



How World Bank Governance Indicators May Predict the Assessment of Elections by International Observers: Evidence From the OSCE-Region

Hans Schmeets

To cite this article: Hans Schmeets (2018) How World Bank Governance Indicators May Predict the Assessment of Elections by International Observers: Evidence From the OSCE-Region, Nordic Journal of Human Rights, 36:1, 55-73, DOI: [10.1080/18918131.2018.1446623](https://doi.org/10.1080/18918131.2018.1446623)

To link to this article: <https://doi.org/10.1080/18918131.2018.1446623>



© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 01 May 2018.



Submit your article to this journal [↗](#)



Article views: 411



View related articles [↗](#)



View Crossmark data [↗](#)

How World Bank Governance Indicators May Predict the Assessment of Elections by International Observers: Evidence From the OSCE-Region

Hans Schmeets

Maastricht University and Statistics Netherlands

ABSTRACT

This paper outlines election monitoring in the OSCE-region, focusing on the analyses and interpretation of the aggregated findings of over 28,000 international election observers producing over 170,000 election observer report forms in 97 elections. Based on 86 elections, to which at least 100 observers were deployed, this paper reveals that: (1) the observers' overall assessments of the voting and counting processes are strongly correlated; (2) the assessment of the voting process is rather stable, whereas the counting process deteriorated from 1996 onwards; (3) observers' assessments of the election process tend to be more positive in societies with higher degrees of political openness and stability; and (4) particularly elections in Tajikistan, Belarus, Kyrgyzstan, Azerbaijan, and Kazakhstan are assessed negatively.

KEYWORDS

Election Observation; OSCE; Voting; Counting; Democracy; Governance

1. Introduction

One third of the vote count observations in the Russian presidential election conducted on 4 March 2012 were evaluated in negative terms. The Organization for Security and Cooperation in Europe (OSCE) concluded as follows in its preliminary statement:

On election day, voting was assessed positively overall; however, procedural irregularities were observed. The process deteriorated during the count which was assessed negatively in nearly one-third of the polling stations observed due to procedural irregularities.¹

The statement attracted great international interest and is based on the findings of 260 international observers deployed throughout the Russian Federation. But what made the Organization for Security and Cooperation in Europe/Office for Democratic Institutions and Human Rights (OSCE/ODIHR) decide that the assessment was quite positive on election day, considering its rather drastic shift to a negative assessment when it came to the counting procedures? Obviously, the findings need to be analysed in view of other elections. Furthermore, one may argue that the authorities organised a perfect election,

CONTACT Hans Schmeets  hshs@cbs.nl

¹OSCE/ODIHR, 'International Election Observation Russian Federation, Presidential Election, 4 March 2012: Statement of Preliminary Findings and Conclusions' (2012) <<http://www.osce.org/odihr/88667>> accessed 14 February 2018.

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

completely in line with international standards and regulations of election law, but that many other aspects of a liberal democracy were not met, such as adherence to the rule of law and freedom of speech.

In this paper, I will focus on the findings of international observers in the OSCE-region from 1996 onwards. The main research questions reads: ‘What are the trends in the assessment of the consecutive elections in the OSCE-region based on election observers’ findings on polling day and how do these findings compare to some other indicators of democratisation in the region?’.

This paper aims to make four contributions: (1) using an extraordinarily rich and unique database, in which the aggregated findings use of some 28,000 international election observers producing over 170,000 election observer report forms in 97 elections in the OSCE-region are included; (2) detail the OSCE/ODIHR standardised approach in monitoring elections; (3) analyse the assessment of polling day – the voting and counting process – within 17 OSCE-countries and 86 elections ($n = 86$), and the developments in the observers’ assessment in the 1996–2010 timespan; (4) analyse the connections with three worldwide governance indicators retrieved from the World Bank.

First, I will elaborate upon the case selection and the hypotheses, followed by a brief outline of the applied methodology in OSCE/ODIHR Election Observation Missions. Subsequently, I will present the results of the observers’ findings of 86 elections in 17 OSCE-countries, followed by conclusions and discussion.

2. Case Selection: OSCE/ODIHR Election Observation Missions

The global rate of observed elections has increased substantially since 1980, and over 80% of all national elections are now monitored by foreign observers.² Election Observation Missions (EOMs) are mostly conducted under the umbrella of intergovernmental organisations such as the OSCE, the Council of Europe (CoE), the European Union (EU), the Organization of American States (OAS), The African Union (AU), and the Commonwealth of Independent States (CIS). All of these organisations assess elections through international observers who visit various polling stations, in order to report the circumstances and procedures.

In 2005, 32 international organisations, including the abovementioned, signed the Declarations of Principles for International Election Observations, in which the key characteristics of an EOM are outlined. However, five years later, in roughly a third of the observed elections, the international organisations disagreed about their overall assessment.³ Such discrepancies between the international bodies undermine the credibility of observation missions. One straightforward explanation is the variation in the applied methodology. Apart from standards related to the fielded number of observers, and consequently the number of polling stations visited, also a standardised observer report form with the appropriate questions is often a shortcoming. Furthermore, the sample designs (i.e. the deployment of the observers in the country and assignment to the polling

²This also includes smaller election assessment missions. See Sarah Birch, *Electoral Malpractice* (OUP 2011); Susan D Hyde, ‘Catch Us If You Can: Election Monitoring and International Norm Diffusion’ (2008) 55 *American Journal of Political Science* 356; Susan D Hyde, *The Pseudo-Democrat’s Dilemma: Why Election Monitoring Became an International Norm* (Cornell University Press 2011).

³Judith Kelley, ‘Election Observers and Their Biases’ (2010) 21(3) *Journal of Democracy* 158.

stations), the data-processing and data-analysis are not standardised at all. This is often even true for the methodology applied within the same organisation. However, the advantage and necessity of using a standardised methodological approach for comparing elections is widely acknowledged, as is the need for more analyses based on the observers' findings.⁴

The methodology used by the OSCE/ODIHR is regarded to be of the highest standard worldwide. This was reflected in the Maastricht (The Netherlands) Expert Meeting on the enhancement of international election observation and assessment of election, on 8–9 April 2011.⁵ One of participants, Bruce George states that the 'ODIHR has shown over the years a very high degree of professionalism in terms of its Core Team and its leadership, and the very high quality of its LTOs and STOs'.⁶ ODIHR's high standard is – among many other aspects – reflected in the large number of international observers fielded to visit polling stations, the deployment plan which includes a high coverage of the regions and is, to a large extent, randomly distributed over the countries or the selected regions within a country,⁷ the quality of the observer report forms, and the processing and analysis of the observers' findings. Many facets have been standardised, which will be outlined below.

