

Pocketsize Plus and Massive Minus.

Explaining the Asymmetry between Reward and Punishment in Economic Voting.

Lössbroek, J.J.H.

Master thesis Comparative Politics, Radboud University Nijmegen

Abstract

The effect of the economy on voting is well-established: incumbent governments are electorally rewarded for economic growth and electorally punished for economic decline. Generally, this relation is implicitly assumed to be symmetric, but case studies are increasingly discovering its asymmetry: the reward for economic success is dwarfed by the punishment for economic malaise, and assuming linearity leads to false conclusions. In addition to corroborating the findings of asymmetry in a cross-country analysis of 22 European countries, this study investigates macro and micro mechanisms explaining why the difference between punishment and reward is greater in some situations than in others. For this purpose, this study uses multilevel logistic regression and instrumental variables probit models to estimate robust results. Three approaches are tested to account for this variation. First, the effects of economic growth, unemployment and inflation are modeled separately instead of assumed to be merely different aspects of the same concept of ‘the economy’; this way, differential effects of economic trends are explored. The results tentatively indicate that the asymmetry between punishment and reward is greater for inflation than for growth and unemployment, but a lack of country level variance prevents proper testing. Second, the salience of these issues varies between and within countries, and more salient issues may generate a greater difference between punishment and reward. Modeling salience of inflation, unemployment and economic growth separately yields no results, but greater economic salience in general does increase the difference, although results are not particularly robust. Third, greater political affinity not only increases the strength of economic voting, but also reduces the difference. Lower scores on indicators of news consumption, political interest and education contribute to a greater asymmetry. In conclusion, economic voting is on average asymmetrical, but the greater one’s political affinity, the more strongly the omnipresent punishment is complemented by rewards for good performance.

Keywords: grievance asymmetry; economic voting; sociotropic perceptions; issue salience; political affinity; elections in Europe

Table of Contents

1. INTRODUCTION	3
2. OVERVIEW OF ECONOMIC VOTING	6
2A. Retrospective versus prospective economic evaluations	7
2B. Sociotropic versus egotropic (pocketbook) voting	8
2C. Subjective versus objective indicators	9
2D. Identifiability of anti-incumbency voting	11
3. EXPLAINING THE ASYMMETRY OF REWARD AND PUNISHMENT	13
3A. Background: thresholds versus win/loss differences	13
3B. Universal explanation: loss aversion	14
3C. Differential explanation 1: splitting up ‘the economy’	15
3D. Differential explanation 2: asymmetry in salience	17
3E. Differential explanation 3: asymmetry political affinity	18
3F. Differential explanations: brief overview of alternative suggestions	21
4. METHODOLOGY, DATA AND OPERATIONALIZATION	23
4A. Methodology	23
4B. Data and operationalization	24
4C. Endogeneity issues: the danger of reverse causality	29
5. EMPIRICAL RESULTS	33
5A. Testing of asymmetry between punishment and reward	33
5B. Splitting up ‘the economy’	34
5C. Issue salience	38
5D. Political affinity	41
6. CONCLUSION AND DISCUSSION	49
7. REFERENCES	53
8A. Tentative literature overview	57
8B. Exact wording of the items in the (English-language) questionnaire	58
8C. Descriptive statistics	60
8D. Post-estimation test statistics of instrumental variables	62
8E. Asymmetry between punishment and reward by country	63
8F. Additional tables macro-economy	64
8G. Additional tables political affinity	65

1. INTRODUCTION

The recent economic crisis has been accompanied by a notable trend in voting behavior: citizens punish their governments for the state of the economy. A study of 35 parliamentary elections in 27 OECD countries has shown that during the economic recession between 2008-2012, only 8 incumbents maintained their leading position (Bouvet & King, 2013). The impact of the economy on incumbency voting behavior has been confirmed repeatedly. In its simplest form, voters hold the government accountable for the state of the economy. In times of growth, the electorate will reward the incumbent politicians by increasingly voting for them. Conversely, when the economy stagnates, the electorate will punish incumbents and increasingly vote for opposition parties instead (Bellucci, Costa Lobo, & Lewis-Beck, 2012; Benton, 2005; Bouvet & King, 2013; Costa Lobo & Lewis-Beck, 2012; Duch & Stevenson, 2005, 2008; Fraile & Lewis-Beck, 2012). Economic voting is part of competency voting: citizens increasingly decide whom to vote for based on government performance rather than on identity or cleavages.

It may be argued that voting against incumbents is part and parcel of any democratic process. True as this may be, strong anti-incumbency sentiments can seriously hamper good policymaking for two reasons: a higher turnover of leaders and a weaker position for governments. First, frequent leader turnover shortens policy cycles: the time a government has for crafting and implementing new laws becomes shorter. However, the creation of appropriate policy simply takes time, and time shortage complicates this process. Second, the position of the government is severely weakened within these shorter time cycles. It is weakened informally when politicians are scorned by public polls, which can reduce their commitment to push through needed yet unpopular reforms. It is also weakened formally since even if incumbents are determined to reform anyway, they are likely to lose mid-term elections. These are often present in one form or another and policy freedom shrinks when the incumbent's parliamentary support is weakened. In short, economic anti-incumbency voting has a profound impact on contemporary politics, especially in times of weak economic performance. Considering the implausibility of substantial, structural economic growth in the next decade, understanding how, why and under what conditions voting behavior in meager times differs from voting behavior in times of plenty is of vital importance.

Economic voting, after all, is not known to be a perfectly symmetric function (Bellucci et al., 2012; Bloom & Price, 1975; Clagett, 1986; Freire & Santana-Pereira, 2012; Headrick & Lanoue, 1991; Lewis-Beck & Stegmaier, 2000; Marsh & Mikheylov, 2012; Nannestad & Paldam, 1997; Nezi, 2012; Van der Brug, Van der Eijk, & Franklin, 2007). A substantial number of authors have argued that there is a strong asymmetry in economic voting: rewards for economic success are dwarfed by punishments for economic malaise. Economic growth does have some impact on pro-incumbency voting, but the effect of economic recession on anti-incumbency voting is far bigger. Nevertheless, the majority of the studies on economic voting are still based on the (implicit) assumption of a symmetric relation between reward and punishment, often without reflecting on this issue. Additionally, most studies that do conceptualize

economic voting as asymmetric accept this as given or are content to provide evidence for the asymmetry, without delving deeper into its causes. Therefore, this study will set to not only show that, on average, punishments exceed rewards for the state of the economy, but also explain which micro and macro characteristics contribute to a greater difference.

Three lines of thought will be explored in this paper. The first line of thought is to subject the central independent variable, economic trends, to closer scrutiny. Many studies assume that different aspects of the economy, most importantly economic growth, unemployment and inflation, are sub-elements of a broader concept 'economic performance'. However, trends in these three factors, while related, tend to be far from identical. Additionally, it is plausible that the relation between punishment and reward is different: voters will want unemployment to be as low as possible and reward governments for it, but once inflation reaches a certain level, little can be gained by a further decrease. Jointly, these differences warrant empirical testing of differential effects of economic trends rather than combining them into a single score.

