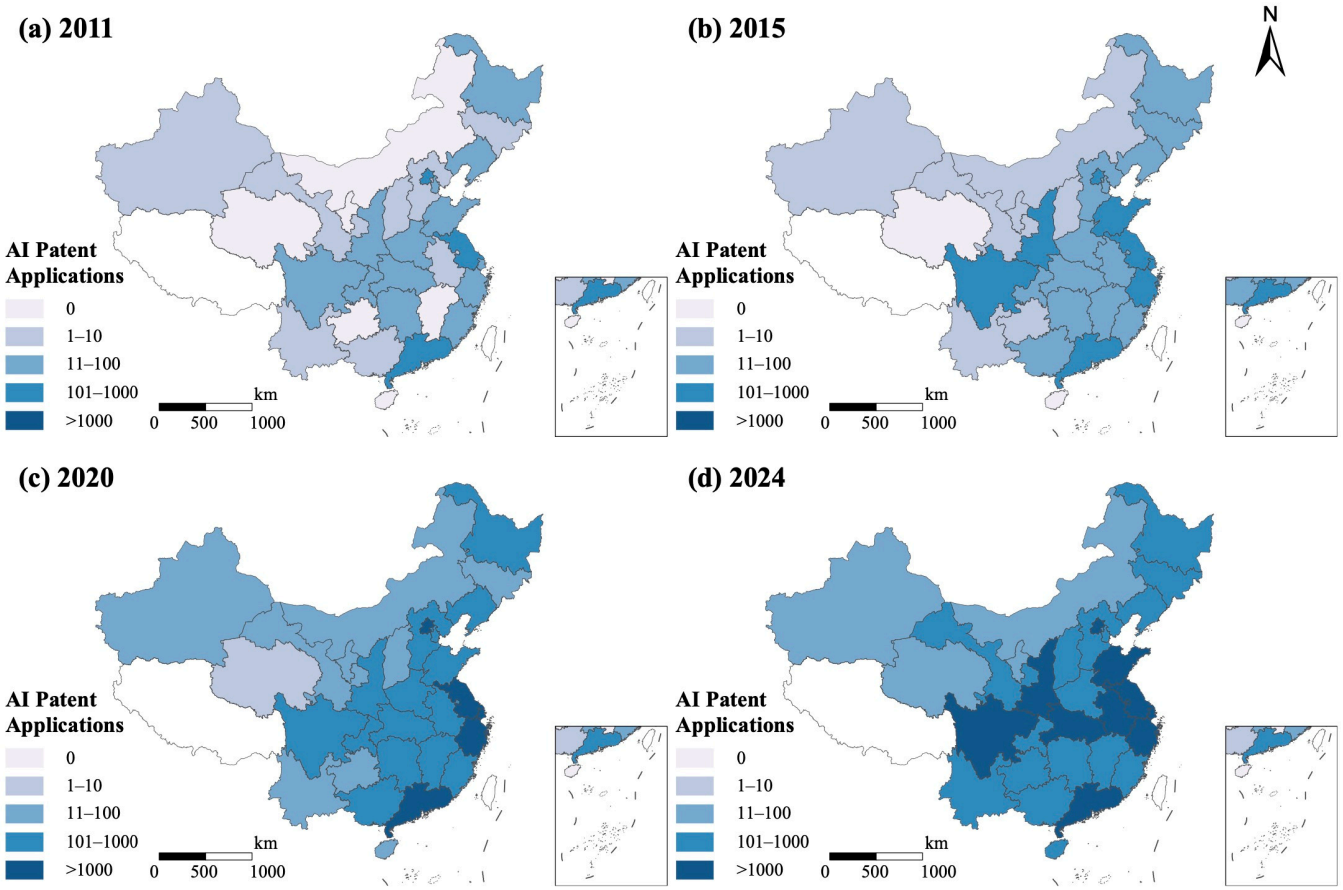


Figure 2. Spatial distribution of strict-definition AI patent applications across provincial-level regions in China (2011, 2015, 2020, and 2024). Notes: Panels report raw application counts; the regression variable was $ln(1 + AI\ patent\ applications)$. Patents were identified with the joint IPC–keyword procedure described in Section 4.1.2. The five class intervals were identical across all panels. A value of zero denotes an observed province–year with no qualifying application, shown in the lightest shade; Tibet, Hong Kong, Macao, and Taiwan are outside the estimation sample and are shown unshaded. The base map follows the standard map of the Ministry of Natural Resources of China, with no modifications to the boundaries.

5.1. Baseline Regression Results

Table 3 reports a sequence of provincial fixed-effects specifications examining the relationship between patent-based AI innovation intensity and carbon emission intensity.

Table 3. Baseline regression results.

Variable	(1) $\ln CI_{it}$	(2) $\ln CI_{it}$	(3) $\ln CI_{it}$	(4) $\ln CI_{it}$
$\ln AI_{it}$	−0.144 ***	−0.079 *	0.067 *	0.062 *
	(0.016)	(0.043)	(0.035)	(0.034)
$(\ln AI_{it})^2$		−0.009 *	−0.012 ***	−0.016 ***
		(0.005)	(0.003)	(0.005)
Controls	No	No	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
Year FE	No	No	No	Yes
Observations	420	420	420	420
R^2	0.6356	0.6565	0.7527	0.7782

Notes: Robust standard errors clustered at the provincial level are given in parentheses; *** and * denote significance at the 1% and 10% levels, respectively.

Table 3 reports the estimates from a sequence of fixed-effects specifications. Column (1) includes the linear AI term and province fixed effects. The coefficient is negative and statistically significant, indicating a negative average association between AI innovation intensity and carbon emission intensity under a linear specification. Column (2) introduces the squared term. Both coefficients are negative and significant at the 10% level. Although this result indicates concavity, it does not support an inverted-U-shaped relationship because the implied turning point lies outside the observed sample range.

Column (3) adds the provincial-level controls. The linear coefficient becomes positive (0.067) and is significant at the 10% level, while the squared coefficient remains negative (−0.012) and is significant at the 1% level. Column (4) further includes year fixed effects to absorb nationwide shocks common to all provinces. In this specification, the linear coefficient is 0.062 and significant at the 10% level, whereas the squared coefficient is −0.016 and significant at the 1% level. The within R^2 is 0.7782.

Column (4) is therefore adopted as the preferred baseline specification. Conditional on the observed provincial characteristics and the two sets of fixed effects, its coefficient pattern is consistent with an inverted-U-shaped relationship: carbon emission intensity initially increases with AI innovation intensity, while the estimated marginal effect becomes negative beyond a certain level. This interpretation applies to the fully specified conditional model and should not be extended to the more parsimonious specifications in Columns (1) and (2).

The signs and significance of the two polynomial terms are not, by themselves, sufficient to establish an inverted U shape within the observed data range. The next section therefore applies the Lind–Mehlum test to examine the endpoint slopes and the location of the estimated turning point. Subsequent analyses assessed the sensitivity of the estimated curvature to alternative variable definitions, sample treatments, and potential endogeneity.

5.2. Formal Test of the Inverted-U-Shaped Relationship

To assess whether the estimated quadratic relationship constitutes an inverted U within the observed data range, we applied the Lind–Mehlum U-test [20,21] to the preferred baseline specification. Table 4 reports the results.

Table 4. Lind–Mehlum U-test results.

Indicator	Value	Description
Sample range	[0.000, 8.311]	Effective sample range of the core explanatory variable
Extreme point	1.948	Lies within the sample range
Slope at lower bound	0.062 **	Significantly positive at the sample minimum
Slope at upper bound	−0.203 ***	Significantly negative at the sample maximum
U-test statistic	1.848 **	Rejects the null of a monotone or U-shaped relationship
U-test *p*-value	0.0374	The inverted-U-shaped relationship is statistically supported
Observations below/above turning point	68/352	Both sides of the estimated turning point are represented in the sample.

Notes: This table reports the Lind–Mehlum test [21] for an inverted-U relationship. The lower- and upper-bound slopes are the first derivatives of the estimated relationship evaluated at the sample minimum and maximum of the core explanatory variable. The endpoint-slope *p*-values are one-sided, as required by the Lind–Mehlum test. *** and ** denote significance at the 1% and 5% levels, respectively.

Table 4 shows that the estimated slope is positive at the lower bound of the sample range (0.062, $p = 0.0374$) and negative at the upper bound (−0.203, $p = 0.0015$). The overall U-test statistic is 1.848, with a *p*-value of 0.0374. The null hypothesis of a monotonic or U-shaped relationship is therefore rejected at the 5% level. Thus, in the preferred specification, the negative quadratic coefficient corresponds to a change in the direction of the estimated

relationship within the observed range rather than merely reflecting concavity outside the sample.

The estimated turning point is 1.948, which lies within the observed range of [0, 8.311]. Because AI innovation intensity is measured as the natural logarithm of one plus the number of strict-definition AI patent applications, this value corresponds to approximately six patent applications. The fitted relationship has a positive slope below this point and a negative slope above it.

Of the 420 province–year observations, 68 fall below the estimated turning point and 352 are located at or above it. Both sides of the fitted curve are therefore represented by observed data, although the observations are concentrated on its downward-sloping segment. Taken together, the endpoint slopes, the overall U-test, and the turning point location provide statistical support for an inverted-U-shaped pattern in the preferred baseline specification. However, this evidence remains relatively close to the 5% significance threshold and should not be interpreted as establishing a causal relationship. The following sections therefore examine whether the estimated curvature is preserved under alternative measurements, sample treatments, and identification strategies.

5.3. Robustness Checks

We next examined whether the baseline result was sensitive to the data-extension procedure, variable measurement, extreme observations, regional sample composition, or explanatory variable timing. Column (5) restricts the sample to 2011–2021, excluding the observed 2022 anchor year and the extrapolated observations for 2023 and 2024. This conservative pre-anchor specification assesses whether the estimated curvature was already present before the emissions series was extended. Column (6) uses an alternative measure of carbon emission intensity, while Column (7) replaces the strict AI patent classification with the broad classification. Column (8) winsorizes the continuous variables at the 1st and 99th percentiles, Column (9) excludes Beijing and Shanghai, and Column (10) introduces a one-period lag in both AI terms. All specifications retain the full set of controls and province and year fixed effects.

