Heliyon 10 (2024) e31457

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Beyond regression: Unpacking research of human complex systems with qualitative comparative analysis

Xinxin Xu [a], Safdar Husain Tahir [b,*], Khuda Bakhsh Khan [c], Mushtaq A. Sajid [d], Muhammad Azhaf Safdar [e]

[a] *Business School, Chengdu University, Chengdu-610106, People's Republic of China*
[b] *Lyallpur Business School, Government College University, Faisalabad-38000, Pakistan*
[c] *Department of Education, Government College University, Faisalabad-38000, Pakistan*
[d] *Mohi-ud-Din Islamic University, Nerian Sharif, AJK-11030, Pakistan*
[e] *Faculty of Pharmaceutical Sciences, Government College University, Faisalabad-38000, Pakistan*

A R T I C L E I N F O

Keywords:
Conventional regression analysis
csQCA
Democracy survival
Cross-national data
Comparative study
fsQCA
Human complex systems

A B S T R A C T

This study underscores the effectiveness of Qualitative Comparative Analysis (QCA) when compared to conventional regression analysis (CRA) in the investigation of complex human systems. Utilizing historical secondary cross-national data from Lipset & Man (1960) spanning 18 countries, where CRA may be impractical, the research emphasizes the superior performance of QCA, specifically utilizing both crisp set QCA and fuzzy set QCA. The dataset includes variables such as democracy survival and its precursors, such as gross national product per capita, urbanization, literacy rate, and industrial labor force. In contrast to conventional regression results indicating an insignificant relationship between democracy survival and its antecedents, crisp set QCA identifies two distinct combinations of antecedents associated with high levels of democracy survival, albeit with limited solution coverage. Meanwhile, fuzzy set qualitative comparative analysis (fsQCA) reveals five combinations of antecedents linked to robust democracy survival, providing adequate solution coverage and consistency. These findings suggest that fsQCA excels in capturing the intricacies of real-life scenarios in human complex systems, offering more robust empirical solutions compared to crisp set QCA and conventional regression. As a result, researchers may find value in integrating fsQCA into their new projects focused on human complex systems.

1. Introduction

I was acquainted with Qualitative Comparative Analysis (QCA) when my postdoc supervisor introduced this methodology in 2018. We found fsQCA as an analytical technique better than other conventional regression-based methods, which suits our top concerning the sustainability of family businesses in the context of the digital economy of human complex systems. Initially, we had reservations about implementing QCA across 80 cases, distributed evenly among four sectors: agriculture, manufacturing, services, and banking. Nonetheless, we dedicated a substantial amount of time and effort to gain a comprehensive understanding of QCA, and we found that it effectively bridged the gap between the realms of qualitative and quantitative research [1] of human complex systems. Qualitative

* Corresponding author.
E-mail address: drsafdargcuf@gmail.com (S.H. Tahir).

https://doi.org/10.1016/j.heliyon.2024.e31457
Received 10 December 2023; Received in revised form 14 May 2024; Accepted 16 May 2024
Available online 16 May 2024
2405-8440/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

Comparative Analysis (QCA) is not only commended for enabling a hermeneutic exchange between theoretical constructs and practical empirical instances [2] but also acknowledged as a methodological progression in the field of research. In this research, we illustrate how QCA can be employed to draw conclusions from a dataset characterized by a limited number of observations (comprising 18 cases), which, in most cases, do not provide an adequate basis for regression analysis. We exemplify this with the Lipset 1960 dataset, highlighting how QCA succeeds where conventional regression analysis (CRA) falls short to draw conclusions of human complex systems.

In 1987, Charles Regin, a sociology professor at the University of California, developed a new research technique, QCA, which is an asymmetric approach [3,4], unlike conventional regression analysis. QCA employs set theory, system theory, fuzzy logic, Boolean algebra, and Quine–McCluskey algorithm to explain the connection between causal antecedents and outcomes [5]. QCA differs from conventional regression analysis (CRA) in that it adheres to the principle of equifinality, which means that multiple combinations of configurational factors can result in the same outcome through distinct pathways [6,7]. Some of configurational factors or input variables may be critical, meaning they are necessary/sufficient conditions that explain the presence/absence of the output. In this paper, we demonstrate how QCA represents a significant advancement as methodology in the context of human complex system and theory of modernization. This advancement in the form of QCA focuses more on research method rather than research approach [8–11] in real life scenarios of complex system.

QCA is a research method that involves three phases, known as the relational move, analytical move, and membership move [8]. Researchers find links between variables or conditions in the relational move, but they do not yet take into account in particular field, company or industry [12]. These relationships are usually represented by researchers using Boolean algebra, where symbols like "+" (AND), "*" (OR), and "~" (NOT) explain how various conditions combine to produce a specific outcome. As an illustration, let's pretend you are researching the elements that contribute to restaurant sector success. Three requirements stand out to you: high quality food, convenient location, and High internet reviews. You might express the relationships as follows using Boolean algebra:

High quality food * Convenient location = Success——————S_1

High Internet Reviews = Success—————————————————————————————— S_2

During the analytical move, scientists investigate different combinations of circumstances that result in the desired outcome. This entails methodically going over various combinations of criteria to determine which are required and/or sufficient for the result. Truth tables are frequently used by researchers to record these various setups and how they affect the result. For instance, keeping with the investigation of the restaurant business, you make a truth table to look at several setups. You may discover that high quality food alone (without convenient location) is not sufficient but necessary for success, and that convenient locations alone (without high quality food) are necessary but not sufficient. However, success occurs when high quality food is paired with a convenient location. These combinations are recorded in the truth table.

Researchers classify examples (such as companies, organisations, or people) into the designated configurations in the membership move according to the circumstances surrounding them. This stage facilitates the categorization of cases as either fitting or not fitting the configurations that were found to produce the desired result. Similarly, to the example discussed above, you classify several restaurants according to their conditions in your study of the restaurant industry. You discover that Restaurant-A meets the requirements for success since it offers both high quality food and has convenient location. Despite having good internet ratings, Restaurant-B doesn't meet any criteria for success because it doesn't provide high quality food or have a convenient location. We categorise each eatery as one of the configurations found in the analytical move. In short, we use all three stages in Qualitative Comparative Analysis (QCA) to examine the linkage between antecedents and outcomes. It provides us a deep understanding regarding configurations of circumstances, which generate the results. This methodology enables us to understand the real-world phenomenon of the human complex system that conventional methods might not suitable for examining configuration.

Research scholars are anxious to explore human complex systems under the context of complexity theory through diversified sources of data qualitative and quantitative types. Fuzzy set qualitative comparative analysis (fsQCA) can handle qualitative and quantitative data and analyses it in three steps through relational, analytical, and membership moves. In the relational move, research scholars talk with participants in a close relationship during the time of collecting qualitative data and try to get deep insight into their problems. Scholars might be able to become their close friends where participants can share all information easily without any fear.

If the data is quantitative, researchers may study the broader context of the data in relational move, such as economic, political, and social science of people [13,14], which are helpful in next moves. This rapport enables them to assign membership values in the subsequent phase (analytical move), which enhances their understanding of data in QCA. These, in turn, help them to draw valid and more generalizable findings from their projects of human complex system. Once they have completed the relational move, researchers move to the analytical move, where they conceptualize the cases under observation, conditions of antecedents, and outcomes in terms of set and subset relationships. The membership move involves assigning threshold values to membership values, a process referred to as calibration. The threshold values may indicate membership fully in, fully out, or neither in nor out [8] from complex system. QCA allows researchers to collect and incorporate sufficient knowledge and data information during analysis, representing a significant advancement in research analysis, typically human complex system and social science research.

The foundation of forming QCA (Qualitative Comparative Analysis) incorporates elements from system theory, fuzzy logic, set theory, Boolean Algebra, and Quine–McCluskey algorithm [15–19]. These components are employed to examine the research phenomena under investigation of human complex system. System theory emphasizes equifinality, which states that different configurations of a system or various paths can lead to the same outcome [20–22]. This differs from regression-based methods that suggest a linear relationship between input and output [23]. Even though regression analysis accounts for non-linear relationships, it implies that the variables extend well beyond that relationship while relying on four assumptions, which include the normal distribution [24, 25]. Fuzzy logic calculates the degree of truth or the conditions of variables, rather than their usual binary values of 1 and 0 [26]. Set theory is used to determine the intersection, union, and complement of sets based on their elementary properties [27,28]. Boolean algebra, on the other hand, deals with binary variables and their logical operations. Finally, the Quine-McCluskey algorithm considers all possible combinations without any order [29]. The components of these operations are employed in the development of QCA [30–33] of human complex system and social sciences.

Charles C. Ragin introduced QCA in 1987 as a novel research approach and methodology in social and politics science research [4], of human complex system in particular.

Initially developed as crisp set QCA, it was later extended to include fuzzy set QCA and multi-value QCA [34]. QCA has effectively bridged the gap between qualitative and quantitative methodologies by highlighting its capacity to incorporate both types of data and strike a balance between contextual understanding and generalization. It enabled researchers to operationalize theoretical concepts by establishing a dialogue between ideas and evidence [35,36].

