Business Data Analytics for Urban Green Innovation: Identifying Multi-Condition Pathways with NCA and fsQCA

Shuokai Wang[1], Bo Zeng[1,*]

[1] School of Economics and Management, Beijing Forestry University, Beijing, China 100091

* Email: zengbo@bjfu.edu.cn (Corresponding Author)

Received 30 December 2025 · Revised 13 February 2026 · Accepted 30 March 2026 · Published 30 June 2026

Abstract

Urban green innovation depends on combinations of digital capabilities, research capacity, finance, and environmental policy, but conventional average-effect models do not directly identify conditions that constrain outcomes or alternative configurations associated with them. This article develops a transparent business data analytics workflow that combines necessary condition analysis (NCA) and fuzzy-set qualitative comparative analysis (fsQCA). Because no observational urban dataset is available for this study, the workflow is demonstrated with 96 synthetic city profiles generated from a disclosed rule rather than presented as empirical evidence about real cities. Four hypothetical conditions—digital analytics readiness, knowledge capacity, green finance, and policy support—are calibrated using stated anchors. Ceiling-envelopment free-disposal-hull NCA, a permutation test, fuzzy necessity analysis, truth-table construction, and Boolean minimization are implemented in a reproducible Python program. The simulated outcome is constructed to require digital readiness and either knowledge capacity or a conjunction of finance and policy support. In the demonstration, the digital condition has an NCA effect size of 0.476 and fuzzy necessity consistency of 1.000. The high-outcome solution recovers two pathways, digital readiness combined with knowledge capacity and digital readiness combined with finance and policy support, with overall consistency and coverage of approximately 0.993. The results illustrate how to reconcile findings across methods and how high- and low-outcome configurations can be analyzed separately. They reproduce a known data-generating rule and do not establish a causal necessity, policy threshold, or pathway for actual cities. The framework provides a replicable starting point for future research using independently collected urban data.

Keywords: urban innovation ecosystems; digital readiness; green finance; environmental governance; configurational methods; necessity analysis; synthetic data

Business Data Analytics for Urban Green Innovation: Identifying Multi-Condition Pathways with NCA and fsQCA

1. Introduction

Cities face the combined challenge of improving economic productivity while reducing pollution, resource use, and greenhouse-gas emissions. Innovation can support this transition when firms, universities, financiers, and local governments develop and implement environmentally beneficial products or processes. Yet the presence of inventive activity does not automatically mean that innovations are adopted, that emissions fall, or that benefits are distributed fairly. Researchers therefore need to distinguish measurable innovation outputs from the broader environmental and social outcomes that policymakers ultimately seek (Liu & Dong, 2021; Shang et al., 2022; Wang, Zhang, et al., 2022).

Urban green innovation is embedded in an ecosystem rather than produced by a single factor. A city may have sophisticated research institutions but limited access to financing for early-stage technologies. Another may have supportive policy instruments and finance but lack the data capability needed to identify demand, evaluate firms, or connect projects with suitable investors. The potential contribution of digital systems is likewise contingent on technical skills, access to reliable data, organizational processes, and safeguards. Evidence from urban and regional research points to interactions among digital development, environmental regulation, finance, and innovation activity, while also showing that their effects vary across settings (Bai & Zhang, 2025; Dian et al., 2024; Hu et al., 2022; Zhang, Wang, et al., 2024).

Most empirical research asks whether a factor is associated with a higher average level of green innovation, often using panel regression or a policy evaluation design. Those questions are valuable but are not identical to asking whether some condition is indispensable for a given level of the outcome or whether different combinations of conditions are compatible with high innovation. A city with weak performance in one dimension may compensate through other resources, provided that a critical bottleneck has been addressed. Necessary condition analysis (NCA) focuses on possible bottlenecks, whereas fuzzy-set qualitative comparative analysis (fsQCA) examines set relations among combinations of conditions (Dul, 2016; Fiss, 2011; Greckhamer et al., 2018).

The distinction matters for decisions. An average association between green credit and patenting does not by itself show that credit is necessary for innovation, nor does a positive correlation guarantee that the same investment will work in a city with a different knowledge base or policy environment. Conversely, a condition may be a prerequisite without being sufficient on its own. A diagnostic workflow that reports both necessity and configurations can help a decision maker ask more precise questions about which combinations deserve further investigation. Its output should not, however, be treated as a causal policy recommendation without suitable theory, case knowledge, identification, and validation.

Existing work has already applied configurational and necessary-condition methods to urban green innovation. Studies of urban innovation ecosystems and structural diffusion have combined these approaches, and recent work investigates digital innovation ecosystems with regional or city-level QCA (Guo et al., 2023; Li et al., 2023; Liu et al., 2025). The present paper does not claim to introduce the first combined use of NCA and fsQCA in this field. Its contribution is a transparent, explicitly simulated business analytics demonstration that exposes the data-generating rule, calculation choices, and differences between the two methods. This design allows readers to see which patterns are recovered because they were built into the data rather than attributing them to observations of actual cities.

The research question is methodological: how can NCA and fsQCA be combined to examine potential bottlenecks and multi-condition pathways for urban green innovation while keeping calibration, uncertainty, and data limitations visible? To answer it, we define four theoretically motivated dimensions—digital analytics readiness, knowledge and research capacity, green finance, and environmental policy support—and

a green innovation outcome. We then generate synthetic city profiles under a known rule, estimate NCA ceilings and fuzzy-set relationships, and compare the results. The construction is illustrative: it supplies no observed city names, official indicators, field measurements, or real intervention effects.

The article contributes a documented pipeline linking conceptual definitions, data preparation, calibration, necessity checks, configurational minimization, and sensitivity analysis. It also illustrates why a significant NCA ceiling and a high fuzzy-set necessity score are not interchangeable and why a solution with high consistency can merely recover a rule imposed during simulation. The remainder of the article reviews the literature, describes the research design and calculations, reports the simulated results, and discusses how the framework could be extended to real urban data.

2. Literature Review and Analytical Framework

2.1 Green innovation as an urban ecosystem outcome

Green innovation can be represented by environmentally relevant patents, new products, process changes, or other outputs. Each measure captures a different stage of innovation and has different shortcomings. Patent counts can be organized by application date and technology classification, but they do not directly measure implementation or environmental performance. An efficiency measure additionally requires a defensible choice of inputs, desirable outputs, and potentially undesirable outputs. Mixing patent counts with efficiency scores without explaining the difference can obscure the meaning of a high-outcome set (Guo et al., 2023; Li et al., 2023; Zhang et al., 2020).

Urban innovation is shaped by relationships among firms, research institutions, investors, suppliers, and local authorities. Studies of regional knowledge spillovers and urban innovation networks show why proximity and collaboration can influence the movement and combination of knowledge (Fritsch & Franke, 2004; Johnston & Huggins, 2016). Intercity linkages and collaborative agglomeration can also alter the reach of an innovation system (Lin et al., 2021; Zeng et al., 2022). These relationships motivate including knowledge capacity in the framework, but no particular level of university activity or network connectivity is assumed to be universally necessary for all green innovations.