The OSCE deploys hundreds of observers, numbering more than 600 in some EOMs, while other organisations, such as the EU, AU, OAS, CIS, and the Commonwealth, deploy substantially fewer observers. Based on the 1990 Copenhagen Commitments, the OSCE observes the elections of its 57 member states during the various phases of an election, before, during, and after election day.⁸ This is common practice for the assessment of elections in countries in transition or new democracies. Another approach

⁴See Robert A Pastor, 'The Role of Electoral Administration in Democratic Transitions: Implications for Policy and Research' (1999) 6(4) *Democratization* 1; David Carroll, 'Declaration of Principles and Code of Conduct: A Q&A with Democracy Program Director David Carroll' (*The Carter Center*, 30 October 2005) <<http://www.cartercenter.org/news/documents/doc2232.html>> accessed 14 February 2018; Daniel Calingaert, 'Election Rigging and How to Fight It' (2006) 17(3) *Journal of Democracy* 138; Jørgen Elklit, 'Is the Degree of Electoral Democracy Measurable? Experiences from Bulgaria, Kenya, Latvia, Mongolia and Nepal' in David Beetham (ed), *Defining and Measuring Democracy* (Sage 1994); Jørgen Elklit and Palle Svensson, 'The Rise of Election Monitoring: What Makes Elections Free and Fair?' (1997) 8(3) *Journal of Democracy* 32; Judith Kelley, 'Assessing the Complex Evaluation of Norms: The Rise of International Election Monitoring' (2008) 62 *International Organizations* 221; Judith Kelley, *Monitoring Democracy: When International Election Observations Works and Why It Often Fails* (Princeton University Press 2012); Fabrice Lehoucq, 'Electoral Fraud: Causes, Types, and Consequences' (2003) 6 *Annual Review of Political Science* 233; Gerardo L Munck, *Measuring Democracy: A Bridge between Scholarship and Politics* (Johns Hopkins University Press 2009); Gerardo L Munck and Jay Verkuilen, 'Conceptualizing and Measuring Democracy: Evaluating Alternative Indices' (2002) 35(1) *Comparative Political Studies* 5; Jonathan Misk, 'Standardising the Principles of International Election Observation' (2012) 43(3) *Vanderderbilt Journal of Transnational Law* 763; Pippa Norris, 'The New Research Agenda Studying Electoral Integrity' (2013) 32(4) *Electoral Studies* 563; Pippa Norris, 'Does the World Agree About Standards of Electoral Integrity? Evidence for the Diffusion of Global Norms' (2013) 32(4) *Electoral Studies* 576; Pippa Norris, Richard W Frank and Ferran Martinez i Coma, 'Assessing the Quality of Elections' (2013) 24(4) *Journal of Democracy* 124; Carolina Van Ham, 'Getting Elections Right? Measuring Electoral Integrity' (2014) 22(4) *Democratization* 714.

⁵Hans Schmeets, 'Analysing Observer Report Forms: An Overview' in Hans Schmeets (ed), *International Election Observation and Assessment of Elections* (Statistics Netherlands 2011) 142.

⁶Bruce George and Simon Kimber, 'Developments in Election Observation' in Schmeets, *International Election Observation* (n 5) 57. See also Audrey Glover, 'The OSCE Office for Democratic Institutions and Human Rights and its Challenges in Relation to Election Observation' in Schmeets, *International Election Observation* (n 5) 102.

⁷For a discussion of the impact of the selection bias, see Max Bader and Hans Schmeets, 'The Problem of Selection Bias in OSCE Election Observation Methodology' (2013) 24 *Security and Human Rights* 1; Max Bader and Hans Schmeets, 'Is International Election Observation Credible? Evidence from Organization for Security and Co-operation in Europe Missions' (2014) 1(2) *Research and Politics* 1.

⁸Glover (n 6); George and Kimber (n 6); Nicolas Kaczorowski, 'The Office for Democratic Institutions and Human Rights' in Schmeets, *International Election Observation* (n 5); Gerald Mitchell, 'History and Mandate of Election Observation: The OSCE/ODIHR Experience' in Schmeets, *International Election Observation* (n 5).

is a limited election mission with some 10 election experts, without regular observers, which is called an Election Assessment Mission (EAM). Such EAMs have become more popular in recent years as a tool to cover the elections in other older, democratic OSCE member states as well.⁹

In order to assess the elections by an EOM, the member states are requested to send international observers. These include long-term observers (LTOs) to cover the whole election period of two to three months, and short-term observers (STOs) who stay for a week around election day. The ODIHR, based in Warsaw, provides the methodology and coordination for such an EOM. In a typical EOM, an EOM core team works on the ground, comprised of a group of some 10 international election experts together with local staff. The STOs visit the polling stations and report their observations by answering questions on their observation forms. Although the polling stations are not randomly selected, the deployment plan assures that it is at least somewhat ‘pseudo-random’. In recent years, the counting station is often randomly selected, and observers select the polling stations for the opening and voting close to this randomly selected counting station to enhance the randomness of the whole sample. Soon after, usually within 24 hours of polls closing on election day, the ODIHR releases a preliminary statement about the elections. Such a statement is a result of negotiations between various actors involved, typically the Parliamentary Assemblies of the OSCE (OSCE PA), the Council of Europe (PACE), and the European Union. The negotiations are occasionally tough. George and Kimber provide detailed examples of the negotiations in which a bad cooperation between ODIHR and the various Parliamentary Assemblies is illustrated, and state that ‘election observation is weakened by this guerrilla warfare – it should end with a cooperative relationship between the various Parliamentary Assemblies and ODIHR’.¹⁰ During election day and the day after, the various delegations meet to discuss the content of the statement. The statistical findings are crucial in the negotiations, including the comparison with other EOMs in the same and in other countries in the OSCE-region, as well regional breakdowns and overviews based on observers’ written comments. The ODIHR also publishes a final election report with recommendations within approximately 6 to 10 weeks of the completion of the electoral process. In this report, the statistical figures are updated and more findings are included based on detailed statistical analysis.

In their statements and through press conferences, the OSCE/ODIHR refrains from saying whether the election is “free and fair”. Many scholars have criticised the use of such words.¹¹ Instead of using the “free” and “fair” terminology, the OSCE/ODIHR always refers to the 1990 OSCE Commitments and checks whether the elections are in line with the specific paragraphs in those commitments, which are to be summarised by seven key words: equal, fair, secret, free, universal, transparent and accountable.¹²

⁹ibid.

¹⁰George and Kimber (n 6) 43.

¹¹Thomas Carothers, ‘The Observers Observed’ (1997) 8(3) *Journal of Democracy* 17; Guy S Goodwin-Gill, *Free and Fair Elections: International Law and Practice* (Inter-parliamentary Union 1994).

¹²OSCE/ODIHR, *Election Observation Handbook* (3rd edn, ODIHR 1998). Meanwhile, the fifth and sixth edition of the Handbook have been released, in which the seven key words were changed to ‘periodic’, ‘genuine’, ‘free’, ‘fair’, ‘universal’, ‘equal’, ‘secret’, ‘honest counting and reporting of results’: see OSCE/ODIHR, *Election Observation Handbook* (6th edn, 2010) 17–24.

These principles can be traced back to the principles of article 25 of the International Covenant on Civil and Political Rights (ICCPR)¹³:

Every citizen shall have the right and the opportunity, without any of the distinctions mentioned in article 2 and without unreasonable restrictions:

- a) To take part in the conduct of public affairs, directly or through freely chosen representatives;
- b) To vote and to be elected at genuine periodic elections which shall be universal and equal suffrage and shall be held by secret ballot, guaranteeing the free expression of the will of the electors;
- c) To have access, on general terms of equality, to public service in his country.