With QCA, researchers can revisit populations (data) at any stage of the research process and compare positive and negative cases, leading to the evolution of populations and causal arguments. Additionally, multi-value QCA has added further value to this methodology, as it captures specific causal contributions of a multi-value condition for every category, giving it a significant advantage over crisp set QCA and fuzzy set QCA [37,38].

The QCA method rests on two fundamental assumptions. In contrast to regression, which allows an individually single factor to influence the outcome, the first QCA assumption posits that in social science research of human complex system, a single factor is typically inadequate to bring about a change or outcome. Instead, achieving the desired result often necessitates various combinations of factors. The second assumption recognizes that multiple combinations of input variables can produce the same outcome, thereby acknowledging the existence of several pathways or conditions leading to a desired result [39].

In this study, we begin by discussing the conceptual differences between conventional regression analysis and QCA, emphasizing the advantages of using QCA over regression. We then present our hypotheses and model specifications before detailing our methods for testing these hypotheses using regression, in this case logistic regression, crisp set qualitative comparative analysis (csQCA), and fuzzy set qualitative comparative analysis (fsQCA). Our findings are discussed, along with the theoretical, methodological, and managerial implications in human complex system of the study. We also discuss the limitations of fsQCA.

1.1. Conceptual differences between CRA and QCA

Table 1 outlines the key conceptual distinctions[1] between conventional regression analysis (CRA) and QCA. CRA focuses on variables with symmetrical and linear relationships, while QCA can analyze relationships that are symmetrical or asymmetric, linear or non-linear [40]. Conventional regression analysis (CRA) is limited to causal relationships in which a higher level of an outcome is the exact opposite of the one explaining the lower level, while QCA allows for more complex relationships [41]. CRA does not allow for case-based modelling human complex system, whereas QCA allows for localized effects and case-based analysis [40]. CRA provides the single best solution based on net-effect of variables of human complex system, while QCA offers multiple solutions, including complex, parsimonious, and intermediate solutions [42].

CRA uses R-square to validate results, while QCA uses coverage based on a truth table. In CRA, the p-value determines significance, while in QCA, consistency explains the strength of the relationship [40]. The way variables operate is different in CRA and QCA, as CRA focuses on individual contributions to an outcome, while QCA analyses the collective contribution of variables. In CRA, research scholars develop hypotheses showing the association between two or more variables, while in QCA, they develop propositions that analyze the combined role of variables in explaining an outcome, including the identification of contrarian cases [43]. While counter-hypothesized relations are difficult to identify in CRA, QCA can explain them easily in variables [42] of human complex system.

In addition, the predictive validity serves to assess a model's ability to anticipate future outcomes. QCA showcases superior predictive validity compared to regression analysis, particularly in the context of small-N datasets within complex human systems, such as the one examined with 18 cases. Regression may not be suitable in such cases due to issues like overfitting, such as multicollinearity. In contrast, QCA employs a sophisticated analytical framework, potentially offering precise predictions of future outcomes by considering the interplay of variable conditions rather than solely focusing on individual variables. Consequently, QCA excels in providing predictive validity, especially in small-N scenarios, by identifying patterns of variable conditions [44]. Likewise, through the integration of quantitative and qualitative methodologies, QCA offers a more profound comprehension of complex systems compared to regression, thereby augmenting the predictive validity of the analysis. Moreover, QCA examines various combinations of variable conditions within specific complex human systems, fostering a deeper understanding that further enhances predictive validity.

Nevertheless, both conventional regression analysis (CRA) and Qualitative Comparative Analysis (QCA) ascertain a causal relationship between input and output elements [45]. In CRA, these elements are denoted as variables, whereas in QCA, they are referred to as conditions of variables as they indicate presence or absence of certain quality of variables [46]. For instance, in the current study "*democracy*" is a variable and "full *democracy*" is the condition of that variable. Therefore, QCA breaks down variables into conditions, categorizing them as fully in, fully out, or neither in nor out [8].

[1] The research study conducts a comparative analysis of logistic regression (LRA) and QCA, highlighting that in datasets with minimal observations, such as the Lipset data from 1960, QCA delivers more precise results.

Table 1
Conceptual difference between CRA and QCA.

Particulars	CRA	QCA
Relation between variables	Symmetrical, Linear	Symmetric, Asymmetric, Linear, Non-Linear
Causal Relation	A higher level of an outcome is the exact opposite of the one, explaining the lower level of an outcome	A higher level of an outcome is not the exact opposite of the one, explaining the lower level of an outcome
Case-based Modelling	Does not allow case-based modeling and identifying localized effects makes a researcher capable of analyzing the cases for an in-depth understanding of the data sample and results	Allowing case-based modeling and identifying localized effects makes a researcher capable of analyzing the cases for an in-depth understanding of the data sample and results
Solution	Unifinality (A single best solution)	Equifinality (Multiple solutions)
Conditions of outcome	Identify important conditions of outcome in terms of measuring mediating, moderating, indirect and direct effects	Identify important conditions of outcome in terms of necessary, sufficient, core and peripheral conditions
Effects	Finding explains the net-effect of variables	Finding explains the net-effect and individual effect variables with contrarian cases
Empirical results	R-square presents the variance of the model by estimating probabilities	Based on the truth table, coverage explains how many cases empirically support a solution
Significance or strength of relationship	*P*-value determines the significance of results on which we accept or reject a hypothesis	Consistency explains empirical support to an outcome, which is the strength of a relationship
Variable's way of operation	All variables in the given model compete each other.	All variables in the model collectively work together to explain an outcome
Hypothesis/ Proposition	The research scholars develop hypotheses, which show an association between two or more variables.	The research scholars develop propositions in which we can analyze combined role of variables to an outcome. In addition, the contrarian cases in the model could be identified.
Counter-hypothesized relation	Counter hypothesized relations are difficult to identify for refining existing theories.	Counter-hypothesized relations are easily explained for refining existing theories.
Predictive validity	A powerful tool when dealing with relationships that are both, especially in the context of large-N datasets.	A powerful tool when dealing with relationships that are asymmetric and nonlinear in the context of small-N dataset.
Variables versus conditions	It deals with variables	It deals with conditions of variables

1.2. Reasons for adopting QCA

The use of QCA in human complex system and social science research is necessary for several reasons. In our daily lives of human complex system, we encounter various phenomena that involve agreement and difference [47], which are critical in shaping our attitudes and character. Comparing two or more events reveals commonalities that act as the cause, while identifying differences helps us understand what is not common, which serves as the effect. Cause and effect can be understood through set-theoretic relationships, such as the union and intersection of sets, which are the foundation of QCA [48]. Charles Ragin developed QCA to combine agreement and difference methods to test complex causal connections systematically and logically [49] in human complex system and social science research. Therefore, to validate daily life problem statements, using QCA as a research methodology in human complex system and social science research is crucial.

Another reason why QCA is essential in human complex system and social science research is that almost all theories in social science are verbal in nature and are formulated using set-theoretic relationships [50]. For example, “small family businesses are risk averse” is based on a set-theoretic relationship where the set of small family businesses (SFBs) is a subset of risk-averse individuals exhibits in Fig. 1.

Similarly, the claim that “religious fundamentalists are politically more conservative on economic and women issues” or “developed countries are democratic” involves set-theoretic connections between sets and subsets, as depicted in Fig. 2.

These statements do not have correlational connections, and their nature is based on set and subset associations [51]. To test these claims accurately, we need a methodology specially designed for research in human complex system and social science, and QCA is better suited for this task than CRA.

A third reason why QCA is a preferable methodology in human complex system and social science research is that it can handle variables with different types of relationships, including symmetric, asymmetric, linear, and non-linear connections [40]. In contrast, CRA can only analyze variables with symmetric and linear associations [51], which are not always representative of real-world in human complex system and social science research data in the context of modernization theory. Data collected from real-world situations of human complex system are seldom symmetric and linear due to the inherent complexities and variations in the world [52]. However, we can transform the data into a balanced form for regression analysis, which makes QCA superior to CRA. Set theory, which underlies QCA, recognizes that asymmetrical connections are fundamentally different from correlational linkages of variables. For instance, the claim that “small family businesses are risk averse” does not necessarily imply that large family businesses are risk-takers. Similarly, the statement “religious fundamentalists are politically more conservative on economic and women issues” does not mean that all fundamentalists are conservative on those issues. Finally, the assertion that “developed countries are democratic” means that the set of developed countries is a subset of democratic nations, but some less-developed countries may also be democratic, which supports the set-theoretic linkages rather than correlational connections. Based on these arguments, it is clear that QCA offers several advantages over CRA in human complex system of social science research.

Fig. 1. Risk Averse Investor subset of SFBs.

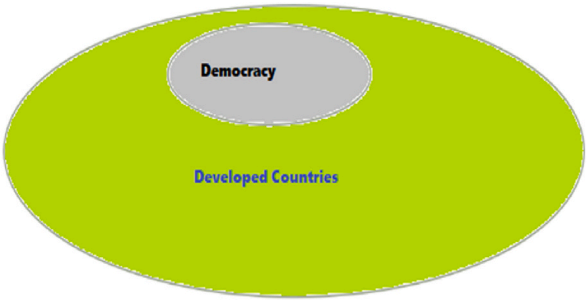

Fig. 2. Democracy as subset of SFB.