The quadratic term remains negative and statistically significant across all six robustness specifications. When the sample is restricted to 2011–2021, the coefficient on the squared AI term is −0.015 and significant at the 1% level. Its implied turning point is 2.026, close to the full-sample estimate of 1.948 reported in Table 4. The Lind–Mehlum test for the restricted sample also rejects the null of a monotone or U-shaped relationship at the 5% level (t = 1.71, p = 0.0487). The estimated slopes at the lower and upper bounds are 0.062 and −0.189, respectively, with 68 observations below the turning point and 262 at or above it. The estimated curvature and interior turning point are therefore not driven by observations from 2022–2024. Because this specification excludes the observed 2022 and reconstructed 2023–2024 values, it constitutes a conservative check on the extension of the emissions series.

Using the alternative measure of carbon emission intensity produces a quadratic coefficient of −0.009, significant at the 5% level. Replacing the strict patent classification with the broad-definition measure yields an estimate of −0.016, significant at the 1% level. After winsorization, the quadratic coefficient remains −0.015 and significant at the 1% level. Excluding Beijing and Shanghai produces an estimate of −0.014, while the coefficient on the one-period-lagged squared term is also −0.014; both are significant at the 1% level.

The estimates obtained from the 2011–2021 sample, the winsorized sample, the restricted regional sample, and the lagged specification are close to the preferred full-sample quadratic coefficient reported in Table 3. The estimated concavity is therefore not attributable solely to the inclusion of post-2021 observations, extreme values, the two leading

patenting regions, or the contemporaneous timing of patent activity. It also remains present under alternative definitions of carbon emission intensity and AI-related patent activity.

The linear coefficient is estimated less precisely in some specifications, particularly when the outcome measure or patent classification is changed. Table 5 should therefore be interpreted as evidence that the negative curvature is stable, rather than as showing that every alternative specification independently establishes an inverted-U-shaped relationship. Formal evidence of a change in slope within the observed range is provided by the Lind–Mehlum test in Table 4. Similarly, the lagged specification reduces concerns about contemporaneous feedback but does not eliminate possible endogeneity, which was examined separately in the subsequent identification analyses.

Table 5. Robustness checks.

Variable	(5) 2011–2021 Sample	(6) Alternative Outcome	(7) Broad AI Measure	(8) Winsorized	(9) Excl. BJ & SH	(10) Lagged AI
Dependent variable	$\ln CI_{it}$	$\ln CI_{it}^{alt}$	$\ln CI_{it}$	$\ln CI_{it}$	$\ln CI_{it}$	$\ln CI_{it}$
$\ln AI_{it}$	0.062 * (0.036)	0.012 (0.027)	0.071 (0.074)	0.060 * (0.034)	0.058 * (0.033)	
$\ln AI_{i,t-1}$						0.058 ** (0.028)
$(\ln AI_{it})^2$	−0.015 *** (0.005)	−0.009 ** (0.004)	−0.016 *** (0.006)	−0.015 *** (0.005)	−0.014 *** (0.004)	
$(\ln AI_{i,t-1})^2$						−0.014 *** (0.004)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	330	420	420	420	392	390
R^2	0.7696	0.9180	0.7753	0.7751	0.7831	0.7710
Implied turning point	2.026	0.637	2.221	1.973	2.075	2.015

Notes: Column (5) restricts the sample to 2011–2021 and excludes all observations from 2022–2024. Column (6) uses an alternative measure of carbon emission intensity. Column (7) uses the broad-definition AI patent measure. Column (8) winsorizes all continuous variables at the 1st and 99th percentiles. Column (9) excludes Beijing and Shanghai. Column (10) uses the one-period-lagged AI terms. Standard errors clustered at the provincial level are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

5.4. Endogeneity Diagnostics

Although the fixed-effects specification absorbs time-invariant provincial characteristics and common year shocks, the estimates may still be affected by reverse causality, omitted time-varying factors, or measurement error in the patent-based AI variable. We therefore conducted Durbin–Wu–Hausman tests using the internal-structure Bartik variables described in Section 5.5 as instruments [50–55]. Because the linear and quadratic AI terms entered the model, their exogeneity was examined jointly and separately. Table 6 reports the results.

The joint test for the linear and quadratic AI terms produces a statistic of 2.134 and a *p*-value of 0.3441. The null hypothesis that the two terms can be treated as exogenous therefore cannot be rejected. The separate tests yield the same conclusion. The test statistic is 0.713 for the linear term, with a *p*-value of 0.3985, and 2.013 for the quadratic term, with a *p*-value of 0.1559. Thus, within this specification, the data do not provide sufficient evidence that either term is endogenous.

Table 6. Durbin–Wu–Hausman endogeneity tests.

Tested Regressor(s)	Test Statistic	Distribution	*p*-Value	Conclusion
Linear and quadratic AI terms	2.134	$\chi^2(2)$	0.3441	Failed to reject exogeneity
Linear AI term	0.713	$\chi^2(1)$	0.3985	Failed to reject exogeneity
Quadratic AI term	2.013	$\chi^2(1)$	0.1559	Failed to reject exogeneity

Notes: The null hypothesis is that the tested regressor or regressors can be treated as exogenous. The tests used the internal-structure Bartik variables as instruments and were based on the 2011–2023 sample of 30 provinces (N = 390). All specifications included the full set of controls, as well as province and year fixed effects. Standard errors were clustered at the provincial level. Because the test was conditional on instrument validity and relevance, the results were interpreted as endogeneity diagnostics rather than evidence that endogeneity is absent.

This result should not be interpreted as proof of exogeneity. The Durbin–Wu–Hausman tests are conditional on instrument relevance and validity and may have limited power under weak identification. They are therefore treated as preliminary diagnostics rather than as a basis for dismissing endogeneity concerns. The following sections examine the issue further using the internal-structure Bartik design and an instrument-free copula-control-function approach.

5.5. Alternative Identification Using the Internal-Structure Bartik Design

5.5.1. Instrument Construction

To obtain variation in provincial AI innovation that is less dependent on contemporaneous local economic conditions, we constructed an internal-structure Bartik instrument. The instrument combined the technological composition of provincial AI patenting before the estimation period with subsequent technology-specific patent growth outside China. It therefore consisted of a fixed provincial exposure component and a time-varying external shift.

AI patents are divided into four branches: algorithms, computer vision, speech processing, and natural language processing. For each province, the baseline share of branch k is calculated over 2005–2010 using

$$s_{ik,0} = \frac{\sum_{\tau=2005}^{2010} P_{ik\tau}^{CN}}{\sum_{l=1}^{4} \sum_{\tau=2005}^{2010} P_{il\tau}^{CN}} \tag{17}$$

where $P_{ik\tau}^{CN}$ denotes the number of patent applications in branch k, province i, and year τ. The four shares sum to one within each province. Accordingly, the exposure measure captures the initial composition of provincial AI innovation rather than its initial scale.

The time-varying shifts were constructed from non-China-origin patent families recorded in EPO PATSTAT. Let P_{kt}^{-CN} denote the number of non-China patent families in branch k and year t. The external shock was defined relative to 2010:

$$g_{kt}^{-CN} = \ln\left(1 + P_{kt}^{-CN}\right) - \ln\left(1 + P_{k,2010}^{-CN}\right) \tag{18}$$

Excluding China-origin patent families prevented patenting by Chinese provinces from entering the external shift mechanically. The use of 2010 as the reference year also separated the predetermined exposure period from the 2011–2023 estimation sample.

The internal-structure Bartik variable was constructed as

$$Bartik_{it}^{struct} = \sum_{k=1}^{4} s_{ik,0}\, g_{kt}^{-CN} \tag{19}$$

The instrument takes a higher value when a province was initially more specialized in AI branches that subsequently expanded more rapidly outside China. Before estimation, the Bartik variable was standardized to have a mean of zero and a standard deviation of one.

Because the outcome equation contains the linear and quadratic AI terms, the logged AI variable is centered at its mean in the 2011–2023 estimation sample:

$$X_{it}^{c} = \ln AI_{it} - \overline{lnAI} = \ln AI_{it} - 4.1952 \quad (20)$$

The two potentially endogenous regressors are X_{it}^{c} and $\left(X_{it}^{c}\right)^{2}$. The standardized Bartik variable and its square serve as the corresponding excluded instruments. Centering does not alter the fitted quadratic relationship or the turning point expressed on the original $ln(1 + AI\ patents)$ scale, but it reduces the mechanical correlation between the linear and quadratic terms.

The identifying variation arises from differences in provinces' initial exposure to subsequent foreign technological developments. For example, a given expansion in computer-vision patents abroad generates a larger predicted shock in provinces that were more specialized in computer vision during the baseline period. Province fixed effects absorb time-invariant differences in technological structure, while year fixed effects account for shocks common to all provinces.

The exclusion restriction requires that, conditional on the controls and fixed effects, foreign branch-specific patent growth affects provincial carbon emission intensity through domestic AI innovation rather than through an independent channel. It also requires that the baseline composition of AI patenting is not systematically associated with unobserved changes in carbon emission intensity after 2010. Defining the exposure component within the AI technology field limits some of the direct channels associated with broader historical education or geographical variables, but it does not establish the exclusion restriction by itself. The Bartik estimates are therefore accompanied by weak-identification diagnostics and randomization inference in the following sections.

5.5.2. Bartik Estimation Results

The external shock series was available through 2023, and thus, the Bartik analysis used the 2011–2023 sample, which comprised 390 observations from 30 provinces. To separate changes arising from the shorter sample from those attributable to the instrument, Column (11) re-estimates the two-way fixed-effects model using the same observations as the 2SLS specification. Column (12) reports the internal-structure Bartik estimate. Table 7 reports the same-sample TWFE and internal-structure Bartik estimates.

The same-sample fixed-effects estimate remains concave. The coefficient on the squared AI term is −0.0160 and is significant at the 1% level, while the corresponding turning point is 1.8501. This indicates that the shorter 2011–2023 sample does not account for the negative curvature found in the baseline analysis. In the Bartik specification, the quadratic coefficient is −0.0240 and remains significant at the 1% level. The turning-point estimate is 3.6549, which lies within the observed range of the AI variable.