The second idea is based on issue salience: voters believe some issues are more important than others. The literature suggests that salience of the economy varies depending on economic performance: weaker performance makes the economy more salient, and, hence, strengthens economic voting. If bad economies coincide with greater importance, naturally, this generates a difference between punishment and reward. For highly salient issues, citizens are likely influenced by both emotional and rational motivations, while hardly salient issues are judged more rationally in a more balanced way. On the other hand, economic concerns tend to be the most salient issues at nearly every moment in every country, unless exceptional circumstances occur. This apparent paradox can be solved, as in the first line of thought, by differentiating between inflation, growth and unemployment. For each issue, the effect of salience will be analyzed. Additionally, even when the economy as a whole is the most important issue, it may matter whether 50%+1 or 80% of the voters believe that it is most important.

Third and finally, three aspects of political affinity will be investigated: news consumption, political interest and level of education. For each aspect of political affinity, higher scores are hypothesized to coincide with less asymmetry. The different dimensions are related, but the causal mechanisms mentioned below are. A difference between punishment and reward may be influenced by a negative information bias: information on economic malaise more easily reaches voters than good news. Citizens with low news consumption, then, will read the bad news on the front page but not the silver lining on page seven. Additionally, politically disinterested citizens will less often mentally connect news to incumbent performance. Hence, they will be more likely to connect primarily bad news to their politicians. Finally, lower educated citizens tend to use less information in casting their vote, which means that the negativity bias may be stronger for them. These three aspects of political affinity can explain differences in the strength of asymmetry within countries, and, through composition effects, also between countries.

In short, at both the micro and macro level, this paper aims to increase our knowledge of explaining variations in differences between reward and punishment. The research question central to this research will be

Research Question: *To what extent do differential macro-economic conditions and micro-level political attitudes account for the asymmetry between reward and punishment of economic voting?*

In addition to the theoretical innovations summarized above, this study will empirically improve on earlier research in multiple ways. First, analyses on economic voting are usually case studies of a single or limited number of countries¹. Using cases studies has several advantages, depending on the purpose of the study. Case studies are especially well-suited for gaining a better understanding of why voting behavior shifted in a specific country, such as the impact of the Irish governmental decision to withdraw from the international bond market in September 2010 on government popularity (Marsh & Mikhaylov, 2012). However, such *sui generis* explanations are less useful in understanding the conditioning of economic voting as a broader process. Additionally, most likely cases are very insightful and useful to assess whether a theoretical mechanism is potentially present in the way it is expected: for a study on reduced clarity of responsibility, the German federal elections of 2009 were a most likely case for finding low economic voting due to the CDU/SPD Grand Coalition and exogeneity of global economic recession affecting German exports (Anderson & Hecht, 2012). On the other hand, most countries do not belong to extreme categories, and the generalization of the findings in such countries requires a larger number of country-level cases. Therefore, large-N cross-country statistical analyses will be at the heart of this paper; specifically, the countries belonging to the EU-25 will be analyzed.

Second, the study will use more sophisticated statistical techniques than most previous studies. Since voting for either opposition or government is mutually exclusive, multilevel logistic regressions will be used as central technique. Different models will be estimated for punishment effects and reward effects, so that both can be compared in terms of strength, significance and direction: which predictors are more important for punishment than for reward? How do different variables moderate the effect of economic evaluations? The use of a micro-macro approach to model both individual and country-level effects, as well as cross-level interactions, is a major step forward compared to many earlier studies. Additionally, as the dependent variable is nonlinear, the commonly used (although rarely explicated) assumption that interaction effects are the same for each respondent would be false. Different interaction effects may even each other out, potentially showing as a non-significant interaction where in reality there is one (i.e. a Type II error). To differentiate between these groups of respondents, marginal instead of general interaction effects will be modeled (Buis, 2010; Norton, Wang, & Ai, 2004). Finally, the relation between economic evaluations and political preference may be endogenous, i.e. strong political preference can change voters' judgment on the economy (Evans & Andersen, 2006; Evans & Pickup, 2010). This would mean the correlational evidence for economic voting found in earlier studies is biased, overestimating the strength of the relation and running the risk of finding a relation absent in reality (i.e. a Type I error). To account for this risk, exogenous instrumental variables will be

¹ Two notable exceptions are the recent books by Duch and Stevenson (2008) and Van der Brug et al. (2007).

used in a two stage probit analysis. Hence, on top of moving forward our theoretical knowledge of the punishment-reward asymmetry, this study will do so in a way that is empirically more robust and reliable than most of its predecessors.

In the second chapter, a general overview of the field of economic anti-incumbency voting will be presented. Before discussing several explanations to account for an asymmetry between reward and punishment, it is necessary to establish how the general effect of economic voting should be conceptualized. Hence, this chapter will also be used to reflect on several decisions regarding what aspects of economic voting are most relevant for this study: retrospective rather than prospective voting, sociotropic rather than egotropic voting, subjective rather than objective data on the economy as central independent variable, and whether anti-incumbency voting is identifiable as a distinct concept. The third chapter more specifically discusses various theoretical mechanisms potentially explaining the asymmetry in judging the incumbent government for the economy. The fourth chapter explains the research design, covers the data that will be used and how the theoretical concepts are operationalized, followed by the fifth chapter in which the statistical results are presented and analyzed. Finally, the sixth chapter will be used for conclusion and discussions.

2. OVERVIEW OF ECONOMIC VOTING

The debate on economic voting can be situated within the broader school of thought on valence politics. Valence politics refers to citizens determining whom to vote for based on how the politicians or parties perform, rather than based on other reasons. In the case of economy voting, performance is the state of the economy: if the economy operates well, citizens will attribute this to their government and reward them with electoral support in the next elections. Conversely, when the economy goes bad, citizens will also attribute this to their government and punish them by voting for parties that are currently in opposition. This mechanism is outcome-based: voters focus not on specific policies or promises, but simply on whether the economy performs well or not. Valence politics can be contrasted with more structural motivations for voting for a certain party, such as social identification and political cleavages. Dealignment and a decline the importance of cleavages allow for greater influence of valence politics c. Also, performance is far more dynamic than identity or cleavages, which entails a greater relevance in explaining differences in election results on a shorter term.