A fourth reason why QCA is a superior research methodology over CRA, is its flexibility in analyzing deductive, hypothetic deductive, abductive, and inductive approaches, making it applicable to a wide range of research problems [53]. Furthermore, QCA provides more generalizable results than CRA [54] in human complex system. In CRA, we simply input the data and rely on the software to generate conclusions, while in QCA, we carefully examine the data in relational move and calibrate it for membership value, analyze it using the software, and then draw conclusions based on both the results and our understanding of the data, which more appropriate to human complex system of social science research. Another advantage of QCA is its ability to integrate both quantitative and qualitative approaches [11] making it a more comprehensive and effective method than CRA. Therefore, QCA should be preferred for human complex system of social science research where flexibility and generalizability are important.

CRA suffer from several issues related to inferential statistics, which make them weak at validating hypotheses, especially in medical and psychological sciences [55]. Many research scholars have criticized the traditional null hypothesis significance testing (NHST), claiming that it is unsuitable for assessing research hypotheses [56–58], in human complex system and it should be banned in assessment of research. Additionally, most researchers and even statisticians do not fully understand the meaning of the p-value. Rejecting the null hypothesis does not always equate to accepting the alternative hypothesis in all cases, which leads to a deep flaw in testing the significance of null hypotheses, particularly in human complex system of social sciences [59]. These complexities have driven many researchers away from CRA, and prompted the adoption of a more modern methodology-QCA in human complex system of social science research. QCA, which combines both quantitative and qualitative approaches, is better equipped to handle the complexity of social science research data [11].

Charles Ragin created QCA as a case study approach specifically for small to medium-sized samples, less than fifty [60], which is not sufficient for CRA's analysis [54]. In human complex system of social science research, QCA can offer thorough and in-depth analysis of such samples, whereas CRA needs bigger samples. The causal factors of the model being evaluated determine the suitable sample size for QCA [61]. A model with four to seven causal conditions, for instance, necessitates a minimum sample size of twelve to thirty examples. But when a model has more than seven causal elements, the outcomes get complicated and difficult to understand in practical applications. In order to solve this problem, we can either divide the huge sample into smaller samples or combine the CRA framework with QCA, which will remove any potential model flaws [62].

There is a notable difference in the model's technique between the approaches taken by CRA and QCA with regard to data calibration [63]. CRA rely on variables that vary in category or value, while QCA works with set-theoretic relationships or causal conditions of social science research. This difference results in CRA estimating the change in the regressand variable due to the average one-unit change in the regressor, while QCA examines the presence/absence of causal conditions or combinations of them to identify connections with the outcome, which captures real world of social science research. This set-theoretic approach has garnered interest among researchers of social science as the heart of configurational methodology [63]. However, QCA requires the researcher's knowledge and understanding of the examined variables in human complex system, their conditions, and the underlying theory context to contribute throughout the study. The social science' researcher of human complex system then calibrates the data into fuzzy sets, selects appropriate solutions, and interprets the results. Despite some criticisms of this subjective bias, the researcher's knowledge and understanding can provide a more comprehensive analysis in human complex system and social science research that is not possible with CRA [64]. The results are more applicable and offer a fine-grained depth of knowledge regarding the connections among variables of interest [40] in human complex system of social science research.

In QCA, it is simple to transform various types of data, such as multimodal, clickstream, and likert scales, into binary or fuzzy sets with values ranging from 0 to 1 [40]. While nominal or categorical data such as gender, which cannot be transformed into fuzzy sets, they can be easily managed in QCA as all variables fall within the 0–1 range. In contrast, CRA, particularly the quantitative approach, struggle to account for the influence of nominal variables, particularly when other independent variables have large values. When dealing with categorical variables, CRA may compare entirely dissimilar items, leading to concerns about the accuracy of the findings.

In CRA, the focus is on estimating the average effect of all predictors on the outcome, and identifying which variables are significant or not. On the other hand, QCA provides multiple configurations that offer different solutions, which can be 2^k-1, where k is the number of predictors that affect the outcome, which more suit to human complex system of social science research. For example, if we have three predictors, α, β, and γ, that affect the outcome ϒ, QCA can offer various combinations of these predictors that may impact ϒ, as shown in Table 2.

The basic difference between QCA and CRA lies in the number and types of solutions they offer. While CRA offer a single best solution, QCA provides multiple solutions, including complex, parsimonious, and intermediate solutions, which are blends of configurations that lead to the same-outcome [65,66]. [67], near to human complex system of life in social science research. In contrast, CRA cannot generate core and peripheral states, limiting the in-depth understanding of results. QCA analyses asymmetric relations between various complex combinations of antecedents that generate to the similar outcome, while CRA estimate the net effect of variables in competing environments. QCA's multiple solutions provide practical and implementable findings for all levels of management, unlike CRA' single best solution. This makes QCA more aligned with real-life phenomena of human complex system, offering a more comprehensive and practical approach for social science research, social changes and modern theory [68].

1.3. INUS and SUIN conditions

The INUS and SUIN conditions are important concepts in QCA, a method used to analyze patterns in data and identify combinations of factors associated with a particular outcome, which makes QCA close to the real world of human complex system, and superior to CRA [69–71]. INUS means "insufficient but necessary components of an unnecessary but sufficient condition", which is although unnecessary but sufficient to generate an outcome. It means the condition which is necessary for an outcome is insufficient on its own. Thus, this condition needs to be joined with other conditions for getting an outcome, otherwise not possible. For example, for a new venture money or finance is necessary but insufficient to do the venture alone. It needs other factors like active managers, comprehensive plans and entrepreneurial courage to achieve success.

Similarly, the "SUIN" condition is very effective in fsQCA. It means "sufficient but unnecessary parts of a necessary condition" indicating those factors, which are sufficient to generate outcomes but unnecessary [70]. For instance, for a healthy life, exercise is an unnecessary but sufficient condition. A healthy life is attained by eating good quality food and enjoying a stressless environment. These concepts of INUS and SUIN conditions enable scholars to examine the complex phenomenon of social science in human-complex systems where they identify combinations of conditions that are strongly linked and associated with a particular outcome. This kind of combination of conditions may not be identified by regression-based conventional methods.

2. Model specification

Some research scholars indicate the key factors of democracy are gross national product (GNP), industrial labour force, urbanization and literacy [72], especially in the study of economic and social sciences. Firstly, higher levels of economic development, measured by GNP per capita, tend to be positively associated with the economic development of democratic institutions and political participation. This is due to the fact that people are more inclined to want a voice in national governance as they get more educated and have steady careers. Second, there is a correlation between increased democracy and urbanization. People are more likely to become politically involved as they relocate to metropolitan regions because they are exposed to a wider range of viewpoints. Thirdly, because of its strong correlation with educational attainment, the literacy rate is a significant predictor of democracy. People are more likely to

Table 2
Combination algorithm of QCA.

Independent Variables: Predictors, Antecedents or Ingredients (n)			Dependent Variable: Recipe, solution and outcome (2^n)
α	β	γ	Corresponding Vector
0	1	0	~α *β ~*γ
1	1	0	α *β ~*γ
1	0	0	α * ~β *~γ
0	1	1	~α *β *γ
0	0	1	~α *~ β *γ
1	1	1	α *β *γ
1	0	1	α * ~β *γ

participate in democracy and be aware of political issues in nations where literacy rates are greater. Finally, because industrialization tends to increase chances for political and educational engagement, an industrial labour force is likewise related with democracy. This is due to the fact that an educated and competent workforce produced by industrialization is more inclined to seek civil rights and political participation. All things considered, Lipset's research indicates that in human complex system of social science, such as economic and social development, is crucial to the growth of democratic institutions and political engagement. Four assertions encapsulate the Lipset theory of socioeconomic modernization and development. We recommend that the conditions[2] conducive to democracy encompass high GNP per capita, high literacy; high urbanization; large share industrial labor force. Additionally, we propose propositions outlining the means by which these conditions can foster the survival of democracy:

P_1: Nearly all cases (i.e., nations) with high scores in any four of the combinations of the following four conditions have high scores in democracy survival: high GNP per capita; high literacy; high urbanization; large share industrial labor force.

P_2: Nearly all cases (i.e., nations) with high scores in any three of the combinations of the following four conditions have high scores in democracy survival: high GNP per capita; high literacy; high urbanization; large share industrial labor force.

P_3: Nearly all cases (i.e., nations) with high scores in any two or fewer of the combinations of the following four conditions have high scores in democracy survival: high GNP per capita; high literacy; high urbanization; large share industrial labor force.

These propositions suggest that as countries become more modernized through economic growth, urbanization, higher literacy rates, and increased industrialization, they are more likely to have stable and enduring democratic systems of government. The Lipset theory has been influential in the study of political development and comparative politics, and has been subject to extensive testing and refinement by scholars of human complex system of social science research over the years.

$$\textbf{Survival of Democracy} = \int \boldsymbol{GNPCAP}, \boldsymbol{URBANIZA}, \boldsymbol{LITERACY}, \boldsymbol{INDLAB} \qquad 1$$

Where: GNPCAP: Gross National Product/Capita, URBANIZA: Urbanization, LITERACY: Literacy, INDLAB: Industrial Labour Force (including mining).