The estimated slope is positive at the sample minimum (0.1757) and negative at the sample maximum (−0.2239). The endpoint-based inverted-U test, however, yields a p-value of 0.1007, marginally above the 10% threshold. The Bartik specification therefore supports negative curvature and produces a turning-point estimate within the sample range, but it does not independently establish an inverted-U shape at conventional significance levels. The turning point is also estimated imprecisely. These results are therefore treated as supplementary evidence and interpreted together with the weak-identification diagnostics reported below.

Table 7. Same-sample TWFE and internal-structure Bartik estimates.

Variable/Statistic	(11) Same-Sample TWFE	(12) Internal-Structure Bartik 2SLS
Centered AI term	−0.0749 **	−0.0260
	(0.0328)	(0.1313)
Squared centered AI term	−0.0160 ***	−0.0240 ***
	(0.0050)	(0.0060)
Turning point in original AI scale	1.8501	3.6549
	(0.8985)	(2.6329)
Slope at sample minimum		0.1757
		(0.0998)
Slope at sample maximum		−0.2239
		(0.1712)
Endpoint-based inverted-U *p*-value		0.1007
Controls & province FE & year FE	Yes	Yes
Observations	390	390

Notes: AI innovation intensity was centered at its 2011–2023 sample mean of 4.1952 before constructing the squared term. Turning points were converted back to the original $ln(1 + AIpatents)$ scale. Standard errors clustered at the provincial level are reported in parentheses. *** and ** indicate significance at the 1% and 5% levels, respectively.

5.5.3. Weak-Identification Diagnostics

Instrument relevance was assessed at the equation-specific and system levels. This distinction is important because the model contains two endogenous regressors and two excluded instruments. A strong relationship in each separate first stage does not necessarily imply that the two endogenous terms are jointly well-identified [56,57]. Table 8 reports the first-stage and weak-identification diagnostics.

Table 8. First-stage and weak-identification diagnostics.

Diagnostic	AI Term	AI Squared	Joint/Model Statistic	*p*-Value
Excluded-instrument F	26.12	49.63		0.0001/<0.0001
Sanderson–Windmeijer F	11.45	19.65		0.0021/0.0001
Kleibergen–Paap rk LM			2.355	0.1249
Kleibergen–Paap rk Wald F			4.213	
Anderson–Rubin Wald F			10.68	0.0003
Stock–Wright LM S			2.42	0.2976

Notes: All diagnostics were cluster-robust. The model contains two endogenous regressors and two excluded instruments and is therefore exactly identified; an overidentification test is not available. Conventional Stock–Yogo critical values do not directly apply to the cluster-robust Kleibergen–Paap statistic.

The separate first stages are statistically significant. The excluded-instrument F statistics are 26.12 for the linear term and 49.63 for the quadratic term, while the corresponding Sanderson–Windmeijer statistics are 11.45 and 19.65. The instruments therefore contain relevant information for each endogenous regressor when considered separately.

The system-level diagnostics are less favorable. The Kleibergen–Paap rk Wald F statistic is 4.213, and the rk LM test does not reject under-identification at conventional levels. The combination of the linear and quadratic endogenous terms is consequently only weakly identified. The Anderson–Rubin test rejects the joint null that both endogenous coefficients are zero (F = 10.68, $p = 0.0003$), whereas the Stock–Wright test does not reject the same null at conventional significance levels ($p = 0.2976$) [58,59]. The divergence between these weak-identification-robust tests further indicates that inference regarding the two endogenous terms should remain cautious. These weak-identification-robust tests concern

the joint contribution of the endogenous terms; they do not, by themselves, establish the sign of the quadratic coefficient or the location of the turning point.

Given the evidence of weak joint identification, the Bartik estimates are treated as supplementary evidence rather than as definitive causal estimates.

5.5.4. Randomization Inference

A randomization exercise was used to determine whether the observed first-stage relevance and quadratic estimate could be reproduced by an arbitrary pairing of provincial exposure structures and external technology shocks. The four-element baseline-share vector was randomly reassigned across provinces, while the foreign shock series and its temporal ordering were held fixed. The Bartik variables and the corresponding first-stage and 2SLS models were then reconstructed. This procedure was repeated 1000 times.

Permuting the complete exposure vector preserves the internal correlation between the four technology shares and ensures that they continue to sum to one. The exercise breaks only the observed correspondence between a province and its actual baseline AI structure. Table 9 reports the randomization-inference results.

Table 9. Randomization inference for the internal-structure Bartik design.

Statistic	Actual Value	Empirical *p*-Value	Tail
First-stage F for the linear AI term	26.118	0.0150	Right
First-stage F for the squared AI term	49.630	0.0100	Right
Studentized 2SLS t-statistic for the squared AI term	−4.003	0.0010	Two-sided

Notes: Provincial baseline AI structures were randomly permuted 1000 times, while the four foreign technology shocks were held fixed. Empirical *p*-values used the finite-randomization correction. The studentized statistic is reported because it is less sensitive than the raw coefficient to variation in instrument strength across permutations.

The actual first-stage statistics lie in the upper tail of their respective permutation distributions. The empirical *p*-values are 0.0150 for the linear AI term and 0.0100 for the squared term. Thus, the relevance of the observed instrument is not readily reproduced by randomly assigning baseline technical structures across provinces. The actual studentized statistic for the quadratic coefficient is −4.003, with a two-sided empirical *p*-value of 0.0010, placing the observed negative curvature in the extreme tail of the randomized distribution.

The randomization exercise strengthens the case that the results depend on the observed alignment between provincial AI specialization and subsequent foreign technology shocks rather than on an arbitrary reassignment of exposure shares. However, it does not prove the exclusion restriction or eliminate the weak-identification problem documented in Table 8. The randomization results and the weak-instrument diagnostics therefore serve different purposes: the former assesses whether the observed share–shock matching is unusual, whereas the latter evaluates the strength of joint identification.

5.6. Instrument-Free Sensitivity Analysis Using 2sCOPE-np

The internal-structure Bartik design provides evidence under an external-shock identification strategy, but its interpretation remains subject to instrument strength and the exclusion restriction. As a complementary exercise, this section presents the results of applying the two-stage nonparametric copula-control-function approach (2sCOPE-np) developed by Hu et al. [60]. This method does not require an excluded instrumental variable. Instead, it uses the conditional distribution of the potentially endogenous regressor to construct a control function that captures its remaining dependence with the structural error.

The 2sCOPE-np results are not intended to replace either the fixed-effects estimates or the Bartik design; rather, they provide evidence under a different set of identifying assumptions. Consistent findings across these approaches are informative because their respective sources of identification are distinct.

5.6.1. Method and Identification Conditions

Let

$$X_{it} = \ln(1 + AI_{it}) \tag{21}$$

where AI_{it} denotes the number of strict-definition AI patent applications in province i and year t. Let $\boldsymbol{K}_{it}$ denote the vector of control variables. In the first stage, X_{it} and the continuous controls were transformed into deviations from their province-specific means. A kernel estimator was then used to estimate the conditional cumulative distribution function:

$$\hat{F}_{\tilde{X}|\tilde{\boldsymbol{K}},t}\left(\tilde{X}_{it} \mid \tilde{\boldsymbol{K}}_{it}, t\right) \tag{22}$$

where year indicators were included among the conditioning variables. The bandwidth was selected using least-squares cross-validation. The copula control function was constructed as

$$\hat{C}_{it} = \Phi^{-1}\left[\hat{F}_{\tilde{X}|\tilde{\boldsymbol{K}},t}\left(\tilde{X}_{it} \mid \tilde{\boldsymbol{K}}_{it}, t\right)\right] \tag{23}$$

where $\Phi^{-1}(\cdot)$ denotes the inverse standard-normal distribution function.

In the second stage, the baseline quadratic fixed-effects model was augmented with the estimated control function:

$$\ln CI_{it} = \beta_1 X_{it} + \beta_2 X_{it}^2 + \rho\hat{C}_{it} + \gamma' K_{it} + \lambda_i + \delta_t + \varepsilon_{it} \tag{24}$$

The coefficient ρ captures the component of dependence between AI innovation intensity and the structural error represented by the copula control function. A statistically significant estimate of ρ would indicate residual dependence consistent with endogeneity under the maintained copula specification. Since X_{it}^2 is a deterministic transformation of X_{it}, the augmented equation contains one control function for the underlying AI variable rather than a separate control function for its squared term.

Although this strategy does not rely on an exclusion restriction, it is not assumption-free. First, the conditional copula must adequately represent the dependence between the endogenous regressor and the structural error, or its endogenous component, after conditioning on the included covariates. Second, the generated control function must not be a linear combination of the other regressors in the outcome equation. Third, the conditional distribution of X_{it} must contain sufficient variation for reliable nonparametric estimation.

The implementation used 420 observations that covered 30 provinces over 2011–2024. AI innovation intensity takes 253 distinct values, and zero-patent observations account for 4.76% of the sample. The estimated conditional probabilities range from 0.028 to 0.965, and none required boundary clipping before the normal-score transformation. The correlations of the generated control function with X_{it} and X_{it}^2 are 0.217 and 0.158, respectively. Moreover, including the control function increases the estimated standard errors of the linear and quadratic terms by only 4.68% and 6.86%, respectively. These diagnostics do not indicate that the augmented model is affected by serious near-collinearity.

The distribution of AI innovation intensity departs from normality, with a skewness of -0.311, kurtosis of 2.431, and a Shapiro–Wilk p-value below 0.001. Nevertheless, the departure from normality is moderate. The results should therefore be interpreted as conditional on the maintained copula restrictions. In particular, failure to detect a statistically significant control-function coefficient cannot establish the absence of endogeneity.

5.6.2. 2sCOPE-np Estimates

Table 10 compares the original two-way fixed-effects estimates with the 2sCOPE-np augmented estimates. Column (1) reproduces the baseline specification. Column (2) adds the nonparametrically estimated copula control function and reports the inference based on 999 province-cluster bootstrap replications.

Table 10. Uncorrected TWFE and 2sCOPE-np estimates.