The first study on economic voting was done by Key in 1964. Initially, the debate was rather straightforward: do economic conditions influence voting behavior or not? Does economic voting exist (Bloom & Price, 1975)? In the next decades, the research field has expanded tremendously: an overview in 2000 counted over three hundred journal articles on economic voting (Lewis-Beck & Stegmaier, 2000), and hundreds more have been published ever since. Although there is no perfect consensus on the existence of economic voting (Evans & Pickup, 2010), nearly all studies find strong effects of the economy on voting behavior. This trend of acknowledging the existence of economic voting has

stretched beyond academics: in the same period, the popular press and politicians themselves have internalized the importance of the economy for voting behavior (Lewis-Beck & Stegmaier, 2000). Additionally, the scope of research on economic voting has broadened geographically. Initially, studies focused on the United States. More recently, Western democracies have been the focus of studies in this field. Nevertheless, economic voting is far from restricted to Western democracies: studies in low-income democracies from Botswana to Uruguay show that economic voting takes place all over the globe (Lewis-Beck & Stegmaier, 2000).

Over time, research on economic voting has covered many different issues. Some topics, however, were more fiercely debated than others, because they concern what is part of the core mechanism of economic voting. Due to the central position of these four topics, most studies explicitly take a position in these debates. It is vital to address these four questions beforehand, as it is impossible to discuss why economic voting is asymmetric before establishing what economic voting entails. Below, the choices made on these debates will be defended. This study will focus on (a) retrospective rather than prospective evaluations, since voters actually have the information needed for retrospective judgments and this is far easier than coming up with reliable predictions on future performance. Also, the most important indicators are (b) sociotropic, not egotropic: when selecting leaders who are responsible for the nation, national conditions are more relevant than personal economic conditions. Additionally, the preferred measurement is (c) subjective rather than objective, since objective economic performance only influences voting decisions through subjective evaluations. Finally, (d) economic incumbency voting is argued to be identifiable as a distinct element of economic voting, as incumbency voting is strong enough to override other voting considerations.

2A. Retrospective versus prospective economic evaluations

The first issue is whether citizens vote retrospectively, basing their decision on the incumbent's performance in the past, or prospectively, basing their decision on how they predict that various candidates or parties would perform if they were to win the next elections. Do voters use information about future or past performance in their voting decisions? Before exploring how and when differences in information consumption contribute to an asymmetry between punishment and reward, it is necessary to establish what type of economic information is relevant to begin with: actual economic information or predictions about the future.

Some have argued that voters try to predict the effect future policies of a candidate or party will have on the economy. In this conception, prospective voters are sophisticated in their capabilities to estimate what their choice would deliver in terms of economics (Lewis-Beck & Stegmaier, 2000). Others have conceptualized voters as retrospective, who judge the incumbents on the state of the economy from the present and / or direct past (Lewis-Beck & Stegmaier, 2000).

It can be argued that both retrospective and prospective considerations influence voting behavior (Clarke & Stewart, 1994). Finally, both measurements have been found to correlate extremely highly (to

the point of multicollinearity) (Lewis-Beck & Stegmaier, 2000). A rare recent study using prospective voting instead of retrospective voting admits this was only done as there was no data on retrospective voting, and argues that prospective voting could still be seen as a way of expressing the opinion about the past economic policies (Hellwig, 2008). Others argue the reverse could also be done: using retrospective evaluations as a ‘crude but nevertheless adequate’ proxy for prospective evaluations (Anderson & Hecht, 2012).

Many mixed results have been found by studies in the 1990’s, (see, for an overview Lewis-Beck & Stegmaier, 2000). Currently, the general consensus is that retrospective voting has more influence than prospective voting. Theoretically, retrospective evaluations are far easier made than prospective evaluations: information is more readily available about the past than about future developments. Since retrospective information is simpler to acquire and more reliable, it is a more plausible predictor of voting than prospective evaluations (Duch & Stevenson, 2010a; Nannestad & Paldam, 1997; Nezi, 2012).² Empirically, too, retrospective economic evaluations has been found to be a stronger predictor of voting decisions (Anderson, 2006; Bartels, 2011; Lewis-Beck & Stegmaier, 2000; Nannestad & Paldam, 1997; Nezi, 2012). A different timing issue is that of lag time. Early studies indicate economic conditions have the strongest effect on the short run (one year) and using a longer time lag actually incorrectly combines two different independent variables (short and long term effects) into a single predictor (Bloom & Price, 1975). For these reasons, the theoretical framework of this study will be built on the assumption that voters base their economic vote on the recent past.

2B. Sociotropic versus egotropic (pocketbook) voting

The second issue is whether people primarily evaluate the changes in the macro economy or in their own micro-economic situation. Since the third chapter is devoted to the question of why responses to economic fluctuations are asymmetrical, it is essential to first establish what economic situation respondents primarily base their vote on.

In addition to sociotropic voting (decision-making based on macro-economic issues), many analysts also include egotropic or pocketbook voting: decision-making based on one’s personal financial situation. Theoretically, it seems plausible that people are more aware of the situation of their own household than the nationwide equivalent. It is easier to compare personal bank accounts to a broad array of macroeconomic characteristics. However, the consensus in the literature seems to be that pocketbook voting is not nearly as important as sociotropic voting or even not relevant at all. Statistically speaking,

² In terms of probability, the chance that an individual choice to vote for a different party leads to a better government is virtually zero. Rational choice theorists would use this balance to explain why many people do not turn out to vote at all. They apply modifications such as ethical norms or social pressure to solve the paradox of voting (Feddersen, 2004), but the mechanism remains that it is in the best interest of any voter to use the information already known about the past to vote instead of spending valuable time to make predictions about the future. Historical experiences provide valuable information for making predictions; in this conceptualization, prospective evaluations turn out to be rather retrospective evaluations in disguise (Duch & Stevenson, 2010a; Nannestad & Paldam, 1997).

egotropic voting effects tend to be dwarfed or become insignificant when sociotropic voting and/or control variables are added to the equation (Anderson & Hecht, 2012; Anderson, 2006; Freire & Santana-Pereira, 2012; Kinder & Kiewiet, 1981; Lewis-Beck & Stegmaier, 2000; Nezi, 2012). Compared to the earlier decision between prospective and retrospective voting, the empirical differences between both options are far stronger in this debate (Anderson & Hecht, 2012). Two case studies have found egotropic voting effects: a study in Denmark that suggests both effects exist, but that the sociotropic vote is substantially larger than the egotropic vote (Nannestad & Paldam, 1997) and a case study in Germany, in which sociotropic voting is absent³ and egotropic voting has a small but significant effect on vote choice (Anderson & Hecht, 2012).

The fact that the effects of egotropic voting disappear when sociotropic voting is controlled for lends support to the idea that both strands are not unrelated. Voters care most about the macro economy, as the impact of the government on the macro economy will be far more direct than the impact on personal financial situations. However, citizens might use changes in their personal financial situation (such as income fluctuations, perceived job insecurity, interest rates) as one among other ways (such as media consumption or personal contacts) to gain information on macro trends (Stevenson & Duch, 2013). In this mechanism, egotropic evaluations will influence voting behavior indirectly, by influencing sociotropic evaluations. Sociotropic evaluations, in turn, are important in deciding whom to vote for. This relation (see Figure 1) between egotropic evaluations, sociotropic evaluations and economic voting is reflected in the empirical findings that the effect of egotropic evaluations disappears when sociotropic evaluations are also modeled. The general picture supports a prevalence of sociotropic voting in empirical reality. Hence, the central mechanism in this study will be conceptualized as the sociotropic vote.

