

# Presentation of TOSMANA

Adding Multi-Value Variables and Visual Aids to QCA

Paper prepared for presentation at the COMPASSS Launching Conference 16-17 Sept. 2003  
in Louvain-La-Neuve and Leuven

Lasse Cronqvist  
Institute for Political Science  
University of Marburg  
35032 Marburg / Germany  
[lasse@staff.uni-marburg.de](mailto:lasse@staff.uni-marburg.de)

## Abstract

In this presentation the TOSMANA (Tool for Small-N Analysis) software is described. TOSMANA is a tool for case-based comparative analysis, implementing existing techniques as *Qualitative Comparative Analysis* (QCA) as well as new approaches for macro-qualitative comparative analysis. The different parts of TOSMANA are explained and some ideas on further development are introduced.

# 1. Creating a new tool for political analysis

The method of *Qualitative Comparative Analysis* as introduced to social science by Ragin (1987) has been used to explore several areas in political and social research (for a listing see DeMeur/Rihoux 2002). One problem arising using QCA is the compulsory use of dichotomous variables. Every variable has to be transformed fitting into a scale of {0;1} indicating if a condition is true (value 1) or not (value 0). This may result in a significant loss of data precision, as cases very different from each other are assigned the same Boolean value while cases close to each other may be assigned different values.

In his book on *Fuzzy-Set Social Science*, Ragin (2000) introduces fuzzy-sets with which it is possible to transform (or directly classify) data in an interval or ordinal scale from 0 (fully out) to 1 (fully in). Thereby it is possible to classify cases as intermediate cases. Using the fuzzy-set method, sufficient and necessary conditions are then found using probabilistic methods. TOSMANA uses multi-value scales as well. But instead of applying probabilistic criteria, the idea of Boolean synthesis is kept. The idea for processing the data as multi-valued Boolean expressions was (once again) found in papers of Electronic Engineers at the University of Berkeley (MVSIS Group), which are using multi-valued synthesis for circuit reduction (see Brayton/Khadri 1999). But as there are differences in the nature of data used in electronic engineering and social science and different purposes for its applications exists, the decision was made to create an own software package. This software includes also major elements of QCA3.0 (Drass/Ragin 1992) and fs/QCA (Ragin/Drass 2002) so that no extra training has to be done to get started with this new tool for researchers familiar with QCA.

This paper is a first summary of the development of TOSMANA, which has now reached a first release with the current version 1.1. I am aware that the software is far from perfect, and that there are still some bugs. Everyone finding such an error is asked to report them to the author.<sup>1</sup>

Good software can only be created when other people are willing to test it and give a feedback of its use. Therefore I would like to thank everyone giving hints on errors and adding new ideas to the development. A special mention has to go to Dirk Berg-Schlosser, Gisèle De Meur, and Sakura Yamasaki for the many explanations and improving comments given during the past development phase.

---

<sup>1</sup> The current release of TOSMANA (version 1.1) can be retrieved at the web site <http://www.tosmana.net> where a manual is also available.

## 2. The Components of the Tosmana Project

When TOSMANA was begun, it was meant as an implementation of multi-valued variables into QCA. But as the realisation of the software went on, many ideas on possible further developments were pursued, so that at the present stage, the TOSMANA project is divided into three major parts (which are, of course, related to each other). The core of the TOSMANA applications contains the GUI (Graphic User Interface) and software structure for the synthesis of Boolean data sets for Qualitative Comparative Analysis as introduced by Ragin (1987) and the extension to multi-valued data sets (Cronqvist 2003b). Furthermore not only the synthesis of data sets is processed by the software, but also tools for editing, revising and classifying data are included. The “Thresholdsetter” is a first example. Finally, an emphasis is laid on didactic purposes. Understanding QCA is difficult for first time users. Therefore, some visual aids are included in the software, indicating the synthesis process in an intuitive way. The “Visualizer” is an example for such a tool, which can present up to four dichotomous variables in a diagram illustrating the idea of Boolean synthesis.

The programming language used to realise TOSMANA easily allows including the synthesis algorithm(s) to a new generation of web applications, which opens new possibilities for the use of an integrated online course on the subject.

### 2.1. (MV) QCA

MVQCA stands for *Multi Value Qualitative Comparative Analysis* and is, as the name suggests, an extension of *Qualitative Comparative Analysis* (QCA). It retains the idea of performing a synthesis of the data set with the result that cases with the same value of the outcome variable are explained by a parsimonious solution. The solution contains one or several *terms* which each explain a number of cases with this outcome, while no cases with a different outcome are explained by any of the terms.