Variable/Statistic	(1) Uncorrected TWFE	(2) 2sCOPE-np
AI innovation intensity, X_{it}	0.062 * (0.034)	0.089 * (0.037)
Squared AI term, X_{it}^2	−0.016 *** (0.005)	−0.016 *** (0.005)
Copula control function, $\hat{C}_{it}$	-	−0.0255 (0.0292)
Turning point	1.948	2.852
Controls & province FE & year FE	Yes	Yes
Observations	420	420

Notes: The dependent variable was the logarithm of carbon emission intensity. Province-clustered standard errors are reported for the uncorrected TWFE model. The 2sCOPE-np column reports province-cluster bootstrap standard errors based on 999 successful replications. Significance in Column (2) was based on finite-sample-adjusted empirical bootstrap *p*-values. *** and * denote significance at the 1% and 10% levels, respectively. The turning points are reported on the ln(1 + AI) scale.

Table 10 compares the conventional two-way fixed-effects estimates with those obtained after introducing the nonparametric copula control function. The correction increases the linear coefficient from 0.062 to 0.089, whereas the quadratic coefficient changes only slightly, from −0.0159 to −0.0155. The negative curvature identified in the baseline model is therefore preserved after the control function is introduced.

The estimated turning point moves from 1.948 in the uncorrected model to 2.852 in the corrected model but remains within the observed range of [0, 8.311]. Based on the corrected point estimates, the slope is positive at the sample minimum (0.089) and negative at the sample maximum (−0.170). The correction therefore affects the estimated location of the turning point more than the direction of the curvature.

The coefficient of the copula control function is −0.0255. This coefficient captures the component of the dependence between AI innovation intensity and the unobserved determinants of carbon emission intensity that is represented by the copula term. Its point estimate alone is not sufficient to establish the presence or absence of endogeneity. Statistical inferences for this coefficient, the quadratic term, and the turning point were therefore based on the province-cluster bootstrap reported in the following subsection.

Overall, the 2sCOPE-np point estimates do not indicate that the negative quadratic coefficient in the baseline model is generated by the omission of the copula control function. The method nevertheless relies on distributional and control-function assumptions that differ from those underlying the Bartik design. The results are consequently interpreted as a complementary sensitivity analysis rather than as definitive causal identification.

5.6.3. Cluster Bootstrap and Leave-One-Province-Out Analysis

Because the copula control function is generated in the first stage, conventional second-stage standard errors do not incorporate all estimation uncertainty. Observations within the same province may also be serially correlated. Inference is therefore based on a province-cluster bootstrap [61]. Provinces, rather than individual province–year observations, were

resampled, and the complete two-stage procedure was re-estimated in each replication. All 999 bootstrap replications converge successfully.

The sensitivity of the results to individual provinces was examined through a leave-one-province-out analysis. The complete 2sCOPE-np model was re-estimated 30 times, sequentially excluding one province and all its yearly observations. Table 11 reports the cluster-bootstrap and leave-one-province-out diagnostics.

Table 11. Cluster-bootstrap and leave-one-province-out diagnostics.

Panel A. Province-cluster bootstrap results.

Statistic	Point Estimate	Bootstrap SE	95% Percentile Interval	Adjusted Empirical *p*-Value
Linear AI term	0.0886	0.0365	[−0.0074, 0.1394]	0.086
Squared AI term	−0.0155	0.0051	[−0.0261, −0.0059]	0.006
Copula control function	−0.0255	0.0292	[−0.0431, 0.0647]	0.844
Turning point	2.8518	1.1269	[−0.2931, 3.8464]	-
Slope at sample minimum	0.0886	0.0365	[−0.0074, 0.1394]	0.086
Slope at sample maximum	−0.1695	0.0707	[−0.3427, −0.0801]	0.006

Panel B. Leave-one-province-out results.

Statistic	Result
Successful estimations/negative quadratic coefficient	30/30
Quadratic coefficient significant at 5%/turning point within the remaining sample	30/30
Positive slope at sample minimum/negative slope at sample maximum	30/30
All pointwise inverted-U conditions satisfied	30/30
Quadratic coefficient: minimum/median/maximum	−0.0183/−0.0156/−0.0131
Turning point: minimum/median/maximum	1.034/2.965/3.717
Significant copula control function at 5%	0/30

Notes: Panel (**A**) reports percentile intervals from 999 province-cluster bootstrap replications. Empirical *p*-values used a finite-sample adjustment. No *p*-value is reported for the turning point because testing whether the turning point equals zero has no substantive interpretation in the present setting. Panel (**B**) reports results from re-estimating the complete 2sCOPE-np model after sequentially excluding each of the 30 provinces.

The bootstrap results indicate that the negative quadratic coefficient is not an artifact of conventional second-stage inference. Its 95% percentile interval remains entirely below zero, and the coefficient is negative in 997 of the 999 bootstrap replications (99.8%). The slope at the sample minimum is positive in 957 replications (95.8%), while the slope at the sample maximum is negative in 997 replications (99.8%). The estimated turning point lies within the sample range in 957 replications (95.8%). Taken jointly, all pointwise conditions for an inverted U shape are satisfied in 956 of the 999 replications (95.7%).

The lower-bound slope is estimated less precisely than the quadratic coefficient. Its adjusted empirical *p*-value is 0.086, and its 90% percentile interval remains above zero [0.0032, 0.1247], although the corresponding 95% interval marginally crosses zero. The slope at the upper bound is significantly negative ($p = 0.006$), and the estimated turning point of 2.852 lies within the observed range. Its 90% percentile interval [0.140, 3.508] also remains within that range. The copula-corrected estimates therefore provide bootstrap-based support for an inverted-U-shaped relationship at the 10% level, although the initial upward-sloping segment is estimated less precisely than the subsequent decline.

The leave-one-province-out results further indicate that the estimates are not driven by a particular province. The quadratic coefficient remains negative and statistically significant in all 30 estimations, ranging from −0.0183 to −0.0131. The corresponding turning points range from 1.034 to 3.717 and remain within the relevant sample range in every case.

The slopes at the lower and upper bounds retain their expected positive and negative signs throughout. The control-function coefficient is insignificant in every leave-one-out estimation, with *p*-values ranging from 0.191 to 0.692.

Taken together, the 2sCOPE-np estimates lead to two conclusions. First, the procedure does not detect statistically significant residual dependence between AI innovation intensity and the outcome disturbance under the maintained copula assumptions. The control-function coefficient is small and statistically insignificant, although this finding should not be interpreted as proof of exogeneity. Second, after the control function is introduced, the negative quadratic coefficient, the opposing endpoint slopes, and the interior turning point are preserved. The bootstrap results support an inverted-U-shaped relationship at the 10% level, and the leave-one-province-out estimates show that this pattern is not driven by any single province. The 2sCOPE-np results therefore provide complementary evidence that the estimated nonlinear pattern is not materially altered by the modeled endogeneity correction.

5.7. Synthesis of Identification Results

The empirical results provide broadly convergent evidence of negative curvature in the relationship between AI innovation intensity and carbon emission intensity. In the preferred two-way fixed-effects specification, the Lind–Mehlum test rejects a monotonic or U-shaped relationship at the 5% level, and the estimated turning point lies within the observed sample range. The robustness checks consistently retain a negative and statistically significant quadratic coefficient under alternative variable definitions, sample treatments, and temporal specifications. These checks primarily establish the stability of the estimated curvature; they should not be interpreted as separate formal tests of an inverted U shape in every specification.

The endogeneity diagnostics do not reject the exogeneity of the linear and quadratic AI terms, although this result cannot establish that endogeneity is absent. The internal-structure Bartik specification produces a significantly negative quadratic coefficient and a turning-point estimate within the sample range. The randomization exercise further shows that the observed first-stage strength and the studentized statistic for the quadratic term are difficult to reproduce by arbitrarily reallocating the baseline technological structures across provinces. Nevertheless, the endpoint-based inverted-U test yields a *p*-value of 0.1007, marginally above the 10% threshold; the turning point is estimated imprecisely, and the Kleibergen–Paap statistic indicates weak joint identification. The Bartik results are therefore treated as supplementary evidence of negative curvature and an economically admissible turning point rather than as definitive causal estimates.

The instrument-free 2sCOPE-np analysis yields a similar nonlinear pattern. The copula control function is statistically insignificant, indicating that the procedure does not detect substantial residual endogeneity under the maintained assumptions. After the control function is introduced, the quadratic coefficient remains significantly negative, the estimated slopes have the expected positive and negative signs at the sample bounds, and the turning point remains within the observed range. Province-cluster bootstrap inference supports the inverted-U-shaped relationship at the 10% level, while all 30 leave-one-province-out estimations preserve the pointwise shape conditions.

Taken together, the evidence is strongest for negative curvature and for the presence of a turning point within the observed data range under the preferred specification. Its precise location varies between estimation methods and is less precisely determined under the supplementary identification strategies. The turning point should therefore be interpreted as a sample-dependent transition range rather than as a universal or mechanically applicable policy threshold.

6. Decomposition and Extended Analyses

6.1. Decomposition of the Energy–Carbon Link

Table 12 reports the exact decomposition of the estimated relationship. Column (1) reproduces the preferred carbon-intensity specification. Columns (2) and (3) use energy intensity and the carbonization factor as the dependent variables, respectively. Each equation includes the full set of controls together with province and year fixed effects.

Table 12. Exact decomposition of the estimated carbon-intensity relationship.

Variable	(1) Carbon Intensity	(2) Energy Intensity	(3) Carbonization Factor
AI innovation intensity	0.062 *	0.048 *	0.014
	(0.034)	(0.027)	(0.032)
Squared AI term	−0.016 ***	−0.012 ***	−0.003
	(0.005)	(0.003)	(0.004)
Controls & province FE & year FE	Yes	Yes	Yes
Observations	420	420	420
R^2	0.7782	0.8361	0.2351

Notes: The dependent variables were $ln\frac{C}{GDP}$, $ln\frac{E}{GDP}$, and $ln\frac{C}{E}$ for Columns (1)–(3), respectively. Robust standard errors clustered at the provincial level are reported in parentheses. *** and * denote significance at the 1% and 10% levels, respectively.

The coefficient identities hold exactly at full estimation precision. The minor discrepancies among the reported four-decimal estimates are due solely to rounding: for the linear term, $0.0621 \approx 0.0484 + 0.0136$, while for the quadratic term, $-0.0159 \approx -0.0125 - 0.0035$. The decomposition indicates that the estimated curvature is concentrated in the energy-intensity component. The squared coefficient in the energy-intensity equation is −0.0125 and significant at the 1% level, accounting for 78.2% of the total quadratic coefficient. The corresponding coefficient in the carbonization-factor equation is −0.0035 and is statistically insignificant, accounting for the remaining 21.8%.