Six of these seven principles are covered by observation on polling day. Voters should have *equal* access to the polling stations and safeguards should be implemented to guarantee that the principle of one person-one vote is not violated, as can be the case with proxy-voting, and that the voter has to cast his vote in person. *Secrecy* can only be guaranteed if the voter casts his or her vote alone, thus permitting only one person per polling booth. *Fairness* refers to voters' exposure to basic information about all political contestants in the election and the fundamental issues to which they are connected. On polling day, it is forbidden to have campaign material in the vicinity of or inside the polling station. Voters should mark their ballots in a *free* manner, without any intimidation or pressure. Furthermore, *universal* suffrage requires that no groups of eligible voters be refused a vote. This also implies that voters' lists should be updated and accurate. *Transparency* is crucial in an election as international observers, NGOs, and party or candidate representatives should have a clear view inside the polling stations.

A shortcoming in research on the assessment of elections has been the translation of the Copenhagen Commitments into concrete dimensions and variables.¹⁴ This translation, however, is made possible by attaching the key words to the questions in the observer report forms. For example, questions such as "voting outside polling booth?", "group or family voting?", and "voters marking ballots in secrecy?" refer to the keyword 'secret'. Schmeets introduced such an overview and demonstrated that the key principles are directly and unambiguously reflected and measured by the questions in the observer report forms.¹⁵ This has an important advantage, as the observers' findings on election day do not only provide information on the various phases in the election process, but also reveal which principles are violated. The linchpin between the Copenhagen Principles and the consequent keywords is the election law which serves as a guide for the phrasing of the questions in the observer report forms. As most election laws in the OSCE-region showed many similarities, an opportunity existed for the production of a standardised form.

¹³Markku Suksi, 'The Electoral Cycle: On the Right to Participate in the Electoral Process' in Veronika U Hinz and Markku Suksi (eds), *Election Elements: On the International Standards of Electoral Participation* (Institute for Human Rights, Åbo Akademi University 2003) 43.

¹⁴Elklit and Svensson (n 4); Elklit (n 4).

¹⁵Johannes JG Schmeets, *Vrije en eerlijke verkiezingen in de OVSE-regio? De ontwikkeling van een meetinstrument* [Free and Fair Elections in the OSCE-region? The Development of a Measurement Tool] (Statistics Netherlands 2002); Schmeets, 'Analysing Observer Report Forms' (n 5).

The OSCE set up an EOM in 1996 to observe the presidential elections in the Russian Federation. The design of the observer forms was a result initiated by the EU to redesign the forms used during the 1995 Russian Duma elections, led by Statistics Netherlands.¹⁶ The redesign revealed to be a starting point for a standard form used by the OSCE. In 1996, the EU was responsible for the deployment of the STOs, while the OSCE/ODIHR was responsible for the LTOs.¹⁷ This design, adopted by the OSCE/ODIHR for the EOM in Bosnia and Herzegovina (BiH) in September 1996 and subsequent EOMs, was characterised by a specific layout in which several sections were distinguished, covering the specific problems and irregularities outside and inside the stations. In addition, the following overall indicator was included, posed by the question: “In general, do you feel the voting process was conducted properly?” This overall indicator became a very important tool in the assessment of the elections based on the STOs’ findings and has been used in almost all OSCE/ODIHR EOMs so far.¹⁸

Evidently, apart from the observers’ findings in a specific EOM, a comparison is needed with the results of previous EOMs in the OSCE-region. For that reason, a database has been created in which the findings of a total of 97 EOMs from 1996 to 2010 are included.¹⁹ This database covers the following 20 OSCE countries (number of missions in parentheses): Albania (6), Armenia (8), Azerbaijan (6), Belarus (4), BiH (7), Croatia (2), Czech Republic (1), Estonia (1), Georgia (8), Hungary (2), Kazakhstan (3), Kyrgyzstan (6), Kosovo (5), Macedonia (14), Moldova (5), Montenegro (4), Russia (4), Serbia (1), Tajikistan (3), and Ukraine (7).²⁰

On average, 331 observers per EOM were on the ground per election, making the total number of observers in all elections surpass 30,000. More and more observers were to be deployed in EOMs since 1996, ranging from slightly more than 200 in 1996 to 400 in 2010.²¹ During election day and night, the STO report forms are handed over, faxed or scanned to the EOM headquarters, and there are immediately processed.²² Severe irregularities – such as ballot box stuffing and intimidation of voters and observers – are checked by matching the irregularities to certain observer teams on the list of teams, and by corroborating the irregularities with the overall indicator score. At the same time, the qualitative comments provided by observers in special forms are processed in order to produce an overview in a database of problems and to identify significant violations. The comments provide further insight into what was seen and heard during the observations. If

¹⁶Hans Akkerboom and Hans Schmeets, ‘Survey Data Collection Methodology for Election Missions: How Professional Form Design Helps to Control a Survey’ (1998) 10(3) *International Journal of Public Opinion Research* 257.

¹⁷ibid; Hans Schmeets, ‘The Observation of the Russian Presidential Elections’ in *Electoral Reform International Services: Report of the European Union Election Unit on the Russian Presidential Elections: 16 June and 3 July 1996* (Electoral Reform International Services 1996) 16.

¹⁸The yes/no question was amended at a workshop organised by ODIHR, 1997, 8–9 February into the questions with four answer options: very good, good, bad, very bad (for more details see Schmeets, *Vrije en eerlijke verkiezingen* [n 15] 82–112). However, especially in some EOMs in the 1990s, this guideline was not always followed. Examples are missions in 1998 with a relatively small number of observers in Slovakia, Montenegro and Moldova in which a five-scale question ranging from poor to excellent was used. These data are not included.

¹⁹The database is summarised in Table 1. The data has been collected from the OSCE/ODIHR reports and statements available online at <<http://www.osce.org/odihr/elections/>> accessed 14 February 2018.

²⁰Note that the Council of Europe was in charge of EOMs in Kosovo and adopted the methodology, including the observer report forms, developed by the OSCE/ODIHR.

²¹Schmeets, ‘Analysing Observer Report Forms’ (n 5).

²²In recent years, data are collected directly from the field with a pen application by which data are transferred to an excel database.

clear quantitative violations (i.e. ballot box stuffing) are not substantiated with comments, these will be investigated and the data will be cleaned accordingly.

A further evidence check helps to verify the overall assessment: clear violations of the various principles should go together with a negative judgement overall. The estimates and correlations are compared to the findings in other EOMs. For example, if intimidation is observed in 5% of the polling stations, it is obvious that this is a serious problem in that particular election, as intimidation – as an indicator of the violation of the ‘free’ principle – is observed at a maximum of 2% of the observations in almost all elections. On the other hand, a 5% ‘family or group voting’ figure is, in EOMs, often rather low and might even be considered as a positive indication vis-à-vis the principle of ‘secrecy’ in an election. Furthermore, comparisons across EOMs reveal clear and consistent patterns. Some problems and irregularities are more strongly correlated with the overall indicator than others. An example of a low-correlated problem is whether voters vote together in a polling booth, typically labelled as ‘family voting’. This irregularity is, however, not seen as a severe violation of the voting procedures and consequently is not a sufficient reason for a negative assessment overall. Examples of high correlated problems are ‘intimidation’, ‘ballot box stuffing’, ‘tension or unrest’. If such problems are witnessed, in a vast majority of the forms the overall assessment is negative.²³ This demonstrates that particular violations have a greater impact on observers’ decisions as to whether the assessment should be either bad or good. It becomes clear, therefore, that the most important suggestion as to whether an election is marred with problems and irregularities is typically this overall indicator.²⁴