While QCA only allows dichotomous variables to be processed, MVQCA also includes multi-value variables in the analysis. In fact, MVQCA is a generalisation of QCA, and each multi-value variable can also be treated as a dichotomous variable<sup>2</sup>. Therefore, data sets used with QCA can also be processed with MVQCA. Using multi-valued variables is not completely new to QCA. In *fs/qca* (Ragin / Drass 2002) the possibility exists to include a multi-value variable coded into several dichotomous dummy variables. But transforming multi-valued

---

<sup>2</sup> The binary scale consisting of the values 0 and 1 is equal to the multi-value scale with values between 0 and 1.

data sets to dichotomous data sets has been shown to be unsatisfactory.<sup>3</sup> A new algorithm to handle multi-value data set synthesis has therefore been used in TOSMANA.<sup>4</sup>

### *The method*

Multi-Value analysis requires that each variable is available on either an ordinal or a nominal scale. As an example, the variable “civil liberties”, which contains information about the degree of civil liberties in a country, can be defined as an ordinal scaled variable with possible values of  $P_{civilliberties}=\{\text{“none”, “few”, “most”, “all”}\}$ , while the variable region, which contains information on the geographical location of a country, is nominally scaled (like  $P_{RegionInEurope}=\{\text{“center”, “northern”, “southern”, “eastern”, “western”}\}$ ). This means that raw interval scaled data has to be transformed into a multi value scale before it can be used with MVQCA. To do so, a set of thresholds is defined, according to which the values are assigned. In TOSMANA, the used values have to be taken from a continuous scale of natural numbers starting with 0 (here {0,1,2,3} would be the scale for the variable *civil liberties*).

The syntax of MVQCA is identical to the syntax of QCA in most respects, but there is one important difference: As variables can have more than two possible values, the use of lowercase and uppercase letters to indicate the presence or absence of a condition cannot be used. Instead, strict set notation is employed. While  $A*B$  in QCA was indicating that A as well B are true, this has to be noted as  $A\{1\}*B\{1\}$  in MVQCA: Dichotomous variables have two possible states FALSE (0) and TRUE (1). Consider Table 1. In this imaginary data set there are four independent variables (A,B,C,D) explaining one dependent variable (O). Three of the dependent variables are of a dichotomous nature, while one variable (B) is a trichotomous variable with the possible values {0,1,2}. The independent variables for case  $C1$  can be written as  $A\{0\}B\{1\}C\{0\}D\{1\}$  while the case  $C2$  is noted as  $A\{1\}B\{2\}C\{0\}D\{1\}$ .

---

<sup>3</sup> This is mainly due to the fact, that many impossible logical combinations are created in such transformed data sets, which slow down set synthesis significantly. Imagine a traffic light: To make this example easy, the traffic light has three possible values: Red, yellow, and green. Encoding the Multi-Value variable *Light* with  $P_{Light}=\{\text{red,yellow,green}\}$  into three binary dummy variables (red,yellow,green) would result in the following table:

<u>MV Value</u>	<u>red value</u>	<u>yellow value</u>	<u>green value</u>
Red	1	0	0
Yellow	0	1	0
Green	0	0	1

Now the point is that not only the upper (real) combinations will be created, but also all other possibilities will be constructed. In reality, a combination like red=1 and green=1 will never occur in regular traffic light usage, but the combination still occurs in the data matrix when including remainder cases in the analysis. This subject has been discussed in depth by Brayton and Khatri (1999: 3)

<sup>4</sup> A brief presentation of the algorithm can be found in a power point presentation available on the Tosmana web site (see Cronqvist 2002)

Every single appearance of a symbol of the type *VariableName{Value}* is called a literal. One can say that a literal occurs if the value of a variable for a case is equal to the value of the literal. (So here, the literal B{1} occurs in the cases C1 and C3). A literal can also be seen as a Boolean function of a case being true if the literal is a subset of the logical expression of the case. As it has already been done implicitly in this chapter, literals can be combined in MVQCA in the way they are combined in QCA: The Boolean *AND* is indicated by  $\otimes$  meaning that both literals have to be true so that the new combined expression is true. Normally the sign  $\otimes$  is left out and the literals are just written next to each other. Also, the Boolean *OR* is indicated by  $\oplus$ . If two literals or combined expressions are combined by  $\oplus$ , their new, combined expression is true for a case if one or both of the original expressions are true. The normal + sign can also be used instead of  $\oplus$ . In Table 1, the combined expression A{0}+B{1} is true for the cases C1, C3, C5, and C7.

Case	A	B	C	D	O
C1	0	1	0	1	0
C2	1	2	0	1	0
C3	1	1	0	0	0
C4	1	0	1	1	0
C5	0	0	0	0	1
C6	1	0	0	0	1
C7	0	2	1	1	1
C8	1	2	1	1	1

**Table 1: Example multi-value set.**

The purpose of MVQCA is the same as that of QCA. The main goal is to find a minimal solution (combined expression) representing all cases having a specified outcome while not explaining any of the other cases (=the Boolean function is FALSE for all cases with another outcome<sup>5</sup>). For the cases with the outcome 1, B{0}C{0}+ B{2}C{1} is such a minimal solution. Similarly to QCA, logical remainders – i.e. logical combinations of independent variables that do not exist in the empirical cases observed – can be included in the analysis or not. In this simple example they are included in the synthesis of the cases with the outcome 1.<sup>6</sup>

It must be mentioned that multi-value scales should not be used carelessly. Using too many fine-graded scales will singularize cases, which makes a minimization pointless. But it will be shown in the example below (and as I have shown elsewhere, Cronqvist 2003b), already the

<sup>5</sup> The scale of the dependent variable may also be of a multi-value type.