3. Methods

We use secondary cross-national data from the book "Political Man: The Social Bases of Politics" [73] to compare the methodological effectiveness of CRA, csQCA, and fsQCA. The original data was collected by Lipset and Man for 18 countries (see Table 3), which were selected to perform the analyses.

The data includes an outcome variable that is binary, with a value of (0) representing "breakdown of democracy" (10 cases) and a value of (1) representing "survival of democracy" (8 cases). The input ingredients used in the analysis are gross national product per capita, urbanization, literacy rate, and industrial labour force. The data were used to compare the effectiveness of the three different methodological approaches in predicting the survival or breakdown of democracy based on these input variables. Overall, this approach demonstrates the utility of using existing cross-national data to compare different methodological approaches in human complex system of social science research and the theory of modernization.

3.1. Conventional regression analysis (CRA)

Table 4 displays the descriptive statistics for five variables in the cross-national data analysed in the study. The variables include Gross National Product (GNP) per capita (GNPCAP), percentage of the population that is urbanized (URBANIZA), literacy rate (LITERACY), percentage of the population engaged in industrial labour (INDLAB), and survival of democracy (SURVIVAL). The sample size for each variable is 18 countries.

In the given sample, the average GNP per capita is dollars (641.33), having maximum (minimum) values are dollars 1098 (320). The standard deviation of GNP per capita is relatively high at dollars 269.42, which suggests that there is considerable diversity in economic development among the countries included in the sample. On the other hand, the average percentage of urban population is 39.94 %, ranging from 15 % to 79 %. The SD of urbanization is 19.45, which infers a moderate level of variability in urbanization among the countries under consideration.

The mean literacy rate in the sample is 84.44 %, with a range of 38 %–100 %. The standard deviation of literacy rate is 19.16, indicating a moderate degree of variability in educational attainment across the countries in the sample. The mean percentage of the population engaged in industrial labour is 28.79 %, with a range of 11 %–50 %. The standard deviation of this variable is 11.85, indicating a moderate degree of variability in the degree of industrialization across the countries in the sample.

The outcome variable, SURVIVAL, ranges from 0 (indicating non-survival) to 1 (indicating survival). In the sample, the mean value of this variable is 0.44, with a standard deviation of 0.51, signifying considerable variability among the countries included in the analysis.

The Model Summary contains information on the adequacy of the logistic regression model's fit for social science research. Table 5 displays the outcome of one stage in the process of constructing the model, where a lower value of −2 log likelihood is an indication of a superior fit [74]. In this instance, the value is 9.343, indicating that the model is a reasonably good fit. The Cox & Snell R Square and Nagelkerke R Square are science of how much of the variability in the dependent variable is accounted for by the model. A higher value

[2] High and large means calibrated score equal or above 0.50.

Table 3
Lipset data.

CASEID	GNPCAP	URBANIZA	LITERACY	INDLAB	SURVIVAL
NET	1008	78.8	99.9	39.3	1
SWE	397	34	99.9	32.3	1
UK	1038	74	99.9	49.9	1
FIN	590	22	99.1	22	1
AUS	720	33.4	98	33.4	0
GER	795	56.5	98	40.4	0
FRA	983	21.2	96.2	34.8	1
CZE	586	69	95.9	37.4	1
EST	468	28.5	95	14	0
IRE	662	25	95	14.5	1
BEL	1098	60.5	94.4	48.9	1
HUN	424	36.3	85	21.6	0
POL	350	37	76.9	11.2	0
ITA	517	31.4	72.1	29.6	0
ROM	331	21.9	61.8	12.2	0
GRE	390	31.1	59.2	28.1	0
SPA	367	43	55.6	25.5	0
POR	320	15.3	38	23.1	0

Table 4
Descriptive statistics.

Variables	n	Min	Max	Mean	SD
GNPCAP	18	320	1098	641.33	269.42
URBANIZA	18	15	79	39.94	19.45
LITERACY	18	38	100	84.44	19.16
INDLAB	18	11	50	28.79	11.85
SURVIVAL	18	0	1	0.44	0.51

Table 5
Analysis of CRA.

Variables	B	S.E.	Wald	df	Sig.	Exp(B)
GNPCAP	0.015	0.011	1.917	1	0.166	1.015
URBANIZA	0.106	0.112	0.902	1	0.342	1.112
LITERACY	0.288	0.457	0.396	1	0.529	1.333
INDLAB	−0.285	0.268	1.134	1	0.287	0.752

−2 Log likelihood (9.343), Cox and Snell R Square (0.575), Nagelkerke R^2 (0.769).

signifies a better ability to explain. In this scenario, the Cox & Snell R Square is 0.575 and the Nagelkerke R Square is 0.769, indicating that the model explains a substantial portion of the variance in the dependent variable [75].

According to Table 5, the logistic regression analysis includes four independent variables: GNPCAP, URBANIZA, LITERACY, and INDLAB, all of which are continuous. The dependent variable is presumably dichotomous, with only two potential outcomes. The table presents data on the coefficients, standard errors, Wald statistic, degrees of freedom, significance, and odds ratios for each independent variable. The column (B) in Table 5 contains the regression coefficients, which are estimations of the effect sizes of each independent variable. The column (S.E) displays the standard errors of the coefficients, which reflect the accuracy of the estimates. The "Wald" column reports the Wald chi-square statistic, a test of the null hypothesis that the coefficient for each independent variable is zero (i.e., that the variable has no impact on the outcome). The "df" column shows the degrees of freedom associated with each Wald statistic.

Overall, it appears that none of the independent variables were statistically significant predictors of the outcome at the standard alpha level of 0.05, as all of the p-values are greater than 0.05 [76].

3.2. csQCA analysis

The original data from Lipset's "Political Man" (1960) was converted into binary variables for the purpose of conducting csQCA analysis. Four conditions were presented in table, which included Gross National Product per Capita (GNPCAP), Literacy, and Industrial Labour Force.

We recoded all the values in Table 3 using QCA software, converting them into a binary format. Each value was assigned either 0 or 1, based on specific definitions relevant to each variable.

For GNPCAP, a value of 0 was assigned if the Gross National Product per Capita was below 600, and a value of 1 was assigned if it was equal to or above 600. Similarly, for Urbanization, a value of 0 was assigned if the population was below 50, and a value of 1 was assigned if it was equal to or above 50 [77]. For Literacy,[3] a value of 0 was assigned if the literacy rate was below 75, and a value of 1 was assigned if it was equal to or above 75. Finally, for Industrial Labour Force, a value of 0 was assigned if the percentage of the labour force engaged in industrial activities (including mining) was below 30, and a value of 1 was assigned if it was equal to or above 30 [78,79]. We get dichotomized variables in this way. As usual, we commence this analysis to ascertain the prerequisite condition for the outcome.

3.2.1. Necessary condition analysis (CNA)

Utilizing the fsQCA software, we obtain consistency and coverage values for each variable (both presence and absence) in relation to survival (both presence and absence), as depicted in Table 6. This table indicates that no individual condition is necessary for the outcome considering the criteria of consistency and coverage value 0.8 simultaneously [80–83].

3.2.2. Sufficiency condition analysis (SCA)

In the next step, we perform sufficiency condition analysis using the fsQCA software. Table 7 serves as a truth table for conducting csQCA analysis. Each condition in the table is represented by either 0 or 1. Additionally, the table includes a column indicating the number of cases (n) that satisfy each combination of conditions.

The subsequent three columns represent consistency measures. Specifically: Raw Consistency corresponds to the proportion of cases that exhibit the outcome (in this scenario, "survival") for each combination of conditions PRI consistency represents a standardized score that accounts for the frequency of the outcome of interest within the sample, signifying the proportional reduction in inconsistency. An important benchmark is that PRI consistency for a model should be equal to or greater than 0.70 [84]. If the PRI falls below this threshold, it suggests that the model may have numerous cases with low scores for the outcome Y despite having high scores for antecedents. Examination of truth Table 7 reveals that only one configuration of variables exceeds this benchmark, achieving a score of 0.7 or higher. It means this model is not good fit for configuration analysis. SYM consistency is the symmetric (or necessary) consistency score, which takes into account the consistency of the absence of the outcome of interest as well as its presence.

Table 8 presents the results of a sufficiency analysis using csQCA. The table includes two conditions and their respective raw coverage, unique coverage, and consistency values. Raw coverage helps us understand how well specific conditions collectively explain an outcome Y, while unique coverage seeks to maximize coverage while maintaining uniqueness in covering elements [85,86]. Coverage refers to the proportion of cases in which the outcome condition (Y) has high scores and these cases also exhibit high scores in the antecedent model. To illustrate, consider a scenario where out of a total sample of 30 cases, out of which 20 cases have high Y scores. Among these 20 cases, 10 also have high scores according to a specific model. In this case, the coverage index for the model would be 0.50. Essentially, it quantifies how well the model captures the cases with high Y scores based on the antecedent conditions [87].

Consistency is the proportion of cases having high scores in the model that have high scores in the outcome [88]. For example, if 12 cases have high scores for a given antecedent model and 10 of these 12 cases have high scores for the outcome, then the consistency index is 0.83.