In previous research based on 13 elections in Russia, BiH, Albania, Armenia, Georgia and the province of Kosovo in the 1996 to 2000 period, it has been determined that the questionnaire for the observation of the voting process includes 24 core problems.²⁵ Based on an optimal scaling analysis by using the ‘Homogeneity analysis by means of alternating least squares’ (HOMALS), it was demonstrated that the 24 problems refer to one latent problem factor, and weights – based on the provided category-quantifications – constructed for these problems.²⁶ These quantifications serve as weights for the impact of the 24 problems on the overall problem score (see Table 1). In particular, ‘disruptions inside the polling station’, ‘intimidation’, and ‘agitation/unrest’ contribute to the overall score (median weights > 3). Other problems – in particular ‘voters refused to vote’ – do not contribute substantially (median = 0.20) to the overall latent variable. The reweighted problems and irregularities were added, resulting in a voting problem-score for each of the thirteen elections: the Standard Election Observation Voting Index (SEOVI). It has been found that the correlation between the overall assessment and the voting problem score is high (Pearson’s $R = 0.90$). This indicates that the overall assessment reflects the 24 problems and irregularities quite well. The high correlation between the problem factor and the overall indicator is also reflected in the correlations on a lower level: the bivariate correlations between the specific problems and irregularities and the overall assessment in the EOMs. The pattern is very consistent: (a) if a problem or irregularity is noted, the overall assessment is more negative, and

²³For examples, see Schmeets, *Vrije en eerlijke verkiezingen* (n 15) 313–23.

²⁴Schmeets, ‘Analysing Observer Report Forms’ 2011 (n 5).

²⁵Schmeets, *Vrije en eerlijke verkiezingen* (n 15).

²⁶ibid; Schmeets, ‘Analysing Observer Report Forms’ (n 5).

Table 1. Selected problems and irregularities in the voting process, OSCE/ODIHR principles, and weights (mean) calculated by HOMALS analyses in EOMs in Russia, Bosnia and Herzegovina, Albania, Armenia, Georgia and the province of Kosovo.

Problems observed	Principle	Weights	
Circumstances outside PS			
1	Access difficult	Equal	0.80
2	Campaigning outside	Fair	1.25
3	Other problems vicinity	General	2.37
Circumstances inside PS			
4	Disruptions inside PS	Equal	4.34
5	Overcrowded	Equal	1.25
6	Intimidation	Free	4.04
7	Agitation / unrest	Free	3.67
8	Campaigning inside PS	Fair	2.04
9	Other circumstances	General	1.66
Procedures			
10	Material not sufficient	Equal	1.07
11	Ballot box sealed	Secret	0.68
12	Ballots signed / receipt	Equal	1.07
13	Multiple voting / ink	Equal	2.03
14	Voters refused to vote	Universal	0.20
15	Identification	Equal	1.23
16	Voters signed List	Equal	1.39
17	ID-entered / ink applied	Equal	1.29
18	Illiterate / handicapped	Equal	0.76
19	Voting outside booth	Secret	2.16
20	Family voting	Secret	0.71
21	Secrecy	Secret	2.57
22	Other procedures	General	1.32
People present			
23	Unauthorised persons	Free	1.24
24	Police inside PS	Free	0.67

Source: OSCE/ODIHR, retrieved from Schmeets, *Vrije en eerlijke verkiezingen* (n 15) 200.

(b) the strength of the correlations is quite similar across the EOMs yet the strength of the correlation shows that some problems are more relevant for the observers' overall judgement than others. For the counting process, this approach was not feasible as, mainly due to variation in the counting procedures, there was no consistent pattern of the questions across both time and countries. For that reason, a sum score was created based on observers' assessment of the organisation, the procedures followed, and the overall impression of the vote count were used: the Standard Election Observation Counting Index (SEOCI). The reliability (Cronbach's alpha) of the SEOCI was high, ranging from 0.85 (Armenia 1998) to 0.94 (Georgia 2000), illustrating that the overall indicator is strongly related to observers' judgement of the organisation of the count by the precinct election commission and whether the procedures were adhered to. Building on this 2002 research, the overall indicators serve as the main indicators for the assessment of the quality of elections. The correlation (Pearsons' R) between the calculated voting (SEOVI) and counting index (SEOCI) is 0.78 and illustrates that problems observed during the day are a good predictor for the problems observed during the count.

This overview, albeit limited to 13 elections, shows that the observers' overall assessments based on the two single questions are good indicators for the problems and irregularities observed during the voting and the vote count on election day.

3. Hypotheses

The aggregated findings, based on over 170 thousand polling stations visited in 86 elections in which at least 100 observers were deployed, will be compared with three indicators of governance in the region: the rule of law, the openness of the system, and the political stability. These indicators, available in the selected timespan, are retrieved from the database compiled by the World Bank and are based on expert opinions and surveys.²⁷ All three indicators are connected to the assessment of whether an election meets international democratic standards, including the 1990 Copenhagen Document in which the commitments of the OSCE member states are manifest, serving as the reference for the preliminary statement released shortly after the election and the final OSCE/ODIHR-reports.

The rule of law captures perceptions of the extent to which agents have confidence in and abide by the rules of society, in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence. The adherence to the rule of law is closely connected to the work of the administrative bodies, the Central Election Commission. Obviously, adherence to the rule of law inhibits large-scale fraud and other severe violations. The 1990 Copenhagen Document states that democracy is an element of the rule of law and that the presence of observers ensures confidence building in society:

(2) [The participating States] are determined to support and advance those principles of justice which form the basis of the rule of law.

(3) They reaffirm that democracy is an inherent element of the rule of law [...]²⁸

The second factor relates to the commitment to international standards of democratic elections, and to the openness of the electoral system in a country. Openness is defined by the political opportunity structures,²⁹ indicating the opportunities to influence governance via conventional means, such as elections. If there are hardly any opportunities to voice political concerns via freedom of expression, the media, and freedom of association, elections would be unable to meet international standards. In such cases, the elections would not be transparent and the principles of a democratic election are jeopardised, especially those elucidated in article 12 in the Copenhagen Document in particular.

The third indicator is political stability, or absence of violence. My reasoning is that the elections process and results will be more often challenged under unstable regimes due to tensions between party or candidate proxies present inside the polling stations. As a consequence, observers would notice more problems and irregularities in such countries.

This leads us to submit three hypotheses. I expect observers to be more positive on the voting and counting process where a country meets the following criteria: (H1) adherence to the rule of law; (H2) openness; and (H3) political stability.

²⁷Daniel Kaufmann, Aart Kraay and Massimo Mastruzzi, 'Governance Matters III: Governance Indicators for 1996, 1998, 2000 and 2002' (2004) 18(2) World Bank Economic Review 253.

²⁸OSCE, Document of the Copenhagen Meeting of the Conference on the Human Dimension of the OSCE, 29 June 1990 (Copenhagen Document) para 2, 3.

²⁹Herbert P Kitschelt, 'Political Opportunity Structures and Political Protests: Anti-Nuclear Movements in Four Democracies' (1986) 16 British Journal of Political Science 57; David S Meyer, 'Protest and Political Opportunities' (2004) 30 Annual Review of Sociology 125; Simon Szreter, 'The State of Social Capital: Bringing Back in Power, Politics, and History' (2002) 31(5) Theory and Society 573.

