Pre-print version.

http://dx.doi.org/10.1016/j.electstud.2015.11.003 0261-3794/

Blais, A., & Galais, C. (2016). Measuring the civic duty to vote: A proposal. *Electoral Studies*, *41*, 60-69.

Measuring the civic duty to vote: A proposal

Andre´ Blais [a 1,] , Carol Galais [b, *]

[a] *Universite de Montreal, C.P.6128, Succ. Centre-ville, Montreal, H3C 3J7, Canada*
[b] *Universitat Oberta de Catalunya, Av. Carl Friedrich Gauss, 5, Parc Mediterrani de la Tecnologia, 08860, Castelldefels, Barcelona, Spain*
*Corresponding author.
E-mail addresses: Andre.blais@umontral.ca (A. Blais), carolgalais@gmail.com (C. Galais).

Abstract

Civic duty is a central concept in the study of turnout, yet little attention has been paid to how it should be measured. After a careful review of previous measures we constructed an original battery of 13 questions that were administered in a survey conducted in seven countries at the time of the 2014 European election. We show that the battery indeed taps the duty construct. We then propose a reduced battery of four questions. We show that the four questions achieve good fit measures and pass several tests of robustness and validity across the seven countries. We invite researchers to implement this battery in future research.

1. Introduction

From a purely utilitarian perspective, voting does not appear to be a rational choice given the extremely low probability that one's decision will be pivotal (Owen and Grofman, 1984; Mueller, 2003). Yet most people vote, which is known as the paradox of voting (Fiorina, 1976; Grofman, 1993). Scholars have tried to solve this paradox by manipulating the costs, benefits and probabilities of one's vote of being decisive, but one of the most widely accepted solutions is to include a normative element, the D term (Dowding, 2005).

From a normative perspective, people vote not because they calculate that the benefits outweigh the costs but because they consider that this is the 'right', 'ethical' thing to do (Blais and Achen, 2010). This is not a new interpretation. In the classic American Voter, Campbell et al. (1960: 105e106) note that turnout is 70% points higher among those with a strong sense of duty than among those with none. In the same vein, Riker and Ordeshook (1968) show that duty (the D term) has a strong impact on the propensity to vote. More recently, Verba et al. (1995: 115) report that civic gratifications, among them civic duty, are the most widespread motivation for voting. Blais (2000: 112) concludes that duty is the overriding motivation for about half of those who vote. Finally, Clarke et al. (2004: 259) find that the variable with the largest effect on turnout is what they call "system benefits", which is analogous to the duty to vote.

In spite of its predictive power and popularity among political behavior scholars, the concept remains fuzzy. The danger exists that it becomes a hodgepodge for all the psychological determinants outside the rational choice perspective. As a result, little attention has been paid to how it should be measured. We hope to fill that gap in this article through a survey conducted in seven European countries which included 13 different questions designed to tap the belief that it is a citizen's moral obligation to vote in a democratic election.

The paper has four goals. First, we propose a clear and simple definition of civic duty and we outline its implications. Second, we review the indicators used in previous research and evaluate their merits and limits. Third, we propose a long (13-question) battery of duty indicators and we show that it satisfies a number of conceptual and empirical criteria. The empirical evidence suggests that the assumption of one underlying concept is justified. Fourth, we propose a reduced battery of four questions that could be used in future research. The reduced battery is submitted to different tests of validity. The results confirm that these four questions can indeed be used to tap the duty to vote.

2. What is the "duty" to vote?

According to the literature, there are three main motivations for casting a ballot (Carlsson and Johansson-Stenman, 2009). First, people may vote for instrumental reasons, that is, with the intention to affect the outcome (Black, 1948; Downs, 1957). Second, individuals may vote for expressive reasons, that is, to express their views (Brennan and Buchanan, 1984; Brennan and Lomasky, 1993; Brennan and Hamlin, 1998, 2000). Finally, some people vote because they believe that the good citizen has a moral obligation to vote and thus not voting is ethically wrong (Tullock, 2000). This last motivation is the one we are interested in. *The civic duty to vote is the belief that a citizen has a moral obligation to vote in elections*.
This definition has a number of implications. Perhaps the most obvious is that civic duty is not instrumental or expressive. Duty does not refer to benefits or costs, nor to the potential outcome of an election. It is not a desire to articulate one's views about the options (parties, candidates and/or the issues). Duty is different. The motivation is *moral*; the dutiful person believes that voting is the right thing to do and abstaining is wrong (Uhlaner, 1986; Mueller, 1989; Coleman, 1990; Knack and Kropf, 1998; Zuckerman and Kotler-Berkowitz, 1998; Blais, 2000; Blais and Achen, 2010).

This definition of civic duty provides a number of guidelines about how duty should (and should not) be measured. The dutiful person should feel compelled to vote even in the absence of instrumental or expressive reasons, and should dismiss instrumental or expressive 'reasons' for not voting. She thinks in ethical terms, such as 'right', 'wrong', 'OK', 'should', 'good' and 'bad'. She should feel 'good' when she votes (and fulfills her duty) and 'bad' otherwise. More precisely, she should feel guilty if she does not fulfill her civic obligations (Knack, 1992) without good cause. Indeed "guilt is usually thought of (…) in connection with the ethics of duty" (Greenspan, 1994: 57). Since she is convinced that voting is the right thing to do, she thinks about voting as a common good, and she cares whether other people also do the right thing, that is, vote.

The literature suggests that sense of duty stems from group loyalty and/or respect for authority (Graham et al. 2011). Hence, the feeling that one has a moral obligation to vote may derive from attachment to the community or respect for the authorities. The most obvious case is when people believe that they need to prove their patriotism by participating in elections. According to Usher, voting can be construed as a patriotic gesture, reflecting "a willingness to participate in the ceremony of democracy, to take one's place in the parade on which most people sense a good society depends" (2011: 23).

Likewise, those who link the duty to vote to support for democracy reason that a vibrant democracy requires a high turnout and that all those who believe that democracy is a 'good' thing should be willing to do their part, that is, they should vote (Downs,1957; Mueller, 1989, 2003, Fowler and Kam, 2007). Similarly, Dennis relates civic duty to diffuse support for the regime, defining it as the citizen's feeling of obligation 'to contribute his own resources of time and effort even when particular elections are anticipated to be unfavorable or trivial to his own interests' (1970: 63). Finally, Dalton (2008) defines citizen duty as adherence to the traditional responsibilities of a good citizen, such as obeying the laws, paying taxes, and

voting.

These guidelines inspire the selection of questions that we introduce below. Before doing so, however, we review previous measures of civic duty.

3. Previous measures of civic duty

4.

The duty to vote has been often interpreted as a citizenship norm (Dalton, 2008). The most commonly asked question is: “There are different opinions as to what it takes to be a good citizen. As far as you are concerned personally, on a scale of 1e7, where 1 is not important at all and 7 is very important, how important is it to always vote in elections?” Many surveys include this question, such
as the first wave of the European Social Survey (2002), the World Values Survey, the Citizens, Involvement and Democracy survey (2000e2001), the General Social Survey (U.S. 2004) and the International Civic and Citizenship Education Study (ICCS).2

The reference to the “good” citizen clearly fits the normative component of duty. The major drawback is that this question asks about people's perceptions of the public norm of citizenship rather than about whether they personally support that norm. As a consequence, there is a strong desirability bias, as the great majority of respondents say that it is very important for the good citizen to vote. Social desirability is the “tendency to respond in self-report items in a manner that makes the respondent look good rather than to respond in an accurate and truthful manner” (Holtgraves, 2004: 161). Of course, a question using value-laden words such as “good”, “should” or “duty” is more prone to be loaded with social desirability. Therefore, correctly measuring sense of civic duty is a huge challenge, once it is recognized that there exists a public norm that the good citizen should vote in an election, and that as a consequence respondents may give voice to the norm even if they have not internalized it.