A useful theoretical distinction is between the potential to innovate and the conditions under which innovative activity can be financed, directed, and used. Knowledge resources contribute technical options; finance helps projects withstand development costs and risks; policy changes incentives and obligations; and digital analytics can improve the organization of information. These mechanisms can interact. For example, a research-intensive city may develop projects despite limited local green credit, while a city with less research capacity may draw on financial support and policy coordination. Such statements are propositions to examine with appropriate data, not empirical findings of the present synthetic example.

2.2 Digital readiness and business data analytics

Digital readiness is more than the existence of a broadband connection or an online portal. It includes the ability to access, integrate, govern, and analyze data in support of decisions made by firms, investors, and public agencies. An analytical capability may improve the matching of green projects to financing, reveal differences across industries, or make monitoring of project performance more timely. It can also be constrained by poor-quality records, incompatible identifiers, limited skills, and restrictions on the use of sensitive information. A nominal digital investment is therefore not equivalent to a verified analytics capability.

Recent studies relate urban digital development and digital innovation ecosystems to environmental or green innovation outcomes, but their measures and designs differ (Bai & Zhang, 2025; Dian et al., 2024; Gao et al., 2024; Wang, Sun, et al., 2022). A digital-economy index may combine infrastructure, industry size, and service adoption, whereas an analytics-readiness indicator may focus more narrowly on data integration, staff

competence, governance, and effective use. Treating these constructs as interchangeable would make the interpretation of an identified pathway ambiguous. In a real application, researchers should define the organizational level, geographic coverage, and time period before choosing indicators.

Digital systems can create risks as well as opportunities. Unequal access to data may favor large firms; data-sharing platforms can introduce privacy or cyber risks; and energy use associated with digital infrastructure may affect environmental outcomes. A city that performs well on patent counts may not improve environmental quality if the inventions are unused or displaced by additional resource consumption. These considerations are consistent with the broader literature on environmental regulation, technological change, and the possibility of rebound effects (Wang, Zhang, et al., 2022; Zhang, Zhu, et al., 2024). The framework consequently treats digital readiness as one potential enabling condition, not as a guaranteed driver of sustainable development.

2.3 Finance, policy, and knowledge complementarities

Green finance can include loans, investment, guarantees, or other instruments designed to support environmentally beneficial activities. Research on pilot policies and green financial development examines whether access to finance and financial incentives are associated with changes in green innovation (Hu et al., 2023; Jing & Liu, 2024; Zhang, Wang, et al., 2024). Some research also considers how finance interacts with environmental regulation and corporate investment (Han & Cai, 2024; Yin et al., 2023). These studies support examining the interplay of conditions, but their estimates should not be transferred mechanically to a new group of cities or used as the numerical inputs of our simulation.

Environmental policy support may include public procurement, standards, enforcement, disclosure, targeted subsidies, and technical assistance. The mechanisms are not equivalent: a binding standard can create demand for cleaner processes, while financial support may reduce the cost of developing them. The Porter-hypothesis literature studies circumstances in which environmental regulation and innovation can coexist, but it does not imply that every regulation increases innovation or improves competitiveness (Ambec et al., 2013; Bitat, 2018; Qiu et al., 2018). The design and enforcement of a policy, local industrial structure, and firms' capabilities can alter its consequences (Zhang et al., 2020; Zhang, Zhu, et al., 2024).

The knowledge dimension encompasses research personnel, technical expertise, access to scientific institutions, and the ability to collaborate or absorb external ideas. Spillovers and cooperation can connect local firms to knowledge produced elsewhere, so local inputs and local outputs need not have a one-to-one relationship (Fritsch & Franke, 2004; Lin et al., 2021; Xie & Wang, 2020). A green innovation ecosystem perspective invites analysis of combinations: knowledge may be valuable in conjunction with digital systems, while finance and a coherent policy environment may provide an alternative route in some contexts. These are theory-informed possibilities, not assumptions that empirical data must confirm.

2.4 Necessary and sufficient relationships

A necessary condition is a constraint that must be present, at a specified level, for an outcome to occur at a specified level; by itself it need not produce that outcome. A sufficient configuration is a combination whose presence is consistently associated with the outcome within the analyzed cases. Different sufficient configurations can coexist. Set-theoretic analysis represents these relationships in terms of membership, whereas NCA evaluates the upper boundary of a two-variable scatterplot (Dul, 2016; Dul, 2024; Pappas & Woodside, 2021). Their outputs should be compared, not assumed to be numerically interchangeable.

NCA estimates an empty ceiling zone and reports its size relative to a defined scope. Its interpretation depends on the choice of ceiling line, the observed ranges, outliers, and a substantive explanation for why an empty region should exist. A permutation test evaluates whether an observed ceiling effect could arise under a specified random reassignment of the outcome values; it does not establish a causal mechanism or demonstrate external validity (Dul et al., 2020; Dul et al., 2023; Dul, 2022). Methodological discussions also

emphasize that the logic of necessity and the formal properties of particular measures should be scrutinized rather than treated as self-evident (Thiem, 2016; Thiem, 2021).

fsQCA permits partial membership in sets such as high finance or high green innovation and then evaluates combinations of conditions using truth tables and set relations. Researchers must justify calibration anchors, report consistency and coverage, handle contradictory configurations, and explain how unobserved combinations enter minimization (Greckhamer et al., 2018; Pappas & Woodside, 2021; Veri, 2018). A pathway with high consistency can have low coverage, meaning that it describes only a small share of the outcome set. Limited diversity is especially important when the number of possible combinations grows quickly relative to the available cases (Cooper & Glaesser, 2016b; Thiem, 2014; Thiem, 2022).

Both methods are sensitive to case selection and measurement. A necessary-condition claim can be uninformative if nearly every case belongs strongly to the proposed condition; a configurational result can be unstable if a few cases determine a truth-table row. Case weighting, temporal change, and the geographic level of analysis can also alter conclusions (Cooper & Glaesser, 2016a; Verweij & Vis, 2021). The integrated workflow should therefore present the substantive meaning of each condition, the distribution of cases, and sensitivity checks before drawing practical conclusions. In our study, the known synthetic construction gives these checks a teaching function rather than an inferential one.

3. Research Design and Methods

3.1 Study design and decision boundary

This is a methodological demonstration using hypothetical data. The 96 cases are labeled C001–C096 and do not correspond to actual Chinese cities or any other identifiable locations. Their four inputs and outcome are generated by the algorithm described below, not collected from statistical yearbooks, patents, financial institutions, surveys, or administrative records. The analysis does not claim that Beijing or any other city follows the reported combinations. Its numerical output illustrates the mechanics of an integrated workflow and allows a reader to check whether the stated rule is recovered.

The intended real-world decision context is a city-level assessment of green innovation readiness. A public or private analyst might wish to identify which combinations of data capability, technical knowledge, finance, and policy deserve case investigation. The framework does not instruct a city to maximize patent counts at any cost. An operational study would require a clear definition of the policy question, a coherent spatial unit, comparable data, outcome validation, and information on sectoral composition and environmental performance. Those requirements are discussed later but are not satisfied by the synthetic illustration.