For measurement of these criteria, we will use the indicators provided by the World Bank: (1) The Rule of Law index, (2) Kaufman's voice and accountability index (for openness), and (3) the political stabilisation index³⁰. Descriptions of the three indicators are based on the QoG codebook:

- (1) The 'Rule of Law' includes several indicators which measure the extent to which agents have confidence in and abide by the rules of society. These include perceptions of the incidence of crime, the effectiveness and predictability of the judiciary, and the enforceability of contracts. Together these indicators measure the success of a society in developing an environment in which fair and predictable rules form the basis for economic and social interactions and the extent to which property rights are protected.
- (2) 'Voice and Accountability' includes a number of indicators measuring various aspects of the political process, civil liberties and political rights. These indicators measure the extent to which citizens of a country are able to participate in the selection of governments. This category also includes indicators measuring the independence of the media, which serves an important role in monitoring those in authority and holding them accountable for their actions.
- (3) 'Political Stability' combines several indicators which measure perceptions of the likelihood that the government in power will be destabilized or overthrown by possibly unconstitutional and/or violent means, including domestic violence and terrorism.³¹

Many other indicators measuring (aspects of) democracy and good governance are available, such as the Freedom House and Polity (e.g. polity4) measures, and the more recently developed measures from Varieties of Democracy.³² Most of these measurements include components that are directly related to elections, such as the turnout rate, which would not serve the aim of this paper very well. Moreover, a comparison of those measurements would be beyond the scope of this paper in which the focus is on the relationship between the three selected World Bank governance indicators and the assessment of international election monitors. Although the World Bank governance indicators are under criticism,³³ they are the most used and cited. The main aim is to explore to which extent the scores on each of the three governance measurements – openness, rule of law, and stability – provide predictions for the judgements by international observers in the OSCE-region. For that sake, no other (control) variables are included in the models. Apart from the three selected measurements, three other indicators are developed by the World Bank: government effectiveness, regulatory quality, and control of corruption. However, these indicators are to a lesser extent linked to the Copenhagen Document, and are considered to be a measurement for state capacity.³⁴

³⁰Kaufmann, Kraay and Mastruzzi (n 27).

³¹Jan Teorell, Marcus Samanni Sören and Bo Rothstein, 'April 6. The QoG Standard Dataset version 6Apr11' (University of Gothenburg: The Quality of Government Institute 2011) <<http://www.qog.pol.gu.se>> accessed 17 September 2012.

³²V-Dem, *Democracy at Dusk? V-Dem Annual Report 2017* (University of Gothenburg 2017).

³³Francis Fukuyama, 'What is Governance?' (2013) 26(3) *Governance* 347; José Antonio Cheibub, Jennifer Gandhi and James Raymond Vreeland, 'Democracy and Dictatorship Revisited' (2010) 143 *Public Choice* 67.

³⁴ibid.

4. Results

4.1 Observers, voting and counting

All results are based on the EOMs covered by at least 100 international observers and for which results were available on both the overall assessment of the voting and counting process (see [Table 2](#)). In total, 11 out of the 97 EOMs were excluded from our analysis; this concerns EOMs in Macedonia (1 November 1998; 24 September 2000; 29 June 2008), Hungary (10 May 1998; 29 May 1998), Estonia (7 March 1999), Czech Republic (19 June 1998), Albania 24 June 2001; 3 July 2005), BiH (12 September 1998) and Armenia (5 March 2003). Based on the 86 EOMs, on average, 8% of the polling stations covering the voting process and 22% of the vote count stations were assessed negatively, i.e. were rated as either bad or very bad overall (see [Table 3](#)).³⁵

The cross-country variation, however, is substantial. In Belarus, the counting process was rated negatively in many polling stations by the observer teams: 62% in 2004, 50% in 2006, 48% in 2008 and also 48% in 2010. The same is true for Tajikistan: 54% in 2005, 55% in 2006 and 56% in 2010. In Azerbaijan too, observers concluded that the counting process was often replete with problems: based on the 22, 52, 55, 41, 23 and 32% for the consecutive elections in the 1998 to 2010 timespan, meaning that on average, the assessments were negative in 38% of the observed polling stations. In addition, most problems related to the voting process were seen in Tajikistan (22%) and Azerbaijan (15%). The tallied (voting plus counting) results demonstrate that the elections in Tajikistan were most often assessed in negative terms, followed by the elections in Belarus, Azerbaijan, Kyrgyzstan and Kazakhstan. On the other hand, the elections in Serbia, Croatia, Montenegro, Ukraine, Kosovo, and BiH received rather positive ratings from the international observers.

This overview also clearly shows that the overall assessment of the voting process is evidently more positive in comparison to the counting process. An explanation for this finding is that observer teams stay on average approximately 30 minutes in a polling station observing the voting process, whereas they stay from the beginning until the end in a polling station to observe the counting procedures. So they may notice more problems and irregularities as they observe the whole process during their stay of at least two to three hours observing the counting procedures than during their observation of the voting process. A further explanation is that it may be easier to manipulate the counting process than the voting process, and consequently more violations are noticed during the vote count. In addition, it is during the count that international fraud can have the highest reward.

These results also imply that observers' findings should be treated in a comparative perspective. In a 2011 research, I introduced a benchmark in order to evaluate observers' findings based on the overall indicators (see [Table 3](#)). This benchmark is used in the observers' briefings and debriefings to clarify the judgements based on the overall assessments in the specific EOM. The best quartile of the voting process overall assessments consists of a group of elections with up to 3% negative reports. The next quartile includes between 3 and 6.5% negative reports, followed by a group ranging between 6.5 and 10%, with the

³⁵In the 2000 and 2002 elections in Macedonia, and the 2005 and 2007 elections in Moldova a five-point scale from poor (0) to excellent (4) was used. The percentages in [Table 3](#) refer to positions 0 and 1.

Table 2. Bad to very bad assessed polling stations during voting and counting process in 17 OSCE-countries, 1996–2010.