Previous research has not paid enough attention to the risk of social desirability bias, and the measurement errors that are entailed. Because of measurement error, there is a greater need for the use of several indicators. In fact this should be the case of any attitude. If we believe that sense of civic duty is a powerful determinant of the decision to vote, we need to develop a battery of indicators, as we do for internal and external efficacy, political trust, or cynicism. As Blalock (1974: 5) told us years ago, “there must be a reliance on more than a single measure of each variable.” This is at the heart of the psychometric theory of measurement, and a large number of studies insist on the importance of using several indicators for tapping attitudinal dimensions (Ansolabehere et al., 2008; Carmines and Zeller, 1979; McIver and Carmines, 1981, Nardo et al., 2005). It is part of a large consensus among method- ologists that in order to properly measure an underlying latent construct ei.e., ensure identification, achieve proper solutions, estimate and therefore avoid latent errors-a minimum of three or four questions are needed per dimension (Kline, 2005, 2011; Baumgartner and Homburg, 1996; Marsh et al., 1998; Ding et al., 1995; Bollen, 1989).

Following this logic, the American National Election Studies (ANES) used, until the late seventies, a battery of four questions intended to tap sense of civic duty.[3] The questions were agree/ disagree statements:

- It isn't so important to vote when your party doesn't have a chance to win.
- A good many local elections aren't important enough to bother with.
- So many people vote in national elections that it does not matter much whether I vote or not.
- f a person doesn't care how an election comes out he shouldn't vote it.

What is most striking here is that the word 'duty' is never used. The approach is indirect, the ANES offering statements that present potentially acceptable reasons for not voting, statements with which those who believe that voting is a moral obligation should disagree. Avoiding a value laden word like duty may be an advantage but it may also be a disadvantage to the extent that people spontaneously think in terms of moral obligations (Blais, 2000). The agree/disagree format used here is also a source of concern. This format is particularly susceptible to social desirability and acquiescence biases (Krosnick, 1999; Schuman and Presser, 1981; Saris et al., 2010). It must be noted, though, that the bias here is les serious since one has to express disagreement in order to be classified as a dutiful person. Finally, the use of negatives (and double negatives in two instances) is particularly problematic (Foddy, 1993)

The author who has most extensively examined sense of civic duty is Blais (2000). He has relied on the following questions, the first five being agree/disagree statements:

1. It is the duty of every citizen to vote
2. It is important to vote even if my party or candidate has no chance of winning.
3. In order to preserve democracy, it is essential to vote.
4. In order to preserve democracy, it is essential that the great majority of citizens vote.
5. If I did not vote, I would feel guilty.
6. If you did not vote, would you feel that you had neglected your duty as a citizen enormously, a lot, a little, or not at all?
7. If you did not vote, would you feel very guilty, somewhat guilty, not very guilty, or not guilty at all?

Democracy emerges here as an important benchmark. The reference to a feeling of guilt is novel, but it is appropriate since people who think in terms of 'right' and 'wrong' should feel guilty if and when they do the 'wrong' thing (see Knack, 1992). The last three questions are interesting in that they invite respondents to think about how they would feel if they were not to vote, that is, if they were not doing the right thing (if you believe that you have a civic duty). Nevertheless, the widespread use of agree/disagree statements is still worrisome. In combination with value-laden words such as duty, guilt and the like, this increases the risk that respondents just pay lip service to the public norm that it is the duty of every citizen to

vote in an election.

More recently, Blais and Achen (2010) made the case for a completely new question: "Different people feel differently about voting. For some, voting is a DUTY. They feel that they should vote in every election however they feel about the candidates and parties. For some, voting is a CHOICE. They feel free to vote or not to vote in an election depending on how they feel about the candidates and parties. For you personally, is voting in an election first and foremost a duty or a choice?" There is a follow-up for those who answer 'duty': "How strongly do you feel that voting in an election is a duty: very strongly, somewhat strongly, or not very strongly?" The question (with the follow-up) has been asked (sometimes with minor modifications) in surveys conducted in Canada, the U.S., Ireland, the Netherlands, Spain, and Taiwan. There is an antecedent in a National Opinion Research (NORC) survey question asked in 1944: "Do you regard voting more as a duty you owe your country or more a right to use if you want to?"[4]

The Blais/Achen question asks directly about duty, it includes normative words like 'should' and explicitly invites respondents to contrast with an expressive motivation (how one feels about the candidates and parties) to vote or abstain. The main justification for this new question is a concern to avoid the social desirability bias associated with the presence of a public norm as well as the acquiescence effect linked to agree/disagree statements. The authors were keen to have a positive 'non duty' option. Indeed, the survey methodology literature suggests that a way to reduce social desirability is suggesting the normalcy of the socially deviant behavior (Oppenheim,1992; Brace, 2004). This is an interesting and useful approach though it is doubtful whether it is appropriate to attempt to measure duty (or any attitude for that matter) through a single indicator.

We seek to fill a big gap in the literature on the attitudinal dimensions related to voting behavior by proposing a battery of questions to tap citizens' belief about whether they have a moral obligation to vote in elections. For this purpose, we first came up with 13 different questions designed to tap civic duty in a more comprehensive manner, paying close attention to the flaws out- lined above, taking great care to minimize the social desirability bias associated with the presence of a public norm. In the next pages we present the 13 duty indicators to capture the underlying construct. Next, we suggest a reduced battery and we test its validity with an original survey conducted across seven European countries.

4. Research design

We were able to include our battery of 13 different questions to measure civic duty in a cross-national internet survey that was conducted immediately after the 2014 European Parliament election in seven different countries: France, Germany, Italy, Spain, Austria, Greece, and Portugal.5 They are all West European countries, but at different degrees of economic development and democratic consolidation.6 The sample sizes go from about one thousand individuals in Portugal and Greece to about four thou- sand in the other countries.7 The design of the battery was inspired by the following criteria. First, special attention was given

to the need to minimize the social desirability bias. Second, questions should reflect the conceptual definition of duty outlined above. Third, and in tension with the previous criterion, the battery should tap different aspects of duty, that is, "the pool of items … should try to do justice to the multiple facets of an attitude domain."(Oppenheim, 1992: 166). Hence, some of the questions are clearly framed in moral terms, that is, they contain words such as 'should' while others are more indirect as they tap views that are related to the sense of duty. Fourth, question wording should be as simple as possible (Robinson et al., 1968: 10). With these criteria in mind, we developed the battery of questions listed in Table 1.

The first question is the duty/choice (CHOICE) question suggested by Blais and Achen, which is about whether the respondent construes voting as a matter of moral duty or personal choice. This question is a bit long but goes directly to the point, and is particularly designed to avoid social desirability by offering a very positive and acceptable non duty option (voting is a matter of personal choice). The second is about the feeling of guilt (GUILT), an emotion that one should experience if she were not to do the right thing. The next is also a feeling (CARE) that one is prone to have if she believes that there is a moral obligation to vote: she would very much like (and hope) that friends and relatives share her view. These three questions are not agree/disagree statements.

