III. Organization of Binary Logic (Set Theory) and Trial of csQCA

III-1. Content of This Chapter

Using the dataset targeted for numerical analysis, we will verify Lipset's hypothesis in a set-theoretic manner. This verification method has already been described in the book by B. Rihoux and C. Ragin (2009) and is not original to the commentator. However, to deepen understanding, basic explanations have been added. Here, we will explain csQCA: crisp set qualitative comparative analysis, and its purpose is to organize the binary logic that forms the basis of QCA.

III-2. Set Theory, Boolean Operations, and Truth Tables

When we say, “Humans are animals,” from a set-theoretic perspective, it means that “the set of humans is within the set of animals.” Conventionally, the set representing the whole is often denoted as U (probably derived from Universe). The set of humans is denoted as A. In logic, a statement like “Humans are animals” is called a proposition. This term is used because it can be questioned whether humans are animals (true) or not (false). This proposition can be expressed in set notation as $A \sqsubseteq U$. Since there are many animals other than humans, it might be more accurate to write $A \sqsubset U$. This categorizes animals by species, but animals can also be categorized by reproductive ability. Thus, a proposition like “Animals that can produce offspring are called females” can be formed. If we denote the set of females as B, then $B \sqsubset U$. A diagram like Figure 6, which expresses such set-theoretic relationships, is called a Venn diagram.

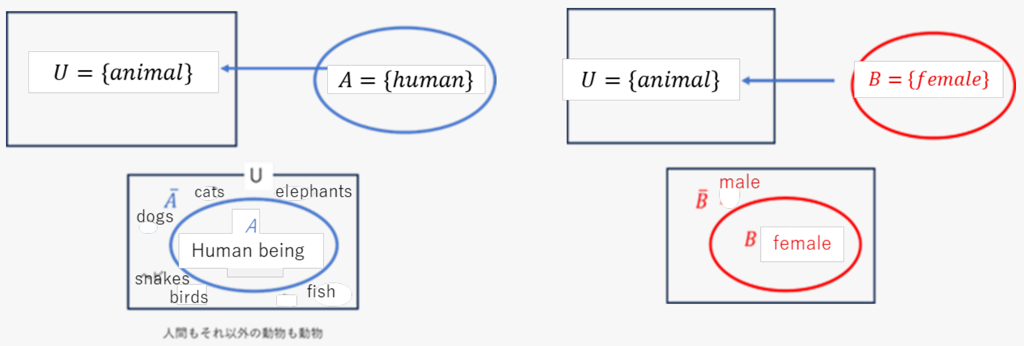


Fig.6 Venn diagrams of propositions of “human beings are animal” and Animals that can produce offspring are called females”

Propositions create subsets A and B within a set U (in this case, animals) and simultaneously create sets of things that are not A and not B. In other words, they create sets that are false with respect to the proposition. These are called the complements of A and B. Within the whole set U, the part that is not A is called the complement of A. In set notation, complements are denoted as$\tilde{A}$ or A^c and $\tilde{B}$ or B^c. When simplifying further, they are represented by lowercase letters a and b. The relationship between A and $\tilde{A}$ is called a complementary relationship.

Figure 7 illustrates the world of the animal kingdom (U) by combining Proposition A: “Humans are animals.” and Proposition B: “The animals that produce offspring are female.”

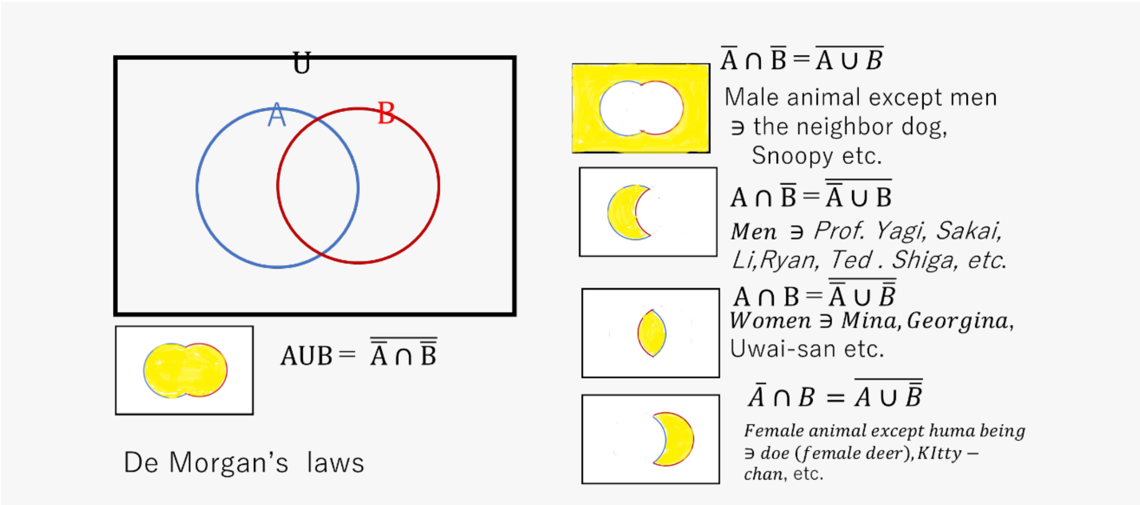


Fig 7. Structure of animal world composed from two preposition

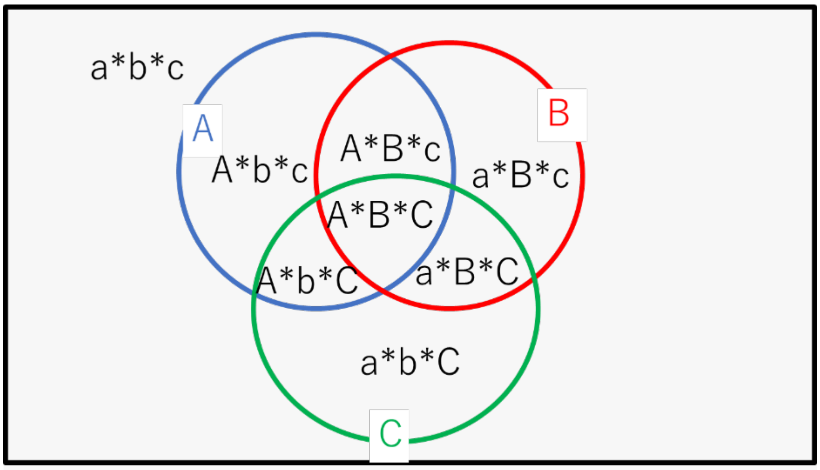


Fig 8. Venn diagram when condition A,B, and C are combined

When two sets are combined, a set that belongs to both sets simultaneously is formed. This is called a product set. Additionally, a set that belongs to either or both of the two sets can also be formed. This is called a sum set. In set operation notation, the product set is represented as $A \cap B$ and the union set is represented a. $A \cup B$. As shown on the right side of Figure 6, the entire set U (in this example, all animals) is completely filled with four subsets. In set operation notation, it can be written as follows. That is, as subsets of U, there are four intersection sets:

$$\tilde{A} \cap \tilde{B} \cup A \cap \tilde{B} \cup A \cap B \cup \tilde{A} \cap B = U$$

The sum of these four intersection sets is U. When described with two binary characteristic values such as type and gender, U has 2^2 subsets of product sets. Figure 8 shows the subsets in the form of a Venn diagram when three conditions A, B, and C are combined. Here, the product sets are described using Boolean operation notation, and the complement sets are represented in lowercase. When three characteristics are combined, U has 2^3 subsets of product sets. When described with k characteristics, U is filled with 2^k subsets of intersection sets (different combination of states or conditions).