Country name	VOTING	COUNTING	Election	Country	VOTING	COUNTING	Election
Albania	12	29	29 June 1997	Kyrgyzstan	9	33	1 September 2007
Albania	8	13	6 July 1997	Kyrgyzstan	11	50	26 July 2009
Albania	10	23	18 February 2007	Kyrgyzstan	7	38	03 October 2010
Albania	8	15	30 June 2009	Kosovo	10	17	29 October 2000
Armenia	16	26	16 March 1998	Kosovo	2	3	17 November 2001
Armenia	14	29	30 March 1998	Kosovo	2	14	26 October 2002
Armenia	13	22	30 May 1999	Kosovo	3	8	23 October 2004
Armenia	10	20	19 February 2003	Kosovo	6	16	17 November 2007
Armenia	10	33	25 May 2003	Macedonia	5	6	18 October 1998
Armenia	6	18	1 April 2007	Macedonia	22	12	10 September 2000
Armenia	5	16	19 February 2008	Macedonia	3	7	15 September 2002
Azerbaijan	13	22	11 October 1998	Macedonia	5	8	14 April 2004
Azerbaijan	24	52	5 November 2000	Macedonia	8	21	28 April 2004
Azerbaijan	24	55	15 October 2003	Macedonia	10	18	13 March 2005
Azerbaijan	13	41	6 November 2005	Macedonia	13	24	27 March 2005
Azerbaijan	6	23	15 October 2008	Macedonia	5	15	5 July 2006
Azerbaijan	11	32	7 November 2010	Macedonia	8	15	1 June 2008
Belarus	10	62	17 October 2004	Macedonia	5	15	22 March 2009
Belarus	10	50	19 March 2006	Macedonia	3	9	5 April 2009
Belarus	5	48	28 September 2008	Moldova	4	6	6 March 2005
Belarus	6	48	19 December 2010	Moldova	4	23	3 June 2007
BiH	3	2	14 September 1996	Moldova	3	9	5 April 2009
BiH	3	7	13 September 1997	Moldova	3	6	29 July 2009
BiH	3	14	05 October 2002	Moldova	2	9	29 November 2010
BiH	7	20	02 October 2004	Montenegro	3	15	21 May 2006
BiH	7	24	01 October 2006	Montenegro	2	13	10 September 2006
BiH	5	9	03 October 2010	Montenegro	3	13	6 April 2008
Croatia	8	5	2 January 2000	Montenegro	2	1	29 March 2009
Croatia	5	5	24 January 2000	Russia	2	7	16 June 1996
Georgia	21	23	31 October 1999	Russia	2	7	3 July 1996
Georgia	16	45	9 February 2000	Russia	5	27	7 December 2003
Georgia	19	31	2 November 2003	Russia	6	24	14 March 2004
Georgia	3	13	4 January 2004	Serbia	3	3	21 January 2007
Georgia	4	15	28 March 2004	Tajikistan	20	54	27 February 2005
Georgia	8	23	5 January 2008	Tajikistan	20	55	6 November 2006
Georgia	8	22	21 May 2008	Tajikistan	25	56	28 February 2010
Georgia	4	21	30 May 2010	Ukraine	6	10	31 October 2004
Kazakhstan	13	28	19 September 2004	Ukraine	7	12	21 November 2004
Kazakhstan	9	28	4 December 2005	Ukraine	3	7	26 December 2004
Kazakhstan	4	39	18 August 2007	Ukraine	8	22	26 March 2006
Kyrgyzstan	10	26	12 February 2000	Ukraine	2	9	30 September 2007
Kyrgyzstan	11	11	27 February 2005	Ukraine	2	5	27 January 2010
Kyrgyzstan	7	35	10 July 2005	Ukraine	2	4	7 February 2010

Source: OSCE/ODIHR and Council of Europe (Kosovo)

Table 3. Quartiles based on overall assessment voting and counting process in EOMs.

Quartile		Voting % bad to very bad	Counting % bad to very bad
1	Best quartile	< 3.0	< 9.0
2	Above average	3.0–6.4	9.0–16.9
3	Below average	6.5–10	17.0–27.4
4	Worst quartile	>10	> 27.5

Source: OSCE/ODIHR, retrieved from Table 8.5.2 in Schmeets, 'Analysing Observer Report Forms' (n 5) 155.

worst quartile comprising those elections with more than 10% negative reports. Similarly, the four election groups are selected based on the overall assessments of vote count observations: the best group consists of elections in which less than 9% of the reports are

negative, the next group falls in the range from 9 to 16.9%, the third group from 17.0 to 27.4%, and the worst over 27.5% negative vote count observations. This benchmark shows that the voting process in Tajikistan and the Southern Caucasus – Azerbaijan, Georgia and Armenia – fall in the worst part (more than 10% negative reports) and consequently are labelled ‘very bad’. Albania, Belarus, Kazakhstan and Macedonia are found in the third quartile (‘bad’). Other countries are characterised as either ‘good’ or ‘very good’. The picture changes for the vote count: Tajikistan (55%) and Belarus (52%) exceed the 27.5% threshold and fall in the worst quartile with more than half of all reports rated negatively. This is also the case for Azerbaijan (38%) and Kazakhstan (32%). The inclusion of more EOMs in recent years does not substantially change this benchmark.³⁶

Both assessments of the voting and counting processes are linked. A linear regression analysis shows that a negative overall assessment of the voting process is associated with a negative overall judgment of the vote count in an EOM ($R^2 = 0.44$).³⁷ In other words, if the observer teams conclude that the process is bad or even very bad in a very limited number of polling stations, they are also more lenient in their overall assessment of the counting process. There are, however, some outliers to this pattern. The four elections in Belarus are characterised by a very negative count originating mainly from a lack of transparency. In these elections, observers were kept at a distance and had no clear view of the procedures. The assessment of the voting procedures, however, was not that negative. The explained variance increases in the linear regression to 56% when those four elections are excluded.³⁸

4.2 World Bank governance indicators

As a next step, the findings of the 86 elections were linked with the three distinguished governance indicators covering the rule of law, openness, and political stability, measured one year before the elections took place. The bivariate associations show that the overall assessment of the voting process is moderately correlated ($p < .01$, 1-tailed) with the rule of law (Pearson’s $R = .37$), openness (.43), and political stability (.35). The correlation between two of the three indicators and the overall assessment of the counting process is substantially higher. For rule of law, we found a correlation coefficient of 0.55, and for openness a coefficient of 0.69, but no significant correlation for political stability (.07).

Based on these bivariate findings, hypotheses H1 and H2 would have been confirmed, and H3 only partially (for voting, but not for counting). However, the three World Bank measurements are correlated. To assess the unique contribution of each indicator, a multiple regression analysis is employed by including all three governance indicators and the year of the election.³⁹ The analysis revealed that the rule of law is not related to the voting, as the relation is not statistically significant ($p > 0.05$). Similarly, the year trend is not significant, which implies that observers’ overall assessment over the timespan did not change. However, both the openness (standardised regression coefficient $\beta = .46$) and

³⁶Based on a total of 121 EOMs, the second quartile for voting slightly decreases from 6.5 to 6.0%; and for counting the third quartile slightly increases from 27.9 to 29.0%.

³⁷In my earlier calculations (Schmeets, ‘Analysing Observer Report Forms’ (n 5) 157, Table 8.5.5), based on 81 instead of 86 EOMs, the explained variance was slightly higher (0.47).

³⁸Note that the rather strong correlations are in line with the finding based on the 13 elections studied in the 1995–2000 timespan: Schmeets, *Vrije en eerlijke verkiezingen* (n 15).

³⁹Due to missing information for political stability in Kosovo (5 EOMs) and Montenegro (2 EOMs), the results are based on a total of 79 EOMs.

stability (.37) of a country are clearly positively related to the voting process (R-square adjusted = .30). The Variance Inflation Factor values are all below 2.0, which is acceptable. Similarly, openness is strongly related to the assessment of the vote count (beta = .62), while the importance of the countries' adherence to the rule of law is of minor importance (.20; $p > 0.05$), and political stability is not relevant for the assessment (.14; $p > 0.05$). The year variable shows that election assessments of the conduct of the counting process are getting more negative over the years. As a next step, we conducted a multivariate analysis, including openness, stability and the year variable in a structural equation model with the voting and counting processes as the two dependent variables, using the software package AMOS (Analysis of Moment Structures). Apart from the expected effects based on the two separate multiple regression models, there is also a positive effect, albeit weak, of stability on the counting process (see Figure 1). The fit of the model is reasonable (Chi-square = 9.8; $df = 4$; $p = 0.044$), indicating that the model fits with the empirical data.⁴⁰ Furthermore, the results show that openness and stability contribute equally to the voting process (.43 and .41 respectively), whereas openness is evidently more strongly related (.66) to the counting process than is stability (.24).