Table 1 defines the conditions and outcome. Each hypothetical input is a normalized index on a 0–100 scale, where larger values represent stronger membership in the relevant construct. These indices are abstract teaching variables. They are not reported as official measurements, their units are not money or emissions, and their relationship to real indicators has not been estimated. The outcome is similarly an artificial green innovation index. Although its label makes the example interpretable, its values result from the stated rule and should not be read as observed green patenting or verified environmental improvement.

Table 1. Constructs in the Synthetic Demonstration

Symbol	Construct	Meaning
D	Digital analytics readiness	Data integration, analytical skills, and the capacity to use information in decisions.
K	Knowledge and research capacity	Technical expertise, research resources, and the ability to access and combine knowledge.

Symbol	Construct	Meaning
F	Green finance	Availability of finance for activities defined as environmentally beneficial.
P	Environmental policy support	Supportive policy design, implementation, and coordination.
Y	Green innovation outcome	A constructed index for illustrating high green innovation; not an observed patent or environmental measure.

Note. All variables are hypothetical 0–100 indices. In fuzzy-set expressions, the same symbols denote calibrated memberships.

3.2 Synthetic case generation

The demonstration uses four binary design positions, one for each condition. Each position is assigned either a low or a high interval, yielding $2^4 = 16$ possible configurations. Six independent synthetic profiles are generated in every configuration, producing 96 cases and ensuring that no binary configuration is absent. For a low position, the raw index is drawn uniformly from 5 to 45; for a high position, it is drawn uniformly from 55 to 95. The fixed random seed is 32418. The deliberate gap around 50 prevents a case from landing exactly at the fuzzy-set crossover.

The full factorial balance is a convenience of the teaching design, not a realistic sampling claim. Actual cities are unlikely to be uniformly distributed across all 16 combinations, and some combinations may be impossible or rarely observed. In a real study, the number of cases in each truth-table row would be an empirical finding with implications for limited diversity and threshold choice. Here every row contains exactly six cases by construction, and the absence of logical remainders should not be generalized beyond this dataset.

Raw inputs are converted to fuzzy membership using the direct logistic function in Equation 1. The anchors are 15 for a membership of 0.05, 50 for a membership of 0.50, and 85 for a membership of 0.95. The same anchor values are used for all five hypothetical indices only because all are constructed on a common artificial scale. In an empirical application, each variable requires anchors justified by substantive standards, institutional knowledge, distributions, or a combination of these sources. Sample percentiles alone do not establish the real-world meaning of full or partial set membership.

$$\mu(x) = 1 / [1 + exp\{-ln(19)\,(x - 50) / 35\}] \quad (1)$$

Let D, K, F, and P denote the calibrated memberships of digital readiness, knowledge capacity, green finance, and policy support; Y denotes the calibrated outcome membership. The synthetic outcome membership is created with Equation 2. For NCA and Figure 1, the same condition labels refer to their raw index values, as specified on the axes. A disturbance ε is drawn uniformly between −0.02 and 0.02, the expression inside the brackets is clipped to the interval 0.015–0.985, and the result is capped at D. The Boolean form underlying the rule is D AND [K OR (F AND P)]. This deliberate cap ensures that the outcome membership never exceeds digital readiness. After generation, the inverse of Equation 1 converts the outcome membership back into a 0–100-style synthetic index for the NCA scatterplot.

$$Y = min\{D, clip[max(K, min(F, P)) + \varepsilon, 0.015, 0.985]\} \quad (2)$$

Equation 2 is not an estimated theory of how actual cities innovate. It embeds both a proposed necessary relation and two sufficient combinations before the analysis begins. Therefore, successful recovery of those patterns is a verification of the implementation under a constructed example, not discovery in independent data. The small disturbance changes membership values but does not remove the digital cap. The full data-generation process, including the order of random draws, is available in the accompanying executable supplement.

3.3 NCA estimation and randomization

For each condition, NCA is conducted on its raw synthetic index and the raw synthetic outcome index. We use the ceiling-envelopment free-disposal-hull (CE-FDH) line because it provides a transparent nonparametric stepwise boundary for this constructed dataset. For a given condition value x, the ceiling is the largest outcome among cases whose condition values do not exceed x. The ceiling zone lies above that step within the observed minimum and maximum of both axes. The reported effect size is the area of this zone divided by the area of the empirical scope, as shown in Equation 3 (Dul et al., 2023; Dul, 2016).

$$d = A_{ceiling} / [(x_{max} - x_{min})(y_{max} - y_{min})] \quad (3)$$

An effect size is a property of the observed scatter and the chosen scope; it is not the proportion of variation explained, a regression coefficient, or a policy return. We preserve all generated points and do not delete ceiling outliers. With actual data, analysts should inspect whether measurement error, exceptional cases, bounded variables, skewness, or a small number of influential points create or destroy the apparent empty region. A second defensible ceiling technique can be used as an additional sensitivity check, but it is not reported as if it had been executed here (Dul, 2022; Thiem, 2021).

A randomization test is performed by permuting the 96 outcome values relative to a condition 10,000 times, holding both marginal distributions fixed. The reported one-sided Monte Carlo permutation p-value is (1 + the number of permuted effect sizes at least as large as the observed size) / 10,001. The random generator uses seed 32419 and is advanced sequentially for the four conditions. Holm adjustment is applied across the four p-values as a transparent multiple-testing sensitivity. Because the data are generated from a known rule rather than sampled from a population, these p-values describe a constructed randomization exercise and do not provide population-level evidence about urban innovation.

For the digitally capped condition, a bottleneck table expresses selected outcome levels as percentages of the observed outcome range and the corresponding minimum observed condition value as a percentage of its observed range. This is an empirical step-function threshold in the artificial sample, not a target for actual cities. The sample minimum may change when cases are added or when the outcome is measured differently. We do not present bottleneck values for other conditions as prescriptions, and we do not interpret a small permutation p-value as proof of causal necessity.

3.4 Fuzzy-set calibration and necessity

All conditions and the outcome are analyzed as fuzzy sets with membership between zero and one. For the purposes of this demonstration, a score above 0.5 indicates closer membership in the high set and a score below 0.5 indicates closer membership in its complement. The complement of a set X is 1 − X. The minimum operator represents intersection and the maximum operator represents union. These operations are explicit analytical choices and do not imply that real-world interactions must follow a minimum or maximum function.

Fuzzy necessity consistency for a condition X and outcome Y is the sum of $\min(X_i,Y_i)$ divided by the sum of Y_i. Necessity coverage divides the same overlap by the sum of X_i. Equation 4 defines both quantities. A high consistency suggests that the outcome set is largely contained in the condition set; coverage helps distinguish a substantive restriction from a nearly ubiquitous condition. The analysis reports both the presence and absence of each single condition. A conventional 0.90 consistency benchmark is used as a descriptive screening rule, not as a universal proof of necessity (Thiem, 2016; Veri, 2018).

$$C_n = \Sigma_i \, min(X_i, Y_i) / \Sigma_i \, Y_i$$
$$Cov_n = \Sigma_i \, min(X_i, Y_i) / \Sigma_i \, X_i \quad (4)$$

NCA effect size and fuzzy necessity use different representations and denominators. NCA in this demonstration uses raw indices and an empirical ceiling area; fuzzy necessity uses calibrated memberships and set overlap. Their conclusions can therefore diverge even when both calculations are correct. Comparing

them requires examining the scatter, calibration, and actual cases rather than choosing the result with the smallest p-value or the largest consistency. In particular, a condition may impose a ceiling on high levels of an outcome while failing to contain the complete fuzzy outcome set at a selected consistency threshold.