<sup>6</sup> A more detailed introduction of MVQCA can be found in Cronqvist 2003b.

exchange of one dichotomous variable by a multi-valued scale with a few (in this case three) values can improve the result of the minimization if the thresholds are set with care.

## **2.2. Additional Tools**

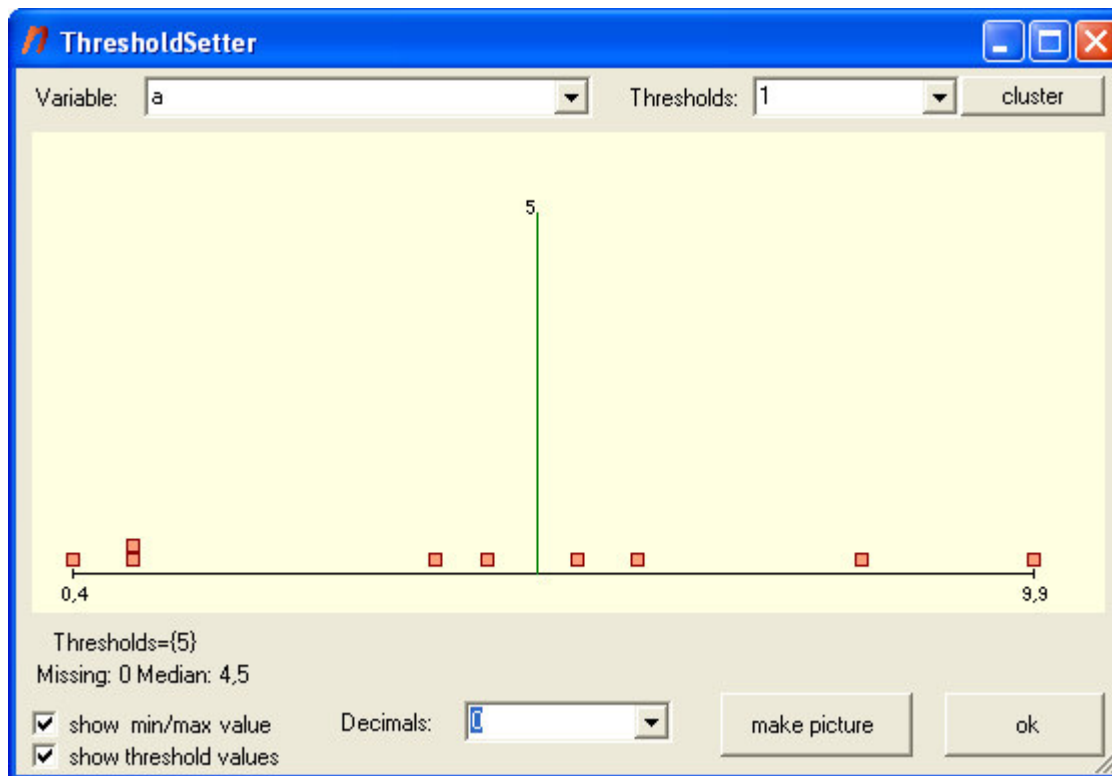
To be able to edit data sets inside the TOSMANA program some effort has been made to make this possible in a convenient way, but there are still a couple of minor bugs to be fixed, and the editing interface still needs improvements. To make migrating to TOSMANA easy, a filter to import CSV files created with Excel has been written as well as a data importer from the SAV format used by SPSS. Also, it is possible to export data sets created with TOSMANA to a plain text file, which can be read by most spreadsheet programs, while TOSMANA saves the data sets for its own use as XML-files.

### *The Thresholdsetter*

As stated above, one main problem using QCA is setting the threshold to transform interval-scaled data into dichotomous or ordinal data. Using a good threshold is essential for the reliability and the theoretical applicability of the solutions found using QCA or MVQCA. Therefore, to facilitate the selection of thresholds several additional aids are included in TOSMANA. The first visual aid is the “Thresholdsetter”. This tool, which can be accessed from the main editor window, allows to manipulate the thresholds within a variable while viewing the data distribution.<sup>7</sup> While the selection, in the end, also has to be made from a theoretical point of view, and should be done very carefully, the thresholdsetter is meant to show the user where the threshold(s) has/have been set at a bad place: Imagine a variable with the values {0.4, 1.0,4.0,4.5,5.4,6.0,8.2,9.9}. As the values are all between 0 and 10, it would seem obvious to place the threshold at the value 5.0. But this would “cut” between the values 4.5 and 5.4 and would separate two very near data points. If this variable is revised with the Thresholdsetter (see Figure 1), this can be discovered and the threshold(s) can be set in another way.