The year variable indicates that the overall assessment for the voting process did not change over the time period, but that the counting process slightly deteriorated (-.25) from 1996 to 2010. This finding is not in line with the rather stable trends ($p > .30$) in the experts' ratings of the adherence to the rule of law ($R = .02$), the openness of the society ($R = -.00$), and in particular the finding that the countries became more politically stable ($R = .29$). The findings based on the specific countries in the 1996–2010 period show no clear downward or upward trend in the three democracy indicators in those countries where at least three EOMs took place. Georgia, Albania, and Macedonia are the exceptions. In Georgia, the rule of law (from -1.33 to -0.21), and political stability (from -1.65 to -0.96) improved gradually over the eight consecutive elections between 1999 and 2010, as did the openness (from -0.41 to -0.17), but this trend is not linear. In Albania, the three indicators improved from -0.84 to 0.11, from -0.92 to -0.35 and from -0.40 to -0.03 between 1997 and 2009. In Macedonia, as well, we also see an improvement between 1998 and 2009: from -0.29 to 0.11, from -0.52 to -0.27, and from -0.50 to -0.29. This improvement is also reflected in the correlations (Pearson's R) with the date of the elections, showing a clear upward trend based on the three joined countries ($n = 23$): 0.79 (openness), 0.58 (rule of law) and 0.29 (stability). We might, therefore, have expected an improvement of the voting and counting assessment. The results reveal, however, that such expectation is only true for the voting process (.43; $p < .05$). For the counting, the trend is not significant (.13; $p > .05$).

5. Conclusions and Discussion

This paper has provided an overview of the aggregated findings of over thirty thousand observers fielded in 20 OSCE member states from 1996 onwards. Out of 97 elections, a total of 86 elections were analysed in which at least 100 observers were present and

⁴⁰The differences between the data and the model are expressed by the Chi-square. The lower this value, given the degrees of freedom (df), the higher the p -value. Although the value does not reach the .05 level, the 0.044 shows that the fit is reasonable, i.e. that the model reflects the data well.

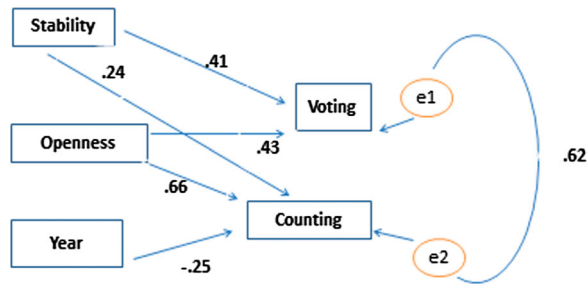


Figure 1. Impact of Stability, Openness and Year on Voting and Counting Process in 86 Elections, 1996–2010

information on the overall assessment of the voting and vote count was available in the OSCE/ODIHR final reports. It has been demonstrated that observers' assessments were rather negative in Tajikistan, Azerbaijan (both during the day and vote count), Belarus, Kyrgyzstan, and Kazakhstan (vote count). Such countries do not meet international standards of democratic elections. Apart from in Belarus, the assessments of the voting and counting processes are strongly correlated in the observed countries in the OSCE-region. Apparently, observers' findings in the morning and afternoon on polling day already serve as a good predictor for what will happen during the evening and night when it comes to the counting of the votes. On average, the voting process is assessed in negative terms, given either a 'bad' or a 'very bad', in 8% of observations, and the counting process in 21% of observations. Considering this gap in the observers' overall assessment between the voting and counting process, it is not particularly unusual for the OSCE/ODIHR to report that '... the process deteriorated during the count' in their post-election statements. Such an interpretation should take into account that in practically all EOMs the vote count is more negatively assessed than the voting process itself. A substantial part of this gap might be explained by the methodology. Observers stay longer, in fact, from start to finish observing the count observing the whole counting process, but spend only approximately 30 minutes on average covering the voting process. As a consequence, the cross-election results presented should be taken into account when it comes to the reporting of a deterioration or improvement. In the 2012 presidential elections in the Russian Federation, the voting process was assessed as bad overall in 5% of the polling stations covered. This result was below the median of 6.5% based on all 96 EOMs. Observers' findings during the counting process were considerably worse: 31% of the reports were negative and fell into the most negative quartile of vote count observations. The poor assessment of the counting process was used as evidence to support the concluding statement saying that '... the process deteriorated during the count'.⁴¹ Such statements on the quality and integrity of the electoral process generally are important signals to the domestic and international audience, and may even lead to regime change.⁴²

For further interpretation based on the findings of the international election observers, the overall assessment has been linked to other indicators retrieved from the World Bank governance measures. Stable trends in the rule of law, openness, and political stability were

⁴¹OSCE/ODIHR, 'International Election Observation Russian Federation' (n 1) 2.

⁴²Daniela Donno, *Defending Democratic Norms. International Actors and the Politics of Electoral Misconduct* (OUP 2013).

noted over the 14-year period. In fact, only three countries became more democratic according to the expert ratings: Georgia, Albania and Macedonia. In these countries, however, only the voting process and not the counting process improved.

In addition, modest to high correlations between the three democracy indicators and the overall assessment were found. Therefore, we might expect that, in general, democracy indicators based on experts' judgement are already rather 'good' predictors for what will happen during election day. The results show that it is especially the openness of a country that matters, considering the variable's strong relationship with the assessment of the voting and counting process, after controlling for various other indicators. Likewise, some support for the impact of stability on the election process was also found. For the rule of law, there was no relation found with the observers' evaluation of the voting and counting process.

This means that the first hypothesis (rule of law) is rejected, while the second and third (openness and political stability) are confirmed. Apparently, having political opportunity structures to voice political concerns via media and in associations, and having freedom of expression are crucial for good elections. If the openness of a society is safeguarded, the contribution of the adherence of the rule of law is rather limited. In addition, greater political stability contributes to an election process that meets international standards laid down in the Copenhagen Commitments. If democratic principles are complied with, it seems that an adherence to the rule of law has no added value for an election. The results may also indicate that the openness of the systems is more a prerequisite for a good count than are the other democratic variables measured during the election process. Transparency is crucial during the count. Indeed, coming to a conclusion as to whether counting the ballots and completing the protocols adhere to the appropriate rules and regulations is only possible in a transparent environment. In Belarus, for example, where observers were often not allowed to approach the tables where the counting took place, the overall assessment suffered following negative appraisals of the transparency in the counting station.

The practical inferences of these findings are twofold. The first concerns the recommendations the OSCE provides in their final reports, often referring to shortcomings in the election law and other aspects related to the rule of law, as well as the openness of the system in a country, including freedom of speech and functioning of the media. Furthermore, recommendations focusing on the reduction of tensions between political actors ultimately targeting at increasing the political stability. The findings suggest that recommendations aimed at enhancing political opportunity structures and political stability in the country are more relevant than those related to improving the rule of law. The second practical inference relates to what might be expected on election day. If the pre-election phase is marred with many problems and irregularities related to the freedom of speech and biased media, international observers might expect a problematic election day, and with the vote count in particular, as demonstrated in the 2012 elections in Russia. In addition, if many tensions and intimidation are observed before election day, observers will also face many problems in the polling stations.