The first condition states that the absence of GNPCAP, in combination with the presence of LITRACY and INDLABF, can generate the outcome. According to the second criterion, the result can also be produced by GNPCAP and LITRACY in addition to URBANISATION and INDLABF being absent. The analysis's findings indicate that, with a solution consistency of 1 and a coverage of 0.375, these two requirements are adequate to provide the result. This indicates that the result is likely to happen if any one of these two circumstances holds true. Nevertheless, the coverage is relatively low.

3.3. FsQCA analysis

We define three inflection points (95th, 50th, and 5th percentiles),[4] when calibrating data for each condition: the point denoting the minimum score for full membership in displaying condition (0.95); the point denoting maximum ambiguity (0.50); and the point denoting full non-membership (0.05) [40,89–92].

We initiate the necessary condition analysis using the fsQCA software to examine the survival of democracy (outcome).

3.3.1. Necessary condition analysis (CNA)

Through the fsQCA software, we acquired consistency and coverage values for each variable (both presence and absence) concerning survival (both presence and absence), as outlined in Table 9.

3.3.2. Sufficiency condition analysis (SCA)

Next, we proceed to conduct a sufficiency condition analysis utilizing the fsQCA software. Table 10 functions as a truth table for the fsQCA analysis. Each condition within the table is denoted by either 0 or 1. Furthermore, the table incorporates a column specifying the number of cases (n) that fulfil each combination of conditions.

[3] This recoding incorporates the local effect of variables, which is not possible in CRA.

[4] The main purpose of this study is to show examples of how QCA is beneficial. However, this calibration method doesn't work when one or more dummy variables are in the study, because the 95th, 50th, and 5th percentiles of (1,0) are (1,0,0) respectively, which is incorrect. This may create a problem in results and XY-plot. We recommend applying fsQCA on non-dummy variables.

Table 6
Necessity Condition Analysis of csQCA.

Variables	Presence		Absence	
	Cons.	Cov.	Cons.	Cov.
GNPCAP	0.750	0.750	0.200	0.750
~GNPCAP	0.250	0.200	0.800	0.250
URBANIZA	0.500	0.800	0.100	0.500
~URBANIZA	0.500	0.308	0.900	0.500
LITERACY	1.000	0.615	0.500	1.000
~LITERACY	0.000	0.000	0.500	0.000
INDLAB	0.750	0.750	0.200	0.750
~INDLAB	0.750	0.750	0.200	0.750

The symbol (~) shows the absence of a condition.

Table 7
Truth Table of csQCA.

GNPCAP	URBANIZA	LITRACY	INDLABF	n	SURVIVAL	Raw-cons	PRI-cons	SYM-cons
1	0	1	0	1	1	1	1	1
0	0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1	1
1	1	1	1	4	0	0.75	0.75	0.75
1	0	1	1	2	0	0.5	0.5	0.5
0	0	1	0	4	0	0.25	0.25	0.25
0	0	0	0	5	0	0	0	0

Table 8
Sufficiency Condition Analysis of csQCA.

Conditions	Raw-Cov	Unique- Cov	Cons
~GNPCAP*LITRACY*INDLABF	0.25	0.25	1
GNPCAP*~URBANIZATION*LITRACY*~INDLABF	0.125	0.125	1
Solution coverage	0.375		
Solution consistency	1		

Where "~" indicates the absence of the variable in the condition.

Table 9
Necessity Condition Analysis of fsQCA.

Variables	Presence		Absence	
	Cons.	Cov.	Cons.	Cov.
GNPCAP	0.615	0.938	0.474	0.310
~GNPCAP	0.548	0.708	0.906	0.502
URBANIZA	0.665	0.937	0.774	0.468
~URBANIZA	0.622	0.865	0.896	0.534
LITERACY	0.706	0.925	0.596	0.335
~LITERACY	0.493	0.740	0.867	0.558
INDLAB	0.646	0.939	0.631	0.393
~INDLAB	0.583	0.787	0.902	0.521
The symbol (~) shows the absence of a condition.				

This table indicates that no individual condition is necessary for the outcome considering the criteria of consistency and coverage value 0.8 simultaneously [80–83].

Table 10 reveals five specific combinations of conditions where the PRI (Proportional Reduction in Inconsistency) exceeds the threshold of 0.7. These combinations are capable of generating the desired outcome. Continuing the analysis with a frequency of 1 and a consistency of 0.9, the results are summarized in Table 11.

Table 11 shows the results of fsQCA analysis. This table includes four conditions and their respective raw coverage, unique coverage, and consistency values. The first condition, GNPCAP*~URBANIZATION*LITRACY, has a raw coverage of 0.307, meaning that this condition is present in 30.7 % of cases, regardless the presence/absence of the outcome. Its unique coverage is 0.103, meaning that only 10.3 % of cases with an outcome are affected uniquely by this configuration. There is a good association between the existence of the condition and the presence of the outcome, as indicated by the condition's consistency score of 0.997.

GNPCAP*URBANIZATION*INDLAB, the second condition, has a unique coverage of 0.239 and a raw coverage of 0.442. It has a

Table 10
Truth Table of fsQCA.

GNPCAP	URBANIZATION	LITERACY	INDLABF	num	SURVIVAL	raw consist.	PRI consists.	SYM consist
1	0	1	0	1	1	1	1	1
1	1	0	1	1	1	1	1	1
1	0	1	1	2	1	0.997	0.991	1
0	0	0	1	1	1	0.989	0.824	1
1	1	1	1	4	1	0.946	0.908	1
0	1	0	0	3	1	0.944	0.629	1
0	0	0	0	3	0	0.888	0.402	1

Table 11
Sufficiency Condition Analysis of fsQCA.

Conditions	Raw-Cov	Unique- Cov	Cons
GNPCAP*∼URBANIZATION*LITRACY	0.307	0.103	0.997
GNPCAP*URBANIZATION*INDLAB	0.442	0.239	0.951
∼GNPCAP*URBANIZATION*∼LITRACY*∼INDLAB	0.306	0.077	0.944
∼GNPCAP*∼URBANIZATION*∼LITRACY*INDLAB	0.224	0.030	0.989
Solution coverage	0.729		
Solution consistency	0.943		

0.951 consistency score. The raw coverage of the third condition, ∼GNPCAP*URBANIZATION*∼LITRACY*∼INDLAB, is 0.306, whereas the unique coverage is 0.077. It has a 0.944 consistency score. The raw coverage of the fourth condition, ∼GNPCAP*∼URBANIZATION*∼LITRACY*INDLAB, is 0.224, while the unique coverage is 0.03. It has a 0.989.

The percentage of cases, regardless of whether the outcome is available, in which the solution conditions are present is displayed in the "Solution coverage". The solution coverage in this instance is 0.729, meaning that 72.9 % proportion of cases having high scores for the outcome condition having high scores in the antecedent model. It means that a substantial portion i.e., 72.9 percent of the outcome is covered by these four solutions. We can compare it with R^2 of conventional regression analysis (CRA) [40].

Similarly, "solution consistency" is the proportion of cases having high scores in the model that have high scores in the outcome. With a solution consistency of 94.3 % in this instance, it indicates a strong relationship between high scores in antecedent conditions and high scores in the outcome.

The study's overall findings can be summarized as follows: Solutions 1–4 illustrate various combinations of conditions, both present and absent, that play a role in sustaining democracy.

For example, the path in solution 1 shows that the combination of the presence enhanced level of gross national product per capita, higher level of literacy rate and absence of higher levels of urbanization are sufficient to generate sustainable democracy. However, the path of solution 2 shows that the presence enhanced level of gross national product per capita in combination with the presence of an industrial labour force and urbanization on a higher level are sufficient for the outcome. Moreover, the path of solution 3 exhibits only the presence of higher levels of urbanization is enough to generate sustainable democracy in the absence of enhanced levels of gross national product per capita, higher levels of urbanization and higher level of industrial labor forces. Similarly, the path of solution 4 exhibits that only a higher level of the industrial labour force in the absence of gross national product per capita, urbanization and literacy rate can contribute toward sustainable democracy. Based on raw coverage, it is evident that among all four solutions, Solution 2 emerges as the most effective and feasible.

3.3.3. Evaluation of propositions

The propositions within the configuration model of the study exhibited in Fig. 3, can be assessed using fsQCA software.

Once we have all possible solutions for the outcome of interest using fsQCA, we may also test for specific propositions. This involves analyzing the extent to which each proposition covers cases within the sample dataset [40]. Additionally, by using this procedure, we are able to identify the relevant cases for every proposition in the dataset. This could be performed by computing specific configuration against each proposition in fsQCA, which creates a model, and matching it to the outcome Y. For instance, we test our second proposition (P_2) i.e., nearly all cases (i.e., nations) with high scores in any three of the combinations of the following four conditions have high scores in democracy survival: high GNP per capita; high literacy; high urbanization; large share industrial labor force. The results of path-2 show this combination is as *GNPCAP*URBANIZATION*INDLAB*. Using fsQCA, we can evaluate this proposition by plotting an XY plot[5] of the function within the context of the data sample, employing the fuzzyand function (x, …) as in Fig. 4. With a consistency of 95 % and a coverage of 42 %, it suggests that proposition P_2 receives partial support. Only 42 % of cases from the dataset supporting this model enable the identification of specific instances across all nations in the dataset. This model holds potential for advancing theory as its consistency level of 95 % surpasses the threshold criteria of 80 % [93].