When considering characteristics that distinguish members of a set, such as being human or female, as conditions or states that the animal satisfies, these conditions can be expressed in binary terms: true (1) when the condition is met and false (0) when it is not. When there are two conditions, the logical conjunction (AND) refers to whether both conditions are met. If both are met, the logical conjunction is true (1); otherwise, it is false (0). This corresponds to the product set (intersection of sets). The logical disjunction (OR) refers to whether at least one of the two conditions is met. If either condition is met, the logical disjunction is true (1); if neither is met, it is false (0). This corresponds to the sum set (union of sets).

For example, let A represent “being human” as true (1) and its negation a:“not being human” as false (0), and let B represent “being female” as true (1) and its negation b:“not being female” as false (0). The logical conjunction, “being human being and female” is true (1) when both “being human” and “being female” are true, and false (0) for all other combinations. Thus, women are true, represented by the logical conjunction as 1, while all other males and animals are false, represented as 0. The logical disjunction “ being human” or “being female” is true (1) if either “being human” or “being female” is satisfied. Therefore, the set of male animals excluding men has a logical disjunction of 0, while all others have a logical disjunction of 1. The logical conjunction is presented as $A \wedge B$, and the logical disjunction as $A \vee B$.

Logic is the condition that distinguishes the set from others and is the boundary line in a Venn diagram. If the condition distinguishing A from others is (α, β) and the condition distinguishing B from others is (β, γ), then

$$A \wedge B \rightarrow (\alpha, \beta, \gamma), \qquad A \vee B \rightarrow (\beta,)$$

Considering sets, if all elements (members) included in the set are represented by a, b, c:

If

$$A = \{a, b\}, \quad B = \{b, c\}$$

then

$$A \cap B = \{b\}, \quad A \cup B = \{a, b, c\}$$

It might be confusing, but it can be understood by thinking, “The more conditions there are, the fewer the number of suitable members.”

Boolean operations (Boolean logic) are methods of calculation for sets and logic, but they yield the same results whether applied to sets or logic. In practical terms, logical expressions can be difficult to read. Therefore, in this explanation, unless it is necessary to emphasize that it is a logical operation, the notation for sets is used.

Some Laws of Boolean Operations

Notation:

- Union(sum): +
- Intersection(product): $*$
- Complement: $\tilde{A}, A^c, a$
- Members: A={α,β,γ} α,β,γ are elements that constitute A.

A$\ni$ α. $A \not\ni \delta$ α is included in the members.is not included in members

- Empty set: ∅ a set with no members

Main Laws:

1. **Commutative Law:**
   - $A+B=B+A$
   - $A*B=B*A$
2. **Associative Law:**
   - $(A+B)+C=A+(B+C)$
   - $(A*B)*C=A*(B*C)$
3. **Distributive Law:**
   - $A*(B+C)=A*B+A*C$
4. **Identity Law:**
   - A+1=1 The union of a part and the whole is the whole
   - A*1=A The intersection of a part and the whole is the part
   - $A+A=A$
   - $A*A=A$
5. **Complement Law:**
   - $A+\tilde{A}$=1
   - $A*\tilde{A}$ =0
6. **Absorption Law:**

- $A + (A * B) = A$
- $A * (A + B) = A$

7. **De Morgan's Laws:**
   - $\widetilde{A + B} = \tilde{A} * \tilde{B}$ The complement of the union is the intersection of the complements
   - $\widetilde{A * B} = \tilde{A} + \tilde{B}$ The complement of the intersection is the union of the complements

Using these laws, if we have the product of A and B and product of A and b (the complement of B) on the left side of the equation, we can simplify the expression as follows:

$$\mathrm{A} * \mathrm{B} + \mathrm{A} * b = A(B + b) = A * 1 = A$$

Figure 9 demonstrates that when there are two conditions (sets) A and B, the union of all combinations of A, B, and their complements constitutes the entire set U, as proven by Boolean operations. The Venn diagram at the top of the figure makes it clear that the first equation of Boolean operations represents the entire set of animals, indicating that this is true. Everything included within this set is an animal, and no animals exist outside of it.

$$\bar{A} \cap \bar{B} \quad \cup \quad A \cap \bar{B} \quad \cup \quad A \cap B \quad \cup \quad \bar{A} \cap B \quad = \quad U$$

(Not A and not B) or (A and not B) or (A and B) or (not A and B) is U

Boolean operation

$$\mathrm{a} * b + A * b + A * B + a * B$$
$$= (a + A) * b + (a + A) * B$$
$$= (a + A) * (b + B) = 1$$

Fig. 9. Demonstration that whole is sum set of all product sets of single conditions.

Truth table

	condition			result
ID	A	B	C	
A*B*C	1	1	1	
A*B*c	1	1	0	
A*b*C	1	0	1	
A*b*c	1	0	0	
a*B*C	1	1	1	
a*B*c	1	1	0	
a*b*C	1	0	1	
a*b*c	0	0	0	

Result: product set and Union

	condition		result	
	A	B	product	union
A-B	1	1	1	1
A-b	1	0	0	1
a-B	0	1	0	1
a-b	0	0	0	0

Fig. 10 An example of procedure to make truth table

Next, the commentator will explain the truth table. A truth table is a table that represents whether certain binary conditions are met or not using 0 and 1. The leftmost column in figure 10 contains the code IDs. For example, in the data of European countries during the interwar period, which is the subject of this analysis, the IDs of each country are recorded. In Table 10, the Boolean expression for the intersection set is included as the ID. To the right of that, the truth values corresponding to each ID are recorded in binary (0 or 1) under the column names of the conditions. Further to the right, there is a result column that shows what happens or what happened under those conditions. In the example of interwar Europe, the data is binary, indicating whether the country maintained democracy (1) or democracy collapsed (0). As an example, when conditions A and B are the values in the condition columns, the logical conjunction and disjunction results are shown.

III-3. How to Use Truth Tables

Truth tables are very convenient once you learn how to use them. By using Excel's sorting function to rearrange the truth table, you can achieve various things. The usage of the truth table is shown in Figure 11. In the truth table shown on the left side of the figure, the condition that leads to the result R being true (1) is only the combination of conditions A∧B∧C. It can be concluded that if A∧B∧C, then the result R is obtained (A∧B∧C→ R R is a necessary condition for A∧B∧C). In the middle example, since the result R=1 is obtained with both A∧B∧C and A∧B∧c, taking their logical or gives A∧B→R = 1 . In Boolean algebra, this can be written as follows:

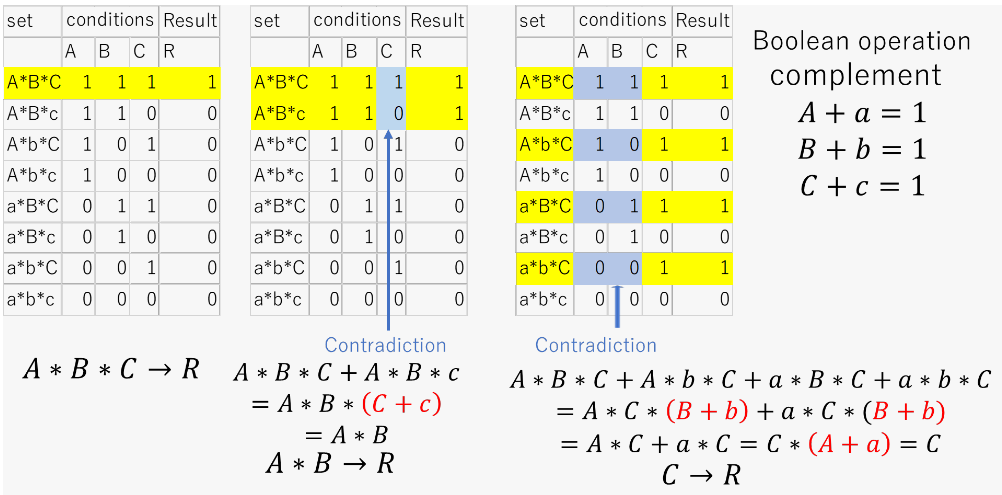
$$A \wedge B \wedge C + A \wedge B \wedge c \rightarrow R$$
$$A \wedge B \wedge (C \vee \tilde{C}) \rightarrow R$$
$$A \wedge B \rightarrow R$$

In the example on the far right, We should reach the conclusion in following route.