---

<sup>7</sup> Please refer to the TOSMANA User Manual for an instruction on how to use the software (see Cronqvist 2003a).



**Figure 1: The Thresholdsetter included in the TOSMANA Software**

It must be mentioned here that all data points shown in the Thresholdsetter are of equal colour and shape, and that there is no difference between the cases with different values on the outcome variable. This ensures that the outcome value is not considered when placing the thresholds.

In addition, it is possible to perform a simple cluster analysis to classify the values into a limited number of categories. This is based on a very simple implementation for the time being, and one of the next enhancements of TOSMANA will be an advisory tool, which will assist the user when deciding how fine a subset should be defined. This seems necessary, because it is difficult to decide in advance how fine the division has to be: If too few thresholds are used, many contradictions may occur. Conversely, if too many thresholds are used, cases are singularized and no useful synthesis can be made.

### *Visualizer*

To allow easy visualisations of a Boolean data set, a “Visualizer” is provided with TOSMANA, which can draw a data set containing up to four variables. The drawing created can be displayed in colours or black-and-white. Within the drawing, a selected solution can be highlighted. In this way, students new to the subject can gain a first “feeling” on how data are

represented within Boolean analysis, and how the reduction is achieved, when two “nearby sub-sets” are merged.

The graphic created with the “Visualizer” can be saved to a bitmap file. Some further options allow the case labels to be written in the diagram, or to show variable labels in varying lengths.

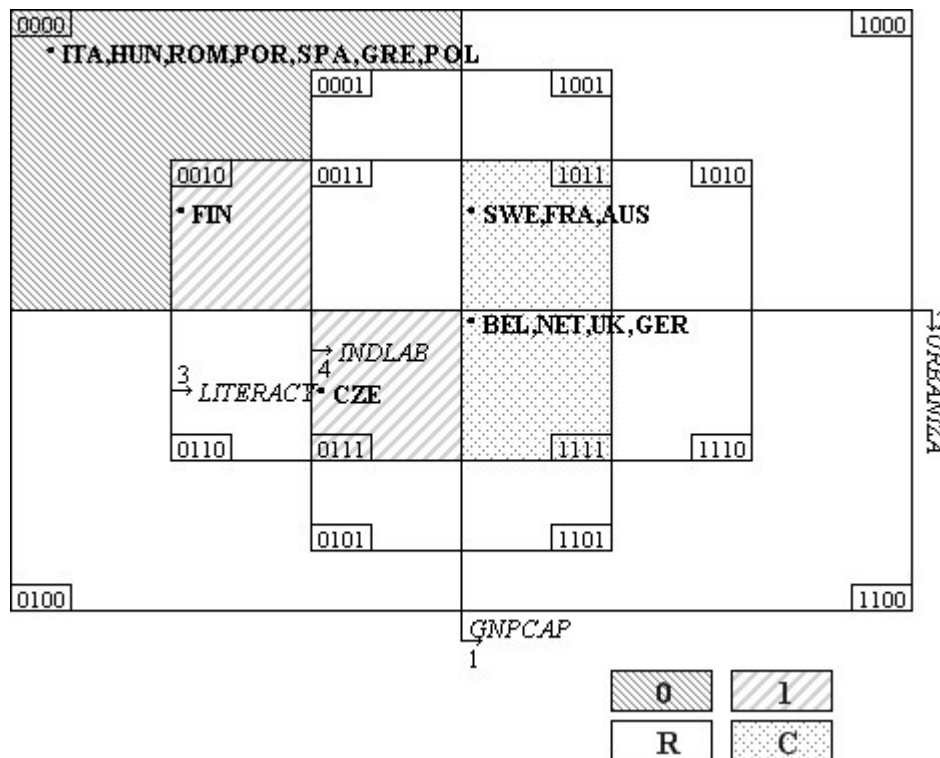


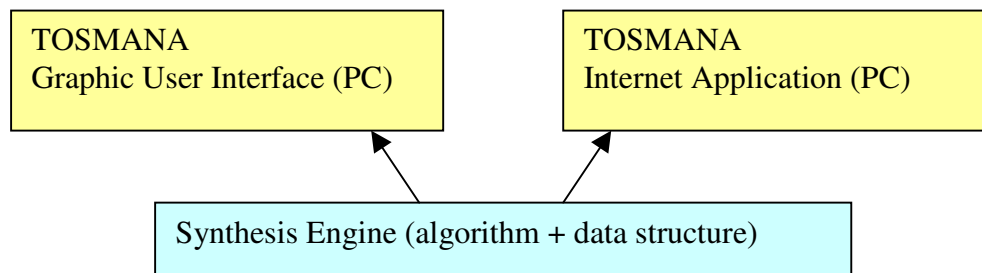
Figure 2: Distribution of cases shown by the Visualizer

### 2.3. Using the full possibilities of technical development – Placing TOSMANA Online

TOSMANA is realised in the new computer language C# (C-Sharp) with use of the .NET Framework. In this way software development has not only become more comfortable and efficient, but it also opens up completely new possibilities in the way QCA can be presented and taught. With the newly developed .NET Framework, Microsoft also realised a new web language (ASP.NET) for their server software, which allows code written for desktop computers to be included in online applications<sup>8</sup>. For TOSMANA, this means that the code for the synthesis of data sets can also be applied to online applications without changes, the computer-user interface only has to be adapted to the web environment.