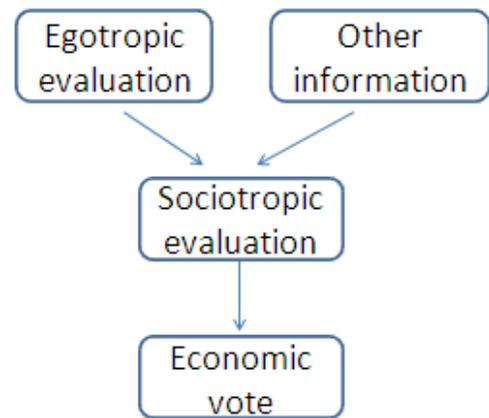


Figure 1. Ego- & sociotropic evaluations

2C. Subjective versus objective indicators

A third distinction can be made between objective and subjective indicators of economic performance. Objective indicators refer to actual macro-economic trends, such as a change in economic growth, inflation or inflation, and are measured at the macro level. Subjective indicators refer to the perception voters have of these macro-economic trends. These indicators are measured at the micro level. There is disagreement on which of these is a better conceptualization of the state of the economy; some

³ The authors provide a convincing combination of explanations for this particular finding: voters realized that Germany’s economic woes were strongly influenced by plummeting exports since the global economic crisis, and did not held their government responsible for this. Moreover, the country was led by a Grand Coalition of Christian democrats and Social democrats, which meant that no new coalition that excluded both incumbent parties was seen as realistic.

authors even argue that these are two different concepts (Duch & Stevenson, 2010b). Chapter 3 explores how different aspects the economy can have differential effects, i.e. contribute to an asymmetry between reward and punishment. Evidently, this requires a more specific conceptualization of what the general relation between economy and voting behavior entails. Is there a direct effect of objective economic indicators, or is this effect indirect, mediated by individual perceptions?

In the early stages of research into economic voting behavior, researchers used objective macro characteristics of the economy, such as inflation or unemployment. These indicators did however not yield the same identical results. The general pattern that emerged confirmed the basic mechanism, but magnitude fluctuated. Researchers looking into this subject could not agree on a clear-cut answer to the question of which indicator was more important than others. On the other hand, the use of objective economic indicators as the primary focus was generally accepted to when determining how voters judged incumbent government: a direct effect of the economy on judgments (see Figure 2) (Lewis-Beck & Stegmaier, 2000). In the 1990's, studies started including not only objective, but also subjective indicators of economic performance (Lewis-Beck & Stegmaier, 2000). This offers several advantages compared to using only objective indicators, which will be summarized below.

First, perceptions could encompass a broad range of different macro indicators, and combine these in a single score (Lewis-Beck & Stegmaier, 2000). For any researcher, it is easy to obtain macro-economic data but hard to estimate how the various indicators of economic performance are weighed by voters. The use of a subjective indicator of general economic performance lets this question be answered by each individual respondent.

Second, even if the importance to difference aspects was known, or only a single aspect was object of study, it takes time for a change in objective performance to trickle down into changes in economic perceptions. For example, consumer confidence has been found to lag substantially behind actual economic resurgence (Bartels, 2011). Using perceptions allows the effect to be lagged according to each respondent. This saves us from resorting to arduous ways to determine what lag term best fits the time it takes for voters to internalize economic information, under the dubious assumption that there is a fixed lag term.

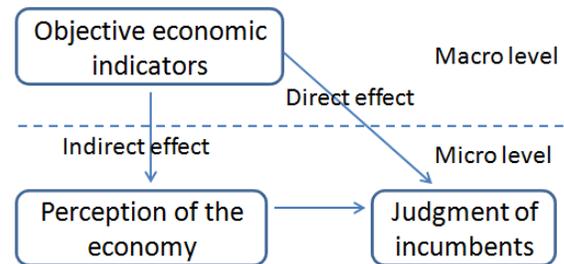


Figure 2. Objective and subjective indicators

Third, economic perceptions are also theoretically preferable to objective standards. A direct effect of economic conditions is a less accurate description of the actual voting process than the indirect effect via economic perceptions. Voters may not know the exact level of GDP per capita growth or inflation level, yet every voter has a perception about the economy. This perception is not independent of objective indicators, but may mediate its effect. In the voting booth, it is this subjective idea about the state of the economy that influences the decision made (Stevenson & Duch, 2013).

Note, however, that subjective indicators are not beyond criticism. Given that variables such as inflation,

and unemployment are the same for all voters in a country, how is it possible that there are such differences in perception? Some have argued that the influence of non-economic factors on perceptions is substantial, and that objective indicators do not suffer from this handicap (Duch & Stevenson, 2010b). It has even been remarked that variation at the micro level can only be caused by measurement errors: this particular study noted that respondents interpreted the survey questions in different ways and that this may guide their answers (Van der Brug et al., 2007). While it is probably true that subjective indicators might be contaminated by non-economic factors, they are still predominantly based on actual economic events (Lewis-Beck & Stegmaier, 2000). The differences at the micro-level are caused by the various ways in which economic information reaches different citizens, and by variation in the value citizens attach to various aspects of the economy (i.e. for some, unemployment may be more important than inflation, for others, the opposite is true). Also, economic trends at the national level can have different effects on regional or local levels: a factory going bankrupt may have a stronger effect on those living next door than on those living further away, both through personal experiences and different local media exposure (Stevenson & Duch, 2013). Moreover, as the third argument in favor of subjective indicators goes, voting decisions are influenced by contaminated perceptions. In order to model actual voting decisions, this influence is not a handicap but rather a closer representation of empirical reality. Finally, if subjective indicators would really only reflect measurement errors and random noise, we would expect that empirical studies in which actual economic trends are controlled for find no effects of economic perceptions: random measurement errors will not be able to predict voting behavior. However, also when controlling for actual economic trends, many studies found that perceptions strongly shape voting decisions (see Appendix 8D).

In short, subjective indicators outperform objective indicators in at least three theoretical aspects: weighting various elements of the economy, correcting for time lags, and more closely reflecting the process that actually goes on in the voting booth. Empirically, they are indeed strong predictors of voting behavior. For all these reasons,⁴ in more recent studies, objective economic indicators are usually either replaced or complemented by subjective economic indicators (Freire & Santana-Pereira, 2012). Since it is possible to use both macro and micro indicators, both will be included in the models. This way, since factual economic trends are controlled for, micro indicators will strongly reflect micro level variation in interpretation and value attached to various aspects of the economy.