1. R=1 is obtained with logical disjunction of

   $(A \wedge B \wedge C) \vee (A \wedge b \vee C) \vee (a \wedge B \wedge c) \vee (a \wedge b \wedge c)$.

2. Ther is no consistency in truth value in condition A and B and there is a consistency in truth value in condition C
3. It can be judged that if C=1, necessarily R=1.

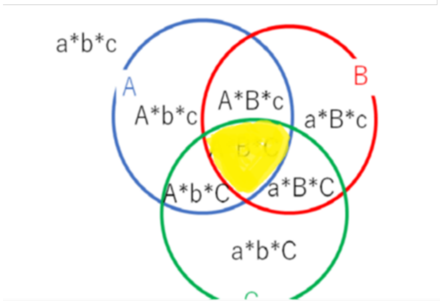


set	conditions			Result
	A	B	C	R
A*B*C	1	1	1	1
A*B*c	1	1	0	0
A*b*C	1	0	1	0
A*b*c	1	0	0	0
a*B*C	0	1	1	0
a*B*c	0	1	0	0
a*b*C	0	0	1	0
a*b*c	0	0	0	0

set	conditions			Result
	A	B	C	R
A*B*C	1	1	1	1
A*B*c	1	1	0	1
A*b*C	1	0	1	0
A*b*c	1	0	0	0
a*B*C	0	1	1	0
a*B*c	0	1	0	0
a*b*C	0	0	1	0
a*b*c	0	0	0	0

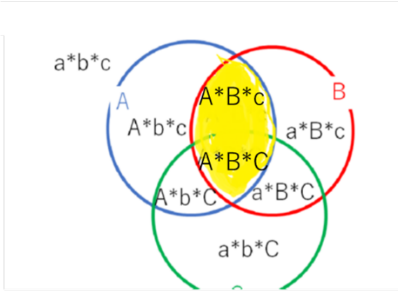
set	conditions			Result
	A	B	C	R
A*B*C	1	1	1	1
A*B*c	1	1	0	0
A*b*C	1	0	1	1
A*b*c	1	0	0	0
a*B*C	0	1	1	1
a*B*c	0	1	0	0
a*b*C	0	0	1	1
a*b*c	0	0	0	0

Fig. 11. How to use truth table



set	conditions			Result
	A	B	C	R
A*B*C	1	1	1	1
A*B*c	1	1	0	0
A*b*C	1	0	1	0
A*b*c	1	0	0	0
a*B*C	0	1	1	0
a*B*c	0	1	0	0
a*b*C	0	0	1	0
a*b*c	0	0	0	0

$$A \wedge B \wedge C \rightarrow R$$

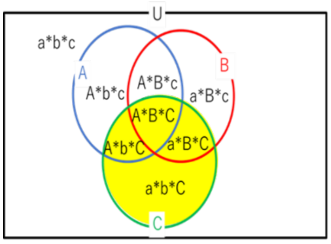


set	conditions			Result
	A	B	C	R
A*B*C	1	1	1	1
A*B*c	1	1	0	1
A*b*C	1	0	1	0
A*b*c	1	0	0	0
a*B*C	0	1	1	0
a*B*c	0	1	0	0
a*b*C	0	0	1	0
a*b*c	0	0	0	0

contradiction

$$(A \wedge B \wedge C) \vee (A \wedge B \wedge c) = A \wedge B \wedge (C \vee c) = A \wedge B$$

$$A \wedge B \rightarrow R$$



set	conditions			Result
	A	B	C	R
A*B*C	1	1	1	1
A*b*C	1	0	1	1
a*B*C	0	1	1	1
a*b*C	0	0	1	1
A*B*c	1	1	0	0
A*b*c	1	0	0	0
a*B*c	0	1	0	0
a*b*c	0	0	0	0

contradiction

$$(A \wedge B \wedge C) \vee (A \wedge b \wedge C) \vee (a \wedge B \wedge C) \vee (a \wedge b \wedge C) = A \wedge C \wedge (B \vee b) \vee a \wedge C \wedge (B \vee b) = (A \wedge C) \vee (a \wedge C) = C \wedge (A \vee a) = C$$

$$C \rightarrow R$$

Venn's diagram　　Truth table　　Boolean operation

Fig. 12. Comarison of analyses by Venn diagram, truth table and Boolean operation

It can be concluded, at first glance, that since only C is common, C=1 is enough condition for R ($C \rightarrow R$) using truth table.

In Figure 12, the analysis of the truth table from Figure 11 is shown side by side with the analysis using Venn diagrams and Boolean operations. Although the representations look different, the content of what is being done is the same. Any method that is easy to understand and minimizes errors is acceptable, so Boolean operations are not particularly meaningful at this point. However, if someone says that they were added just to give some authority, they wouldn't be entirely wrong, but they are not completely meaningless either. This will become clear later. At this point, what is more important is how to use the Excel sheet. In the example table, there are only three conditions, so the truth table has only eight rows. When there are more conditions, it becomes difficult to find the rows where R = 1. In such cases, using Excel's sorting feature to prioritize the R column makes the task easier and reduces errors. There is also a QCR package for R. I took a quick look, but it seemed that the meanings of the individual functions were not clearly defined. With some technique and knowledge, QCR, including fsQCR, can be executed in Excel, so to understand it properly, it is advisable to try executing QCR in Excel once.

III-4. Trial of csQCA Using Lipset's Theory Verification as a Subject

csQCA is an analysis of binary data expressed in 0-1. Combination of explanatory variables is used to explain the dependent variable similarly to multiple regression analysis. Both the explanatory variables and the dependent variable are binary, and the results are also binary, indicating whether they can be explained or not. The data is strictly (or crisply) written in binary, and the results are also binary, hence the name "crisp set QCA." Not only the dataset but also the conclusions are crisp. The specific procedure for csQCA has already been explained in the use of the truth table. Here, we will verify Lipset's hypothesis, which was the subject of numerical analysis in the previous chapter. In csQCA, the data (Table1) must be binary. Therefore, thresholds are set for each data item to binarize them. Figure 13 shows the first step of csQCA: binarization and sorting. The data is from interwar Europe, which was the subject of analysis in Chapter 1. The set thresholds are A: wealth 600, B: urbanization 50.0, C: education 75.0, D: industrialization 30.0, E: political stability 9.9 (government turnover), R: maintenance of democracy 0.