<sup>8</sup> A technical but still understandable introduction to the .NET Framework can be found on the Microsoft website at [http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dnguinnet/html/drguinnet0\\_update.asp](http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dnguinnet/html/drguinnet0_update.asp) .





**Figure 3: The code for the synthesis of the algorithm and the data structure is used by both the online application and the desktop program**

This allows the analysis of data sets on every computer with internet access, as the software is entirely run on the server and the client computer is only used for the human-computer interface part, so that no software has to be installed on the PC of the user.<sup>9</sup> This can encourage non-QCA users to test TOSMANA, as no software has to be installed and so its use becomes more convenient.

Furthermore, and more importantly, teaching QCA can be enriched by direct hands-on learning: An introductory course for QCA can be placed on the web and the examples used there can be directly employed by the students within the same web site. As the design of the web interface is not very much different from the design of the window design used in TOSMANA, the users should have no major problems understanding the desktop version of TOSMANA.

Currently, this is the latest part of the TOSMANA project, and only the interface for Boolean QCA has been adapted for online use and been tested with some data sets, but it is already possible to use own data sets created with TOSMANA with this online tool. The tool can be tested on the web site <http://online-qca.tosmana.net>. Further developments will also allow the usage of multi-value scales as well as improved data handling.

---

<sup>9</sup> It has to be mentioned that there still are restrictions for using TOSMANA in this way. The first is, that capacity is more limited when an internet server is used, so that only “small” data sets (with a maximum of six to eight variables) can be used online. Secondly, ideally every browser should be able to interact with the TOSMANA server, but currently it seems that only the Internet Explorer from version 6 is showing the interface without problems.

### 3. Using the possibilities of Tosmana – An Example

In one of the first published applications of QCA in political science, Dirk Berg-Schlosser and Gisèle De Meur (1994) tested several major hypotheses on the conditions of the breakdown or survival of democratic political systems in Interwar Europe. One of the hypotheses tested was the influence of the four socio-economic variables discussed by Lipset (wealth, industrialization, education, and urbanization) in his book *Political Man* (1963). According to his proposition that a high level of socio-economic development (“modernisation”) enhances the chances of a sustainable democracy (Berg-Schlosser/De Meur 1994: 255), the following four variables are used to express this. *Wealth* is measured by the gross national product per capita (ca. 1930). *Urbanization* is measured as the portion of the population living in towns with more than 20000 inhabitants. *Industrialization* is measured as the share of the industrial labour force (incl. mining) of the active population. The raw data set is presented in Table 2.<sup>10</sup>

CASEID	GNPCAP	URBANIZA	LITERACY	INDLAB	Outcome
AUS	720	33,4	98	33,4	0
BEL	1098	60,5	94,4	48,9	1
CZE	586	69	95,9	37,4	1
FIN	590	22	99,1	22	1
FRA	983	21,2	96,2	34,8	1
GER	795	56,5	98	40,4	0
GRE	390	31,1	59,2	28,1	0
HUN	424	36,3	85	21,6	0
ITA	517	31,4	72,1	29,6	0
NET	1008	78,8	99,9	39,3	1
POL	350	37	76,9	11,2	0
POR	320	15,3	38	23,1	0
ROM	331	21,9	61,8	12,2	0
SPA	367	43	55,6	25,5	0
SWE	897	34	99,9	32,3	1
UK	1038	74	99,9	49,9	1

**Table 2: Lipset’s Indicators. Data taken from Berg-Schlosser (2002).**

In their article, Berg-Schlosser and De Meur then set the threshold values, which are used to categorise the data into four dichotomised variables, so that they “seem to be in line with Lipset’s basic arguments” (Berg-Schlosser / De Meur 1994: 256). The threshold for the *gross national product per capita* is set to \$600; the threshold for *urbanization* is set to 50%; the threshold for *literacy* is set to 75% and finally the threshold for *industrialization* is set to 30%. Using these thresholds the following dichotomized data was created (see Table 3).

<sup>10</sup> In this paper a revised data set (see Berg-Schlosser 2002) is used, so that slight changes to the calculations in Berg-Schlosser / De Meur (1994) appear.