2D. Identifiability of anti-incumbency voting

Fourth and finally, the question arises whether it is possible to isolate anti-incumbency effects from other voting considerations. Obviously, the choice a voter makes encompasses more than just the decision of whether or not to support the current government. Is this mechanism of reward and punishment strong enough to override other motivations, such as ideological preference, and

⁴ They are, on the other hand, more vulnerable for endogeneity issues than objective indicators. This will be dealt with in section 4C.

identifiable if the voter can choose between a multitude of political parties? In order to analyze anti-incumbency voting behavior, it should be plausible that this is theoretically distinct.

In general, the effect of economic influences on voting behavior is known to be rather strong. Since other motivations that influence voting decisions, such as identity and ideology, tend to become less important over time, the power of economic influences is great enough to be recognizable among other factors (Lewis-Beck & Stegmaier, 2000). In the same sense, it is plausible that political parties are clustered according to their incumbency status. Most countries are characterized by multi-party systems, allowing for subchoices within the categories of pro-incumbency and anti-incumbency voting. However, while problematic for estimating probabilities to vote for a specific party, this does prevent us from estimating effects on the decision to vote for either government or opposition.

Additionally, even in hypothetical perfect two-party systems, respondents have the option to not vote at all – an option used by a substantial and increasing share of the electorate (Anderson & Hecht, 2012). Some would argue that abstaining from voting is an expression of dissatisfaction with the incumbents (Freire & Santana-Pereira, 2012; Scotto, 2012). However, abstaining in itself is neither a punishment nor a reward when it does not influence the chance of being re-elected. Lower turnout does not shrink the electoral pie, nor does it change the division of seats unless there is a strong pattern of abstaining related to the intention of voting pro or anti incumbents. A case study in Germany suggests that the impact of economic considerations on the decision whether or not to vote is rather small. This lends credibility to the assumption that abstention effects do not (or do only to a very small degree) contaminate effects of reward and punishment (Anderson & Hecht, 2012). Therefore, the analysis will focus on the distinction between pro- and anti-incumbency voting, acknowledging that the voting decision is more complex, but under the assumption that this enables capturing the effects of theoretical interest.

In conclusion, central in this study will be voters who judge their governments on the way the national economy performed in the recent past. Economic performance will be measured as the satisfaction individual voters experience towards the economy, and the result of these votes will be conceptualized as the degree to which citizens decide to support the incumbent or opposition party / parties. These choices are theoretically most plausible, and unsurprisingly most popular in recent studies. Although the literature on economic voting is too vast to summarize in a single table, Appendix 8A provides a tentative overview of how the most important studies cited in this paper have conceptualized economic voting. This table shows a clear majority of studies conceptualize economic voting in the same way as this study, which maximizes comparability.

3. EXPLAINING THE ASYMMETRY OF REWARD AND PUNISHMENT

3A. Background: thresholds versus win/loss differences

Many authors note that there is a ‘grievance asymmetry’: the punishment in times of economic stagnation exceeds the reward in times of growth (Bellucci et al., 2012; Bloom & Price, 1975; Freire & Santana-Pereira, 2012; Headrick & Lanoue, 1991; Lau, 1985; Lewis-Beck & Stegmaier, 2000; Marsh & Mikhaylov, 2012; Nannestad & Paldam, 1997; Nezi, 2012; Van der Brug et al., 2007). However, this finding is not uncontested (Duch & Stevenson, 2008), the degree of asymmetry varies substantially between countries (Van der Brug et al., 2007), and the theoretical explanations are far from clear. Perhaps most strikingly, the vast majority of studies still implicitly assumes that the economic vote function is symmetric.

The first indications of a non-linear relation between economic conditions and voting were found decades ago (Clagett, 1986). Such non-linearity was conceptualized in two different ways, but in both ways, the core idea is that the effect of a change in economic performance on voting behavior is not the same for every level of economic performance. In contrast, the effect of the economy depends on the level of performance. Two different, but not necessarily mutually exclusive conceptualizations of non-linearity have been studied: threshold models and asymmetry models of reward for growth and punishment for malaise.

The first proponent of the former conceptualization was Mueller (1970) who argued that economic changes only matter when they are extreme. Regardless of the direction of economic fluctuation, voting was claimed to be influenced by the economy only when changes exceeded a certain threshold. Minor changes in the economy are, supposedly, not enough to either raise voter awareness of economic fluctuation, or allow for non-economic issues to dominate the electoral choice instead (Mueller, 1970). Threshold approaches are still used; for example, a case study in Greece suggested thresholds lying at 2.5% GDP growth, 1% unemployment change and 17% inflation per year need to be passed in order for economic voting to become relevant. Changes of

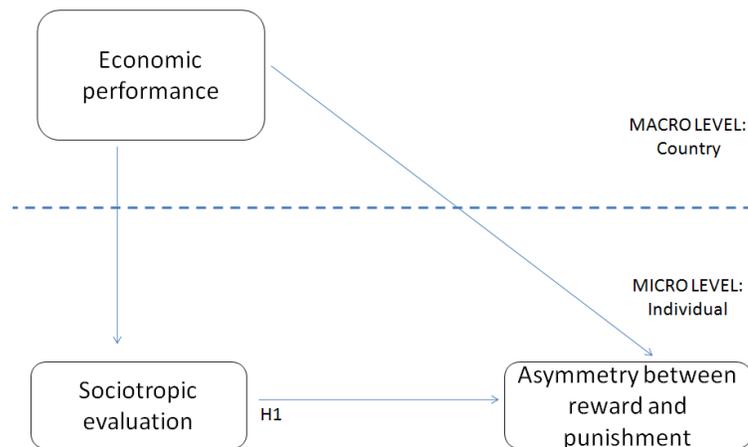


Figure 3. Basic model

economic conditions that did not pass these thresholds, would not significantly alter voting behavior (Nezi, 2012). A key weakness of such an approach is that any threshold will be arbitrary (Nannestad & Paldam, 1997): what theoretical mechanism would explain why a change of 2.5% is fundamentally different from a change of 2.4 or 2.6%?

A more fruitful and more popular approach, then, is to conceptualize the non-linearity of the vote function as consisting of both a reward and a punishment effect, which can have different strengths. This asymmetry of economic voting was introduced by Bloom and Price (1975) in a study on American voting behavior, and has become a more popular conceptualization than Mueller's threshold model (Nannestad & Paldam, 1997; Van der Brug et al., 2007). An extreme example of this asymmetry would be to conceptualize the relationship as full punishment effects and zero reward effects (Headrick & Lanoue, 1991). More commonly, however, reward effects are found to be existent, only smaller than punishment effects. The asymmetry has been found in case studies of diverse countries such as Denmark (Nannestad & Paldam, 1997), Greece (Nezi, 2012), Ireland (Marsh & Mikhaylov, 2012), the United Kingdom (Headrick & Lanoue, 1991) and the United States (Bloom & Price, 1975; Clagett, 1986). Studies that explicitly tested the asymmetry but did not find it are rather rare: it was not found in a case study on Portugal (Freire & Santana-Pereira, 2012). These results call for further understanding of the asymmetry between reward and punishment: under what conditions, and for which voters, is the explanation for electoral reward different than the explanation for electoral punishment?