The table on the left of figure 13 is the binarized dataset, and by using Excel's sorting function to prioritize the columns from E to A in descending order, the table on the right is obtained. Countries with the same conditions in the truth table are grouped together and color-coded. From the top, the group with 11111 includes Belgium, Czechoslovakia, the Netherlands, and the United Kingdom, which are countries that

Binarization

	conditions					result
Case Id	A	B	C	D	E	F
AUT	1	0	1	1	0	0
BEL	1	1	1	1	1	1
CZE	1	1	1	1	1	1
EST	0	0	1	0	1	0
FIN	1	0	1	0	1	1
FRA	1	0	1	1	1	1
GER	1	1	1	1	0	0
GRC	0	0	0	0	0	0
HUN	0	0	1	0	0	0
IRL	1	0	1	0	1	1
ITA	0	0	0	0	1	0
NLD	1	1	1	1	1	1
POL	0	0	1	0	0	0
PRT	0	0	0	0	0	0
ROU	0	0	0	0	1	0
ESP	0	0	0	0	0	0
SWE	1	0	1	1	1	1
UK	1	1	1	1	1	1

Sorting →

	conditions					result
Case Id	A	B	C	D	E	F
BEL	1	1	1	1	1	1
CZE	1	1	1	1	1	1
NLD	1	1	1	1	1	1
UK	1	1	1	1	1	1
GER	1	1	1	1	0	0
FRA	1	0	1	1	1	1
SWE	1	0	1	1	1	1
AUT	1	0	1	1	0	0
FIN	1	0	1	0	1	1
IRL	1	0	1	0	1	1
EST	0	0	1	0	1	0
HUN	0	0	1	0	0	0
POL	0	0	1	0	0	0
ITA	0	0	0	0	1	0
ROU	0	0	0	0	1	0
GRC	0	0	0	0	0	0
PRT	0	0	0	0	0	0
ESP	0	0	0	0	0	0

Fig. 13. First step of csQCA: binarization and sorting

maintained democracy. The second group, 11110, includes Germany, a country where democracy collapsed. Following are 10111: France, Sweden; 10110: Austria; 10101: Finland, Ireland; 00110: Estonia; 00100: Hungary, Poland; 00001: Italy, Romania; 00000: Greece, Portugal, Spain. It is noteworthy that no group includes both countries that maintained democracy and those where democracy collapsed. Theoretically, there are 2^5 = 32 combinations of conditions. Table 10 includes the names of each country for all combinations in the truth table. The consistency in Table 10 is the proportion of countries with same outcomes. If there is low consistency in the results, that combination of the conditions cannot be used as condition to explain the results. The Ø in the table represents an empty set, meaning there were no countries belonging to that condition. In logical terms, an empty set is called a logical remainder. Of the theoretically possible combinations, only nine cases were observed. Of course, since there are only data for 18 countries, this is unavoidable. However, it seems that Denmark, Switzerland, Norway, etc., should be included in the analysis. It is interesting to see whether consistency is maintained even with such a dataset.

Table 10. Truth table of all combination of conditions with countries belonging the condition

Truth table

Consistency $c = \frac{number\ of\ F = 1\ or\ F = 0}{number\ of\ countries}$

	condition					result			
set	A	B	C	D	E	F	country	n	c
A*B*C*D*E	1	1	1	1	1	1	BEL,CZE,NLD,UK	4	1.00
A*B*C*D*e	1	1	1	1	0	0	GER	1	1.00
A*B*C*d*E	1	1	1	0	1		Logical Remainder		
A*B*C*d*e	1	1	1	0	0		Logical Remainder		
A*B*c*D*E	1	1	0	1	1		Logical Remainder		
A*B*c*D*e	1	1	0	1	0		Logical Remainder		
A*B*c*d*E	1	1	0	0	1		Logical Remainder		
A*B*c*d*e	1	1	0	0	0		Logical Remainder		
A*b*C*D*E	1	0	1	1	1	1	FRA,SWE	2	1.00
A*b*C*D*e	1	0	1	1	0	0	AUT	1	1.00
A*b*C*d*E	1	0	1	0	1	1	FIN,IRL	2	1.00
A*b*C*d*e	1	0	1	0	0		Logical Remainder		
A*b*c*D*E	1	0	0	1	1		Logical Remainder		
A*b*c*D*e	1	0	0	1	0		Logical Remainder		
A*b*c*d*E	1	0	0	0	1		Logical Remainder		
A*b*c*d*e	1	0	0	0	0		Logical Remainder		

a*B*C*D*E	0	1	1	1	1		Logical Remainder		
a*B*C*D*e	0	1	1	1	0		Logical Remainder		
a*B*C*d*E	0	1	1	0	1		Logical Remainder		
a*B*C*d*e	0	1	1	0	0		Logical Remainder		
a*B*c*D*E	0	1	0	1	1		Logical Remainder		
a*B*c*D*e	0	1	0	1	0		Logical Remainder		
a*B*c*d*E	0	1	0	0	1		Logical Remainder		
a*B*c*d*e	0	1	0	0	0		Logical Remainder		
a*b*C*D*E	0	0	1	1	1		Logical Remainder		
a*b*C*D*e	0	0	1	1	0		Logical Remainder		
a*b*C*d*E	0	0	1	0	1	0	EST	1	1.00
a*b*C*d*e	0	0	1	0	0	0	HUN,POL	2	1.00
a*b*c*D*E	0	0	0	1	1		Logical Remainder		
a*b*c*D*e	0	0	0	1	0		Logical Remainder		
a*b*c*d*E	0	0	0	0	1	0	ITA.ROU	2	1.00
a*b*c*d*e	0	0	0	0	0	0	GRC, PRT,ESP	3	1.00

minimization

set	A	B	C	D	E	F	country	n	c
A*B*C*D*E	1	1	1	1	1	1	BEL,CZE,NLD,UK	4	1.00
A*b*C*D*E	1	0	1	1	1	1	FRA,SWE	2	1.00
A*b*C*d*E	1	0	1	0	1	1	FIN,IRL	2	1.00

$$(A \wedge B \wedge C \wedge D \wedge E) \vee (A \wedge b \wedge C \wedge D \wedge E) \vee (A \wedge b \wedge C \wedge d \wedge E) \longrightarrow R$$

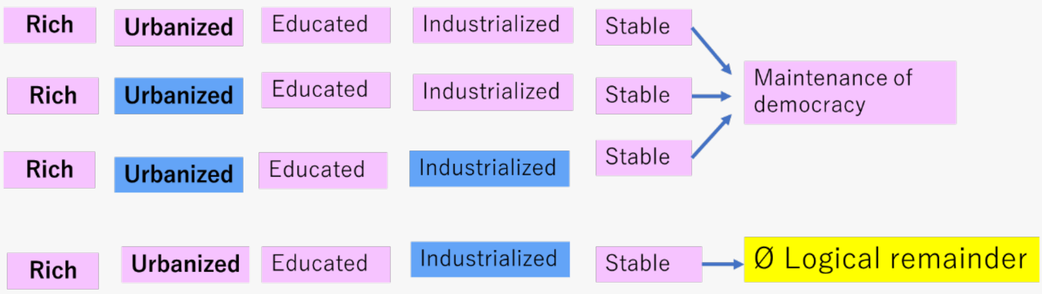


Fig. 14. Second step of csQCA: minimization of initial solution

From Table 10, extracting only the countries that maintain democracy, we can create the truth table shown in Figure 14. By listing the IDs of sets in the table, we get the following initial solution:

$$A * B * C * D * E + A * b * C * D * E + A * b * C * d * E \subseteq R$$

$$(A \wedge B \wedge C \wedge D \wedge E) \vee (A \wedge b \wedge C \wedge D \wedge E) \vee (A \wedge b \wedge C \wedge d \wedge E) \longrightarrow R$$

Verbalizing this, we get a very lengthy result: "Countries that are wealthy, urbanized, highly educated, industrialized, and politically stable maintained democracy. Even if not urbanized, countries that are wealthy, highly educated, industrialized, and politically stable also maintained democracy. Furthermore, even if not urbanized and not industrialized, countries that are wealthy, highly educated, and politically stable maintained democracy." (Flowchart in Figure 14). Reading this makes us irritated.