Three statements offer possible justifications for a duty to vote which refer to the object towards which one feels an obligation. One is about showing love for one's country (COUNTRY), the second is about showing support for democracy (DEMOCRACY), and the third is about showing loyalty to one's party (PARTY). Some people may feel a duty towards each of these three objects (country/democracy/party), though we suspect that parties are a more con- tested institution.[8] But party identification captures a psychic feeling that goes well beyond one's preference in a given election (Fiorina, 1976).These three questions are short, clear, and simple.

Three statements provide justification for not voting, statements with which the dutiful person should disagree with. The first is the enunciation of a general principle of the right to abstain (RIGHT, see Lomasky and Brennan, 2000). The other two statements refer to circumstances which would appear to justify abstention, and which should strike the dutiful citizen as weak, unacceptable excuses for not fulfilling one's moral obligation. The first is that it is OK to abstain in unimportant elections (UNIMPORTANT) and the second is that those with no opinion have no reason to vote (OPINION).

These are also agree/disagree statements, but the risk of social desirability is reduced since one has to disagree with the statement in order to score high on civic duty.

Three other questions tap duty in a more indirect fashion. The first is that "voting is just like paying taxes, you just have to do it" (TAXES). The statement picks up another dimension of civic duty, that is, that some people may think that there are many things that you have to do in life whether you like them or not while others may believe that so called duties should always be called into question. It also indirectly refers to the state as a source of authority from which the social norm about voting may stem. The second statement is even more

indirect. It is that "only those who vote have the right to criticize the government" (CRITICIZE, see Lomasky and Brennan, 2000). Those who agree with the statement are saying that they have little sympathy for those who abstain, possibly because they believe that abstaining is wrong. And indeed one previous study found that responses to this statement loaded on a civic duty factor (Blais and Young, 1999). The third refers to compulsory voting (COMPULSORY). The normative argument behind compulsory voting is that "voting is a necessary attribute of citizenship; it is a public trust, a moral obligation, a duty of citizenship" (Birch, 2009: 42). If voting is a moral duty, why not make it a legal responsibility?

We finally include a straightforward assertion that it is everyone's duty to participate actively in politics (PARTICIPATE). This is a clear reference to civic duty though duty is here defined more broadly as one of active political participation, not strictly confined to electoral participation. It is not clear whether civic duty is construed in narrow or broad terms. If Dalton (2008) is right and there are two opposing conceptions of duty, then responses to this question may be at variance with the other twelve questions which are all dealing exclusively with voting and elections.

We believe that these 13 questions capture the core of civic duty, that is, the belief that voting is morally "good" and abstaining "wrong", that voting is a public good, and thus that there is an ethical obligation to vote in an election. At the same time, each question taps a specific aspect, either directly or indirectly, positively or negatively, with a special emphasis on different considerations and in contrast with instrumental or expressive motivations.[9]

We also strove to have a variety of formats in the battery of questions. Nine of the 13 questions are agree/disagree questions, which are easier for respondents but more susceptible to acquiescence bias (Krosnick, 1999; Schuman and Presser, 1981; Saris et al., 2010), and four are not.10 Among the agree/disagree statements, three are framed in such a way that one has to disagree in order to be considered dutiful. The questions are short, but one (CHOICE), which we use as our starting point, is longer because it contrasts two opposite (and respectable) conceptions of voting. Some measure duty directly while others are indirect questions. Nine of the 13 questions are agree/disagree questions, which are easier for respondents but more susceptible to acquiescence bias (Krosnick, 1999; Schuman and Presser, 1981; Saris et al., 2010), and four are not.10 Among the agree/disagree statements, three are framed in such a way that one has to disagree in order to be considered dutiful. The questions are short, but one (CHOICE), which we use as our starting point, is longer because it contrasts two opposite (and respectable) conceptions of voting. Some measure duty directly while others are indirect.

As we are concerned with measuring an attitudinal dimension -the duty to vote-, our research fits into the category of construct validity studies, which assess the accuracy of an instrument used to tap a concept (Linn, 2000; Stewart-Brown et al., 2009). In this respect, confirmatory factor analysis (CFA) does better in testing both the validity and reliability of an instrument than other traditional analytic methods (e.g. correlations or multiple regression) because it accounts for measurement error and yields more parsimonious results (Said, 2011; Brown and Moore, 2012.) Moreover, CFA allows for evaluating the equivalence of measurement models

across distinct groups, testing for measurement invariance and yielding powerful evidence regarding external validity. In sum, CFA determines to what extent a series of observed indicators are related to an unobserved, latent, factor while eliminating measurement error and verifying the dimensionality, reliability and construct validity of a measurement model (Brown, 2006).

Hence, the next section assesses the appropriateness of the battery of 13 questions aimed to tap civic duty by means of a CFA. Subsequently, a reduced version of the battery is presented to gain parsimony while still limiting social desirability and maintaining internal consistency and item heterogeneity. This translates into keeping the "best" indicators -those with highest factor loadings- while tapping diverse aspects of civic duty and limiting the number of agree/disagree questions, especially when agreement ethe easiest option-indicates duty. This second -collapsed- measurement model will also be tested for convergent validity using CFA. In addition, we will perform predictive and concurrent validity tests.[11] With these purposes, we first use the reduced civic duty measurement model to predict turnout in the 2014 European election in each of the 7 countries. Second, we explore how the duty construct relates to other relevant variables. More specifically, we predict duty to vote using age and interest in politics (two important correlates of turnout) as well as a socialization variable that poses the question as to whether the respondent's parents told her that voting was a duty when she was a child. These analyses are replicated in the seven countries under study.

Finally, we test the measurement invariance of the collapsed version of the battery by means of a multiple group confirmatory factor analysis (MGCFA, Davidov et al., 2011, Baumgartner and Steenkamp, 1996). MGCFA ascertains whether factor loadings and intercepts are the same across groups, in this case countries, whether there is response bias in a group and, ultimately, if the same construct is being measured in all groups and is equally helpful in predicting turnout.

Table 1
Thirteen duty indicators.

Question wording	Answer options	Abbreviation
Different people feel differently about voting. For some voting is a duty. They feel that they should vote in every election however they feel about the candidates and parties. For some, voting is a choice. They feel free to vote or not to vote in an election depending on how they feel about the candidates and parties. For you personally, is voting in an election first and foremost a duty or a choice? (if duty) How strongly do you feel that voting in an election is a duty:	0 = choice/don't know, 1 = Duty: not very strongly 2 = Duty: somewhat strongly 3 = Duty, very strongly	CHOICE
How guilty would you feel if you did not vote in an election?	0/I would not feel guilty at all/10 I would feel extremely guilty	GUILT
How much do you care whether your friends and relatives do or do not vote?	0 I do not care at all/10 I care a great deal	CARE
I see voting as a way to show support for democracy	0 fully disagree/10 fully agree	DEMOCRACY
I see voting as a way to show love for my country	0 fully disagree/10 fully agree	COUNTRY
I see voting as a way to show loyalty to my party	0 fully disagree/10 fully agree	PARTY
In a democracy, people should have the right to vote, but also the right to abstain	0 fully disagree/10 fully agree	RIGHT
It is OK to abstain in unimportant elections	0 fully disagree/10 fully agree	UNIMPORTANT
It is OK to abstain if you have no opinion in an election	0 fully disagree/10 fully agree	OPINION
Voting is like paying taxes, you just have to do it	0 fully disagree/10 fully agree	TAXES
Only those who vote have the right to criticize the government	0 fully disagree/10 fully agree	CRITICIZE
In some countries, voting is compulsory. How favorable or opposed are you to make voting compulsory in your country?	0 I am totally opposed/10 I am totally favorable	COMPULSORY
It is everyone's duty to participate actively in politics	0 fully disagree/10 fully agree	PARTICIPATE

4. Results

4.1. *Construct validity of the extended version of the duty battery*

We first estimate a model to measure civic duty using all our 13 questions. In order to achieve identifiability (that is, the number of unknowns should not exceed the number of equations) we need to set a regression weight to one for a reference indicator, which in our case is CHOICE. The reasons are several: CHOICE represents the most serious attempt to avoid social desirability when tapping the duty to vote, its format is not an agree/disagree question, and it explicitly mentions the word "duty". Fixing its measurement weight to 1 gives this factor the same variance as the common variance portion of the measured variable, although it does not affect the relative loadings or the overall model fit (Maruyama, 1997).