3.5 Truth table, minimization, and coverage

A truth table is formed from the 16 combinations of the four conditions. Each case is assigned to the row corresponding to whether each condition membership is above or below 0.5. Row membership is the minimum of the relevant condition memberships or their complements; let R_i denote case i's membership in that row. For each row, raw sufficiency consistency is Σmin(row,Y)/Σrow. Proportional reduction in inconsistency (PRI) additionally considers overlap with the complement of the outcome; Equation 5 gives the formula used here. These fuzzy consistency calculations use the memberships of all cases, not only the cases whose closest binary row is being described.

$$PRI = [\Sigma_i \, min(R_i, Y_i) - \Sigma_i \, min(R_i, Y_i, 1 - Y_i)] / [\Sigma_i \, R_i - \Sigma_i \, min(R_i, Y_i, 1 - Y_i)] \quad (5)$$

For the high-outcome analysis, a row is coded as supporting the outcome if it has at least two assigned cases, raw consistency of at least 0.80, and PRI of at least 0.75. All rows contain six cases in this design, so every row meets the frequency requirement. Rows that fail either consistency criterion are coded as non-supporting for that analysis. The same procedure is run independently for the complement outcome, 1 − Y, to make the distinction between high and low outcome explicit. No claim of causal asymmetry follows merely from finding different expressions; the low-outcome structure must be interpreted against the construction and cases.

The program enumerates Boolean implicants that cover at least one supporting row and no observed non-supporting row, identifies prime implicants, and searches for solutions with the fewest terms and then the fewest literals. Since all 16 rows are observed and meet the frequency threshold, no logical remainders or directional assumptions are used. The resulting expressions are complex solutions in the sense that simplification uses only observed supporting rows. Under this unusually complete design there is no distinction arising from the use of remainders between complex and parsimonious solutions; real datasets with limited diversity require additional, explicitly defended decisions (Cooper & Glaesser, 2016b; Greckhamer et al., 2018; Thiem, 2022).

For each pathway, raw coverage is Σmin(path,Y)/ΣY and consistency is Σmin(path,Y)/Σpath. Unique coverage is the share of the outcome overlap not simultaneously covered by the other pathway or pathways. Solution membership is the maximum of the component pathway memberships. Because pathways can overlap, adding raw coverages does not give solution coverage. The distinction between raw, unique, and overall coverage helps avoid describing every formally sufficient term as equally important (Fiss, 2011; Veri, 2018).

3.6 Sensitivity and implementation

The main sensitivity checks increase the raw-consistency threshold from 0.80 to 0.85, raise the frequency threshold from two to three, and change the calibration anchors from 15/50/85 to 10/50/90 for all conditions and the outcome. Each check is run separately, with PRI held at 0.75. The frequency check is weak by design because all 16 rows have six cases; it cannot demonstrate robustness to sparsity. The alternative anchors test a change in the slope of the common calibration curve while preserving the crossover, and they should not be interpreted as substantively validated thresholds.

All calculations are implemented in Python 3.12.14 with NumPy 2.3.5 using a disclosed implementation of the CE-FDH area, randomization, fuzzy-set metrics, and four-condition Boolean minimization. The figure is rendered with Matplotlib 3.10.8. We do not claim that R's NCA package or a commercial fsQCA program was used. The supplement generates the synthetic cases from the seed and prints the NCA and complete

truth-table results. Its purpose is computational transparency, not validation against an independently collected dataset.

The implementation checks that there are 96 cases, each binary row has six cases, the digital cap holds, and the observed conditions avoid exact crossover scores. For the Boolean minimization, the recovered expressions are checked against all observed row codes. These checks address whether the program implements the declared exercise. They cannot establish whether the assumed constructs, invented values, or deterministic rule describe a city. Empirical applications should also perform independent software verification, case-level inspection, missing-data analysis, and substantive review of influential observations.

4. Results of the Synthetic Demonstration

4.1 Ceiling effects and fuzzy necessity

Figure 1 shows the generated relationship between digital readiness and the outcome in raw index units. The stepwise line is the CE-FDH ceiling; the shaded region is the empty area above the ceiling within the empirical scope. The shape is not an unexpected discovery: digital readiness was imposed as a cap on outcome membership, and the same monotonic calibration function was used to convert both series to raw-scale indices. Most cases lie on or below a rising boundary because of this construction. The figure should be interpreted as a visual check of the example, not as a graph of observed cities.

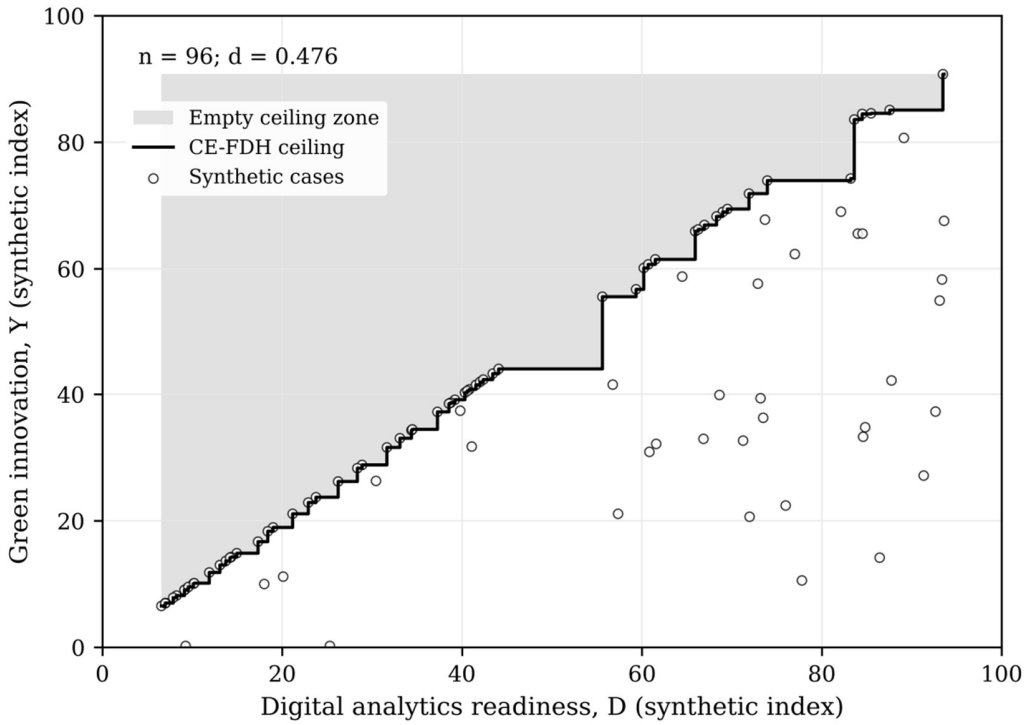


Figure 1. CE-FDH Ceiling for Digital Analytics Readiness and Green Innovation

Note. CE-FDH ceiling for digital analytics readiness and the synthetic green innovation index (n = 96). The shaded area is the empty ceiling zone within the empirical scope; d = 0.476. CE-FDH = ceiling envelopment–free disposal hull.