Immediately upon reading this, one might want to summarize it as "Wealthy, highly educated, and politically stable countries could maintain democracy" This is called minimization. The resulting conclusion is referred to as a parsimonious solution. In essence, it means to explain in a clear and concise manner using the minimum necessary words. This might be what theorizing is about. Behind this

minimization lies an analytical technique called the use of logical remainders for simplification. We try to simplify the equation obtained directly from the truth table using Boolean algebra:

$$(A \wedge B \wedge C \wedge D \wedge E) \vee (A \wedge b \wedge C \wedge D \wedge E) \vee (A \wedge b \wedge C \wedge d \wedge E) \longrightarrow R \qquad \text{i}$$

Using distributive law inversely, we can combine first and second term in left side as follow.

$$(A \wedge C \wedge D \wedge E) \wedge (B + b) \vee (A \wedge b \wedge C \wedge d \wedge E) \longrightarrow R \qquad \text{ii}$$

$$(A \wedge C \wedge D \wedge E) \vee (A \wedge b \wedge C \wedge d \wedge E) \longrightarrow R \qquad \text{iii}$$

Further minimization is impossible from the logical formula.

Fortunately, $A * B * C * d * E$ (the third row of Table 10) is empty set. This means that $A \wedge B \wedge C \wedge d \wedge E$ is logical remainder.

If

$$A \wedge B \wedge C \wedge d \wedge E \longrightarrow R$$

We can add $A \wedge B \wedge C \wedge d \wedge E$ in left side of formula iii and progress the minimization as follow.

$$(A \wedge C \wedge D \wedge E) \vee (A \wedge b \wedge C \wedge d \wedge E) \vee (A \wedge B \wedge C \wedge d \wedge E) \longrightarrow R$$

$$(A \wedge C \wedge D \wedge E) \vee \{(A \wedge b \wedge C \wedge d \wedge E) \vee (A \wedge B \wedge C \wedge d \wedge E)\} \longrightarrow R$$

$$(A \wedge C \wedge D \wedge E) \vee \{(A \wedge C \wedge d \wedge E) \wedge (B \vee b)\} \longrightarrow R$$

$$(A \wedge C \wedge D \wedge E) \vee (A \wedge C \wedge d \wedge E) \longrightarrow R$$

$$(A \wedge C \wedge E) \wedge (D \vee d) \longrightarrow R$$

$$A \wedge C \wedge E \longrightarrow R$$

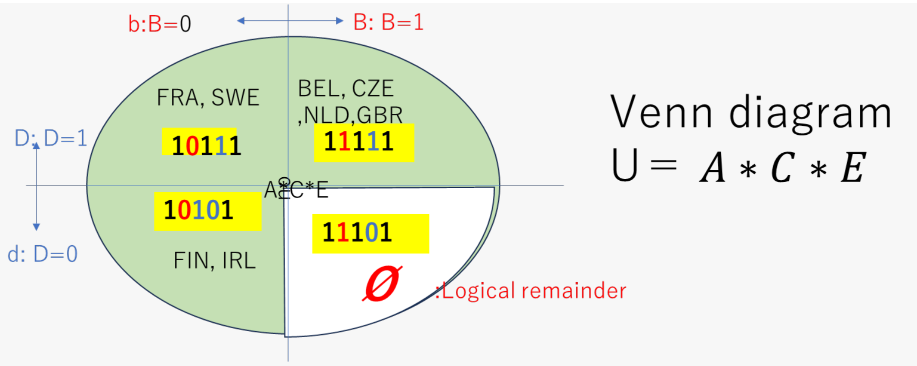


Fig. 15 Venn diagram of $U = A \cap C \cap E$

This is the technique for minimization using logical remainder. However, the explanation should not end here. The moment someone shouted, “Summarize it as ‘countries that are wealthy, highly educated, and politically stable have maintained democracy,’” they said, “Ignore things like $A * B * C * d * E$ that don’t exist, and simplify it by adding it to the logical formula since it’s obviously a country that maintains democracy.” While this might be true, it is logically reckless.

In Figure 15, a Venn diagram is drawn. The entire circular set U in the Venn diagram is $A * C * E$. This set includes elements with characteristics such as $B * D, B * d, b * D$ and $b * d$. Since there is no intersection (overlap) between them, so they are divided into four parts and illustrated. The lower right part is an empty set, so it cannot be colored the same as the others and is left white. The claim of the person angrily insisting is, “Color it green as a country that maintains democracy.”

Whether it should be colored green is a subtle issue. After all, there is no data, and it is impossible to add data now. One reference is the data of $b * d$ in the third quadrant, to the left of the white part. The fact that this is green is one reference, and since $b * d$ is green, there is some basis to think that B*d is also green. It is probably green with a fairly high probability. In this way, adding logical remainders with evidence, regardless of logical validity, seems to be a technique of analysis. As evidence, it is also acceptable to bring in some other examples.

There is one problem. In writing this explanation, I referred to Berg-Schlosser D. and De Meur (1994), which apparently states that logical remainders are easy to use and convenient, they recommend active use of logical remainder. The reason I say “apparently” is because I have only read the Japanese translation by Ishida et al. (2016), 「質的比較分析（QCA）と関連手法」. I will explain the reason why I did not read the original separately. Analysts should not cherry-pick convenient logical remainders just because it leads to conclusions that support their claims. This is against research ethics. The assertion that software should be used because it makes this easy is outrageous. It is important to consider the necessity and basis for using those logical remainders. Of course, the burden of proof lies with those who oppose the claim, so one could retort, “If you have complaints, bring counterexamples corresponding to those logical remainders.” However, it has been about 100 years, and there are hardly any similar cases to those in Europe at that time. While falsifiability is a necessary condition for scientific propositions, making claims based on practically unfalsifiable grounds is not commendable and lacks persuasiveness. In fact, simplification can be achieved without overtly using logical remainders. It is a method of evaluating logical consistency.

The method is shown in Figure 16. Among the three tables in Figure 16, the top table examines$A \wedge C \longrightarrow R$. First, $A \wedge C$ meaning selecting cases where both A and C are 1 (in the Excel sheet, prioritize A and C and sort them in descending order). Count the number of countries that meet this condition, dividing them into countries that maintain democracy and those that collapse. Using the following formula, calculate the proportion of countries that maintain democracy relative to the total number of countries:

$$c = \frac{total\ number\ of\ countries\ maintained\ democracy}{total\ numers\ of\ countries\ belonging\ the\ conditon}$$

This proportion represents the consistency of the claim when $A \wedge C \longrightarrow R$. Since csQCA is crispy and does not accept intermediate values other than 1, combinations with a consistency other than 1 are not accepted (rejected). The consistency of $A \wedge C \longrightarrow R$ is 0.80, so it is rejected. Similarly, $C \wedge E \longrightarrow R$ is also rejected with a consistency of 0.89. The only accepted combination is $A \wedge E \longrightarrow R$ with a consistency of 1. Although it might be obvious without calculation, the method of quantifying and comparing consistency is central to fsQCA (csQCA can be considered a special case of fsQCA).