CASEID	GNPCAP	URBANIZA	LITERACY	INDLAB	Outcome
SWE	1	0	1	1	1
FIN	0	0	1	0	1
BEL	1	1	1	1	1
NET	1	1	1	1	1
FRA	1	0	1	1	1
UK	1	1	1	1	1
CZE	0	1	1	1	1
AUS	1	0	1	1	0
GER	1	1	1	1	0
ITA	0	0	0	0	0
HUN	0	0	1	0	0
ROM	0	0	0	0	0
POR	0	0	0	0	0
SPA	0	0	0	0	0
GRE	0	0	0	0	0
POL	0	0	1	0	0

**Table 3: Data from Table 2 as dichotomous variables**

The Truth-table for this data set is given in Table 4.

CASEID	GNPCAP	URBANIZA	LITERACY	INDLAB	Outcome
SWE,FRA,AUS	1	0	1	1	1
FIN,HUN,POL	0	0	1	0	0
BEL,NET,UK,GER	1	1	1	1	1
CZE	0	1	1	1	1
ITA,ROM,POR,SPA,GRE	0	0	0	0	0

**Table 4: Truth table: Lipset's Indicators**

If a Boolean synthesis to the cases with the outcome 0 (excluding logical remainders) is performed, the following solution is found:

$$(S1a) \text{GNPCAP}\{0\}\text{URBANIZA}\{0\}\text{LITERACY}\{0\}\text{INDLAB}\{0\}$$

$$(\text{ITA,ROM,POR,SPA,GRE})$$

Five cases are explained by this solution.

If the logical remainders are included, the following solution is found:

$$(S1b) \text{LITERACY}\{0\}$$

$$(\text{ITA,ROM,POR,SPA,GRE})$$

For the outcome 1, the following solution is found explaining only one case (excluding logical remainders):

$$(S2) \text{GNPCAP}\{0\}\text{URBANIZA}\{1\}\text{LITERACY}\{1\}\text{INDLAB}\{1\}$$

$$(\text{CZE})$$

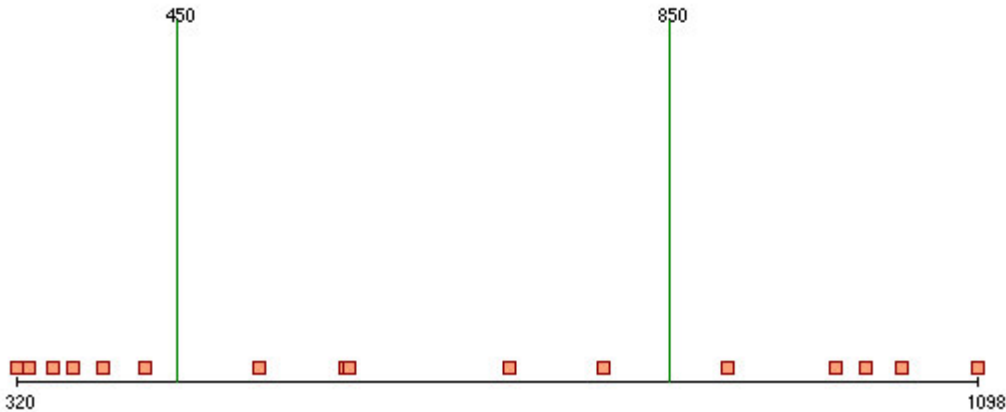
For both outcome values, only one logical configuration represents the outcome, so that without including logical remainders no further synthesis can be performed.

In this Boolean table many contradictions occur. These are combinations in which cases have the same values for the independent variables (causes), but show different outcomes. This indicates that the variables used with these thresholds cannot explain the different outcomes, as the independent variables are equal, but the outcome is different. When this occurs, it can mean that other variables must be found to explain the phenomenon. But it can also be the result of badly chosen thresholds. In this example, seven of the fifteen cases are in contradictory configurations, which means that almost half of the cases are not covered by the solutions found.

In his paper contrasting multiple macro-qualitative and macro-quantitative methods, Berg-Schlosser (2002) argues that this high number of contradictions might be due to the “relatively crude dichotomization procedure” (Berg-Schlosser 2002: 13), which accounts for a loss of precision. He therefore decided to split the *gnp/cap* variable into four parts using three thresholds at {400,600,800} and thus creating four dichotomous variables representing this new classification. As a result, he obtained a truth table with only two contradictory cases.

We will now review the *gnp/cap* variable using the TOSMANA software, and we see that using a single threshold at \$600 is an unfortunate choice, as the data points can only be separated with a great loss of data precision.<sup>11</sup>

Using a simple cluster analysis with the Thresholdsetter provided by TOSMANA, we can cut the variable into three subsets which seems to be a reasonable division into countries with low, middle, and high values of *gnp/cap* (see Figure 4).