Table 2 summarizes NCA and fuzzy necessity for all four conditions. Digital readiness has a ceiling effect size of 0.476 and a raw permutation p-value of approximately 0.0001; the Holm-adjusted value is 0.0004. Its fuzzy necessity consistency is 1.000 and necessity coverage is 0.702. Since the generated outcome is capped at digital readiness, perfect fuzzy necessity is expected. The coverage value shows that high digital readiness

alone is not sufficient to place every case equally strongly in the high-outcome set. The absence of digital readiness has necessity consistency 0.373 and does not meet the descriptive 0.90 screen for the high outcome.

Table 2. Synthetic NCA and Fuzzy-Set Necessity Results

Condition	NCA effect d	Raw p-value	Holm p-value	Fuzzy necessity consistency	Fuzzy necessity coverage	Absent-set consistency
D: digital	0.476	0.0001	0.0004	1.000	0.702	0.373
K: knowledge	0.147	0.0085	0.0255	0.835	0.613	0.513
F: finance	0.108	0.0440	0.0730	0.708	0.507	0.558
P: policy	0.126	0.0365	0.0730	0.706	0.509	0.578

Note. NCA uses raw synthetic indices and CE-FDH with empirical scope. Fuzzy metrics use calibrated scores. Holm adjustment covers four permutation tests; n = 96 and 10,000 permutations.

Knowledge capacity has an NCA effect size of 0.147 and a raw permutation p-value of 0.0085, which becomes 0.0255 after Holm adjustment. Its fuzzy necessity consistency is 0.835, below the chosen 0.90 screening level. The two results should not be collapsed into a single yes-or-no judgment: the ceiling statistic describes the geometry of a raw-score scatter, whereas fuzzy necessity describes how much of the calibrated outcome set lies inside the knowledge set. The known rule permits high outcomes through the finance–policy combination even when knowledge is not high. Finite-sample geometry can nevertheless produce an empty upper-left region for knowledge.

Finance and policy have NCA effect sizes of 0.108 and 0.126, respectively. Their unadjusted permutation p-values are approximately 0.0440 and 0.0365, but both Holm-adjusted values are 0.0730. Their fuzzy necessity consistencies are 0.708 and 0.706. These numbers do not establish that finance or policy is universally unnecessary in real innovation systems. Within the synthetic rule, neither is required in every high-outcome pathway because knowledge can provide an alternative. The example demonstrates why reporting the full pattern and its construction is more informative than declaring a single significant factor.

The CE-FDH bottleneck calculation for digital readiness gives minimum observed condition levels of 25.03%, 56.33%, 68.22%, and 88.53% of its observed range when target outcome levels are set at 30%, 50%, 70%, and 90% of the observed outcome range. Both percentages refer to their own empirical ranges, not their raw values or population percentiles. These steps are produced by the particular simulated cases and the ceiling rule. They should not be interpreted as a finding that a real city must achieve any of these numeric targets.

4.2 High-outcome truth table and pathways

Table 3 presents the full high-outcome truth table. The condition order is D, K, F, P; 1 denotes membership above the crossover and 0 denotes membership below it. All 16 rows contain six assigned synthetic cases. Five rows satisfy both the 0.80 raw-consistency and 0.75 PRI thresholds: 1011, 1100, 1101, 1110, and 1111. Every row containing D and K is included, as is the row containing D, F, and P without high K. No logical remainders are used in the minimization.

Table 3. Complete High-Outcome Truth Table

D K F P	Cases	Raw consistency	PRI	Supporting row
0 0 0 0	6	0.535	0.000	No
0 0 0 1	6	0.539	0.000	No
0 0 1 0	6	0.566	0.000	No
0 0 1 1	6	0.599	0.000	No
0 1 0 0	6	0.606	0.000	No
0 1 0 1	6	0.632	0.000	No
0 1 1 0	6	0.647	0.000	No
0 1 1 1	6	0.640	0.000	No
1 0 0 0	6	0.669	0.003	No
1 0 0 1	6	0.724	0.008	No
1 0 1 0	6	0.720	0.016	No
1 0 1 1	6	0.993	0.978	Yes
1 1 0 0	6	0.994	0.976	Yes
1 1 0 1	6	0.994	0.979	Yes
1 1 1 0	6	0.995	0.983	Yes
1 1 1 1	6	0.998	0.993	Yes

Note. D, K, F, P are in that order; 1 denotes membership above 0.5. A row supports the high outcome when $n \geq 2$, consistency ≥ 0.80 and PRI ≥ 0.75.

The minimized high-outcome solution is D·K + D·F·P. The first term combines high digital readiness with high knowledge capacity; the second combines high digital readiness with both high green finance and high policy support. The terms are alternative, overlapping descriptions of the constructed high-outcome set. A case can belong to both terms, and no inference is made about which term a real city should pursue. Table 4 reports the fuzzy consistency and coverage of each term and of the overall solution.

Table 4. Synthetic High- and Low-Outcome Solutions

Outcome	Path or solution	Raw coverage	Unique coverage	Consistency
High Y	D · K	0.835	0.462	0.995
High Y	D · F · P	0.532	0.159	0.994
High Y	Overall solution	0.993	—	0.993
Low Y	~D	0.768	0.363	1.000
Low Y	~K · ~F	0.469	0.063	0.997
Low Y	~K · ~P	0.474	0.071	0.995
Low Y	Overall solution	0.996	—	0.996

Note. A centered dot denotes intersection; ~ denotes set complement. Coverage and consistency are rounded. Overall solution coverage uses the union of the terms.

For D·K, raw coverage is 0.835, unique coverage is 0.462, and consistency is 0.995. For D·F·P, raw coverage is 0.532, unique coverage is 0.159, and consistency is 0.994. Overall solution consistency and coverage are approximately 0.993 and 0.993. The sum of raw pathway coverages exceeds overall coverage because a portion of the fuzzy outcome set belongs to both terms. The very high fit values are expected: the outcome was generated from essentially the same expression with a small disturbance. They are not evidence of predictive performance on external cases.

The sample's five positive binary rows also show why the role of a condition depends on the configuration. K appears in the first pathway but is absent from the second term; F and P appear together in the second but are not required in the first. This does not prove that knowledge and finance are economically substitutable, or that policy always complements finance. It shows how a declared set of interactions can be summarized in a truth table and how alternative combinations can be represented without fitting an additive average-effect model.

4.3 Low-outcome analysis and sensitivity

Analyzing the complement outcome separately yields ~D + ~K·~F + ~K·~P, where the tilde denotes the complement of a condition. The solution consistency is 0.996 and coverage is 0.996. The three terms have raw coverages of approximately 0.768, 0.469, and 0.474, respectively; their unique coverages are approximately 0.363, 0.063, and 0.071. These components overlap, so their raw coverages again must not be summed as if they represented mutually exclusive groups of cities.