These results suggest that the estimated curvature operates predominantly through changes in the amount of energy used per unit of output. At relatively low levels of AI innovation, the expansion of computing infrastructure, equipment, and complementary production capacity may increase energy requirements. As innovation becomes more developed, improvements in production scheduling, resource allocation, and energy management become more important, reducing energy use per unit of output. By contrast, the estimates provide less evidence that the curvature is driven by changes in the average carbon content of energy consumption.

The contribution percentages are accounting shares of the estimated quadratic coefficient rather than causal mediation proportions. In particular, the carbonization factor is derived from the same emissions and energy accounts and is not statistically independent of carbon intensity. The decomposition should therefore be interpreted as distinguishing two physical margins—the quantity of energy used and the carbon content of that energy—rather than as estimating separate causal mediators.

The preceding decomposition addresses the accounting relationship between emissions and energy use, but carbon intensity and energy intensity contain GDP in the denominator. To examine whether the estimated curvature is generated mechanically by this common denominator, the models are re-estimated in levels. The dependent variables are the logarithms of total carbon emissions and total energy consumption, while the logarithm of GDP is included on the right-hand side. GDP therefore no longer appears in the denominator of either dependent variable. Table 13 reports the level specifications that exclude the common GDP denominator.

Table 13. Level specifications excluding the common GDP denominator.

Variable	(1) Total Carbon Emissions	(2) Total Energy Consumption
AI innovation intensity	0.045	0.040
	(0.028)	(0.026)
Squared AI term	−0.011 ***	−0.010 ***
	(0.004)	(0.003)
ln*GDP*	−0.900 **	0.001
	(0.382)	(0.293)
Controls & province FE & year FE	Yes	Yes
Observations	420	420
R^2	0.4300	0.7207

Notes: The dependent variables were the logarithms of total carbon emissions and total energy consumption. Both specifications controlled for ln*GDP*. Robust standard errors clustered at the provincial level are reported in parentheses. *** and ** denote significance at the 1% and 5% levels, respectively.

The squared AI coefficient remains negative when the common GDP denominator is removed. It is −0.0107 in the total-emissions equation and −0.0097 in the total-energy-consumption equation, and both coefficients are significant at the 1% level. The persistence of negative curvature in these level specifications indicates that the baseline result is not solely an artifact of dividing emissions and energy consumption by GDP.

At the same time, the level regressions do not eliminate the physical relationship between energy consumption and calculated carbon emissions. They serve a narrower purpose: ruling out the shared GDP denominator as the sole source of the nonlinear result. Taken together, the exact decomposition and the level specifications show that the estimated curvature is primarily associated with the energy-use margin, while avoiding the interpretation of energy intensity as an independent causal mediator.

6.2. Moderating Role of Human Capital

The environmental implications of AI innovation may depend on a region's capacity to absorb, adapt, and apply new technologies. To examine this possibility, human capital was introduced as a moderator of the linear and quadratic AI terms. The specifications include the two interaction terms, their constituent variables, the full set of controls, and province and year fixed effects. Table 14 reports the results.

Table 14. Moderating role of human capital in the nonlinear relationship.

Variable	Coefficient	Variable	Coefficient
$\ln AI_{it}$	0.779 **	$\ln AI_{it} \times \ln HR$	0.184 **
	(0.374)		(0.088)
$(\ln AI_{it})^2$	−0.134 ***	$(\ln AI_{it})^2 \times \ln HR$	−0.030 ***
	(0.039)		(0.010)
ln*HR*	−0.359	Controls	Yes
	(0.319)		
Year FE	Yes	Province FE	Yes
Observations	420	R^2	0.7938

Notes: The dependent variable was the logarithm of carbon emission intensity. ln*HR* denotes the logarithm of human capital. The model included the full set of controls and province and year fixed effects. Robust standard errors clustered at the provincial level are reported in parentheses. R^2 is the within R^2. *** and ** denote significance at the 1% and 5% levels, respectively.

The interaction between the linear AI term and human capital is positive (0.184) and statistically significant at the 5% level, whereas the interaction between the squared AI term and human capital is negative (−0.030) and significant at the 1% level. These results

indicate that human capital moderates the slope and curvature of the estimated relationship. The standalone coefficient on human capital is not statistically significant, but this does not contradict the presence of moderation, which is captured by the interaction terms.

Following the established framework for moderation in U-shaped relationships [20], the direction of the turning point shift is determined by the sign of $\alpha_1\alpha_{12} - \alpha_2\alpha_{11}$. Based on the estimated coefficients, this expression equals 0.001286, which is positive, implying that the turning point moves to the right as human capital increases. Meanwhile, the significantly negative coefficient on the interaction with the squared AI term indicates that higher human capital makes the curve more concave. The initial upward-sloping segment therefore becomes stronger, while the decline after the turning point becomes steeper.

These findings suggest that human capital does not exert a uniformly emission-reducing effect. Instead, it reinforces the stage-dependent nature of the relationship. Provinces with stronger human-capital foundations may experience a more pronounced expansion-related increase in carbon emission intensity at relatively low levels of AI innovation. At higher levels, however, their greater capacity to absorb and deploy advanced technologies strengthens the efficiency-enhancing effects of AI innovation, resulting in a steeper reduction in carbon emission intensity.

6.3. Regional Heterogeneity

China's regions differ substantially in industrial structure, technological capacity, energy endowments, and the maturity of AI adoption. The relationship between AI innovation and carbon emission intensity may therefore vary across regions. To examine this possibility, the sample was divided into eastern, central, and western regions. A pooled fixed-effects model was estimated by interacting the regional indicators with the linear and squared AI terms. This approach permits formal tests of regional differences while retaining the full sample, common controls, and province and year fixed effects. Table 15 reports the implied regional coefficients and the corresponding cross-regional tests.

The joint test of all regional interaction terms produces an F-statistic of 1.61 and a p-value of 0.1988. Coefficient homogeneity across the three regions therefore cannot be rejected. The pairwise tests likewise show no regional differences at the 5% level. The central–eastern comparison provides only weak evidence of a difference at the 10% level ($F = 2.70$, $p = 0.0839$), whereas the western–eastern and central–western comparisons are statistically insignificant. The regional estimates should therefore be interpreted primarily as descriptive differences in coefficient patterns rather than as conclusive evidence of systematic regional heterogeneity.

For the eastern region, the linear coefficient is statistically insignificant, whereas the quadratic coefficient is negative (-0.0094) and significant at the 5% level. The implied turning point lies below the observed range and therefore has no meaningful interpretation as an interior threshold. The point estimates imply a negative marginal effect throughout the observed range, although the slope at the lower bound is imprecisely estimated. This pattern is consistent with the comparatively mature stage of AI development in the eastern region, where efficiency gains may already offset expansion-related pressures. This interpretation nevertheless remains suggestive because the relevant cross-regional differences are not statistically significant.

The central region exhibits the strongest negative curvature. Its linear coefficient is positive (0.2701) and significant at the 5% level, while its quadratic coefficient is negative (-0.0433) and significant at the 1% level. The corresponding turning point is 3.116 and lies within the observed range. These estimates indicate a positive slope at relatively low levels of AI innovation and a negative slope at higher levels, consistent with a transitional stage in which the energy requirements associated with technological expansion coexist

with efficiency gains generated by industrial upgrading and wider technology diffusion. However, the joint difference between the central and eastern regions is significant only at the 10% level and should not be treated as conclusive evidence of regional divergence.

Table 15. Regional heterogeneity in the nonlinear relationship.

Variable/Statistic	Eastern Region	Central Region	Western Region
Linear AI term	−0.0060 (0.0345)	0.2701 ** (0.1237)	0.0561 (0.0360)
Squared AI term	−0.0094 ** (0.0046)	−0.0433 *** (0.0152)	−0.0163 ** (0.0066)
Implied turning point	Outside sample range	3.116 (0.483)	1.726 (0.820)
Number of provinces	13	6	11
Observations	182	84	154
Comparison	F-Statistic	*p*-Value	Inference
Central versus eastern	2.70	0.0839	Weak evidence of a difference
Western versus eastern	0.79	0.4653	No significant difference
Central versus western	1.48	0.2444	No significant difference
All regional interaction terms	1.61	0.1988	Fails to reject coefficient homogeneity

Notes: Region-specific coefficients are linear combinations obtained from a pooled interaction model in which the eastern region is the reference group. The dependent variable is the logarithm of carbon emission intensity. All specifications include the full set of controls and province and year fixed effects. Robust standard errors clustered at the provincial level are reported in parentheses. The turning points are calculated from the corresponding regional linear and quadratic coefficients. The eastern turning point is not reported because its point estimate lies outside the observed range and therefore does not represent an interior threshold. The eastern region comprises Beijing, Tianjin, Hebei, Liaoning, Jilin, Heilongjiang, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong, and Hainan. The central region comprises Shanxi, Anhui, Jiangxi, Henan, Hubei, and Hunan. The remaining 11 provinces constitute the western region. *** and ** denote significance at the 1% and 5% levels, respectively.

The western region also has a negative and statistically significant quadratic coefficient (−0.0163), while its positive linear coefficient is not statistically significant. Its implied turning point is 1.726 and lies within the sample range. These estimates indicate concavity and an interior point estimate, although the initial positive slope is imprecisely estimated. Moreover, the western coefficients do not differ significantly from either the eastern or central coefficients. The western estimates should therefore not be interpreted as establishing a statistically distinct regional pattern.

Overall, the quadratic coefficients are negative and statistically significant in all three regions, indicating that the concavity observed in the full sample is not driven by a single regional subsample. The fitted turning points follow the descriptive ordering of the eastern region below the observed range, followed by the western region and then the central region. This ordering is consistent with differences in technological maturity and development stage, but the formal tests do not establish statistically distinct regional relationships. Hypothesis 4 therefore receives limited support: the point estimates are broadly consistent with the predicted stage-dependent ordering, but the cross-regional coefficient differences are not conclusively identified.