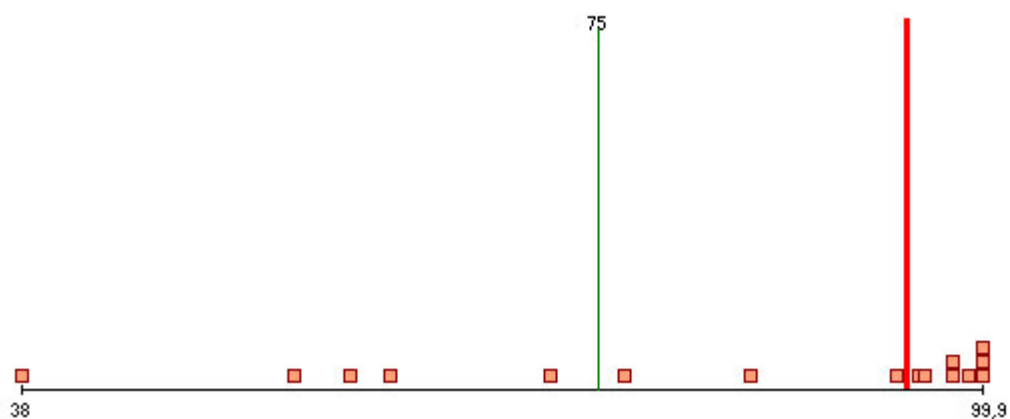
**Figure 4: The variable *gnp/cap* divided into three subsets with the threshold values of 450 and 850**

This divides the variable into three parts. Two of them contain data points relatively close to each other. The five “low” values are quite similar as well as the five “high” values. The cases in between are not really closely related, but it seems more appropriate to include them into one group than dividing them further, as the scale would become too fine-grained.

We will use these subsets for a MVQCA synthesis below, but first the other variables are reviewed. For the variable *literacy*, the choice of the threshold 75% also creates a problem.

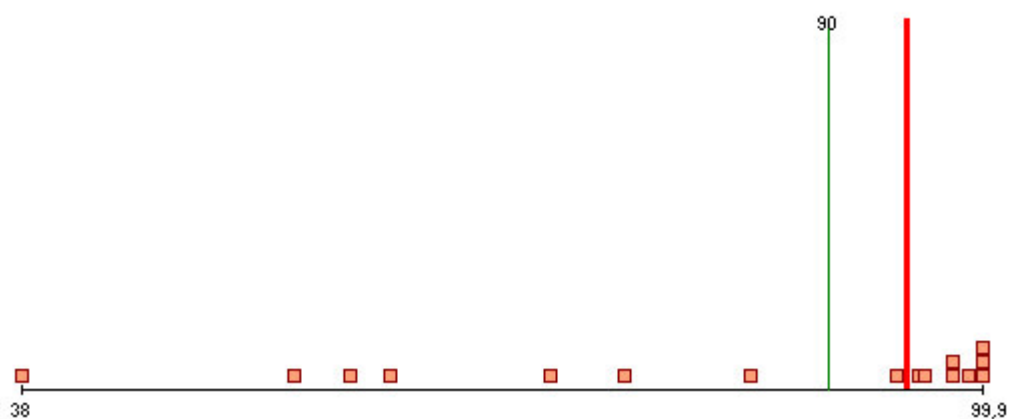
---

<sup>11</sup> This is mainly due to the fact that the values are spread on a scale between 320 and 1098 and there is no real visible division into two parts.



**Figure 5: Distribution of literacy values with a threshold of 75% and the mean of 95.15%**

Here, Italy is assigned the Boolean value of 0 with an original value of 72,1%, while Poland with a value of 76,9% is taken as highly literate. In fact, viewing the data shows a division into two parts: One half having a very high literacy rate (above 90%) and the other half below.



**Figure 6: Distribution of literacy values with a threshold of 90%.**

This choice of threshold has, of course, to be checked for its theoretical justification (as every threshold should be), but there seems to be no artificial cut between the data points any more. It also can be considered further, whether the countries with low *literacy* should be divided into subsets of *very low* and *low*.

Using these thresholds provides a much clearer matrix:

CASEID	GNPCAP	URBANIZA	LITERACY	INDLAB	Outcome
BEL,UK,NET	2	1	1	1	1
FRA,SWE	2	0	1	1	1
GER,CZE	1	1	1	1	C
AUS	1	0	1	1	0
FIN	1	0	1	0	1
ITA	1	0	0	0	0
HUN,GRE,SPA,POL,ROM,POR	0	0	0	0	0

**Table 5: Data matrix with multi-value scales. Thresholds: GNPCAP: {450,850}, URBANIZA: {50%}, LITERACY: {90%}, INDLAB: {30}**

In the dichotomised data set there were seven cases with contradictory configurations, now there is only one contradictory configuration representing two cases. For the outcome 0, this can be minimized to the solution (using logical remainders):

$$(S3) \text{ LITERACY}\{0\} + \text{GNPCAP}\{1\}\text{URBANIZA}\{0\}\text{INDLAB}\{1\} \\ (\text{ITA+HUN,GRE,SPA,POL,ROM,POR}) + (\text{AUS})$$

This solution is similar to the one calculated with dichotomous scales (S1b) but more cases are covered by this solution: Due to the change of thresholds, more cases with the outcome 0 have been categorized as cases with a *literacy* value of 0, while *Austria* is no longer contradicting another case, but is placed in a logical configuration of its own due to the changed *gnp/cap* variable.