sets ID	condition					Result		
	A	B	C	D	E	R	country	n
A*B*C*D*E	1	1	1	1	1	1	BEL,CZE,NLD,UK	4
A*B*C*d*E	1	1	1	0	1	Ø	Logical Remainder	
A*b*C*D*E	1	0	1	1	1	1	FRA,SWE	2
A*b*C*d*E	1	0	1	0	1	1	FIN,IRL	2
A*B*C*D*e	1	1	1	1	0	0	GER	1
A*B*C*d*e	1	1	1	0	0	Ø	Logical Remainder	
A*b*C*D*e	1	0	1	1	0	0	AUT	1
A*b*C*d*e	1	0	1	0	0	Ø	Logical Remainder	

$$A * C * E + A * C * e$$
$$= A * C * (E + e)$$
$$= A * C$$
$$\text{consistency}(A \wedge C \longrightarrow R)$$
$$c = \frac{4 + 2 + 2}{4 + 2 + 2 + 1 + 1} = 0.80$$

Reject

sets ID	condition					Result		
	A	B	C	D	E	R	country	n
A*B*C*D*E	1	1	1	1	1	1	BEL,CZE,NLD,UK	4
A*B*C*d*E	1	1	1	0	1	Ø	Logical Remainder	
A*b*C*D*E	1	0	1	1	1	1	FRA,SWE	2
A*b*C*d*E	1	0	1	0	1	1	FIN,IRL	2
a*B*C*D*E	0	1	1	1	1	Ø	Logical Remainder	
a*B*C*d*E	0	1	1	0	1	Ø	Logical Remainder	
a*b*C*D*E	0	0	1	1	1	Ø	Logical Remainder	
a*b*C*d*E	0	0	1	0	1	0	EST	1

$$A * C * E + a * C * E$$
$$= C * E * (A + a)$$
$$= C * E$$
$$\text{consistency}(C \wedge E \longrightarrow R)$$
$$c = \frac{4 + 2 + 2}{4 + 2 + 2 + 1} = 0.89$$

Reject

sets ID	condition					Result		
	A	B	C	D	E	R	country	n
A*B*C*D*E	1	1	1	1	1	1	BEL,CZE,NLD,UK	4
A*B*C*d*E	1	1	1	0	1	Ø	Logical Remainder	
A*B*c*D*E	1	1	0	1	1	Ø	Logical Remainder	
A*B*c*d*E	1	1	0	0	1	Ø	Logical Remainder	
A*b*C*D*E	1	0	1	1	1	1	FRA.SWE	2
A*b*C*d*E	1	0	1	0	1	1	FIN, FIN	2
A*b*c*D*E	1	0	0	1	1	Ø	Logical Remainder	
A*b*c*d*E	1	0	0	0	1	Ø	Logical Remainder	

$$A * C * E + A * c * E$$
$$= A * E * (C + c)$$
$$= A * E$$
$$\text{consistency}(A \wedge E \longrightarrow R)$$
$$c = \frac{4 + 2 + 2}{4 + 2 + 2} = 1.00$$

Accept

Fig 16. Minimizing of condition by consistency

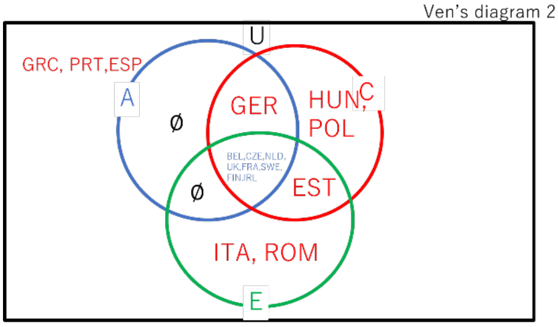


Fig. 17 Venn diagram of A, C, E

	condition					result			
set	A	B	C	D	E	R	country	n	c
A*B*C*D*e	1	1	1	1	0	0	GER	1	1.00
A*b*C*D*e	1	0	1	1	0	0	AUT	1	1.00
a*b*C*d*E	0	0	1	0	1	0	EST	1	1.00
a*b*C*d*e	0	0	1	0	0	0	HUN,POL	2	1.00
a*b*c*d*E	0	0	0	0	1	0	ITA.ROU	2	1.00
a*b*c*d*e	0	0	0	0	0	0	GRC, PRT,ESP	3	1.00

$A * C * D * e$

	condition					result			
set	A	B	C	D	E	R	country	n	c
A*B*C*D*e	1	1	1	1	0	0	GER	1	1.00
A*b*C*D*e	1	0	1	1	0	0	AUT	1	1.00

$a * b * d$

	condition					result			
set	A	B	C	D	E	res	country	n	c
a*b*C*d*E	0	0	1	0	1	0	EST	1	1.00
a*b*C*d*e	0	0	1	0	0	0	HUN,POL	2	1.00
a*b*c*d*E	0	0	0	0	1	0	ITA.ROU	2	1.00
a*b*c*d*e	0	0	0	0	0	0	GRC, PRT,ESP	3	1.00

Fig. 18. Conditions for collapse of democracy, from initial solution to intermediate solution

Figure 17 shows the relationship of sets A, C, and E in a Venn diagram. Since $A \cap C \cap E$ is a subset of $A \cap E$, when $A \wedge E \longrightarrow R$ holds, $A \wedge C \wedge E \longrightarrow R$ also holds. The explanation based on consistency is more convincing than an explanation with logical remainders. $A \cap E$ *includes* $A \cap c \cap E$ as a subset. Since $A \cap c \cap E$ is an empty set *(Ø)*, $A \wedge c \wedge E$ is a logical remainder. In fact, even in simplification by confirming consistency, there is an implicit interpretation of logical remainders. In other words, it can be overturned if there is a counterexample. It is confirmed that$A \wedge C \wedge E \longrightarrow R$ has no counterexamples in the data.

Next, we will examine the conditions that lead to the collapse of democracy. The condition is represented as R=0 or r. Returning to the truth table of conditions in Table 10, we create a list of countries where democracy has collapsed, as shown in the left table of Figure 18. This is the initial solution. These can be divided into two groups.

One group consists of countries with A=1, C=1, D=1, and E=0, represented by the condition $A \wedge C \wedge D \wedge e$, which includes Germany and Austria. The other group consists of countries with A=0, B=0, and D=0, represented by the condition $a \wedge b \wedge d$ which includes Estonia, Hungary, Poland, Italy, Romania, Greece, Portugal, and Spain. Thus, the intermediate solution is:

$$(A \wedge C \wedge D \wedge e) \vee (a \wedge b \wedge d) \longrightarrow r$$

We then simplify each of the two terms in this intermediate solution to find the most parsimonious solution. The process of simplifying $A \wedge C \wedge D \wedge e$ from aggregation to

$A \rightarrow r$ r: R=0

set	A	B	C	D	E	R	country	n
A*B*C*D*e	1	1	1	1	0	0	GER	1
A*B*C*d*e	1	1	1	0	0		Logical Remainder	
A*B*c*D*e	1	1	0	1	0		Logical Remainder	
A*B*c*d*e	1	1	0	0	0		Logical Remainder	
A*b*C*D*e	1	0	1	1	0	0	AUT	1
A*b*C*d*e	1	0	1	0	0		Logical Remainder	
A*b*c*D*e	1	0	0	1	0		Logical Remainder	
A*b*c*d*e	1	0	0	0	0		Logical Remainder	
A*B*C*D*E	1	1	1	1	1	1	BEL,CZE,NLD,UK	4
A*B*C*d*E	1	1	1	0	1		Logical Remainder	
A*B*c*D*E	1	1	0	1	1		Logical Remainder	
A*B*c*d*E	1	1	0	0	1		Logical Remainder	
A*b*C*D*E	1	0	1	1	1	1	FRA,SWE	2
A*b*C*d*E	1	0	1	0	1	1	FIN,IRL	2
A*b*c*D*E	1	0	0	1	1		Logical Remainder	
A*b*c*d*E	1	0	0	0	1		Logical Remainder	