In short, models showing that economic voting is asymmetrical are theoretically preferable to threshold models due to the clear demarcation at zero instead of arbitrary country-specific levels of economic change. Additionally, empirical results supporting an asymmetric function are more robust throughout different countries and periods. Therefore, in this study, the non-linearity of economic voting effects will be conceptualized as an asymmetry.⁵ This is reflected in the first hypothesis stated below:

H1 The electoral reward for improving economic conditions is smaller than the electoral punishment for declining economic conditions

3B. Universal explanation: loss aversion

It is surprising that there is so little insight in the causes of the asymmetry, given that there is such broad support for the asymmetrical effects of the economic on voting. A relatively simple explanation for this asymmetry can be found by borrowing from psychologists the concept of loss aversion (Abdellaoui, 2007; De Martino, Camerer, & Adolphs, 2010; Hammerstein & Hagen, 2005). A potential gain of X has a smaller impact on people than a potential loss of the same amount X.

⁵ This does not necessarily prevent testing of threshold models, but it would require far more complex models. Moreover, clear criteria for establishing the cut-off point are absent and data-mining should be avoided. Hence, the (relatively) parsimonious asymmetry conceptualization will be the focus of the paper.

Loss aversion is a convincing mechanism that has been found again and again in psychological and economic experiments. Nevertheless, loss aversion is not the first candidate for explaining differences between reward and punishment as it is problematic for several reasons. First, there is a great disagreement on how loss aversion should be conceptualized. Many different definitions are used in studies on loss aversion, yielding strongly different results. Moreover, some regard loss aversion to be an universal effect, others as a coefficient that varies among individuals. Regardless, the fact remains that estimated coefficients vary greatly (Abdellaoui, 2007). Also, on a more practical level, it would require tremendous efforts to test respondents on their loss aversion, which makes it unrealistic to gather data in large-N cross-country studies.

It can be argued that loss aversion is caused by biological or social factors, or a combination of both. On the one hand, neurological research has shown that the strength of loss aversion is influenced by biological differences between people, such as the functioning of the amygdala (De Martino et al., 2010). Such biological differences tend to be rather stable over time, which makes the concept unhelpful in explaining longitudinal variance in the asymmetry between punishment and reward. Additionally, biological factors do not vary substantially between the small neighboring countries in the European Union, which makes loss aversion also comparatively useless for explaining such variance. Alternatively, social factors might lead to different loss aversion scores (Abdellaoui, 2007). Without a clear mechanism explaining which social factors influence the scores on loss aversion, however, it is more fruitful to use micro characteristics as demographic controls rather than running the risk of data-mining. Hence, there will be no separate hypothesis on loss aversion. Social mechanisms explaining varieties of asymmetry need to be identified to explain these differences. Nevertheless, in section 3D, the salience hypothesis will resemble to some degree the negativity bias of loss aversion. In the next sections, two mechanisms that can explain variations of asymmetry between and within countries will be discussed: issue salience and information.

3C. Differential explanation 1: splitting up ‘the economy’

In the early decades of research into economic voting, different macro indicators for economic performance have been used. Authors found that using inflation or unemployment yielded different results. Modeling lagged effects only diversified the outcomes, although the general negative effect of bad economic performance persisted (Lewis-Beck & Stegmaier, 2000). Although no consensus was reached or even approached, the field moved on nevertheless. The rise of subjective measurements of economic performance ‘solved’ the question which macro indicator was more important: subjective measurements of the state of the economy were thought to encapsulate various indicators, implicitly weighed on importance by the respondents (Lewis-Beck & Stegmaier, 2000).

Subjective measurements were an elegant step forward for multiple reasons (see section 2C) but are limited in their precision. These early studies indicate that different economic indicators may have different effects. This paper argues that this should not be interpreted in terms of one indicator being a

better reflection of voter judgments than another indicator, but as different indicators simply having different effects on voter judgments.

Many studies assume that different aspects of the economy have comparable results (Bellucci et al., 2012), but it is theoretically plausible that different indicators not only have different effects. In this case, the asymmetry between reward and punishment likely varies between indicators. In total, four types of indicators are good candidates for further inspection: economic growth, unemployment, inflation and the 2008 banking crisis. All four aspects are familiar to voters throughout Europe, enabling cross-country testing. Below, a tentative argument shall be made for the asymmetrical effects each indicator could have.

The clearest issue is the 2008 banking crisis: naturally, this has not yielded a reward before 2008. On the other hand, it has generated great disappointment in governments recently (Marsh & Mikhaylov, 2012). For this issue, the asymmetry between punishment and reward is maximal: there can be strong punishment, but there is no reward. On the continuum between full asymmetry and full symmetry,

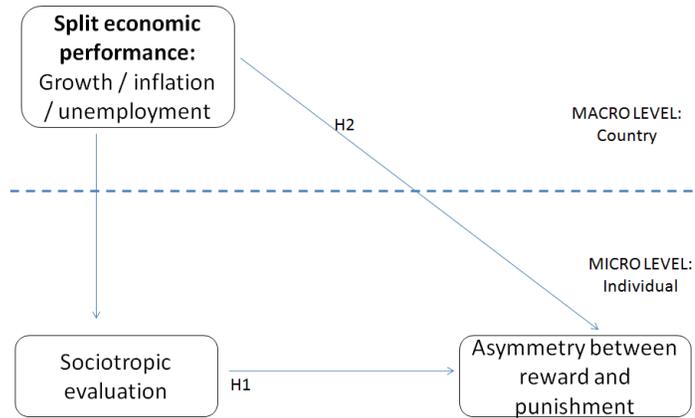


Figure 4. Different effects of economic indicators

the next aspect of economic performance is inflation. Inflation is noticed only when it is substantial; hence, this issue is a good candidate for a strong (yet not perfect) asymmetry as well. An American case study suggests that the fluctuations in inflation have a strongly asymmetric effect, because inflation rates can be high or low, but are very rarely negative (Lanoue, 1988, in Headrick & Lanoue, 1991). In fact, positive inflation rates may even be preferable over zero inflation for two reasons: inflation is negatively correlated to inflation-corrected interest rates, and a low amount of inflation can help smoothing wage adjustments.⁶ Therefore, low inflation rates potentially yield minor rewards, but these will be far smaller than the punishments.



Figure 5. Asymmetry between different economic indicators

⁶ Wages tend to be rather inflexible downwards. By keeping wages in certain sectors constant while positive inflation exists, it is possible to overcome this rigidity.