After conducting multiple bayesian imputations, we estimated the parameters by means of an asymptotically distribution-free model.[12] The resulting unbiased estimates estandardized coefficients-are displayed in Fig. 1. The coefficients are all significant ($p < .00$).[13]
The model fit measures are satisfactory. The root mean squared error of approximation (RMSEA) should ideally be lower than .05 to indicate a close fit of the model to the data, between .05 and .08 to indicate a reasonable error of approximation, or at least lower than .1 to be considered acceptable (Browne and Cudeck, 1993:144; Steiger, 1989).[14] Although the comparative fit index (CFI) is not close to the ideal value of 1 -suggesting that there is room for improvement-, the RMSEA value (.78) is within the limits of what is customarily accepted as a reasonable error of approximation. In the same vein, the Cronbach's Alpha obtained when we use all the 13 indicators to construct an additive scale is quite good (.87), although the item-specific statistics suggest that the scale would improve even more if we dropped RIGHT, PARTY and UNIMPOR- TANT.[15] This is consistent with the low measurement weights for the three indicators depicted in Fig. 1.

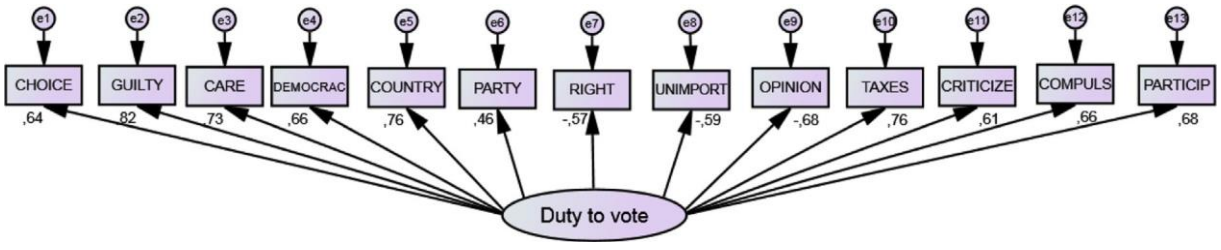

Fig. 1. Full (13 indicators) model. Standardized regression weights. Degrees of freedom ¼ 65; X^2 ¼ 87010.02; Probability level ¼ .000, RMSEA ¼ .078, CFI ¼ .676. N ¼ 22,072.

Our next task is to propose a shorter battery of questions.

4.2. *Constructing a short battery*

How many indicators should the reduced version include? The literature suggests that at least four indicators per latent construct are necessary to ensure identification, achieve proper solutions, and estimate latent errors (Baumgartner and Homburg, 1996; Bollen, 1989). This short version of a duty battery should ideally reflect the great variety of formats presented so far, because "the greater the variety of questions, of course, the wider the generalization which can be made about the coverage of a particular scale" (Stouffer, 1950:140).

In the process of selecting these four items we aimed to maintain internal consistency while maximizing item heterogeneity in both content and format and minimizing the social desirability bias. First of all, we decided to keep CHOICE because it is the most direct indicator (it contains the word "duty"), it is not an agree/disagree question, and it has been explicitly designed to minimize social desirability. More- over, its standardized regression weight in the full model (Fig. 1) is higher than .6, which is the threshold for considering that an indicator is aligned with a factor (Rakowski et al., 1997).

Next, we wanted to keep at least one of the three other questions that are not agree/disagree statements (CARE, GUILT, COMPULSORY) in order to avoid having more than two agree/ disagree statements in our short battery. We dropped CARE and COMPULSORY because they have lower measurement weights and kept GUILT, thus satisfying the internal consistency criterion. As stated before, guilt is a central emotion associated with the failure to fulfill one's duty.

With CHOICE and GUILT we already have two of the four questions to be included in the short battery. Among the nine agree/ disagree statements, six are positive in the sense that agreeing with the statement means having a higher score on duty and three are negative, whereby an agreement entails the absence of duty. Given our concern with social desirability, we were keen to keep no more than one positive statement. In order to arrive at four indicators as well as maintaining as much consistency as possible, we selected the "best" positive and the "best" negative agree/disagree statement to complete our short list. Among the positive statements the one with the highest measurement weight is COUNTRY, which has the added advantage of tapping a widespread justification for civic duty ('I owe this to my country' as the old NORC question put it). Among the three negative statements, the question with the highest loading is OPINION, which has the additional advantage of contrasting with an expressive view of voting (if elections are seen as an opportunity for people to express their opinions there is no reason to vote if one has no opinion to convey).

The short battery of questions thus consists of: CHOICE, OPINION, GUILT, and COUNTRY. As shown in Fig. 1, all four indicators have relatively strong loadings on the underlying construct. As regards format heterogeneity, one of the questions (CHOICE) is direct and has been explicitly designed to reduce social desirability. Two of the questions (CHOICE and GUILT) are not agree/disagree statements and one of the agree/disagree statements (OPINION) is reversed, a higher duty score being associated with disagreement with the statement. The questions thus tap diverse aspects of duty by means of diverse

question formats.

We believe that the reduced scale provides a satisfactory coverage of the meaning of civic duty to vote. We have one question (CHOICE) that asks directly whether the person feels that she has a duty to vote in all elections. We have one question (GUILTY) that taps the emotion that one should feel if she fails to live up to the norm that she adheres to. The third question (COUNTRY) deals with an underlying attitude (support for the community) that may motivate people to conclude that they have a duty to vote. The fourth question (OPINION) allows us to determine whether people subscribe to an expressive view of voting (you vote to express your views and so you don't vote if you have no views), which is the most widespread alternative to the ethical perspective entailed by the belief that one has a moral obligation to vote.

The combination of these items allows us to achieve our goals of obtaining sufficient internal consistency and sufficient item het- erogeneity while reducing the risk of social desirability.

4.3. *Validating the short battery*

The following step consists of testing the unidimensionality of this short battery by means of a CFA analysis. We estimated this reduced model, the standardized regression weights of which are displayed in Fig. 2.

The model meets our parsimony goal, with only four questions. Besides, our model fit measures improve; the RMSEA is now .069. Still, the chi-square is significant, but the relative chi-square (that is, the chi-square index divided by the degrees of freedom) has clearly improved, from 1338.6 (Fig. 1) to 1064 (Fig. 2). Finally, the Comparative Fit Index (CFI) is now clearly beyond the threshold of .95 and closer to the ideal value of 1.16

In the next step of our analysis, we wish to determine whether this latent construct has a similar association with turnout across countries. For this, we include turnout in the model and test it separately in each country. We add voting behavior in the 2014 European election (1 ¼ voted/0 ¼ abstained) as an observed dependent variable to the estimation, and test both the measurement and the structural models separately in each country (Brown, 2006: 269). This is aimed to test the predictive validity of civic duty. The results are displayed in Table 2.