In this particular exercise the low-outcome expression is the Boolean complement of the embedded high-outcome rule, apart from the effects of fuzzy membership and the small disturbance. It is therefore inappropriate to claim that the result independently proves causal asymmetry. The exercise does, however, show how a high-outcome expression and its complement can have different conjunctions and disjunctions. With real observations, separate analysis may reveal contradictions or patterns that cannot be deduced simply by reversing the sign of an average association; such claims would require case evidence and a carefully specified causal design.

The high-outcome solution remains D·K + D·F·P when the raw-consistency threshold is increased to 0.85. It also remains unchanged when the row frequency threshold increases from two to three, but this comparison is mechanically weak because each row has six cases. Under the alternative 10/50/90 calibration anchors, the same five rows and two terms are recovered; solution consistency and coverage remain approximately 0.993, although the individual raw and unique coverage values change. These checks describe stability under a small set of preselected choices, not general robustness to misspecification, noisy data, or changes in the data-generating rule.

5. Discussion

5.1 What combining the two methods reveals

The demonstration makes the relationship between prerequisite and configuration visible. Digital readiness is a common component of both recovered high-outcome terms and is also a strong ceiling condition in the synthetic NCA. That agreement is expected because the data-generating process imposes a digital cap. In a real study, an analogous pattern could motivate investigation of whether a minimum data capability is genuinely required for the particular innovation outcome. It would not, by itself, establish that investing in a digital system causes green innovation or that a single threshold applies across cities.

The results for knowledge provide a more cautionary example. A statistically unusual NCA ceiling does not automatically mean that high knowledge is a universal fuzzy-set prerequisite, and a fuzzy consistency below 0.90 does not mean that the scatter has no useful boundary. The methods summarize different aspects of the data. A responsible interpretation would examine whether the apparent ceiling is driven by a few influential

cases, whether the outcome and condition were measured on meaningful scales, and whether alternative pathways are supported by detailed case knowledge (Dul et al., 2023; Dul, 2022; Thiem, 2021).

The simulated finance and policy findings also demonstrate the importance of joint exposure. Their combination appears in one high-outcome term, while neither condition is strongly necessary on its own according to the fuzzy screen. In an actual urban context, this could prompt inquiry into whether finance is usable only when institutions provide credible standards, project selection, or demand. It could equally reveal a spurious pattern caused by shared regional characteristics or measurement choices. Set membership is a descriptive and theoretical representation, not a substitute for an identification strategy.

5.2 Designing a study with real cities

A real application should begin with a defined population, period, and decision objective. Researchers might select a documented set of cities with comparable administrative boundaries and analyze an outcome measured after the conditions. The choice of country, time span, exclusions, and treatment of municipal boundary changes should be stated before examining preferred solutions. Multiple observations from the same city are not automatically independent cases; a panel-QCA design needs to explain how it handles time and dependence (Liu et al., 2025; Verweij & Vis, 2021).

The outcome could be defined as environmentally relevant patent applications per unit of population or per number of firms over a stated period, using a documented technology-classification system and allocation rule. Patent applications, grants, citations, commercial adoption, and environmental improvements are different outcomes. A study should not label an outcome "efficiency" unless it explicitly defines inputs and outputs and chooses a defensible efficiency method. If the objective is emissions reduction rather than invention, the outcome and causal interpretation must change accordingly (Guo et al., 2023; Liu & Dong, 2021; Shang et al., 2022).

Digital readiness could combine data infrastructure, availability of skilled staff, organizational data governance, and practical analytical use. Knowledge capacity could be represented by research personnel, research expenditure, research institutions, or documented innovation collaboration. Green finance could use lending or investment measures with careful definitions of what qualifies as green. Policy support could distinguish regulatory enforcement, subsidies, procurement, and implementation capability. Different indicators may be correlated, but combining them indiscriminately could create an index that has no clear interpretation or double-counts the same activity (Bai & Zhang, 2025; Hu et al., 2022; Johnston & Huggins, 2016; Zhang, Wang, et al., 2024).

Measurement design should also avoid embedding the outcome in a condition. If digital readiness includes the number of green patent applications and the outcome is green patent applications, the apparent necessity may arise mechanically. Indicators recorded after the outcome can similarly introduce post-outcome bias. Researchers should document denominators, prices and inflation adjustments, missing values, extreme observations, and changes in administrative classifications. Where source data are confidential, they can provide reproducible transformation rules and a synthetic example while arranging appropriate access to the underlying evidence for independent review.

5.3 Calibration and case selection

Calibration determines what it means to be inside or outside a set. A value of 0.95 should represent a substantively defensible level of strong membership, 0.50 the qualitative crossover, and 0.05 strong nonmembership. When no authoritative standard exists, researchers can use multiple sources of evidence, examine the consequences of alternative anchors, and report why each choice fits the theory (Greckhamer et al., 2018; Pappas & Woodside, 2021). A percentile is a convenient description of a sample distribution but does not by itself imply that a city has an adequate research base, policy regime, or financing system.

The simulated example uses identical anchors and balanced intervals to reveal method behavior. Actual data will have different units and distributions, and some values may be zero, highly skewed, or unavailable. A single monotonic calibration function may not suit a condition whose relationship with the underlying concept is non-monotonic. Analysts should inspect the calibrated distribution, the number of cases near 0.5, and the consequences of alternative thresholds. Cases exactly at 0.5 require a documented procedure; silently changing values to force binary assignment can affect the truth table.

Case selection affects both NCA ceilings and fsQCA solutions. An apparent empty zone can disappear when a previously excluded high-performing case with a low condition value is included. A pathway may be supported by only one or two cases, and a frequency threshold can change whether that row is analyzed or treated as a remainder. Researchers should report all possible combinations, case counts, and the coding of contradictory rows. They should also distinguish results that use only observed rows from those that incorporate simplifying assumptions about unobserved configurations (Cooper & Glaesser, 2016b; Thiem, 2014; Thiem, 2022).

5.4 From patterns to policy decisions

For business and public decision makers, a useful first product is a diagnosis of the city's current profile, the uncertainty around its indicators, and the cases most comparable to it. A second step is to examine whether a proposed intervention addresses a credible mechanism. If finance and policy appear together in a pathway, the team should investigate access to finance, project eligibility, local institutions, and whether a policy is implemented effectively. A set-theoretic expression does not specify the amount to invest, the appropriate instrument, or the direction of causality (Han & Cai, 2024; Jing & Liu, 2024; Yin et al., 2023).

Evaluation can then proceed through a staged design. A city might pilot a data-sharing mechanism, a guarantee product, or a procurement program and specify outcomes before the pilot begins. Comparisons with an appropriate baseline, attention to spillovers, and a plan for unintended effects can strengthen interpretation. A configurational analysis can help select cases or formulate hypotheses for these designs, while a causal evaluation can test whether a particular intervention changes outcomes. Neither method should be used to claim effects that it was not designed to identify.

Policy priorities also involve distribution. A program that raises total green patents may benefit established firms while excluding smaller enterprises, or it may transfer pollution across administrative boundaries. Policymakers should consider environmental outcomes, cost, equity, and implementation feasibility alongside the innovation indicator. The digital condition should include data protection, governance, and accessibility rather than only computational infrastructure. Any recommended action requires local institutional and technical knowledge beyond the synthetic exercise presented here.