For economic growth and unemployment, it is more likely that these are relevant in both good and bad times. There may still be an asymmetry between reward and punishment, but this is likely smaller than for other indicators. This has yet to be empirically tested. A case study of Denmark in the 1980's has tried testing the effect of some different indicators, but found no differences in results, possibly due to the restricted variance in a single country (Nannestad & Paldam, 1997). The European Election Survey has a far greater N, and will be more suitable for empirical testing. For this study, the European Election Survey of 2004 will be used, which (naturally) includes no information on the banking crisis. This leaves three variables to form the second hypothesis: economic growth, unemployment and inflation.

H2 The difference between punishment and reward is greater for inflation than for unemployment and economic growth

3D. Differential explanation 2: asymmetry in salience

A second differential school of thought is built on the salience-hypothesis, which holds that voters cannot reasonably judge incumbents on all aspects of their output, and therefore only take a few salient dimensions into account (Singer, 2011). Economic concerns are more salient during a crisis, and higher salience strengthens economic voting (Bloom & Price, 1975; Van der Brug et al., 2007). This would explain greater coefficients for economic malaise leading to punishments, and economic growth leading to rewards.

However, while this general idea seems plausible, it needs to be adjusted to empirical reality. After all, over time, the economy has always been an important issue for voters (Wlezien, 2005). Case studies show that at every point in time, ‘the economy’ consistently ranks as the most important issue, toppled only incidentally by country-specific issues. However, different voters may differently value different economic issues (Jonung & Wadensjö, 1979).

Indeed, different economic problems such as unemployment and banking crisis changed in prominence (Anderson & Hecht, 2012; Nezi, 2012). Nezi analyzes Greek voters in 2004 (before crisis) and 2009 (during crisis). Jointly, an umbrella of economy-related concerns was mentioned by nearly three out of four voters as the most important issue. However, the composition of this umbrella changed drastically over time: the economy doubled from 20 to 40 percent, unemployment was over 50 percent in 2004 but less than half of

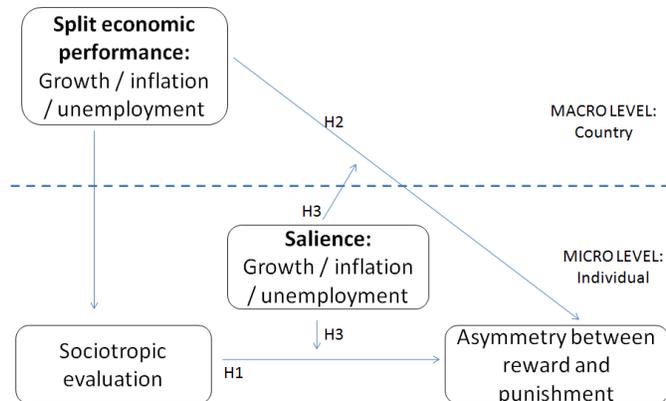


Figure 6. Asymmetry in issue salience

this in 2009, worries about inflation were marginalized, and the financial crisis was (obviously) only present in 2009 (Nezi, 2012).

One step forward, then, is to relax the assumption of all economic issues being mere sub-indicators of the general concept of economic performance, and instead conceptualizing them as related but separate issues on which the government can deliver good or bad performance. If different aspects of the economy change in salience over time, and this change is related to economic performance, it is logical to hypothesize a relation between the economic situation and salience of different economic indicators. The mechanisms discussed in section 3B on loss aversion suggest that negative connotations coincide more strongly with greater salience than positive connotations. Positive changes will be processed rationally by voters, where negative changes will move voters on both rational and emotional levels. A combination of emotional and rational triggers will have a greater impact on citizens than only rational triggers. Meanwhile, unlike the mechanism of loss aversion in itself, salience can vary substantially between neighboring countries and over time within countries. Therefore, salience is a useful variable and can be interacted with economic performance on the various issues. In total, four different economic indicators were substantially salient at one point in time: economic growth, unemployment, inflation and the 2008 economic crisis. As the data concerns 2004, the former three issues will be analyzed:

H3 The difference between punishment and reward is greater for more salient issues

3E. Differential explanation 3: asymmetry political affinity

Citizens vary in the degree they have affinity for politics. For reasons discussed below, it is plausible that the asymmetry between punishment and reward is smaller for citizens with greater political affinity. All groups of citizens are likely to punish bad performance, but those with a high political affinity will be more likely to also punish good performance. In this section, three elements of political affinity are discussed: news consumption, political interest and education, each of which is expected to interact with sociotropic economic judgments. While these aspects are related, they are fundamentally different and influence voting behavior at a different step in the causal chain, as will be illustrated afterwards.

An additional asymmetry in the voting decision can be found in the information that citizens use to judge the current situation. Standard theory on economic voting uses the implicit assumption that the transformation of actual economic events into citizen's minds happens in a symmetric way: all types of information have the same rate at which they reach citizens. However, this assumption is questionable at best (Headrick & Lanoue, 1991).

Citizens are almost never fully informed about governmental performance and the economic situation (Lupia, 1994; Marsh & Mikhaylov, 2012): they never have access to all information and the 'complete picture'. For practical reasons, they tend to accept that they will cast their vote based on incomplete

information. A rational choice perspective would argue that the expected benefit of the additional information (a near-zero chance that the different vote choice yields a more preferable election outcome) would be outweighed by the expected costs (sacrificing valuable spare time to better inform oneself) (Nannestad & Paldam, 1997). Alternatively, citizens may simply believe they are well-enough informed on the issues they care about most. Either way, the information used for the voting decision will resemble the information that is most readily available to voters.

Not only is the information citizens have about the economy incomplete, it also varies per person (Stevenson & Duch, 2013). This pattern of information gathering would have little impact on voting decisions if readily available information were representative of all information, but this seems implausible. Negative events are often regarded to make for more interesting news than positive events. Voters, then, will be better informed about negative developments than positive economic achievements. In short, voters are more likely to use negative information than positive information in their decision.

However, some would argue that citizens vary in the degree to which they are informed about politics and society, and differences in media consumption have been previously shown to affect voting behavior (Palmer & Whitten, 2011). The effect of negative information on prime-time television will become smaller for media consumers who read more newspapers and watch news on television more often, since positive developments will have more possibilities to reach the consumer. Unsophisticated citizen U will only read the headlines and the front page, where the most interesting (and, hence, more often negative) news can be found. Highly sophisticated citizen H will also read the subsequent pages in the newspaper, where the less spectacular news and backgrounds are covered, increasing the likelihood of reading good news. This provides theoretical support for the notion that the grievance asymmetry will be smaller for voters who spend more time consuming news.