The measurement and regression weights reveal some heterogeneity across countries. OPINION is less related to the duty construct in Greece, Italy and Spain. The duty latent construct exhibits an uneven relationship with turnout across the seven countries under study. The standardized regression weights shown in Table 2 suggest that the duty measure is more successful at predicting turnout in Germany, Portugal and Austria than in Greece or Italy.17 All in all, the battery seems to work somewhat better in the north than in the south of Europe. Nevertheless, the model achieves a good model fit in every country, confirming its robustness. The RMSEAS, for instance, are excellent to good, never exceeding the threshold of .08. The CFI, in turn, is never lower than .94, and sometimes even exceeds .95, especially

in France and Germany.

An additional way to check the validity of our duty measure is to analyze its relationship with other relevant variables.[18] With this aim, we have estimated the duty to vote (measured using an additive scale built using the four aforementioned indicators) in our seven European countries. The results are displayed in Table 3.

The three variables included in the estimation manage to predict between 15% and 26% of the variation of the phenomenon under study, being especially successful in Germany and less so in Italy or Portugal. According to Blais (2000), the older and those more interested in politics should exhibit higher levels of duty to vote. It is indeed the case for interest in politics, with very similar effects everywhere. We see some cross-country variations for age, which has a weaker effect in Italy and no significant effect in Portugal. But perhaps the most interesting variable here is the one referring to political socialization. One of the survey questions asked: "When you were a child, did your parents tell you that voting is a duty?" The answer options have been recoded as yes = 1 and no = 0. The duty scale should correlate strongly with this variable referring to the individual primary socialization. The results for this variable point to a positive, significant effect in all our 7 countries, although the effect is almost twice as strong in Ger- many as in Portugal. In spite of these cross-country differences, that again point to consistent specific patterns for the southern Euro- pean countries, we can conclude that the sources of the duty to vote are reasonably similar in the countries under study.

We can further ascertain the robustness of the model displayed in Table 2 by testing it for measurement invariance through multiple group confirmatory factor analysis (MGCFA). This procedure allows us to check if the model is able to properly explain the relationship between the indicators and the duty construct in each of the seven countries, that is, whether path coefficients (measurement and regression weights) and intercepts are invariant across countries. There are three levels of measurement invariance: configural, metric and scalar. Configural invariance is achieved when the same pattern of fixed and free factor loadings and regression coefficients is found across groups, without equality constraints. In other words, when the items exhibit the same configuration of salient and non-salient factor loadings across different groups, and so we can conclude that the underlying dimension exists in all groups and can be measured using the suggested indicators. Metric invariance is about demonstrating that all items have similar factor loadings (slopes) everywhere. When we add the constraint that the intercepts for any item should be the same in all the groups we are testing scalar invariance. When scalar invariance is confirmed we can safely compare factor means; otherwise, differences between two groups could be due to systematic biases in how citizens from different countries respond to some items (Steenkamp and Baumgartner, 1998).

When some parameters, that is, weights or intercepts, are not invariant for some items, it is possible to test for partial metric and/ or scalar invariance, freeing the parameters of those items that are not invariant across groups. According to some scholars (Byrne et al., 1989; Steenkamp and Baumgartner, 1998), full metric invariance is not necessary to safely compare factor means across groups, as long as one item in the construct, besides the reference item, is

invariant.

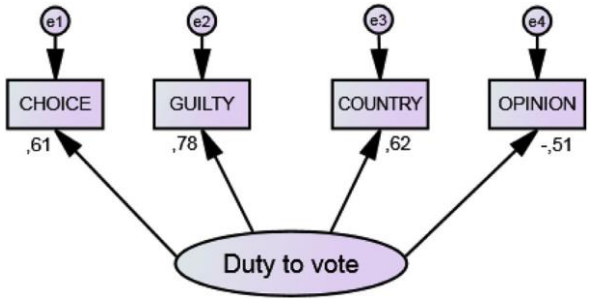

Fig. 2. Reduced (4 indicators) model. Standardized regression weights. Degrees of freedom = 2, X^2 = 2129.18, Probability level = .000, RMSEA:.069, CFI:.98. N = 22,072.

Table 2
Standardized regression weights and model fit measures. Unconstrained.

	France	Germany	Italy	Spain	Austria	Greece	Portugal	All
Choice	.65	.54	.54	.70	.59	.63	.70	.59
Guilty	.81	.77	.76	.76	.76	.70	.80	.78
Country	.52	.66	.70	.61	.61	.61	.61	.61
Opinion	—.57	—.60	—.42	—.46	—.64	—.38	—.54	—.52
Duty to vote / turnout European elections	.57	64	49	54	.62	.36	.63	.57
N	4013	4014	4003	4010	4002	1017	1013	22072
RMSEA	.053	.054	.065	.062	.060	.066	.067	.048
CFI	.971	.970	.940	.966	.968	.942	.965	.975
X2	575.8	595.3	855.9	770.6	716.5	228.7	235.3	2585.7
df	5	5	5	5	5	5	5	5

All coefficients and Chi-squares are significant at .001 level for 5 DF. Method: ADF. Bayesian imputation applied.

Table 3
OLS duty estimations. Concurrent validity tests.

	(1) France	(2) Germany	(3) Italy	(4) Spain	(5) Austria	(6) Greece	(7) Portugal
Political interest	.28***	.31***	.28***	.31***	.32***	.29***	.30***
	(.01)	(.01)	(.01)	(.01)	(.01)	(.02)	(.02)
Parents said voting is	.11***	.13***	.09***	.10***	.10***	.09***	.07***
a duty	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.02)
Age	.16***	.10***	.04* (.02)	.13***	.07***	.14**	—.01
	(.02)	(.02)		(.02)	(.02)	(.04)	(.04)
_cons	.28***	.24***	.31***	.21***	.25***	.27***	.31***
	(.01)	(.01)	(.01)	(.01)	(.01)	(.02)	(.02)
R-squared	.217	.259	.148	.209	.201	.189	.169
N	3937	3961	3935	3929	3895	987	964

Standard errors in parentheses. All variables have been recoded so as to range between 0 and 1. Listwise deletion of missing data.
+ p < .1, *p < .05, **p < .01, ***p < .001.

Table 4
Multiple group analysis for seven European countries. Metric invariance tests.

Model	RMSEA	LO 90	HI 90	PCLOSE	CFI
1.Configural invariance	.023	.023	.023	1.000	.983
2.Metric invariance I (measurement weights)	.026	.025	.026	1.000	.967
3.Metric invariance II (also regression weights)	.026	.026	.026	1.000	.962
4.Scalar invariance (all weights and intercepts)	.046	.046	.046	1.000	.823
5.Partial scalar invariance	.039	.039	.039	1.000	.880

Method: Maximum likelihood.

Table 4 summarizes the fit indices for several types of measurement invariance. The first four rows correspond to configural, metric and scalar invariance. The last is an alternative model releasing some constrains (i.e. partial scalar invariance).

In order to test invariance, it is customary to compare configural, metric and scalar invariance models using chi-square difference tests. But since this test is too sensitive to samples larger than 1000 (Brown, 2006), this procedure yields almost always significant results. Instead, we can use change in goodness-of-fit indexes to assess whether or not we have achieved measurement invariance (Cheung and Rensvold, 2002; Chen, 2007). According to this approach, an « increase of RMSEA by .02 and a decrease of CFI by .02 would indicate non-invariance of factor loadings/intercepts » (Comça, 2010:646).