5.5 Limitations and future research

The principal limitation is that the data are synthetic and the demonstrated relationships are imposed. The analysis cannot estimate the prevalence of any pathway among actual cities, compare regions, establish causal effects, or validate a policy program. The unusually high consistency and coverage values largely reflect the construction of the outcome. A research team seeking empirical publication should collect and validate an independent dataset and report its source, measurement, case selection, and uncertainty. The present article is a methodological demonstration, not a completed empirical study of urban green innovation.

Second, the four conditions simplify an ecosystem that may also depend on industrial structure, energy prices, market demand, intercity spillovers, and institutional capacity. Adding more conditions is not costless: the number of possible configurations increases exponentially, potentially creating many unobserved rows. Researchers should select conditions based on theory and the number and diversity of cases, rather than adding variables solely because data are available. They should also assess alternative outcome definitions and explore whether a small number of cases drive the reported paths.

Third, a nonparametric ceiling is sensitive to measurement boundaries and extreme observations. The permutation test preserves marginal distributions but its interpretation depends on the null and the random reassignment mechanism. Holm adjustment addresses multiplicity for the four tests but does not cure bias, confounding, or poor construct validity. Future work can compare credible ceiling specifications, investigate influential observations, and assess how alternative sample definitions affect necessity claims (Dul et al., 2020; Dul et al., 2023; Dul, 2022).

Fourth, the demonstration uses a single time period and no spatial structure. Real green innovation may accumulate over years, and finance, policy, and patents can influence one another in both directions. A temporal design could define condition windows before outcome windows and examine how pathways evolve. Spatially informed comparisons could examine whether intercity collaboration and spillovers challenge a purely local interpretation (Lin et al., 2021; Shang et al., 2022; Verweij & Vis, 2021; Zeng et al., 2022). These extensions require data and methods beyond the present demonstration.

Finally, independent verification of the custom implementation against established NCA and QCA software would strengthen a future empirical application. The provided code exposes each computation and recovers the rule embedded in the synthetic data, but that does not establish equivalence for all possible ties, missing observations, logical remainders, or ambiguous truth tables. The code deliberately stops when a truth-table row has too few cases, rather than silently applying a simplifying assumption. Any empirical use should include independent software checks and case-based validation.

6. Conclusion

This article presents a reproducible workflow for considering potential necessary conditions and alternative multi-condition pathways in urban green innovation. It distinguishes the geometric ceiling logic of NCA from fuzzy-set necessity and sufficiency, specifies calibration and minimization choices, and demonstrates how the results can be compared without treating them as interchangeable.

In the 96-case synthetic example, the methods recover a digitally capped outcome and two pathways that were specified in advance. This is evidence that the example's calculations reproduce its construction, not evidence that real cities require the same conditions. A substantive assessment of urban green innovation requires independently collected data, transparent measurement, case investigation, and validation of any policy or causal interpretation.

Acknowledgement

We acknowledge the researchers whose published work informed this study. Any remaining errors are our responsibility.

Author Contributions

Conceptualization, S.K.W. and B.Z.; Methodology, S.K.W. and B.Z.; Formal Analysis, S.K.W. and B.Z.; Investigation, S.K.W.; Data Curation, S.K.W.; Software, S.K.W.; Validation, S.K.W. and B.Z.; Resources, S.K.W. and B.Z.; Visualization, S.K.W.; Writing – Original Draft, S.K.W.; Writing – Review & Editing, S.K.W. and B.Z.; Supervision, B.Z.; Project Administration, S.K.W. and B.Z. We have read and approved the final manuscript and take responsibility for the work.

Funding

This research received no external funding.

Conflicts of Interest

We declare no conflicts of interest.

Ethical Approval and Consent

The analysis uses synthetic cases and does not involve human participants, personal data, or animals.

Data Availability Statement

The synthetic case generator, parameters, calibration, NCA permutation test, truth tables, and minimization are provided in the accompanying Supplement S1. Figure 1 is also provided as a separate image. No empirical or proprietary city dataset was used.

Declaration of Generative AI Use

We used OpenAI ChatGPT (GPT-6) to assist with literature discovery, drafting, analytical implementation, and document preparation. We take full responsibility for the accuracy of the references, calculations, and final manuscript.

References

Ambec, S., Cohen, M. A., Elgie, S., & Lanoie, P. (2013). The Porter hypothesis at 20: Can environmental regulation enhance innovation and competitiveness?. Review of Environmental Economics and Policy, 7(1), 2–22. https://doi.org/10.1093/reep/res016

Bai, X., & Zhang, Z. (2025). Does the digital innovation ecosystem promote green technology innovation? Evidence from 264 cities in China. Technology in Society, 83, 102977. https://doi.org/10.1016/j.techsoc.2025.102977

Bitat, A. (2018). Environmental regulation and eco-innovation: The Porter hypothesis refined. Eurasian Business Review, 8(3), 299–321. https://doi.org/10.1007/s40821-017-0084-6

Cooper, B., & Glaesser, J. (2016a). Analysing necessity and sufficiency with qualitative comparative analysis: How do results vary as case weights change?. Quality & Quantity, 50(1), 327–346. https://doi.org/10.1007/s11135-014-0151-3

Cooper, B., & Glaesser, J. (2016b). Qualitative comparative analysis, necessary conditions, and limited diversity. Field Methods, 28(3), 300–315. https://doi.org/10.1177/1525822x15598974

Dian, J., Song, T., & Li, S. (2024). Facilitating or inhibiting? Spatial effects of the digital economy affecting urban green technology innovation. Energy Economics, 129, 107223. https://doi.org/10.1016/j.eneco.2023.107223

Dul, J. (2016). Necessary condition analysis (NCA): Logic and methodology of "necessary but not sufficient" causality. Organizational Research Methods, 19(1), 10–52. https://doi.org/10.1177/1094428115584005

Dul, J. (2022). Problematic applications of necessary condition analysis (NCA) in tourism and hospitality research. Tourism Management, 93, 104616. https://doi.org/10.1016/j.tourman.2022.104616

Dul, J. (2024). A different causal perspective with necessary condition analysis. Journal of Business Research, 177, 114618. https://doi.org/10.1016/j.jbusres.2024.114618

Dul, J., Hauff, S., & Bouncken, R. B. (2023). Necessary condition analysis (NCA): Review of research topics and guidelines for good practice. Review of Managerial Science, 17(2), 683–714. https://doi.org/10.1007/s11846-023-00628-x

Dul, J., van der Laan, E., & Kuik, R. (2020). A statistical significance test for necessary condition analysis. Organizational Research Methods, 23(2), 385–395. https://doi.org/10.1177/1094428118795272

Fiss, P. C. (2011). Building better causal theories: A fuzzy set approach to typologies in organization research. Academy of Management Journal, 54(2), 393–420. https://doi.org/10.5465/amj.2011.60263120

Fritsch, M., & Franke, G. (2004). Innovation, regional knowledge spillovers and R&D cooperation. Research Policy, 33(2), 245–255. https://doi.org/10.1016/s0048-7333(03)00123-9

Gao, J., Hua, G., Randhawa, A., Huo, B., & Li, Z. (2024). Towards a green digital city: Digital economy, digital innovation, and urban haze pollution. Journal of Urban Technology, 31(4-5), 157–189. https://doi.org/10.1080/10630732.2024.2419349