7. Discussion

7.1. Main Findings

The results support a conditional, stage-dependent relationship between patent-based AI innovation intensity and carbon emission intensity. In the preferred two-way fixed-effects specification, the Lind–Mehlum test supports an inverted-U-shaped pattern within

the observed sample range. The robustness exercises consistently preserve the negative quadratic coefficient, although not every alternative specification independently establishes an inverted U shape. The Bartik and 2sCOPE-np analyses yield similar negative curvatures, but weak joint identification in the Bartik model and the maintained assumptions of the copula approach preclude a definitive causal interpretation.

The accounting decomposition locates most of the estimated curvature in the energy use component rather than the carbonization factor. This finding describes how the fitted relationship is allocated within the carbon-accounting identity; it does not establish energy intensity as an independently identified causal mediator. Human capital also changes the shape of the relationship, strengthening the initial upward slope and subsequent decline.

Regional differences are not statistically conclusive. The central region exhibits the strongest negative curvature and an interior turning point, although the central–eastern comparison is only weakly significant at the 10% level. The western point estimates also indicate concavity, but they are not jointly distinguishable from those of the eastern or central regions. The regional evidence therefore indicates descriptive differences in coefficient patterns rather than statistically established heterogeneity.

7.2. Interpretation and Contributions

7.2.1. Stage-Dependent Expansion and Efficiency Effects

Previous studies provide mixed evidence on the environmental consequences of AI, industrial intelligence, and related digital technologies. Some report lower carbon intensity or improved environmental performance, whereas others find heterogeneous, delayed, or nonlinear effects [5–11]. The present findings suggest that part of this variation may arise because the expansion costs and efficiency benefits of AI do not emerge simultaneously.

As a general purpose technology, AI requires complementary investment before its productivity effects are fully realized [32,33]. Early diffusion may involve computing infrastructure, equipment replacement, organizational restructuring, data system construction, and the expansion of electricity-intensive applications. These adjustments can raise energy demand before measurable efficiency gains appear. The energy requirements of large-scale computing and the possibility of rebound effects reinforce this initial expansion pressure [2–4,34].

At higher levels of adoption, a different set of effects may become more important. AI can improve production scheduling, predictive maintenance, load management, resource allocation, and the identification and diffusion of cleaner technologies [12,13,35–37]. The estimated inverted-U pattern is therefore consistent with a transition from an expansion-dominated stage to an efficiency-dominated stage.

This interpretation differs from a mechanical extension of the environmental Kuznets curve [62,63]. The explanatory variable is technological innovation intensity rather than income, and the proposed mechanisms concern the changing balance between complementary investment and efficiency gains during technological diffusion. The turning point should accordingly be understood as a sample-specific transition range, not as a fixed technological or policy threshold. The contribution lies in distinguishing the two countervailing effects and formally testing whether the resulting curve changes direction within the observed range [20,21].

7.2.2. Energy-Use and Carbonization Margins

Carbon emission intensity can be decomposed into energy consumption per unit of output and carbon emissions per unit of energy. This accounting structure has long been used to distinguish the contribution of energy efficiency from that of the energy mix and associated emission factors [23–25].

The estimates show that the negative curvature in carbon emission intensity is primarily associated with the energy use component. By contrast, the coefficients in the carbonization factor equation are smaller and statistically insignificant. The result suggests that within the present sample, the changing relationship between AI innovation and carbon emission intensity operates mainly through the amount of energy required to produce a unit of output rather than through a systematic change in the average carbon content of energy consumption.

This distinction is important for interpreting the empirical evidence. Carbon emissions are calculated from energy use, and carbon intensity and energy intensity share GDP as a denominator. Treating energy intensity as a conventional mediator would therefore combine an accounting relationship with an estimated behavioral relationship. The decomposition avoids this problem: it does not estimate a causal indirect effect but identifies where the fitted curvature appears within the carbon-accounting identity.

The result is consistent with research showing that AI and digital technologies can affect energy intensity through industrial process optimization, energy management, and cleaner production [13,36,37]. However, it does not demonstrate that these applications constitute an independently identified causal mechanism. Establishing such a mechanism would require more disaggregated evidence on technology adoption and energy use at the firm, facility, or industrial process level.

7.2.3. Human Capital as a Boundary Condition

The moderation results are consistent with the view that the consequences of new technologies depend on the capacity to absorb and apply them. Absorptive capacity theory and the human capital literature emphasize that prior knowledge and skills influence the adoption, adaptation, and productive use of external technologies [26–28,38,39].

The positive interaction with the linear AI term suggests that provinces with stronger human capital foundations may experience a more pronounced expansion-related effect at relatively low levels of AI innovation. At the same time, the negative interaction with the squared term indicates that the subsequent decline in carbon emission intensity becomes steeper. Human capital therefore appears to strengthen the stage dependence of the relationship rather than exerting a uniformly emission-reducing influence.

One possible explanation is that skilled regions can mobilize complementary investment and implement AI applications on a larger scale. This capacity may initially increase computing and infrastructure requirements, but it can also improve the conversion of AI-related knowledge into production efficiency, energy management, and organizational change. The insignificant standalone coefficient of human capital is not inconsistent with this interpretation: its empirical role lies primarily in altering the conditional relationship between AI innovation and carbon intensity.

The estimated interactions should nevertheless be interpreted with appropriate caution. Although the point estimates imply a small rightward movement of the turning point, the more clearly supported result concerns the stronger negative curvature. The evidence therefore establishes human capital as a boundary condition for the shape of the relationship, rather than a precisely estimated determinant of when the transition occurs.

7.3. Regional Heterogeneity

The nonlinear relationship is statistically clearest in central China. This finding is compatible with the industrial characteristics of the region, where manufacturing transformation, digital retrofitting, and energy efficiency upgrading may be occurring concurrently. Under these conditions, the expansion-related costs of AI adoption and its subsequent efficiency gains are more likely to be visible within the sample [8,11,31].

The eastern region has an insignificant linear coefficient and a significantly negative quadratic coefficient. Its fitted vertex lies below the observed range, so the point estimates imply a declining relationship throughout the sample, although the slope at the lower bound is imprecisely estimated. One possible interpretation is that AI applications and energy-management practices are already more extensively integrated in these provinces. This explanation remains tentative because the formal cross-regional differences are not statistically significant.

The western region has a positive but statistically insignificant linear coefficient, a significantly negative quadratic coefficient, and an interior point estimate of the turning point. However, the pairwise tests do not reject equality between the western coefficients and those of either the eastern or central region. The western estimates therefore support concavity within the region but do not establish a statistically distinct regional relationship. Because the analysis covers only 30 provinces, all regional comparisons should be interpreted cautiously.

The broader implication is that the environmental consequences of AI innovation are unlikely to be determined by innovation volume alone. Their realization may also depend on the industrial settings in which the technology is deployed and on the complementary capabilities available to convert technological knowledge into improvements in energy use. Further research using city-, industry-, or firm-level data is required to identify these regional mechanisms more directly.

7.4. Limitations and Future Research

Several limitations remain. First, the explanatory variable was based on strict-definition AI patent applications. The strict classification improves the measurement precision by excluding patents that are only loosely related to core AI technologies, but it does not broaden the conceptual coverage of the indicator. Patent applications capture the observable scale of AI-related knowledge production; they do not directly measure computing infrastructure, data availability, skilled labor, institutional support, or the depth of commercial application. The results should therefore be understood as applying to the patent-observable dimension of regional AI innovation rather than to the entire innovation environment. Future research could combine patents with measures of computing capacity, data resources, R&D inputs, talent, and institutional support.

Second, the empirical strategies do not eliminate all endogeneity concerns. The fixed-effects model controls for time-invariant provincial characteristics and common annual shocks, but it cannot remove all time-varying omitted factors or reverse causality.

The internal-structure Bartik design uses predetermined AI technology shares and subsequent non-China technology shocks. This construction limits several direct channels associated with broader historical or geographical instruments, but the exclusion restriction remains an identifying assumption. Randomization inference shows that the observed province–shock matching is not readily reproduced by arbitrary reassignment, but it cannot establish the exclusion restriction. Moreover, the low Kleibergen–Paap statistic limits the precision of the Bartik estimates [53–57].

The 2sCOPE-np approach provides evidence under a different set of assumptions, but it is not assumption-free. Its identifying information comes from the conditional distribution of the potentially endogenous variable, and an insignificant control-function term may reflect either limited endogeneity or limited power [60]. For these reasons, the Bartik and copula estimates are treated as complementary checks rather than substitutes for a stronger quasi-natural experiment.

Third, the accounting decomposition identifies how the estimated coefficients are allocated between energy intensity and the carbonization factor, but it does not establish a behavioral sequence comparable with causal mediation. Firm- or facility-level data on computing investment, electricity use, production processes, and technology adoption would provide a stronger basis for distinguishing specific mechanisms.

Fourth, provincial carbon emissions for 2023 and 2024 were extrapolated using province-specific emissions-to-energy conversion coefficients anchored to the 2022 CEADs observations. The backcasting exercise and the conservative 2011–2021 specification reduce concerns that the principal result is driven by this extension procedure, but they cannot eliminate measurement error in the extrapolated values. Future research should replace these observations when updated CEADs provincial inventories become available [42,43,47].

Finally, the provincial level of aggregation conceals substantial variation between cities, industries, and firms. It also limits the precision of the regional subsample estimates. Whether the identified curvature and decomposition pattern extend to other countries remains an open question. Comparable international evidence would help determine which results reflect general features of AI diffusion and which are specific to China's industrial and energy systems.

8. Conclusions

8.1. Research Conclusions

Using panel data for 30 Chinese provinces from 2011 to 2024, this study examined the nonlinear relationship between patent-based AI innovation intensity and carbon emission intensity. The preferred two-way fixed-effects specification yields a positive linear coefficient and a negative quadratic coefficient. The Lind–Mehlum test supports an inverted-U-shaped relationship within the observed sample range. The negative curvature remains under alternative outcome measures, patent classifications, sample treatments, and lag structures, although each of these robustness specifications should not be interpreted as an independent formal test of an inverted U shape.

The supplementary identification analyses support a cautious interpretation. The internal-structure Bartik model retains negative curvature and an interior point estimate of the turning point, but joint identification is weak, and the endpoint test falls just short of the 10% significance level ($p = 0.1007$). The 2sCOPE-np estimates also preserve the negative quadratic coefficient, while the copula-control-function term is statistically insignificant. These findings strengthen confidence in the estimated shape but do not establish a definitive causal effect or a universal transition threshold.