The first row in Table 4 refers to configural invariance. All the measures (and particularly, RMSEA and CFI) suggest that the duty factor exists most probably in all the countries and that it

can be measured using these four indicators. The next model restricts measurement weights to be equal. According to the cut-off in change of goodness-of-fit, restraining the duty to vote indicators to have the same factor loadings in all the seven countries does not yield results significantly different from the unconstrained model (configural invariance). The third row adds another restriction: now the regression coefficient from the duty to vote construct to turnout in the EU election has also to be equal across groups. Again, changes in the model fit measures are within the rule of thumbs. Therefore, metric invariance is definitely achieved.

The fourth model adds to the previous the restriction that intercepts should be equal across countries, which is known as scalar invariance. Here the model fit measures are not as good. RMSEA increase exceeds the cut-off point of .02 (.023 / .046), and CFI worsens by more than .02. (.983 / .823). This means that the model does not achieve full scalar invariance. The next row frees some parameters to let them vary across countries, in order to test for partial invariance.[19] In this case, the variation in the goodness-of-fit measures with regards the unconstrained model is smaller than when testing for full scalar invariance. RMSEA increase is clearly within the limits of the cut-off point, and the variation in CFI is closer to the maximum of .02. Therefore, we can conclude that these values meet the partial scalar invariance test, and that we may compare the means of the duty to vote across countries (Byrne, 2008; Davidov, 2008).

5. Conclusion

Although the concept of a civic duty to vote has been used in electoral behavior research for more than 50 years, there is no consensus about how to measure it. Several different questions have been used in prior research, with little systematic attention to the quality of these questions. We reviewed prior research on civic duty and we identified serious flaws in the literature, the most important being a dismissal of an inherent social desirability bias when asking people about their citizen duty. Most studies rely on one single indicator, in contrast with other relevant attitudes which are usually measured through a battery of questions. Finally there are too many agree/disagree statements, which is particularly problematic since the risk of social desirability is quite high in this instance.

Our goal in this study has been to propose a simple and clear definition of sense of civic duty, to specify the implications and the dimensions of the concept, and to propose a large and a short battery of questions to adequately tap civic duty in its diverse aspects while paying due attention to question wording issues, most especially the social desirability associated with the presence of a public norm. The battery was administered in seven different European countries at the time of the 2014 European election. The empirical evidence suggests that we can measure the underlying construct of the duty to vote with these indicators.

We understand that few research teams may be willing to devote that much space in their questionnaire to measure a single concept. With this constraint in mind we set up to construct a short battery of questions designed to maximize internal consistency and item heterogeneity while minimizing social desirability. This led us to propose a battery of four questions: CHOICE, GUILT, OPINION, and COUNTRY.

The reduced battery achieved very good fit measures and passed several tests that confirm its validity across seven European countries. Although the impact of some indicators and the relationship between the duty construct and turnout vary somewhat across countries, we constrained the weights and intercepts of our four duty indicators to be equal in the seven

countries. The results indicate that configural and metric invariance are confirmed for our model measuring duty and predicting turnout in the seven countries. This means that the duty construct has the same structure and indicators, as well as the same metric in all groups. Hence, we conclude that these four questions adequately tap the duty to vote in West European countries, and most likely in established democracies writ large. Full scalar invariance is not achieved, although partial scalar invariance is. If we release the intercepts of some indicators in southern countries, the resulting model does not differ significantly from the unconstrained model, and therefore we should be able to compare the factor means.

Our four question battery provides researchers with a tool that will allow them to conduct cross-national research on the civic duty to vote. Further research may use our measure to examine the role of contextual, cultural, and individual factors in the development of the civic duty to vote. Hopefully, with better measures of the concept, more systematic research can be undertaken about why for some people construe the act of voting as a moral obligation while for others it is a matter of personal choice.

References

Ansolabehere, S., Rodden, J., Snyder, J.M., 2008. The strength of issues: using mul- tiple measures to gauge preference stability, ideological constraint, and issue voting. Am. Political Sci. Rev. 102 (02), 215e232.

Baumgartner, H., Homburg, C., 1996. Applications of structural equation modeling in marketing and consumer research: a review. Int. J. Res. Mark. 13 (1996), 139l (of).

Birch, S., 2009. The case for compulsory voting. Public Policy Res. 16 (1), 21e27. Black, D., 1948. On the rationale of group decision-making. J. Political Econ. 23e34. Blais, A., Achen, C.H., 2010. Of Group Decision-making.Ng Voting: Why Theory and

Voter Turnoutr Turnout 2010) of group decisi. http://www.princeton.edu/csdp/events/Achen031110/Achen031110.pdf. Consulted on 20/5/2013.

Blais, A., Young, R., 1999. Why do people vote? An experiment in rationality. Public Choice 99 (1e2), 39e55.

Blais, A., 2000. To Vote or Not to Vote? The Merits and Limits of Rational Choice.

University of Pittsburgh Press, Pittsburgh.

Bollen, K.A., 1989. Structural Equations With Latent Variables. John Wiley, New York.

Brace, I., 2004. Questionnaire Design. Kogan Page, London.

Brennan, G., Buchanan, J., 1984. Voter choice: evaluating political alternatives. Am.

Behav. Sci. 28 (2), 185e201.

Brennan, G., Hamlin, A., 1998. Expressive voting and electoral equilibrium. Public Choice 95 (1e2), 149e175.

Brennan, G., Hamlin, A., 2000. Democratic Devices and Desires. Cambridge Uni- versity Press.

Brennan, G., Lomasky, L., 1993. Democracy and Decision: the Pure Theory of Elec- toral Politics. Cambridge University Press, Cambridge.

Brown, T., 2006. CFA with Equality Constraints, Multiple Groups, and Mean Struc- tures. Confirmatory Factor Analysis for Applied Research 236e319.

Brown, T.A., Moore, M.T., 2012. Confirmatory factor analysis. In: Hoyle, Rick H. (Ed.), Handbook of Structural Equation Modeling. The Guilford Press,

pp. 361e379.

Browne, M.W., Cudeck, R., 1993. Alternative ways of assessing model fit. In: Bollen, K., Long, J. (Eds.), Testing Structural Equation Models. Sage, Newbury Park, CA, pp. 136e162.