$$c = \frac{1+1}{1+1+4+2+2} = \frac{2}{10} = 0.20$$

$C \rightarrow r$ r: R=0

set	A	B	C	D	E	R	country	n
A*B*C*D*e	1	1	1	1	0	0	GER	1
A*B*C*d*e	1	1	1	0	0		Logical Remainder	
A*b*C*D*e	1	0	1	1	0	0	AUT	1
A*b*C*d*e	1	0	1	0	0		Logical Remainder	
A*B*C*D*E	1	1	1	1	1	1	BEL,CZE,NLD,UK	4
A*B*C*d*E	1	1	1	0	1		Logical Remainder	
A*b*C*D*E	1	0	1	1	1	1	FRA,SWE	2
A*b*C*d*E	1	0	1	0	1	1	FIN,IRL	2
a*B*C*D*e	0	1	1	1	0		Logical Remainder	
a*B*C*d*e	0	1	1	0	0		Logical Remainder	
a*b*C*d*e	0	0	1	0	0	0	HUN,POL	2
a*b*C*D*e	0	0	1	1	0		Logical Remainder	
a*B*C*D*E	0	1	1	1	1		Logical Remainder	
a*B*C*d*E	0	1	1	0	1		Logical Remainder	
a*b*C*d*E	0	0	1	0	1	0	EST	1
a*b*C*D*E	0	0	1	1	1		Logical Remainder	

$$c = \frac{1+1+2+1}{1+1+4+2+2+2+1} = \frac{5}{13} = 0.38$$

$D \rightarrow r$ r: R=0

set	A	B	C	D	E	R	country	n
A*B*C*D*e	1	1	1	1	0	0	GER	1
A*b*C*D*e	1	0	1	1	0	0	AUT	1
A*B*C*D*E	1	1	1	1	1	1	BEL,CZE,NLD,UK	4
A*b*C*D*E	1	0	1	1	1	1	FRA,SWE	2
a*B*C*D*e	0	1	1	1	0		Logical Remainder	
a*b*C*D*e	0	0	1	1	0		Logical Remainder	
a*B*C*D*E	0	1	1	1	1		Logical Remainder	
a*b*C*D*E	0	0	1	1	1		Logical Remainder	
A*B*c*D*e	1	1	0	1	0		Logical Remainder	
A*b*c*D*e	1	0	0	1	0		Logical Remainder	
A*B*c*D*E	1	1	0	1	1		Logical Remainder	
A*b*c*D*E	1	0	0	1	1		Logical Remainder	
a*B*c*D*e	0	1	0	1	0		Logical Remainder	
a*b*c*D*e	0	0	0	1	0		Logical Remainder	
a*B*c*D*E	0	1	0	1	1		Logical Remainder	
a*b*c*D*E	0	0	0	1	1		Logical Remainder	

$$c = \frac{1+1}{1+1+4+2} = \frac{2}{8} = 0.25$$

$e \rightarrow r$ r: R=0

set	A	B	C	D	E	R	country	
A*B*C*D*e	1	1	1	1	0	0	GER	1
A*b*C*D*e	1	0	1	1	0	0	AUT	1
a*B*C*D*e	0	1	1	1	0		Logical Remainder	
a*b*C*D*e	0	0	1	1	0		Logical Remainder	
A*B*c*D*e	1	1	0	1	0		Logical Remainder	
A*b*c*D*e	1	0	0	1	0		Logical Remainder	
a*B*c*D*e	0	1	0	1	0		Logical Remainder	
a*b*c*D*e	0	0	0	1	0		Logical Remainder	
A*B*C*d*e	1	1	1	0	0		Logical Remainder	
A*b*C*d*e	1	0	1	0	0		Logical Remainder	
a*B*C*d*e	0	1	1	0	0		Logical Remainder	
a*b*C*d*e	0	0	1	0	0	0	HUN,POL	2
A*B*c*d*e	1	1	0	0	0		Logical Remainder	
A*b*c*d*e	1	0	0	0	0		Logical Remainder	
a*B*c*d*e	0	1	0	0	0		Logical Remainder	
a*b*c*d*e	0	0	0	0	0	0	GRC, PRT,ESP	2

$$c = \frac{1+1+2+2}{1+1+2+2} = \frac{6}{6} = 1.00$$ Accept

Fig. 19 Intermediate solution $A \wedge C \wedge D \wedge e \rightarrow r$ to final solution consistency calculation, is shown in Figure 19. We examined A→r, C→r, D→r, and e→r. Predictably, A→r, C→r, and D→r are not possible. In fact, their consistency is extremely low. On the other hand, e→r has a consistency of 1 and can be adopted as the most parsimonious solution.

Next, the process from intermediate solution to the most parsimonious solution for $a \wedge b \wedge d \longrightarrow r$ is shown in Figure 20. This development yields the most parsimonious solution:

set	A	B	C	D	E	R	country	
a*B*C*D*e	0	1	1	1	0		Logical Remainder	
a*b*C*D*e	0	0	1	1	0		Logical Remainder	
a*B*c*D*e	0	1	0	1	0		Logical Remainder	
a*b*c*D*e	0	0	0	1	0		Logical Remainder	
a*B*C*d*e	0	1	1	0	0		Logical Remainder	
a*b*C*d*e	0	0	1	0	0	0	HUN,POL	2
a*B*c*d*e	0	1	0	0	0		Logical Remainder	
a*b*c*d*e	0	0	0	0	0	0	GRC, PRT,ESP	2
a*B*C*D*E	0	1	1	1	1		Logical Remainder	
a*b*C*D*E	0	0	1	1	1		Logical Remainder	
a*B*c*D*E	0	1	0	1	1		Logical Remainder	
a*b*c*D*E	0	0	0	1	1		Logical Remainder	
a*B*C*d*E	0	1	1	0	1		Logical Remainder	
a*b*C*d*E	0	0	1	0	1	0	EST	1
a*B*c*d*E	0	1	0	0	1		Logical Remainder	
a*b*c*d*E	0	0	0	0	1	0	ITA.ROU	2

$$c = \frac{2+2+1+2}{2+2+1+2} = \frac{7}{7} = 1.00$$ Accept

set	A	B	C	D	E	R	country	n
a*b*C*D*e	0	0	1	1	0		Logical Remainder	
a*b*c*D*e	0	0	0	1	0		Logical Remainder	
a*b*C*d*e	0	0	1	0	0	0	HUN,POL	2
a*b*c*d*e	0	0	0	0	0	0	GRC, PRT,ESP	3
a*b*C*D*E	0	0	1	1	1		Logical Remainder	
a*b*c*D*E	0	0	0	1	1		Logical Remainder	
a*b*C*d*E	0	0	1	0	1	0	EST	1
a*b*c*d*E	0	0	0	0	1	0	ITA.ROU	2
A*b*C*D*e	1	0	1	1	0	0	AUT	1
A*b*c*D*e	1	0	0	1	0		Logical Remainder	
A*b*C*d*e	1	0	1	0	0		Logical Remainder	
A*b*c*d*e	1	0	0	0	0		Logical Remainder	
A*b*C*D*E	1	0	1	1	1	1	FRA,SWE	2
A*b*c*D*E	1	0	0	1	1		Logical Remainder	
A*b*C*d*E	1	0	1	0	1	1	FIN,IRL	2
A*b*c*d*E	1	0	0	0	1		Logical Remainder	