This research also focused on the huge cross-country variation in the assessment of the elections, which ranged from 1 and 2% in the 2009 elections in Montenegro to 25 and 56% in the 2010 elections in Tajikistan. Obviously, in some countries, elections do meet the international standards of democracy more so than in others. Alongside Belarus, bad

elections are most often conducted in the Eurasia region, followed by the South Caucasus countries of Armenia, Azerbaijan and Georgia. Of course, such conclusions based on the findings of international observers on polling day can be challenged. Elections and their observation is not a one-day event. The pre-election period is also an important part contributing to making a full assessment of the elections, comprising assessments as to the balance of media coverage, the existence of a level playing field, and the transparency of campaign and party financing. This study did not take into account the aggregation of the result protocols in the tabulation centre, which has also been a part of the OSCE/ODIHR election observation methodology in recent years. In addition, this did not cover the post-election phase during which – among other aspects – the handling of appeals is observed. It is possible that the authorities organised a perfect election day, completely in line with international standards and the regulations in the election law, but the pre- and post-election day phases are marred with severe problems and irregularities.⁴³ Moreover, an excellent election is no safeguard for a democratic society either.⁴⁴

There are, obviously, more conceivable explanations as to why an election does or does not meet international standards. We also acknowledge that the authorities may have become more experienced in handling international observer teams over the years. More sophisticated ways to rig the elections might have been developed and put to use nowadays⁴⁵ and observers may not as yet have been capable of deterring such fraud and the authorities may even try to manipulate the EOM itself.⁴⁶ We could argue, however, that observers are also more experienced, and keener to detect violations on election day. We might also consider that other available indicators would be more appropriate than those retrieved from the World Bank. We looked at the polity4 indicators on political regime characteristics and transitions and concluded that they did not serve our goals very well.⁴⁷ Most of the indicators include components that are directly related to elections, and as such the World Bank measurements are more appropriate. However, it would be a fruitful alley for further research to explore the relationships with other democracy measures, including the over 350 indicators developed within the Varieties of Democracy.⁴⁸ The V-Dem programme is a rather new approach to conceptualise and to measure democracy, and is related to distinct measures of electoral, liberal participatory, deliberative and egalitarian democracy, as well as to more-specific components of democracy.⁴⁹ A further research line, elaborating on earlier work,⁵⁰ would be creating a

⁴³For an overview of the electoral cycle, see Jørgen Elklit, 'Electoral Institutional Change and Democratization: You Can Lead a Horse to Water, but You Can't Make It Drink' (1999) 6(4) *Democratization* 28; Suksi (n 13); Jørgen Elklit and Andrew Reynolds, 'A Framework for the Systematic Study of Election Quality' (2005) 12(2) *Democratization* 147; Jørgen Elklit and Andrew Reynolds, 'Judging Elections and Election Management Quality by Process' (2005) 41(3) *Representation* 189; Norris, 'The New Research Agenda' (n 4); Pippa Norris, *Why Electoral Integrity Matters* (Cambridge University Press 2014); Van Ham (n 4).

⁴⁴Andreas Schedler, 'Elections Without Democracy: The Menu of Manipulation' (2002) 13(2) *Journal of Democracy* 36; Andreas Schedler, *The Politics of Uncertainty. Sustaining and Subverting Electoral Authoritarianism* (OUP 2013); Jonathan Hartlyn and Jennifer McCoy, 'Observer Paradoxes: How to Assess Electoral Manipulation' in Andreas Schedler (ed), *Electoral Authoritarianism: The Dynamics of Unfree Competition* (Lynne Rienner 2006); Staffan Lindberg, *Democracy and Elections in Africa* (Johns Hopkins University Press 2006); Munck (n 4).

⁴⁵For an overview of various types of fraud, see Lehoucq (n 4); Hyde, *The Pseudo-Democrat's Dilemma* (n 2); Alberto Simpser, *Why Governments and Parties Manipulate Elections: Theory, Practice, and Implications* (Cambridge University Press 2013).

⁴⁶Calingaert (n 4).

⁴⁷For discussions, see Cheibub, Gandhi and Vreeland (n 33); Munck (n 4); V-Dem (n 32).

⁴⁸V-Dem (n 32).

⁴⁹ibid.

⁵⁰Schmeets, *Vrije en eerlijke verkiezingen* (n 15).

new Standard Election Observation Voting Index and a Standard Election Observation Counting Index based on over 100 EOMs. Such approach will enable the validity and reliability of the two overall opinion indicators in more detail and will cover more elections, in more countries in the OSCE-region and within a longer period. Furthermore, it would be useful to include the more recently conducted EOMs in the analyses. Adding some 25 elections, would show whether the results still hold if the 15-year timespan is extended by 6 more years. However, as the benchmark and the correlation between voting and counting hardly change by including more elections, and the governance indicators are rather stable, I do not expect that the relationships will yield in other conclusions.

The elections are also often assessed by other organisations, such as the CIS and various NGOs. A comparison of these findings and those from the OSCE would be another research avenue. One specific explanation for the disagreement between the international organisations is possibly the fact that the OSCE is covering all election cycles, while the parliamentary groups often observe only the election day. This may particularly lead to discrepancies in the assessment of the various organisations when it comes to the pre- and post-election events. Apart from the coverage of the whole process, the methodology used by the organisations may possibly explain, at least to some extent, the judgement of what happened on election day. An observer report form guides observers through the process and as such will have an impact on observers' findings. For many years the forms used by OSCE/ODIHR were to a large extent standardised, while this was, e.g. for the EU EOMs, not the case. Comparing elections by various observer groups, ultimately calls for a harmonised, standard approach.

It has been demonstrated that some of the countries studied still have a long way to go in their transition towards a democracy that meets international standards and commitments. Yet, many international election experts in the 1990s, including the Warsaw-based OSCE/ODIHR, were wondering what they might do after EOMs were no longer needed in these countries.⁵¹ In 1997, Carothers predicted that EOMs would stop within a decade.⁵² At present, more than two decades later, there is no discussion about the end of election monitoring. On the contrary, as shown in this paper, the member states are willing to send more and more observers, and at least until 2010, were willing to do so even taking into consideration that practically all states have to cover the costs of their observers themselves. The capability of having such a large number of observers on the ground is a huge advantage for the OSCE/ODIHR in comparison to most other international organisations, such as the EU and OAS, who deploy remarkably fewer observers. EOMs with a low number of observers suffer from reliable statistical information of what was seen and heard on the ground. The margins of the findings, such as the overall assessment indicators, are substantial in small EOMs, as the standard errors increase if the number of observations drop. Considering these margins around the estimates, i.e. the 95% confidence intervals, it is advisable to deploy a minimum of 400 to 500 observers to cover the consecutive steps in the observation. Such numbers will result in a minimum of 200 to 250 observations for the vote count, enabling to make assessments with an acceptable level of accuracy and to use the benchmark as a tool providing a solid quality assessment of this crucial part in the election process.

⁵¹Peter Eiger, 'Improving OSCE Election Observation' (2009) 4 Security and Human Rights 264.

⁵²Carothers (n 11).

Following this argument, the ODIHR should request such STO-numbers for each EOM, a practice which at the current time is not at all common. The political dimension certainly plays a role when it comes to the number of observers requested for an election. In 2014, the OSCE/ODIHR asked 900 observers to observe the Early Presidential Election on 25 May in Ukraine. The member states sent some 1200 observers, a number far beyond what had been demanded. It shows that the need for more EOMs and observers is widely acknowledged. In some countries this need is greater than in others, as has been demonstrated in this paper.