[5] This XY plot may contains error due to dummy outcome variable.

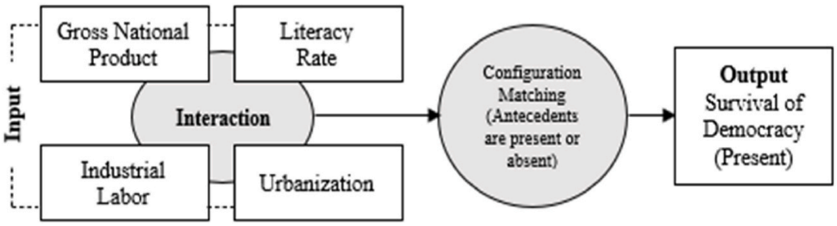

Fig. 3. Configuration model.

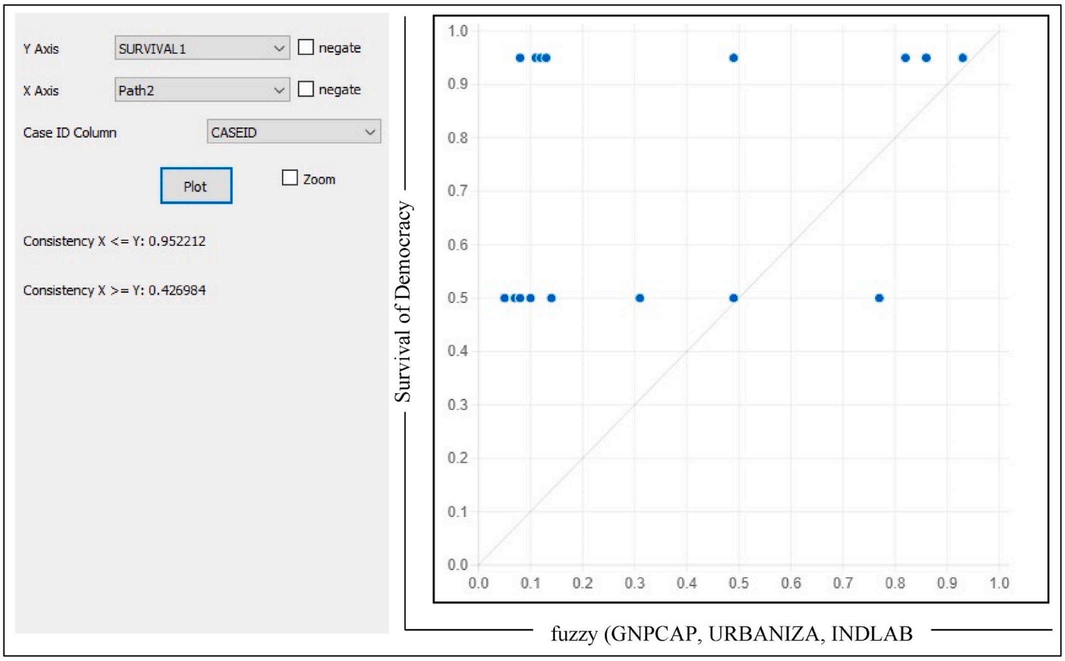

Fig. 4. XY-plot.

4. Conclusions and discussion

The model summary of CRA in human complex system of social science research, which is a statistical method used to model the relationship between a categorical dependent variable (usually binary) and one or more independent variables. The value of likelihood test (9.343) suggests that the model fits the data reasonably well. The Cox & Snell R Square explains about 57.5 % of the variance in the dependent variable. The Nagelkerke R Square is 0.769, which indicates that the model explains about 76.9 % of the variance in the dependent variable. Thus, we can conclude that the model is good fit.

However, the results suggest that none of the variables in the model have a statistically significant relationship with the dependent variable at the 0.05 level, although the coefficients for LITERACY and GNPCAP are positive and suggest a positive association with the dependent variable. It's important to note that the non-significance of these variables could be due to a lack of statistical power, and further analysis may be needed to fully understand the relationship between these ingredients and the output variable. It is, therefore, difficult to predict the outcome through conventional regression analysis (CRA) in human complex system.

In logistic regression analysis, like in this particular study involving four independent variables and one binary dependent variable, it's typically advised to aim for a sample size ranging from 40 to 80 observations, allocating 10 to 20 observations for each variable [88, 94] as a general rule of thumb. This guideline is intended to ensure that the model possesses adequate statistical power to detect meaningful effects. Given the total of 18 observations in the current study, it falls short of ensuring the validity of the results. Consequently, conventional regression analysis (CRA) cannot be conducted with a dataset of 18 observations. In addition to this, Conventional regression analysis (CRA) presents various concerns regarding the validity of results, including issues such as the assumption of linearity in relationships, multicollinearity, homoscedasticity, and potential convergence problems during estimation. These challenges raise doubts about the accuracy and reliability of the findings. Failing to address these above said issues undermines the ability to ensure that your regression analysis yields valid and reliable results.

We now shift our focus to examining the outcomes derived from crisp set qualitative comparative analysis (csQCA). Table 8 presents two solutions generated by the csQCA analysis.

i. The first solution (~GNPCAPLITRACYINDLABF) has a raw coverage of 0.25, which means that 25 % of cases meet the conditions of this solution.
ii. The second solution (GNPCAP ~ URBANIZATIONLITRACY*~INDLABF) has a raw coverage of 0.125, which means that 12.5 % of cases meet the conditions of this solution.

In this case, the solution coverage stands at 0.375, indicating that only 37.5 % of cases with high scores for the outcome condition also exhibit high scores in the antecedent model. This implies that a mere 37.5 % portion of the outcome is accounted for by these two solutions, which low from minimum threshold i.e., 0.5 [95]. Consequently, we remain unable to confidently predict the results. In addition, neither solution has a particularly high raw coverage or unique coverage, suggesting that the conditions alone may not be sufficient for the outcome and that additional conditions or factors may need to be considered. Yet, these results are better than CRA.

However, in contrast to CRA and csQCA, the fuzzy set qualitative comparative analysis (fsQCA) presents four solutions or combinations of conditions that are deemed sufficient for achieving the outcome of interest. These solutions exhibit raw coverage values ranging from 0.224 to 0.442 and unique coverage values ranging from 0.013 to 0.239. Additionally, these solutions demonstrate a high level of consistency with the outcome of interest, with consistency scores ranging from 0.944 to 0.989. The overall solution coverage (0.729) and consistency (0.943) are good indicating that model is fit and outcome is well explained by the inputs. Therefore, we can now confidently forecast the outcome of the predictors.

Upon reviewing the comparative results of CRA, csQCA, and fsQCA, it seems that fsQCA could be deemed the most appropriate research method for studying complex human systems in social science research, within the framework of modern theory. The results of CRA indicates that none of the predictors is significant to the outcome. This implies that neither of the predictors—such as gross national product per capita, urbanization, literacy rate, and industrial labor force—can be relied upon to achieve sustainable democracy. While the results from csQCA offered valuable insights to some extent, they did not provide the same degree of depth and detail as fsQCA. The fsQCA results offer a comprehensive package, encompassing four configurational solutions or pathways through which any participating country can achieve sustainable democracy. Should one solution or pathway not be applicable to any nation, there are three other options available to choose from based on the particular circumstances. As a result, fsQCA examines intricate causal links and offers a detailed explanation of the phenomenon of interest. Researchers can bridge the gap between qualitative and quantitative methodologies and enhance the theoretical value of their work by integrating QCA with traditional statistical analysis [11]. Contrast to QCA, it's important to note that CRA only offers the single best solution. Conventional regression analysis (CRA) is not as advantageous as qualitative comparative analysis (QCA), particularly when working with small-N datasets [96,97]. Hence, we recommend that future research on human complex system of economic development within the framework of contemporary theory take into account the application of fsQCA. It does, however, have some limits much like any other study methodology.

4.1. Limitations of the study

- The data with a limited sample size might not be reliable. Future research may be conducted on large-scale datasheets.
- We calibrated the data on standardized procedure i.e., 95th, 50th, and 5th percentiles, which don't work on dummy variables. The next study may be conducted non-dummy variables.
- This analysis depends heavily on the quality of the data used. The analysis's conclusions could not be trustworthy if the data is erroneous, contradictory, or inadequate.
- To evaluate the validity and reliability of the results, fsQCA does not offer any statistical tests. Because of this, the researcher's only means of evaluating the results' robustness is through qualitative analysis.

Data availability statement

The complete dataset is available at this source: "Handbooks used to prepare the tables are Flora et al., 1983, 1987; League of Nations, Statistical Yearbook, Geneva, various years; Mitchell, Brian R., European Historical Statistics 1750–1975, London: MacMillan, 1891; Statistisches Reichsrat, Statistisches Handbuch der Weltwirtschaft, Berlin 1936". The data used in this study is also available in Table 3.