For the outcome 1 the change is even more important. The solution found is now

$$(S4) \text{ GNPCAP}\{2\} + \text{LITERACY}\{1\}\text{INDLAB}\{0\} \\ (\text{BEL,UK,NET+FRA,SWE}) + (\text{FIN})$$

Due to the new classification of the *gnp/cap* variable, five cases with the outcome 1 have been identified as cases with a very high value on this variable, and Finland is identified as a special case with a medium level of *gnp/cap* low industrialization, but high literacy.

These results thus test and differentiate Lipset's more sweeping proposition that economic wealth improves the chance for a stable democracy, at least for the cases and the period considered.

One possibility provided by TOSMANA is to try different levels of the thresholds for a variable. By this, a "relative factor" can be incorporated in the analysis: In a global perspective, Portugal for example has a high gross national product per capita (relative to most African states), but if an analysis of West-European states is made, this information seems worthless, as all these states have a high gnp per capita. Thus, if the *gnp/cap* variable is used in a European context, a different threshold may be chosen to differentiate among these cases to identify some less "consolidated" democracies in Lipset's sense.

## **4. Conclusion**

In TOSMANA, a number of new ideas have been introduced to make the use of QCA more diversified and productive. Surely, there are still many meaningful expansions possible. Further ideas are always welcome. With the inclusion of multi-valued variables one of the main problems of QCA, the strict use of dichotomous variables, has been addressed. Although many contradictions can disappear using multi-valued scales and data can be represented in subsets closer to reality, there are still a number of questions surrounding the use of these variables. The main challenge is to find a suitable set of thresholds, which divides a variable into subsets in a way which makes these neither appear as purely artificial groupings nor making the number too large to allow a meaningful synthesis. Some of the additional tools incorporated into TOSMANA, as the example given above has demonstrated, allows for such more flexible and transparent procedures and a better critical and more differentiated “dialogue between theory and data”.

## 5. References

- Berg-Schlosser, Dirk and Gisèle De Meur (1994): "Conditions of Democracy in Interwar Europe: A Boolean Test of Major Hypotheses" In *Comparative Politics*, 26:3, 252-279.
- Berg-Schlosser, Dirk (2002): Macro-quantitative vs. macro-qualitative methods in the social sciences – testing empirical theories of democracy. Paper presented at the International Sociological Association World Congress, Brisbane.
- Brayton, Robert K. And Sunil P. Khatri (1999): Multi-value Logic Synthesis. Paper. University of Berkeley, CA. Paper. Internet: [http://www-cad.eecs.berkeley.edu/Respep/Research/mvsis/doc/mvsis\\_main.pdf](http://www-cad.eecs.berkeley.edu/Respep/Research/mvsis/doc/mvsis_main.pdf) [cited 2003-09-10].
- Cronqvist, Lasse (2002): How MVQCA works. A short Introduction to the Ideas of the Algorithm used in TOSMANA. Powerpoint- Presentation. Internet: <http://www.tosmana.net/resources> [cited 2003-09-10].
- Cronqvist, Lasse (2003a): Tosmana – Tool for small-n analysis. Version 1.0. Marburg: University of Marburg. Internet: <http://www.tosmana.net/> [cited 2003-09-10].
- Cronqvist, Lasse (2003b): *Using Multi-Value Logic Synthesis in Social Science*. Paper prepared for the 2nd General Conference of the European Consortium for Political Research (ECPR), Marburg, September 2003. Also available on <http://www.tosmana.net/resources>
- De Meur, Gisèle and Benoit Rihoux (2002): L'analyse quali-quantitative comparée (AQQC/QCA): Approche, techniques et applications en sciences humaines. Louvain-La-Neuve.
- Drass, Kriss A. and Charles C. Ragin (1992): Qualitative Comparative Analysis 3.0. Evanston, Illinois: Institute for Policy Research, Northwestern University. Internet: <http://www.fsqca.com> [cited 2003-09-10].
- Lipset, Seymour M. (1963): *Political Man. The Social Bases of Politics*. New York: Doubleday.
- MVSIS Group: Multi-Valued Logic Synthesis. Software Release. Online in Internet: <http://www-cad.eecs.berkeley.edu/Respep/Research/mvsis/> [cited: 2003-09-10].
- Ragin, Charles C. (1987): The Comparative Method. Moving Beyond Qualitative and Quantitative Strategies. Berkeley – Los Angeles – London: University of California Press.
- Ragin, Charles C. (2000): Fuzzy-Set Social Science. Chicago - London: The University of Chicago Press.



Ragin, Charles C. and Kriss A. Drass (2002): Fuzzy-Set/Qualitative Comparative Analysis 0.963. Tucson, Arizona: Department of Sociology, University of Arizona. Online in Internet: <http://www.fsqca.com> [cited: 2003-09-10].