Browne, Michael W., 1984. Asymptotically distribution-free methods for the anal- ysis of covariance structures. Br. J. Math. Stat. Psychol. 37 (1), 62e83.
Byrne, B.M., 2008. Testing for multigroup equivalence of a measuring instrument: a walk through the process. Psicothema 20 (4), 872e882.
Byrne, B.M., Shavelson, R.J., Muthe´n, B., 1989. Testing for the equivalence of factor covariance and mean structures: the issue of partial measurement invariance. Psychol. Bull. 105 (3), 456.
Campbell, A., Converse, P.E., Miller, W.E., Stokes, D.E., 1960. The American Voter. Wiley, New York.
Carlsson, F., Johansson-Stenman, O., 2009. Voting Motives, Group Identity, and Social Norms rapport nr.: Working Papers in Economics 366.
Carmines, E.G., Zeller, R.A. (Eds.), 1979. Reliability and Validity Assessment, vol. 17. Sage.
Chen, F.F., 2007. Sensitivity of goodness of fit indexes to lack of measurement invariance. Struct. Equ. Model. 14 (3), 464e504.
Cheung, G.W., Rensvold, R.B., 2002. Evaluating goodness-of-fit indexes for testing measurement invariance. Struct. Equ. Model. 9 (2), 233e255.
Clarke, H.D., Sanders, D., Stewart, M.C., Whiteley, P., 2004. Political Choice in Britain. Oxford University Press, Oxford.
Coleman, J.S., 1990. Foundations of Social Theory. Harvard Univ. Press, Cambridge. Comça, M., 2010. How to compare means of latent variables across countries and waves: testing for invariance measurement. An application using eastern European societies. Sociologia 42 (6), 639e669.
Cook, C., Heath, F., Thompson, R.L., Thompson, B., 2001. Score reliability in Webor internet-based surveys: unnumbered graphic rating scales versus likert-type scales. Educ. Psychol. Meas. 61 (4), 697e706.
Cook, T.D., Campbell, D.T., 1979. Quasi-experimentation: Design and Analysis Issues for Field Settings. Houghton Mifflin, Boston.
Dalton, R.J., 2008. The Good Citizen: How a Younger Generation Is Reshaping American Politics. CQ Press, Washington, Dc.
Davidov, E., 2008, March. A cross-country and cross-time comparison of the human values measurements with the second round of the European social Survey. Surv. Res. Methods 2 (1), 33e46.
Davidov, E., Schmidt, P., Billiet, J. (Eds.), 2011. Cross-cultural Analysis: Methods and Applications. Routledge.
Dennis, J., 1970. Support for the institution of elections by the mass public. Am. Political Sci. Rev. 64 (3), 819n Pol.
Ding, L., Velicer, W.F., Harlow, L.L., 1995. Efffects of estimation methods, number of indicators per factor and improper solutions on structural equation modeling fit indices. Struct. Equ. Model. 2, 119e143.
Dowding, K., 2005. Is it rational to vote? Five types of answer and a suggestion. Br. J. Polit. Int. Relations h 7 (3), 442e459.
Downs, A., 1957. An Economic Theory of Democracy. Harper Collins Publishers, New York.
Fiorina, M.P., 1976. The voting decision: instrumental and expressive aspects. J. Polit. 38, 390e413.
Foddy, W., 1993. Constructing Questions for Interviews. Cambride University Press, Cambridge.
Fowler, J.H., Kam, C.D., 2007. Beyond the self: social identity, altruism, and political participation. J. Politics (3), 813e827 o 69.

Graham, J., Nosek, B.A., Haidt, J., Iyer, R., Koleva, S., Ditto, P.H., 2011. Mapping the moral domain. J. Personality Soc. Psychol. 101 (2), 366.
Greenspan, P.S., 1994. Guilt and virtue. J. Philos. 91, 57e70.
Grofman, B., 1993. Is turnout the paradox that ate rational choice theory. In: In- formation, Participation, and Choice: an Economic Theory of Democracy in Perspective, pp. 93e103.
Holtgraves, T. (2004). Social desirability and self-reports: Testing models of socially desirable responding. *Personality and Social Psychology Bulletin*, *30*(2), 161-172.
Huntington, S.P., 1993. The Third Wave: Democratization in the Late Twentieth Century, vol. 4. University of Oklahoma press.
Kline, R.B., 2011. Principles and Practice of Structural Equation Modeling, third ed.
The Guilford Press, New York.
Kline, R.B., 2005. Principles and Practice of Structural Equation Modeling, second ed.
Guilford Press, New York, NY.
Knack, S., 1992. Civic norms, social sanctions, and voter turnout. Ration. Soc. 4 (2), 133e156.
Knack, S., Kropf, M.E., 1998. For shame! The effect of community cooperative context on the probability of voting. Polit. Psychol. 19, 585,ical.
Krosnick, J.A., 1999. Survey research. Annu. Rev. Psychol. 50 (1), 537e567.
Linn, M.C., 2000. Designing the knowledge integration environment. Int. J. Sci. Educ.
22 (8), 781e796.
Lomasky, L.E., Brennan, G., 2000. Is there a duty to vote? Soc. Philos. Policy 17 (01), 62e86.
Marsh, H.W., Hau, K.T., Balla, J.R., Grayson, D., 1998. Is more ever too much? The number of indicators per factor in confirmatory factor analysis". Multivar. Behav. Res. 33, 181e220.
Maruyama, G., 1997. Basics of Structural Equation Modeling. Sage. McIver, J., Carmines, E.G., 1981. Unidimensional Scaling, vol. 24. Sage.
Mueller, D.C., 2003. Public Choice III. Cambridge University Press, Cambridge. Mueller, D.C., 1989. Public Choice I I. Cambridge University Press, Cambridge.
Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., Giovannini, E., 2005. Handbook on Constructing Composite Indicators. OECD Library. STD/DOC 3. Available at: http://www.keepeek.com/Digital-Asset-Management/oecd/ economics/handbook-on-constructing-composite-indicators_ 533411815016#page2.
Oppenheim, A.N., 1992. Questionnaire Design, Interviewing and Attitude Mea- surement. Pinter, London.
Owen, G., Grofman, B., 1984. To vote or not to vote: the paradox of nonvoting. Public Choice 42 (3), 311e325.
Rakowski, W., Andersen, M.R., Stoddard, A.M., et al., 1997. Confirmatory analysis of opinions regarding the pros and cons of mammography. Health Psychol. 16, 433P.
Revilla, M.A., Saris, W.E., Krosnick, J.A., 2014. Choosing the number of categories in agree-disagree scales. Sociol. Methods Res. 43, 73e97.
Riker, W.H., Ordeshook, P.C., 1968. A theory of the calculus of voting. Am. Political Sci. Rev. 25e42.
Said, H., 2011. Confirmatory factor analysis (CFA) for testing validity and reliability instrument in the study of education. Aust. J. Basic Appl. Sci. 5 (12), 1098e1103.
Saris, W.E., Revilla, M., Krosnick, J.A., Shaefer, E.M., 2010. Comparing questions with agree/disagree response options to questions with item-specific response op- tions. Surv. Res. Methods 4, 61e79.
Sauger, N., Dehousse, R., Gougou, F., 2015. "Comparative Electoral Dynamics in the
European Union in 2014 (CED-EU14): a Data User's Guide ", Les Cahiers euro- pe´ens de Sciences Po, n◦ 01. Available at. http://www.cee.sciences-po.fr/en/
publications/les-cahiers-europeens/2015/doc/1174/raw. Consulted on 10/20/ 2015.

Schuman, H., Presser, S., 1981. Questions and Answers in Attitude Surveys: Exper- iments on Question Form, Wording, and Context. Sage.
Schwitzgebel, E., Rust, J., 2014. The moral behavior of ethics professors: relation- ships among self-reported behavior, expressed normative attitude, and directly observed behavior. Philos. Psychol. 27 (3), 293e327.
Steenkamp, J.B.E., Baumgartner, H., 1998. Assessing measurement invariance in cross-national consumer research. J. Consumer Res. 25 (1), 78e107.
Steiger, J.H., 1989. Causal Modeling: a Supplementary Module for SYSTAT and SYGRAPH. Systat, Evanston, IL.
Stewart-Brown, S., Tennant, A., Tennant, R., Platt, S., Parkinson, J., Weich, S., 2009. Internal construct validity of the Warwick-Edinburgh mental well- being scale (WEMWBS): a rasch analysis using data from the Scottish health education population survey. Health Qual. Life Outcomes 7 (1), 15e22.
Stouffer, S.A., 1950. An overview of the contributions to scaling and Scale Theory. reprinted. In: Maranell, Gary M. (Ed.), Scaling: a Source- Book for Behavioral Scientists. Aldine, Chicago, pp. 131e141, 1974.
Tullock, G., 2000. Some further thoughts on voting. Public Choice 104 (1), 181e182. Uhlaner, C.J., 1986. Political participation, rational actors, and rationality. Polit.
Psychol. 7, 551e573.
Usher, D., 2011. What Exactly Is a Duty to Vote? Queen, D. (2011). What exactly is a duty to 1266, 2011. Available on: http://qed.econ.queensu.ca/working_papers/ papers/qed_wp_1266.pdf. Consulted on 11/04/2014.
Verba, S., Schlozman, K.L., Brady, H.E., 1995. Voice and Equality: Civic Voluntarism in American Politics. Harvard University Press, Cambridge.
Yuan, K.H., Bentler, P.M., 1997. Improving parameter tests in covariance structure analysis. Comput. Stat. Data Analysis 26 (2), 177e198.
Zuckerman, A.S., Kotler-Berkowitz, L.A., 1998. Politics and society political diversity and uniformity in households as a theoretical puzzle. Comp. Polit. Stud. 31 (4), 464e497.