The accounting decomposition attributes most of the estimated quadratic coefficient to the energy use component, while the carbonization component is smaller and statistically insignificant. Human capital makes the estimated relationship more concave, although its effect on the precise location of the turning point is not separately established. Regional interaction tests do not reject complete coefficient homogeneity, although the point estimates show the strongest curvature in central China and the central–eastern comparison is weakly significant at the 10% level.

The results therefore do not imply that increasing the number of AI patents automatically reduces carbon emission intensity. The environmental consequences of AI innovation depend on the stage of regional development, the extent to which technological knowledge is converted into lower energy use per unit of output, and the availability of complementary regional capabilities.

From a sustainable development perspective, this study provides a quantifiable and stage-sensitive framework for assessing the environmental implications of patent-based AI innovation. The findings caution against assuming that AI innovation is intrinsically green:

its association with carbon emission intensity varies across stages, as initial expansion may increase energy demand, whereas subsequent emission-reduction gains depend on the conversion of technological knowledge into greater production and energy-use efficiency. The decomposition, moderation, and regional analyses further indicate that this conversion is conditioned by energy use, human capital, and local implementation conditions. The framework therefore links innovation measurement to the dynamic monitoring of carbon emission intensity and supports a more integrated assessment of technological development, energy use, and carbon mitigation. Accordingly, the framework provides a quantitative basis for dynamic and context-sensitive sustainability assessment and policy monitoring across regions at different stages of AI innovation.

8.2. Policy Implications

The policy implications follow directly from the empirical sequence of the analysis: the estimated relationship is stage-dependent, the principal physical margin is energy intensity, the relationship's shape is conditioned by human capital, and the regional point estimates vary in magnitude, although the cross-regional differences are not statistically conclusive.

First, AI and low-carbon policies should be evaluated through a stage-sensitive framework. The inverted-U-shaped relationship indicates that the environmental consequences of AI innovation may differ between the expansion and application stages. Policy evaluation should therefore extend beyond patent counts, investment volumes, or short-term changes in emissions. Regional monitoring systems should track whether additional innovation is accompanied by changes in electricity demand, computing efficiency, energy use per unit of output, and the diffusion of energy-saving applications. The estimated turning point should not be converted into a uniform administrative threshold because its location varies between specifications and reflects the characteristics of the present sample. It is more appropriately used as evidence that policy priorities must change as regional AI systems develop.

Second, policy intervention should focus on the energy use margin identified by the decomposition. The finding that most of the estimated curvature is associated with energy intensity provides a direct basis for integrating energy performance requirements into AI infrastructure and application policy. New data centers and computing facilities should be subject to transparent energy-efficiency standards, utilization requirements, carbon monitoring, and renewable-energy procurement arrangements. At the application stage, support should be directed toward verifiable reductions in energy use through industrial process control, predictive maintenance, building management, transport scheduling, and power system optimization. Patent production and infrastructure expansion should be assessed alongside measurable improvements in energy productivity. Although the carbonization component is not statistically significant in the decomposition, continued decarbonization of electricity supply remains necessary to prevent growing computing demand from creating high-carbon lock-in.

Third, human capital policy should be developed together with energy and industrial policy. The moderation results show that human capital strengthens both sides of the nonlinear relationship rather than producing a uniformly negative effect on emissions. Expanding the supply of technical personnel alone is therefore insufficient. Training programs should combine expertise in algorithms and data systems with knowledge of industrial processes, energy management, environmental engineering, and carbon accounting. Cooperation between universities, research institutes, energy providers, and industrial firms should focus on the implementation and evaluation of low-carbon applications. Such coordination can improve the conversion of technological knowledge into operational energy savings while limiting the energy costs associated with infrastructure expansion.

Fourth, implementation should reflect the regional evidence rather than follow a uniform geographical template. In the central region, where the point estimates indicate the strongest curvature and the difference from the eastern region is only weakly significant at the 10% level, policy should prioritize the application of AI to industrial retrofitting, production scheduling, equipment management, and supply-chain energy efficiency.

The objective should be to shorten the transition from innovation expansion to measurable efficiency improvement. In the eastern region, where the estimates do not show an interior inverted U shape, policy should place greater emphasis on evaluating the actual energy performance of established digital infrastructure and disseminating applications with demonstrated environmental benefits. In the western region, the estimates are consistent with a nonlinear relationship, but the cross-regional differences are not statistically conclusive. Policy should therefore proceed through locally evaluated programs that combine renewable energy resources, efficient computing infrastructure, technical training, and industrial absorptive capacity rather than assuming that resource advantages alone will generate low-carbon outcomes.

Author Contributions: Conceptualization, S.S. and S.W.; methodology, S.S. and S.W.; software, S.S.; validation, S.S. and S.W.; formal analysis, S.S.; investigation, S.S.; resources, S.S.; data curation, S.S.; writing—original draft preparation, S.S.; writing—review and editing, S.S. and S.W.; visualization, S.S.; supervision, S.W.; project administration, S.S. and S.W.; funding acquisition, S.S. and S.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China under Grant Nos. 71773118 and 72373136.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data supporting the findings of this study have been openly published in Figshare and are available at https://doi.org/10.6084/m9.figshare.32915594 (accessed on 7 July 2026).

Acknowledgments: During the preparation of this manuscript, the authors used DeepSeek-V4-Flash to assist with the translation of selected text from Chinese into English and minor language editing. The authors have reviewed, revised, and verified all output and take full responsibility for the integrity and accuracy of the content of this publication.

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

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial intelligence
CEADs	China Emission Accounts and Datasets
EDGAR	Emissions Database for Global Atmospheric Research
EKC	Environmental Kuznets curve
FE	Fixed effects
2SLS	Two-stage least squares
LIML	Limited-information maximum likelihood
IPCC	Intergovernmental Panel on Climate Change
IV	Instrumental variable
KP	Kleibergen–Paap
SW	Sanderson–Windmeijer

References

1. Vinuesa, R.; Azizpour, H.; Leite, I.; Balaam, M.; Dignum, V.; Domisch, S.; Felländer, A.; Langhans, S.D.; Tegmark, M.; Fuso Nerini, F. The Role of Artificial Intelligence in Achieving the Sustainable Development Goals. *Nat. Commun.* **2020**, *11*, 233. [CrossRef] [PubMed]
2. Strubell, E.; Ganesh, A.; McCallum, A. Energy and policy considerations for deep learning in NLP. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, 28 July–2 August 2019*; Korhonen, A., Traum, D., Màrquez, L., Eds.; Association for Computational Linguistics: Stroudsburg, PA, USA, 2019; pp. 3645–3650. [CrossRef]
3. Masanet, E.; Shehabi, A.; Lei, N.; Smith, S.; Koomey, J. Recalibrating Global Data Center Energy-Use Estimates. *Science* **2020**, *367*, 984–986. [CrossRef] [PubMed]
4. De Vries, A. The Growing Energy Footprint of Artificial Intelligence. *Joule* **2023**, *7*, 2191–2194. [CrossRef]
5. Liu, J.; Liu, L.; Qian, Y.; Song, S. The Effect of Artificial Intelligence on Carbon Intensity: Evidence from China's Industrial Sector. *Socio-Econ. Plan. Sci.* **2022**, *83*, 101002. [CrossRef]
6. Meng, X.; Xu, S.; Zhang, J. How Does Industrial Intelligence Affect Carbon Intensity in China? Empirical Analysis Based on Chinese Provincial Panel Data. *J. Clean. Prod.* **2022**, *376*, 134273. [CrossRef]
7. Wang, L.; Chen, Q.; Dong, Z.; Cheng, L. The Role of Industrial Intelligence in Peaking Carbon Emissions in China. *Technol. Forecast. Soc. Change* **2024**, *199*, 123005. [CrossRef]
8. Ding, T.; Li, J.; Shi, X.; Li, X.; Chen, Y. Is Artificial Intelligence Associated with Carbon Emissions Reduction? Case of China. *Resour. Policy* **2023**, *85*, 103892. [CrossRef]
9. Zhong, W.; Liu, Y.; Dong, K.; Ni, G. Assessing the Synergistic Effects of Artificial Intelligence on Pollutant and Carbon Emission Mitigation in China. *Energy Econ.* **2024**, *138*, 107829. [CrossRef]
10. Liu, B.; Yang, X.; Zhang, J. Nonlinear Effect of Industrial Robot Applications on Carbon Emissions: Evidence from China. *Environ. Impact Assess. Rev.* **2024**, *104*, 107297. [CrossRef]
11. Dong, M.; Wang, G.; Han, X. Impacts of Artificial Intelligence on Carbon Emissions in China: In Terms of Artificial Intelligence Type and Regional Differences. *Sustain. Cities Soc.* **2024**, *113*, 105682. [CrossRef]
12. Zhou, W.; Zhang, Y.; Li, X. Artificial Intelligence, Green Technological Progress, Energy Conservation, and Carbon Emission Reduction in China: An Examination Based on Dynamic Spatial Durbin Modeling. *J. Clean. Prod.* **2024**, *446*, 141142. [CrossRef]
13. Li, H.; Lu, Z.; Zhang, Z.; Tanasescu, C. How Does Artificial Intelligence Affect Manufacturing Firms' Energy Intensity? *Energy Econ.* **2025**, *141*, 108109. [CrossRef]
14. Li, Y.; Zhang, Y.; Pan, A.; Han, M.; Veglianti, E. Carbon Emission Reduction Effects of Industrial Robot Applications: Heterogeneity Characteristics and Influencing Mechanisms. *Technol. Soc.* **2022**, *70*, 102034. [CrossRef]
15. Chen, P.; Gao, J.; Ji, Z.; Liang, H.; Peng, Y. Do Artificial Intelligence Applications Affect Carbon Emission Performance? Evidence from Panel Data Analysis of Chinese Cities. *Energies* **2022**, *15*, 5730. [CrossRef]
16. Li, Z.; Wang, J. The Dynamic Impact of Digital Economy on Carbon Emission Reduction: Evidence City-Level Empirical Data in China. *J. Clean. Prod.* **2022**, *351*, 131570. [CrossRef]
17. OECD. *OECD Patent Statistics Manual*; OECD Publishing: Paris, France, 2009. [CrossRef]
18. Griliches, Z. Patent Statistics as Economic Indicators: A Survey. *J. Econ. Lit.* **1990**, *28*, 1661–1707.
19. World Intellectual Property Organization. *WIPO Technology Trends 2019: Artificial Intelligence*; WIPO: Geneva, Switzerland, 2019. [CrossRef]
20. Haans, R.F.J.; Pieters, C.; He, Z.-L. Thinking about U: Theorizing and Testing U- and Inverted U-Shaped Relationships in Strategy Research. *Strateg. Manag. J.* **2016**, *37*, 1177–1195. [CrossRef]
21. Lind, J.T.; Mehlum, H. With or Without U? The Appropriate Test for a U-Shaped Relationship. *Oxf. Bull. Econ. Stat.* **2010**, *72*, 109–118. [CrossRef]
22. Zhang, L.; Mu, R.; Zhan, Y.; Yu, J.; Liu, L.; Yu, Y.; Zhang, J. Digital Economy, Energy Efficiency, and Carbon Emissions: Evidence from Provincial Panel Data in China. *Sci. Total Environ.* **2022**, *852*, 158403. [CrossRef] [PubMed]
23. Dong, F.; Yu, B.; Hadachin, T.; Dai, Y.; Wang, Y.; Zhang, S.; Long, R. Drivers of Carbon Emission Intensity Change in China. *Resour. Conserv. Recycl.* **2018**, *129*, 187–201. [CrossRef]
24. Zheng, X.; Lu, Y.; Yuan, J.; Baninla, Y.; Zhang, S.; Stenseth, N.C.; Hessen, D.O.; Tian, H.; Obersteiner, M.; Chen, D. Drivers of Change in China's Energy-Related CO_2 Emissions. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 29–36. [CrossRef] [PubMed]
25. Ang, B.W. The LMDI Approach to Decomposition Analysis: A Practical Guide. *Energy Policy* **2005**, *33*, 867–871. [CrossRef]
26. Cohen, W.M.; Levinthal, D.A. Absorptive Capacity: A New Perspective on Learning and Innovation. *Adm. Sci. Q.* **1990**, *35*, 128–152. [CrossRef]
27. Nelson, R.R.; Phelps, E.S. Investment in Humans, Technological Diffusion, and Economic Growth. *Am. Econ. Rev.* **1966**, *56*, 69–75.
28. Benhabib, J.; Spiegel, M.M. The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-Country Data. *J. Monet. Econ.* **1994**, *34*, 143–173. [CrossRef]