CRediT authorship contribution statement

Xinxin Xu: Writing – review & editing, Validation, Software, Methodology, Formal analysis, Investigation, Project administration, Visualization. **Safdar Husain Tahir:** Conceptualization, Data curation, Formal analysis, Methodology, Software, Writing – original draft, Writing – review & editing. **Khuda Bakhsh Khan:** Funding acquisition, Investigation, Resources, Supervision, Visualization. **Mushtaq A. Sajid:** Funding acquisition, Investigation, Resources, Validation, Visualization. **Muhammad Azhaf Safdar:** Data curation, Funding acquisition, Methodology, Project administration, Supervision, Validation.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:Safdar Husain Tahir reports a relationship with Government College University Faisalabad that includes: employment. Xinxin Xu reports a relationship with Chengdu University Business School that includes: employment. Khuda Bakhsh Khan reports a relationship with Government College University Faisalabad that includes: employment. Mushtaq A Sajid reports a relationship with Mohi-ud-Din Islamic University that includes: employment. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Abbreviations

csQCA crisp-set Qualitative Comparative Analysis
fsQCA fuzzy set Qualitative Comparative Analysis
GNP Gross National Product
GNPCAP Gross National Product per Capita
INDLABF Industrial Labor Force
INUS: Insufficient but Necessary parts of an Unnecessary but Sufficient condition
CRA Logistic Regression Analysis
QCA Qualitative Comparative Analysis
SD Standard Deviation
SUIN Sufficient but Unnecessary part of an Insufficient but Necessary condition

References

[1] B. Rihoux, Bridging the gap between the qualitative and quantitative worlds? A retrospective and prospective view on qualitative comparative analysis, Field Methods 15 (4) (2003) 351–365.
[2] E. Fritzsche, Making hermeneutics explicit: how QCA supports an insightful dialogue between theory and cases, Int. J. Soc. Res. Methodol. 17 (4) (2014) 403–426.
[3] A. Marx, B. Rihoux, C. Ragin, The origins, development, and application of Qualitative Comparative Analysis: the first 25 years, European Political Science Review 6 (1) (2014) 115–142.
[4] C.C. Ragin, Comment: lucas and szatrowski in critical perspective, Socio. Methodol. 44 (1) (2014) 80–94.
[5] F. Castelló-Sirvent, A fuzzy-set qualitative comparative analysis of publications on the fuzzy sets theory, Mathematics 10 (8) (2022) 1322.
[6] L. Dell'Aguzzo, Dealing with causal complexity and equifinality in international relations: a fuzzy-set QCA analysis of conflict resolution, in: SAGE Research Methods Cases Part 2, SAGE Publications Ltd, 2019.
[7] S. Verweij, C. Zuidema, 17. Qualitative Comparative Analysis for analyzing spatial planning processes, Handbook on Planning and Complexity (2020) 334.
[8] S. Pagliarin, S. La Mendola, B. Vis, The "qualitative" in qualitative comparative analysis (QCA): research moves, case-intimacy and face-to-face interviews, Qual. Quantity (2022) 1–19.
[9] C.Q. Schneider, C. Wagemann, Qualitative comparative analysis (QCA) and fuzzy-sets: agenda for a research approach and a data analysis technique, Comp. Sociol. 9 (3) (2010) 376–396.
[10] E. Thomann, M. Maggetti, Designing research with qualitative comparative analysis (QCA): approaches, challenges, and tools, Socio. Methods Res. 49 (2) (2020) 356–386.
[11] J. Meuer, C. Rupietta, A review of integrated QCA and statistical analyses, Qual. Quantity 51 (2017) 2063–2083.
[12] S. Pagliarin, S. La Mendola, B. Vis, The "qualitative" in qualitative comparative analysis (QCA): research moves, case-intimacy and face-to-face interviews, Qual. Quantity 57 (1) (2023) 489–507.
[13] X. Basurto, J. Speer, Structuring the calibration of qualitative data as sets for qualitative comparative analysis (QCA), Field Methods 24 (2) (2012) 155–174.
[14] S. Vaisey, QCA 3.0: the "Ragin Revolution" Continues, SAGE Publications Sage, CA: Los Angeles, CA, 2009.
[15] C.Q. Schneider, C. Wagemann, Set-theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis, Cambridge University Press, 2012.
[16] T. Greckhamer, et al., Using qualitative comparative analysis in strategic management research: an examination of combinations of industry, corporate, and business-unit effects, Organ. Res. Methods 11 (4) (2008) 695–726.
[17] C.Q. Schneider, C. Wagemann, Reducing complexity in Qualitative Comparative Analysis (QCA): remote and proximate factors and the consolidation of democracy, Eur. J. Polit. Res. 45 (5) (2006) 751–786.
[18] A. Thiem, Unifying configurational comparative methods: generalized-set qualitative comparative analysis, Socio. Methods Res. 43 (2) (2014) 313–337.
[19] P.A. Mello, Qualitative Comparative Analysis: an Introduction to Research Design and Application, Georgetown University Press, 2021.
[20] D.J. Teece, Dynamic capabilities as (workable) management systems theory, J. Manag. Organ. 24 (3) (2018) 359–368.
[21] M. Kapsali, Equifinality in project management exploring causal complexity in projects, Syst. Res. Behav. Sci. 30 (1) (2013) 2–14.
[22] R.L. Lyman, The concept of equifinality in taphonomy, Journal of Taphonomy 2 (1) (2004) 15–26.
[23] D. Maulud, A.M. Abdulazeez, A review on linear regression comprehensive in machine learning, Journal of Applied Science and Technology Trends 1 (4) (2020) 140–147.
[24] H. White, I. Domowitz, Nonlinear regression with dependent observations, Econometrica: J. Econom. Soc. (1984) 143–161.
[25] J.W. Osborne, E. Waters, Four assumptions of multiple regression that researchers should always test, Practical Assess. Res. Eval. 8 (1) (2019) 2.
[26] A. Duşa, User manual for the QCA (GUI) package in R, J. Bus. Res. 60 (5) (2007) 576–586.
[27] I. Drugus, A universal algebraic set theory built on mereology with applications, Logica Universalis (2022) 1–31.
[28] A. Lex, et al., UpSet: visualization of intersecting sets, IEEE Trans. Visual. Comput. Graph. 20 (12) (2014) 1983–1992.
[29] L. Bedhouche, et al., Efficient association rules minimization using a double-stage quine-McCluskey-based approach, in: International Conference on Computational Collective Intelligence, Springer, 2023.
[30] T. Sridevi, S.S. Fathima, Digital image watermarking using fuzzy logic approach based on DWT and SVD, Int. J. Comput. Appl. 74 (13) (2013).
[31] M. Zamani, M. Movahedi, in: Machine Learning Techniques for Intrusion Detection, 2013 arXiv preprint arXiv:1312.2177.
[32] M.N.M. Adnan, et al., Content based news recommendation system based on fuzzy logic, in: 2014 International Conference on Informatics, Electronics & Vision (ICIEV), IEEE, 2014.
[33] A.E. Andaya, et al., Meat marbling scoring using image processing with fuzzy logic based classifier, Int. J. Sci. Technol. Res 8 (8) (2019) 1442–1445.
[34] C.C. Ragin, The distinctiveness of comparative, Sociological Worlds: Comparative and Historical Readings on Society (2000) 362.
[35] B. Rihoux, B. Lobe, The case for qualitative comparative analysis (QCA): adding leverage for thick cross-case comparison, The Sage handbook of case-based methods (2009) 222–242.