Endnotes:

[1] Skype: ablais

[2] A somewhat similar approach is followed by Schwitzgebel and Rust (2014) in a survey of US professors, both inside and outside philosophy, in which the respondents were asked to rate various actions, one of them being ‘regularly voting in public elections’ on a 1e 9 scale, from very morally bad to morally neutral to very morally good. Interestingly, voting received a very high mean rating of 7.4, as high as ‘donating one tenth of one's income to charity’ and slightly higher than ‘regularly donating blood’. This is an intriguing approach which has the advantage of directly asking whether voting is good morally and is thus in line with our definition of civic duty. It remains to be seen whether such a question can be put to the general public.

[3] Only the fourth question has been asked in some of the more recent ANES surveys.

[4] 59% chose ‘duty’ and 36% ‘right’. This pioneer measure goes directly to the heart of the matter, which is whether voting is a duty that is owed (a moral obligation) or not. Yet the question assumes that the opposite of duty is a right. It also specifies that is a duty to your country and thus assumes a patriotic component, which may not be always present. Finally, the question is dichotomous and does not allow for the measurement of the intensity of feelings.

[5] Why 13 questions? According to Stouffer (1950:140), ten to twelve questions is the ideal number to test scalability.

[6] Greece, Spain and Portugal belong to the third wave of democratization (Huntington, 1993), which means that most traits of their political culture regarding elections and support for democracy are relatively recent.

[7] The survey was an initiative of the Centre d'Études Europeénes of SciencePo, Paris. The online survey was conducted by TNS Sofres between May 28 and June 12, 2014, immediately after the European election. The samples were drawn from existing online panels, randomly selecting email addresses with a simple stratification by region, and correcting sociodemographic biases with sex, age and social status quotas. Sample sizes are only about 1000 people in Greece and Portugal because of the difficulty to get larger balanced samples in these two countries. Average response rate is above 30% (from 18% in Austria to 57% in Portugal), and drop out during the questionnaire is about 5% in all countries, except in Austria and Greece where it reaches 10%. More details can be found in Sauger et al. (2015).

[8] These questions have an expressive tone (I see voting as a way to show …), which may seem inconsistent with our defining duty as being moral rather than instrumental or expressive. But we construe an expressive motivation as referring to a willingness to convey a specific preference in a given election. The motive to demonstrate support for the country or for the democratic ideal is something different, which belongs to the realm of values and norms.

[9] We refer the reader to Appendix I for more detail on the distribution of these indicators.

[10] Although it has been proven that 5 labeled categories for agree/disagree questions provide higher quality data than formulations with 7 or 11 categories (Revilla et al., 2014) we decided to include horizontal integer scales ranging from 0 to 10 with equal spacing "radio buttons", where lower values indicated (and where labeled as such) strong disagreement and higher values indicated strong agreement. Previous research has found that this type of radio button scale is an acceptable psychometrical tool for measuring attitudes on the Internet and that they take less time to complete than other scale formats, such as sliders (Cook et al., 2001). Finally, from a statistical point of view, the main advantage of this kind of format is that it generates data on the interval scales which meet important requirements for the applicability of parametric procedures.

[11] Predictive validity is defined as "the approximate validity with which we infer that a relationship between two variables is causal or that the absence of a relationship implies the absence of cause" (Cook and Campbell, 1979: 37).

[12] Our final N is 22,072 individuals. Although the missing data did not represent more than 5% of the total sample (and therefore listwise deletion was justified), we decided to apply multiple bayesian imputation because the presence of missing data does not allow computing modification indices or applying other estimating methods than maximum likelihood. Due to bayesian imputation, Amos has produced 10 alternative complete datasets, which affects our Chi-square values, as they are ten times larger than if computed over the original dataset. As for the asymptotically distribution-free function (ADF), it has no distributional assumptions (Browne, 1984). Moreover it is believed to work better than maximum likelihood with large samples such as ours (Yuan and Bentler, 1997). Note that we have tried alternative estimation methods (maximum likelihood, unweighted least squares) with different treatments for the missing data. All these alternative models did not yield significantly different model fit measures, measurement or regression weights.

[13] Additional empirical evidence of convergence of the proposed items on the same construct is provided by the correlation matrix (see Appendix II). This correlation matrix also supports

the selection of CHOICE as the reference item, as it is the indicator with the highest number of correlations (11) above .3.

[14] Regarding the model fit, we notice a significant Chi-Square. Nevertheless, it is also known that when structural equation models (SEM) are run with a sample over 400 observations the Chi-square is almost always statistically significant (Brown, 2006).

[15] All the internal consistency reliability measures (inter-item correlations, item-total correlations, Chronbach's Alpha variation if items are deleted) pointed out that these items were less connected to the others. Modification indices (not shown) and additional analyses suggested that the measure would benefit from correlating the error terms of the reversed and agree-disagree questions, which in turn points to possible sub-dimensions or metric factors. We are not exploring this avenue in the present research.

[16] Additional analyses show that this reduced measurement model works even better in Austria and Germany (RMSEA of .05 and .06 respectively). The Cronbach's Alpha for a scale built using this shortened version of the duty battery and for the merged sample is .71.

[17] If we build an additive scale with these four indicators and use it to predict turnout by means of a logistic regression -controlling for age, sex, education and interest in politics-we see that going from no duty at all to maximum duty increases the likelihood to vote from about 8% to about 97% everywhere except in Greece, Italy and Spain. In these three countries the probability to turnout when the duty scale is at its minimum is 50%, 25% and 22%, respectively. The same exercise with the complete scale yields probabilities to show up at the polls between 4% and 98% for those with minimum and maximum duty respectively. Again, the exceptions are Greece, Italy and Spain where those exhibiting a minimal sense of duty still have a probability to vote of about 43%, 16% and 13%. Although the causes and effects of the duty to vote should be further studied in other research, the results for the southern Mediterranean countries suggest that in these countries people also attend the polls for other reasons.

[18] This is equivalent to a concurrent validity test. The models displayed in Table 3 allow us to measure the correlation of the duty construct with other concepts without reproducing seven different correlation matrices.

[19] More specifically, we relaxed the invariance assumption for the following intercepts: COUNTRY for Italy, Spain and Portugal, OPINION in France and Turnout for Greece. If we keep the constraint for OPINION, the RMSEA goes up to .04, which still points to non-significant differences with the unconstrained model. Note that the intercept of GUILT and CHOICE are still invariant across countries, which should allow us to compare factor means across countries (Byrne et al., 1989; Steenkamp and Baumgartner, 1998).