Greckhamer, T., Furnari, S., Fiss, P. C., & Aguilera, R. V. (2018). Studying configurations with qualitative comparative analysis: Best practices in strategy and organization research. Strategic Organization, 16(4), 482–495. https://doi.org/10.1177/1476127018786487

Guo, J., Fu, Y., & Sun, X. (2023). Green innovation efficiency and multiple paths of urban sustainable development in China: Multi-configuration analysis based on urban innovation ecosystem. Scientific Reports, 13(1), 12975. https://doi.org/10.1038/s41598-023-40084-x

Han, X., & Cai, Q. (2024). Environmental regulation, green credit, and corporate environmental investment. Innovation and Green Development, 3(3), 100135. https://doi.org/10.1016/j.igd.2024.100135

Hu, Y., Dai, X., & Zhao, L. (2022). Digital finance, environmental regulation, and green technology innovation: An empirical study of 278 cities in China. Sustainability, 14(14), 8652. https://doi.org/10.3390/su14148652

Hu, Y., Wang, C., Zhang, X., Wan, H., & Jiang, D. (2023). Financial agglomeration and regional green innovation efficiency from the perspective of spatial spillover. Journal of Innovation & Knowledge, 8(4), 100434. https://doi.org/10.1016/j.jik.2023.100434

Jing, R., & Liu, R. (2024). The impact of green finance on persistence of green innovation at firm-level: A moderating perspective based on environmental regulation intensity. Finance Research Letters, 62, 105274. https://doi.org/10.1016/j.frl.2024.105274

Johnston, A., & Huggins, R. (2016). The spatio-relational nature of urban innovation systems: Universities, knowledge intensive business service firms, and collaborative networks. Journal of Urban Technology, 23(1), 29–52. https://doi.org/10.1080/10630732.2015.1090192

Li, F., Zhang, H., Zhang, D., & Yan, H. (2023). Structural diffusion model and urban green innovation efficiency—a hybrid study based on DEA-SBM, NCA, and fsQCA. Sustainability, 15(17), 12705. https://doi.org/10.3390/su151712705

Lin, S., Chen, Z., & He, Z. (2021). Intra-city industrial collaborative agglomeration, inter-city network connectivity and green technology innovation. Sustainability, 13(16), 8835. https://doi.org/10.3390/su13168835

Liu, Q., Zhang, Y., & Sun, X. (2025). Digital innovation ecosystems and regional green technological innovation: Evidence from China's panel-QCA analysis. Journal of Innovation & Knowledge, 10(5), 100789. https://doi.org/10.1016/j.jik.2025.100789

Liu, Y., & Dong, F. (2021). How technological innovation impacts urban green economy efficiency in emerging economies: A case study of 278 chinese cities. Resources, Conservation and Recycling, 169, 105534. https://doi.org/10.1016/j.resconrec.2021.105534

Pappas, I. O., & Woodside, A. G. (2021). Fuzzy-set qualitative comparative analysis (fsQCA): Guidelines for research practice in information systems and marketing. International Journal of Information Management, 58, 102310. https://doi.org/10.1016/j.ijinfomgt.2021.102310

Qiu, L. D., Zhou, M., & Wei, X. (2018). Regulation, innovation, and firm selection: The Porter hypothesis under monopolistic competition. Journal of Environmental Economics and Management, 92, 638–658. https://doi.org/10.1016/j.jeem.2017.08.012

Shang, H., Jiang, L., Pan, X., & Pan, X. (2022). Green technology innovation spillover effect and urban eco-efficiency convergence: Evidence from chinese cities. Energy Economics, 114, 106307. https://doi.org/10.1016/j.eneco.2022.106307

Thiem, A. (2014). Navigating the complexities of qualitative comparative analysis. Evaluation Review, 38(6), 487–513. https://doi.org/10.1177/0193841x14550863

Thiem, A. (2016). Standards of good practice and the methodology of necessary conditions in qualitative comparative analysis. Political Analysis, 24(4), 478–484. https://doi.org/10.1093/pan/mpw024

Thiem, A. (2021). The logic and methodology of “necessary but not sufficient causality”: A comment on necessary condition analysis (NCA). Sociological Methods & Research, 50(2), 913–925. https://doi.org/10.1177/0049124118782548

Thiem, A. (2022). Beyond the facts: Limited empirical diversity and causal inference in qualitative comparative analysis. Sociological Methods & Research, 51(2), 527–540. https://doi.org/10.1177/0049124119882463

Veri, F. (2018). Coverage in fuzzy set qualitative comparative analysis (fsQCA): A new fuzzy proposition for describing empirical relevance. Comparative Sociology, 17(2), 133–158. https://doi.org/10.1163/15691330-12341457

Verweij, S., & Vis, B. (2021). Three strategies to track configurations over time with qualitative comparative analysis. European Political Science Review, 13(1), 95–111. https://doi.org/10.1017/s1755773920000375

Wang, X., Zhang, T., Nathwani, J., Yang, F., & Shao, Q. (2022). Environmental regulation, technology innovation, and low carbon development: Revisiting the ekc hypothesis, Porter hypothesis, and jevons’ paradox in China's iron & steel industry. Technological Forecasting and Social Change, 176, 121471. https://doi.org/10.1016/j.techfore.2022.121471

Wang, X., Sun, X., Zhang, H., & Xue, C. (2022). Digital economy development and urban green innovation capability: Based on panel data of 274 prefecture-level cities in China. Sustainability, 14(5), 2921. https://doi.org/10.3390/su14052921

Xie, X., & Wang, H. (2020). How can open innovation ecosystem modes push product innovation forward? An fsQCA analysis. Journal of Business Research, 108, 29–41. https://doi.org/10.1016/j.jbusres.2019.10.011

Yin, X., Chen, D., & Ji, J. (2023). How does environmental regulation influence green technological innovation? Moderating effect of green finance. Journal of Environmental Management, 342, 118112. https://doi.org/10.1016/j.jenvman.2023.118112

Zeng, J., Pagàn-Castaño, E., & Ribeiro-Navarrete, S. (2022). Merits of intercity innovation cooperation of environment-friendly patents for environmental regulation efficiency. Technological Forecasting and Social Change, 180, 121404. https://doi.org/10.1016/j.techfore.2021.121404

Zhang, G., Wang, Y., Zhang, Z., & Su, B. (2024). Can green finance policy promote green innovation in cities? Evidence from pilot zones for green finance reform and innovation in China. Journal of Environmental Management, 370, 122816. https://doi.org/10.1016/j.jenvman.2024.122816

Zhang, J., Kang, L., Li, H., Ballesteros-Pérez, P., Skitmore, M., & Zuo, J. (2020). The impact of environmental regulations on urban green innovation efficiency: The case of xi'an. Sustainable Cities and Society, 57, 102123. https://doi.org/10.1016/j.scs.2020.102123

Zhang, W., Zhu, B., Li, Y., & Yan, D. (2024). Revisiting the Porter hypothesis: A multi-country meta-analysis of the relationship between environmental regulation and green innovation. Humanities and Social Sciences Communications, 11(1), 232. https://doi.org/10.1057/s41599-024-02671-9