$$c = \frac{2+3+1+2+1}{2+3+1+2+1+2+2} = \frac{9}{13} = 0.69$$

$$d \rightarrow r$$

set	A	B	C	D	E	R	country	
a*b*C*d*e	0	0	1	0	0	0	HUN,POL	2
a*b*c*d*e	0	0	0	0	0	0	GRC, PRT,ESP	3
a*b*C*d*E	0	0	1	0	1	0	EST	1
a*b*c*d*E	0	0	0	0	1	0	ITA.ROU	2
A*b*C*d*e	1	0	1	0	0		Logical Remainder	
A*b*c*d*e	1	0	0	0	0		Logical Remainder	
A*b*C*d*E	1	0	1	0	1	1	FIN,IRL	2
A*b*c*d*E	1	0	0	0	1		Logical Remainder	
a*B*C*d*e	0	1	1	0	0		Logical Remainder	
a*B*c*d*e	0	1	0	0	0		Logical Remainder	
a*B*C*d*E	0	1	1	0	1		Logical Remainder	
a*B*c*d*E	0	1	0	0	1		Logical Remainder	
A*B*C*d*e	1	1	1	0	0		Logical Remainder	
A*B*c*d*e	1	1	0	0	0		Logical Remainder	
A*B*C*d*E	1	1	1	0	1		Logical Remainder	
A*B*c*d*E	1	1	0	0	1		Logical Remainder	

$$c = \frac{2+3+1+2}{2+3+1+2+2} = \frac{8}{10} = 0.80$$

Fig. 20. Intermediate solution $a \wedge b \wedge d \rightarrow r$ to final solution

$$a \longrightarrow r$$

The logical OR of the two parsimonious solutions is:

$$a \vee e \longrightarrow r$$

In everyday language, this means that if there is “poverty” or “political instability,” democracy will collapse. If either “poverty” or “political instability” exists, democracy will collapse.

Verify the plausibility of this proposition using a Venn diagram. Figure 21 is a Venn diagram of A, C, D, and e, while Figure 22 is a Venn diagram of a, b, and d. Theoretically, 4 simple conditions make of $2^4 = 16$ combinations of conditions subsets, but in Figure 21, $A * c * d *$e *and* $a * C * D * E$ are hidden behind other sets. Both $A * c * d *$e *and* $a * C * D * E$ are empty sets, and $A * c * d *$e *belongs to e* $(A * c * d * e \subset e)$. e consists of eight subsets, and seven countries (HUN, POL, AUT, GER, GRC, PRT, ESP)

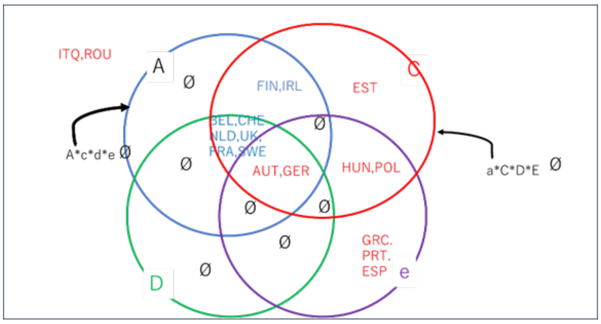


Fig. 21 Venn diagram of A, B,C, and e

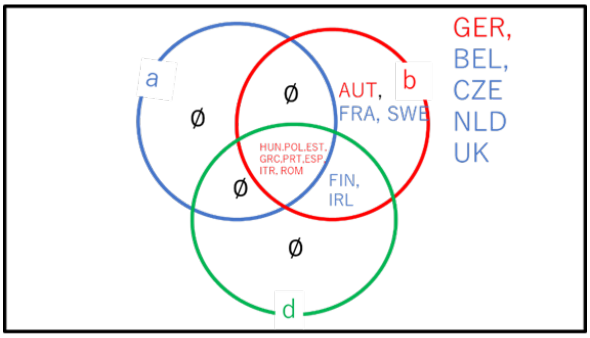


Fig. 21 Venn diagram of a, b, and d

belong to e. All of these are countries where democracy has collapsed. There is no logical problem in concluding $e \longrightarrow r$ meaning “if politics is unstable, democracy will collapse,” as the most parsimonious solution, but it should be noted that five out of the eight subsets are empty sets. Additionally, e includes seven countries where democracy has collapsed. There are ten countries where democracy has collapsed. When s concluded, it explains 70% of the countries. In contrast, concluding $e \wedge C \longrightarrow$ r *explains 40% of the total, and* e ∧ D ⟶ r explains only 20%. Even $e \wedge C \wedge D \longrightarrow R$ explains only 20%. If that is true, which part of the result to emphasize in the conclusion should be left to the analyst’s judgment, but in general cases without special analytical purposes, it is common sense to conclude what applies to a wider range. The ratio of how much of the result can be explained is called coverage. Comparing the coverage, the final solution is concluded as $e \longrightarrow r$. Looking at Figure 22, a includes eight countries where democracy has collapsed, and no countries where democracy is maintained. Most parsimonious solution is $a \longrightarrow r$.

From the perspective of consistency, this is also logically unproblematic. However, including logical residues, $a \wedge b \longrightarrow r$ *and* $a \wedge d \longrightarrow r$ also hold. Originally, $a \wedge b \wedge d \longrightarrow r$ was established, so the final conclusion is chosen from the four solutions. In this case,$a \longrightarrow r$, $a \wedge b \longrightarrow r$, $a \wedge d \longrightarrow r$, and $a \wedge b \wedge d \longrightarrow r$ are established, and the coverage is the same. Explaining phenomena with as few factors as possible might be one of the principles of science. If so, the general idea would be to conclude $a \longrightarrow r$, with a as the core condition a as the core condition, and b and d as peripheral conditions, with $a \longrightarrow r$ as the final conclusion.

In conclusion, it becomes:

$$a \vee e \rightarrow r$$

Let’s write this conclusion alongside the conclusion for countries maintaining democracy:

$$A \wedge E \rightarrow R, \qquad a \vee e \rightarrow r$$

Have you noticed? It follows De Morgan’s laws. De Morgan’s laws state:

$$\widetilde{A * B} = \tilde{A} + \tilde{B}$$

It might be a bit confusing, so let’s write it properly using set notation:

$$A \cap E \subseteq R, \qquad \tilde{A} \cup \tilde{E} \subseteq \tilde{R}$$

De Morgan's law states:

$$\tilde{A} \cup \tilde{B} = \widetilde{A \cap B}$$

In other words, the negation of $(A \land E)$ will always be the negation of (R). This means:

$$A \land E \Leftrightarrow R$$

$A \land E$ and R are equivalent, being necessary and sufficient conditions for each other.

In everyday language, if a country is prosperous and politically stable, democracy will always be maintained. Conversely, if these conditions are not met, democracy will inevitably collapse.

What I am trying to say here is not that this will always be the case in csQCA. Nor am I suggesting that you should aim for this in your analysis. Such situations are rare and feel quite unnatural. I suspect that the original data might have been constructed to produce these results. After all, there is no explanation of how the degree of democracy maintenance was evaluated. Once you start doubting, there's no end to it. However, I won't delve deeper into this. The purpose of this explanation is to ultimately demonstrate what can be done with fsQCA2 and how to do it. In the next chapter, I will explain fsQCA.

The reference book I am using explains mvQCA, which involves dividing the data into three or more categories instead of binary values for csQCA. This means that for some items in the original data, instead of categorizing them as simply large or small, they are divided into large, medium, and small, or even more detailed categories. After that, you just need to write the data divided into three categories in the truth table as 0, 1, and 2. Nothing else changes. If necessary, I can provide an explanation, but for now, I will explain fsQCA.