29. Qian, Y.; Liu, J.; Shi, L.; Forrest, J.Y.-L.; Yang, Z. Can Artificial Intelligence Improve Green Economic Growth? Evidence from China. *Environ. Sci. Pollut. Res.* **2023**, *30*, 16418–16437. [CrossRef] [PubMed]
30. Li, J.; Chen, L.; Chen, Y.; He, J. Digital Economy, Technological Innovation, and Green Economic Efficiency—Empirical Evidence from 277 Cities in China. *Manag. Decis. Econ.* **2022**, *43*, 616–629. [CrossRef]
31. Zheng, J.; Mi, Z.; Coffman, D.; Milcheva, S.; Shan, Y.; Guan, D.; Wang, S. Regional Development and Carbon Emissions in China. *Energy Econ.* **2019**, *81*, 25–36. [CrossRef]
32. Bresnahan, T.F.; Trajtenberg, M. General Purpose Technologies "Engines of Growth"? *J. Econom.* **1995**, *65*, 83–108. [CrossRef]
33. Brynjolfsson, E.; Rock, D.; Syverson, C. The Productivity J-Curve: How Intangibles Complement General Purpose Technologies. *Am. Econ. J. Macroecon.* **2021**, *13*, 333–372. [CrossRef]
34. Sorrell, S.; Dimitropoulos, J.; Sommerville, M. Empirical Estimates of the Direct Rebound Effect: A Review. *Energy Policy* **2009**, *37*, 1356–1371. [CrossRef]
35. Acemoglu, D.; Aghion, P.; Bursztyn, L.; Hemous, D. The Environment and Directed Technical Change. *Am. Econ. Rev.* **2012**, *102*, 131–166. [CrossRef] [PubMed]
36. Liu, L.; Yang, K.; Fujii, H.; Liu, J. Artificial Intelligence and Energy Intensity in China's Industrial Sector: Effect and Transmission Channel. *Econ. Anal. Policy* **2021**, *70*, 276–293. [CrossRef]
37. Zhou, W.; Zhuang, Y.; Chen, Y. How Does Artificial Intelligence Affect Pollutant Emissions by Improving Energy Efficiency and Developing Green Technology? *Energy Econ.* **2024**, *131*, 107355. [CrossRef]
38. Acemoglu, D. Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality. *Q. J. Econ.* **1998**, *113*, 1055–1089. [CrossRef]
39. Lucas, R.E. On the Mechanics of Economic Development. *J. Monet. Econ.* **1988**, *22*, 3–42. [CrossRef]
40. Audretsch, D.B.; Feldman, M.P. R&D Spillovers and the Geography of Innovation and Production. *Am. Econ. Rev.* **1996**, *86*, 630–640.
41. Krugman, P. Increasing Returns and Economic Geography. *J. Polit. Econ.* **1991**, *99*, 483–499. [CrossRef] [PubMed]
42. Shan, Y.; Guan, D.; Zheng, H.; Ou, J.; Li, Y.; Meng, J.; Mi, Z.; Liu, Z.; Zhang, Q. China CO_2 Emission Accounts 1997–2015. *Sci. Data* **2018**, *5*, 170201. [CrossRef] [PubMed]
43. Shan, Y.; Huang, Q.; Guan, D.; Hubacek, K. China CO_2 Emission Accounts 2016–2017. *Sci. Data* **2020**, *7*, 54. [CrossRef] [PubMed]
44. Guan, Y.; Shan, Y.; Huang, Q.; Chen, H.; Wang, D.; Hubacek, K. Assessment to China's Recent Emission Pattern Shifts. *Earth's Future* **2021**, *9*, e2021EF002241. [CrossRef]
45. Shan, Y.; Liu, J.; Liu, Z.; Xu, X.; Shao, S.; Wang, P.; Guan, D. New Provincial CO_2 Emission Inventories in China Based on Apparent Energy Consumption Data and Updated Emission Factors. *Appl. Energy* **2016**, *184*, 742–750. [CrossRef]
46. IPCC. *2006 IPCC Guidelines for National Greenhouse Gas Inventories, Volume 2: Energy*; Eggleston, H.S., Buendia, L., Miwa, K., Ngara, T., Tanabe, K., Eds.; Institute for Global Environmental Strategies: Hayama, Japan, 2006. Available online: https://www.ipcc-nggip.iges.or.jp/public/2006gl/vol2.html (accessed on 8 August 2026).
47. Xu, J.; Guan, Y.; Oldfield, J.; Guan, D.; Shan, Y. China Carbon Emission Accounts 2020–2021. *Appl. Energy* **2024**, *360*, 122837. [CrossRef]
48. Janssens-Maenhout, G.; Crippa, M.; Guizzardi, D.; Muntean, M.; Schaaf, E.; Dentener, F.; Bergamaschi, P.; Pagliari, V.; Olivier, J.G.J.; Peters, J.A.H.W.; et al. EDGAR v4.3.2 Global Atlas of the Three Major Greenhouse Gas Emissions for the Period 1970–2012. *Earth Syst. Sci. Data* **2019**, *11*, 959–1002. [CrossRef]
49. Cameron, A.C.; Miller, D.L. A Practitioner's Guide to Cluster-Robust Inference. *J. Hum. Resour.* **2015**, *50*, 317–372. [CrossRef]
50. Durbin, J. Errors in Variables. *Rev. Int. Stat. Inst.* **1954**, *22*, 23–32. [CrossRef] [PubMed]
51. Wu, D.M. Alternative Tests of Independence between Stochastic Regressors and Disturbances. *Econometrica* **1973**, *41*, 733–750. [CrossRef]
52. Hausman, J.A. Specification Tests in Econometrics. *Econometrica* **1978**, *46*, 1251–1271. [CrossRef]
53. Goldsmith-Pinkham, P.; Sorkin, I.; Swift, H. Bartik Instruments: What, When, Why, and How. *Am. Econ. Rev.* **2020**, *110*, 2586–2624. [CrossRef]
54. Borusyak, K.; Hull, P.; Jaravel, X. Quasi-Experimental Shift-Share Research Designs. *Rev. Econ. Stud.* **2022**, *89*, 181–213. [CrossRef] [PubMed]
55. Adão, R.; Kolesár, M.; Morales, E. Shift-Share Designs: Theory and Inference. *Q. J. Econ.* **2019**, *134*, 1949–2010. [CrossRef]
56. Sanderson, E.; Windmeijer, F. A Weak Instrument F-Test in Linear IV Models with Multiple Endogenous Variables. *J. Econom.* **2016**, *190*, 212–221. [CrossRef] [PubMed]
57. Kleibergen, F.; Paap, R. Generalized Reduced Rank Tests Using the Singular Value Decomposition. *J. Econom.* **2006**, *133*, 97–126. [CrossRef]
58. Anderson, T.W.; Rubin, H. Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations. *Ann. Math. Stat.* **1949**, *20*, 46–63. [CrossRef]
59. Stock, J.H.; Wright, J.H. GMM with Weak Identification. *Econometrica* **2000**, *68*, 1055–1096. [CrossRef]

60. Hu, X.; Qian, Y.; Xie, H. *Correcting Endogeneity via Nonparametric Copula Control Functions*; NBER Working Paper No. 33607; National Bureau of Economic Research: Cambridge, MA, USA, 2025. [CrossRef]
61. Cameron, A.C.; Gelbach, J.B.; Miller, D.L. Bootstrap-Based Improvements for Inference with Clustered Errors. *Rev. Econ. Stat.* **2008**, *90*, 414–427. [CrossRef]
62. Grossman, G.M.; Krueger, A.B. Economic Growth and the Environment. *Q. J. Econ.* **1995**, *110*, 353–377. [CrossRef]
63. Stern, D.I. The Rise and Fall of the Environmental Kuznets Curve. *World Dev.* **2004**, *32*, 1419–1439. [CrossRef]

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.