[36] T. Ide, P.A. Mello, QCA in international relations: a review of strengths, pitfalls, and empirical applications, Int. Stud. Rev. 24 (1) (2022) viac008.
[37] T. Haesebrouck, The added value of multi-value qualitative comparative analysis, in: Forum: Qualitative Social Research, Freie Universität Berlin, 2016.
[38] S.C.T. Barbosa, Comparative analysis applied to research on the implementation of public policy: the multi-value qualitative comparative analysis (MVQCA) method, Revista Alcance (Online) 29 (1) (2022) 35–52.
[39] C.C. Ragin, S.E. Mayer, K.A. Drass, Assessing discrimination: a Boolean approach, Am. Socio. Rev. (1984) 221–234.
[40] I.O. Pappas, A.G. Woodside, Fuzzy-set qualitative comparative analysis (fsQCA): guidelines for research practice in information systems and marketing, Int. J. Inf. Manag. 58 (2021) 102310.
[41] J. Seawright, Qualitative comparative analysis vis-a-vis regression, Stud. Comp. Int. Dev. 40 (1) (2005).
[42] Y. Liu, et al., Applying configurational analysis to IS behavioural research: a methodological alternative for modelling combinatorial complexities, Inf. Syst. J. 27 (1) (2017) 59–89.
[43] I.O. Pappas, User experience in personalized online shopping: a fuzzy-set analysis, Eur. J. Market. 52 (7/8) (2018) 1679–1703.
[44] A. Sukhov, M. Friman, L.E. Olsson, Unlocking potential: an integrated approach using PLS-SEM, NCA, and fsQCA for informed decision making, J. Retailing Consum. Serv. 74 (2023) 103424.
[45] J. Ho, C. Plewa, V.N. Lu, Examining strategic orientation complementarity using multiple regression analysis and fuzzy set QCA, J. Bus. Res. 69 (6) (2016) 2199–2205.
[46] C.Q. Schneider, B. Grofman, It might look like a regression equation… but its not! An intuitive approach to the presentation of QCA and fs/QCA results, in: Conference on "Comparative Politics: Empirical Applications of Methodological Innovations", , Sophia University, Tokyo (Japan), 2006.
[47] N. Roig-Tierno, T.F. Gonzalez-Cruz, J. Llopis-Martinez, An overview of qualitative comparative analysis: a bibliometric analysis, Journal of Innovation & Knowledge 2 (1) (2017) 15–23.
[48] C.C. Ragin, How to lure analytic social science out of the doldrums: some lessons from comparative research, Int. Sociol. 21 (5) (2006) 633–646.
[49] C.C. Ragin, Using qualitative comparative analysis to study causal complexity, Health Serv. Res. 34 (5 Pt 2) (1999) 1225.
[50] C.C. Ragin, P. Pennings, Fuzzy sets and social research, Socio. Methods Res. 33 (4) (2005) 423–430.
[51] C.C. Ragin, Redesigning Social Inquiry: Fuzzy Sets and beyond, University of Chicago Press, 2009.
[52] D.B. Pedersen, Integrating social sciences and humanities in interdisciplinary research, Palgrave Communications 2 (1) (2016) 1–7.
[53] Y. Park, O.A. El Sawy, P. Fiss, The role of business intelligence and communication technologies in organizational agility: a configurational approach, J. Assoc. Inf. Syst. Online 18 (9) (2017) 1.
[54] C.C. Ragin, Fuzzy-set Social Science, University of Chicago Press, 2000.
[55] D. Colquhoun, An investigation of the false discovery rate and the misinterpretation of p-values, R. Soc. Open Sci. 1 (3) (2014) 140216.
[56] J.W. Schneider, NHST is still logically flawed, Scientometrics 115 (2018) 627–635.
[57] D. Szucs, J.P. Ioannidis, When null hypothesis significance testing is unsuitable for research: a reassessment, Front. Hum. Neurosci. 11 (2017) 390.
[58] A.-M. Krypotos, et al., A primer on Bayesian analysis for experimental psychopathologists, J. Exp. Psychopathol. 8 (2) (2017) 140–157.
[59] J. Gill, The insignificance of null hypothesis significance testing, Polit. Res. Q. 52 (3) (1999) 647–674.
[60] M. Halme, M. Korpela, Responsible innovation toward sustainable development in small and medium-sized enterprises: a resource perspective, Bus. Strat. Environ. 23 (8) (2014) 547–566.
[61] A. Marx, Towards more robust model specification in QCA results from a methodological experiment, American Sociological Association (2006) 2006. Philadelphia, PA.
[62] P.C. Fiss, D. Sharapov, L. Cronqvist, Opposites attract? Opportunities and challenges for integrating large-N QCA and econometric analysis, Polit. Res. Q. (2013) 191–198.
[63] S. Fainshmidt, et al., The Contributions of Qualitative Comparative Analysis (QCA) to International Business Research, Springer, 2020, pp. 455–466.
[64] S. Kraus, D. Ribeiro-Soriano, M. Schüssler, Fuzzy-set qualitative comparative analysis (fsQCA) in entrepreneurship and innovation research–the rise of a method, Int. Enterpren. Manag. J. 14 (2018) 15–33.
[65] A. Thiem, A. Dusa, QCA: a package for qualitative comparative analysis, Rom. Jahrb. 5 (1) (2013) 87.
[66] C.Q. Schneider, C. Wagemann, Doing justice to logical remainders in QCA: moving beyond the standard analysis, Polit. Res. Q. (2013) 211–220.
[67] E. Moschouli, et al., Factors affecting the cost performance of transport infrastructure projects, Eur. J. Transport Infrastruct. Res. 18 (4) (2018).
[68] J.B. Davidsson, P. Emmenegger, Defending the organisation, not the members: unions and the reform of job security legislation in W estern E urope, Eur. J. Polit. Res. 52 (3) (2013) 339–363.
[69] E. Thomann, J. Ege, Qualitative Comparative Analysis (QCA) in Public Administration, Oxford Research Encyclopedia of Politics, 2020.
[70] C.Q. Schneider, Two-step QCA revisited: the necessity of context conditions, Qual. Quantity 53 (2019) 1109–1126.
[71] J. Møller, S.-E. Skaaning, Set-theoretic methods in democratization research: an evaluation of their uses and contributions, Democratization 26 (1) (2019) 78–96.
[72] S.M. Lipset, The indispensability of political parties, J. Democr. 11 (2000) 48.
[73] S.M. Lipset, P. Man, The Social Bases of Politics, The Johns Hopkins UniversityPress, Baltimore, 1960.
[74] R.J. Harvey, W.D. Murry, D.T. Stamoulis, Unresolved issues in the dimensionality of the myers-briggs type indicator, Educ. Psychol. Meas. 55 (4) (1995) 535–544.
[75] A. Ozdemir, T. Altural, A comparative study of frequency ratio, weights of evidence and logistic regression methods for landslide susceptibility mapping: sultan Mountains, SW Turkey, J. Asian Earth Sci. 64 (2013) 180–197.
[76] S. Kwak, Are only p-values less than 0.05 significant? A p-value greater than 0.05 is also significant, Journal of Lipid and Atherosclerosis 12 (2) (2023) 89.
[77] D. Berg-Schlosser, L. Cronqvist, Macro-quantitative vs. macro-qualitative methods in the social sciences—an example from empirical democratic theory employing new software, Historical Social Research/Historische Sozialforschung (2005) 154–175.
[78] B. Rihoux, G. De Meur, Crisp-set qualitative comparative analysis (csQCA), Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques 51 (2009) 33–68.
[79] B. Rihoux, Configurational comparative methods (QCA and fuzzy sets): complex causation in cross-case analysis, in: Handbook of Research Methods and Applications in Political Science, Edward Elgar Publishing, 2016, pp. 383–399.
[80] S.M. Rasoolimanesh, et al., The combined use of symmetric and asymmetric approaches: partial least squares-structural equation modeling and fuzzy-set qualitative comparative analysis, Int. J. Contemp. Hospit. Manag. 33 (5) (2021) 1571–1592.
[81] S.M. Rasoolimanesh, N. Valaei, S. Rezaei, Guideline for application of fuzzy-set qualitative comparative analysis (fsQCA) in tourism and hospitality studies, in: Cutting Edge Research Methods in Hospitality and Tourism, Emerald Publishing Limited, 2023, pp. 137–156.
[82] M.M. Raza, et al., Board characteristics and firm performance: a configurational analysis, JISR management and social sciences & economics 21 (1) (2023) 69–91.
[83] S. Husain Tahir, et al., Impact of economic freedom on air pollution: configuration analysis of Asia-Pacific region, Environ. Sci. Pollut. Control Ser. 28 (35) (2021) 47932–47941.
[84] M. Baumgartner, Qualitative comparative analysis and robust sufficiency, Qual. Quantity 56 (4) (2022) 1939–1963.
[85] C.C. Ragin, Set relations in social research: evaluating their consistency and coverage, Polit. Anal. 14 (3) (2006) 291–310.
[86] T. Haesebrouck, An alternative update of the two-step QCA procedure, Qual. Quantity 53 (2019) 2765–2780.
[87] F. Veri, Coverage in fuzzy set qualitative comparative analysis (fsqca): a New fuzzy proposition for describing empirical relevance, Comp. Sociol. 17 (2) (2018) 133–158.
[88] C.Q. Schneider, Realists and idealists in QCA, Polit. Anal. 26 (2) (2018) 246–254.
[89] M. Ahmed, S.H. Tahir, G. Ahmad, Fostering inclusive growth: exploring gender dynamics and loan sizes in microfinance institutions, Annals of Social Sciences and Perspective 4 (2) (2023) 285–305.

[90] M.R. Ullah, S.H. Tahir, H. Shahzadi, From pixels to profits: configurations of digital alchemy for SMEs excellence, Journal of Development and Social Sciences 4 (3) (2023) 708–719.
[91] M. Zulfiqar, et al., Digitalized world and carbon footprints: does digitalization really matter for sustainable environment? Environ. Sci. Pollut. Control Ser. 30 (38) (2023) 88789–88802.
[92] M.R. Ullah, et al., Digital pathways to success: the transformative power of digitalization and digital capabilities on SMEs' financial performance, iRASD Journal of Economics 5 (2) (2023) 465–485.
[93] A.G. Woodside, Embracing the complexity turn in management research for modeling multiple realities, The complexity turn: Cultural, management, and marketing applications (2017) 1–19.
[94] E. Vittinghoff, et al., Predictor selection, Regression methods in biostatistics: linear, logistic, survival, and repeated measures models (2012) 395–429.
[95] G. Zhang, W. Wang, Y. Liang, Understanding the complex adoption behavior of cloud services by SMEs based on complexity theory: a fuzzy sets qualitative comparative analysis (fsQCA), Complexity 2021 (2021) 1–17.
[96] W. Olsen, Qualitative comparative analysis (QCA): a classic mixed method using theory, in: Systematic Mixed-Methods Research for Social Scientists, Springer, 2022, pp. 131–156.
[97] J. Seawright, Qualitative comparative analysis vis-à-vis regression, Stud. Comp. Int. Dev. 40 (2005) 3–26.