H4a The difference between punishment and reward is greater for citizens that consume less news

A related, but conceptually different dimension of political affinity can be made between voters who are strongly or weakly interested in politics (Palmer & Whitten, 2011). Many people are barely involved in politics, and only a select number of issues stimulates them strongly enough to mentally connect their opinion on an issue to their opinion on incumbents (Lau, 1985). For politically more interested citizens, there will be more issues that pass this threshold and that will be mentally connected to their judgment of incumbents. For politically less interested voters, there will be a (stronger) bias towards negative trends surpassing the threshold than for their more interested counterparts, and therefore they will be less likely to reward incumbents for economic success. Hence, the corresponding hypothesis will be:

H4b The difference between punishment and reward is greater for politically disinterested citizens

The final step in the described chain of information processing is the usage of the information in making voting decisions. A case study of six European countries has provided tentative support for the idea that higher and lower educated citizens do not differ substantially in the information they consume.

Nevertheless, the information has a differential effect on their voting behavior, because the high educated are more likely to use the additional information in calibrating their vote (Duch & Stevenson, 2010a). This means that the expected effect of education is in the same direction as news consumption and political interest.

H4c The difference between punishment and reward is greater for lower educated citizens



Figure 7. Causal chain of political affinity moderating economic voting

It is highly plausible that the three aspects of political affinity discussed above are related to each other. Theoretically, it is not hard to think of mechanisms linking one’s education to news consumption and political interest later in life, or how political interest en media consumption can mutually reinforce each other. Empirical studies show that such relations indeed exist: for example, that more educated citizens are often more politically interested (Sunshine Hillygus, 2005), spend more time reading or watching news (Althaus, 2002; Stevenson & Duch, 2013), and that political interest and new consumption are

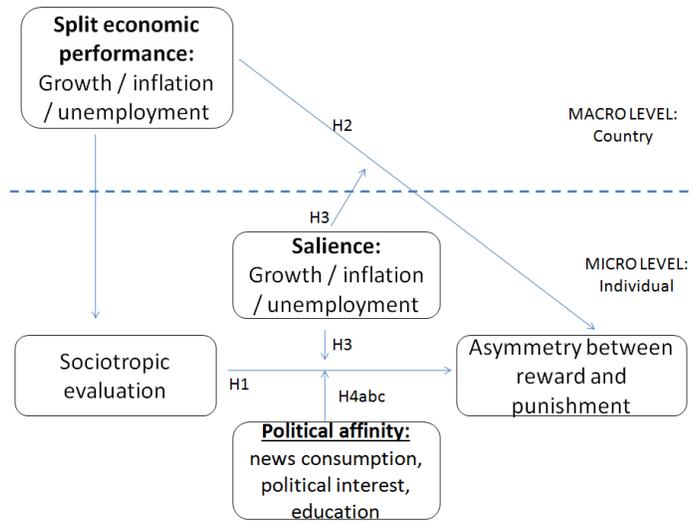


Figure 8. Full theoretical model

positively correlated (Strömbäck & Shehata, 2010). Nevertheless, the three aspects all need to be tested, as they are theoretically distinct: each influences the voter at a different step in the chain between macro economic developments and voting decision (as visualized in Figure 7). Even if H4a would be rejected and good and bad news reached all people in the same way, they may be a difference in how often they mentally connect this news to the performance of their incumbents. Also, even if people connect issues to the performance of the incumbent with the same frequency, they may still vary in the number of issues that they consider when casting their vote. Due to these substantial differences, findings based on a conceptualization in which education serves as a proxy for information (Krause, 1997) risk overestimating the effect of political affinity by wrongly treating effects of two different mechanisms in as a single effect.⁷ Therefore, all three aspects of political affinity will be tested.

⁷ Multiple specifications in principal component analysis, principal factor analysis and reliability analysis have been tried, but no specification leads to results that suggest the three different aspects of political sophistication can be combined into an index. Although every analysis hinted at unidimensionality rather than multidimensionality (Kaiser’s criterium >1 was always met by a single dimension), scores were below one or more critical values each

3F. Differential explanations: brief overview of alternative suggestions

Next to the explanations analyzed above, which are all theoretically well-founded and empirically testable, some other explanations have been suggested. However, these were either theoretically less convincing, impossible to be theoretically tested, or both. In order to be complete, these are all briefly outlined and discussed below.

Headrick and Lanoue argue that in some time periods, British voters paid more attention to economic changes than in other periods (Headrick & Lanoue, 1991). However, these time periods are rather arbitrarily selected, and substantive argumentation for this is missing. This explanation seems to be an explanation of last resort to ponder on unexplained variance, rather than an empirically generalizable finding.

A comparable explanation of last resort is the notion that the relation between economic conditions and voting has changed over time, including changes in the asymmetry. However, the notion that this shift is caused by ‘various changes in the nature of the electorate, parties or campaigns’ (Clagett, 1986, p. 625) is too vague and undefined to allow for empirical testing.

Next, there is the ‘mechanical argument’ based on ‘electorate potential’, which is the group of people who would consider voting for the party at some point. Parties in government are often parties that have done well in the previous elections. They capture a large share of their electorate potential, while opposition parties often capture a smaller share of the potential. Therefore, on average, government parties have relatively more to lose than opposition parties, and this creates an asymmetry (Van der Brug et al., 2007). This seems plausible, but it is hard to operationalize, and still cannot explain which voters in an electorate potential are more likely than others to punish or reward in the voting booth.

The ‘figure ground hypothesis’ argues that the effect of information depends on the contrast it poses with the existing mindset: optimistic people are more strongly influenced by negative information, and pessimistic people are more influenced by positive information. Since more people have a positive mindset, the impact of malaise is greater than the impact of prosperity (Lau, 1985). This hypothesis has never been tested, and justly so, as it is unclear how ‘intrinsic optimism’ can be properly modeled.

Finally, the effect of economic changes on voting behavior may be different for citizens who trust the government and for citizens who distrust the government (Lau, 1985). However, trust in government is probably related to economic performance, and modeling interactions between two variables would generate results that are problematic to interpret.

time. Since rather lenient critical values were used (communalities $>.2$, item loadings $>.4$, Cronbach’s alpha $>.6$), the empirical results confirm the theoretical argument to test each indicator separately.

Summed up together, the hypotheses that will be investigated are:

- H1 The difference between electoral punishment for declining economic conditions and electoral reward for improving economic conditions is greater than zero
- H2 The difference between punishment and reward is greater for inflation than for unemployment and economic growth
- H3 The difference between punishment and reward is greater for more salient issues
- H4a The difference between punishment and reward is greater for citizens that consume less news
- H4b The difference between punishment and reward is greater for politically disinterested citizens
- H4c The difference between punishment and reward is greater for lower educated citizens

4. METHODOLOGY, DATA AND OPERATIONALIZATION

4A. Methodology

In this chapter, the strategy for dissecting reward effects from punishment effects will be outlined. The goals of the are twofold: to separate punishment effects from reward effects, and to deal with the complications that arise when interactions are tested in logistic models. Both will be dealt with below. Before doing so, a baseline of effects will be established by using multilevel logistic regression analysis. Since each citizen can only vote for a single party, logistic regressions are common approaches for modeling voting behavior (Anderson, 2006; Costa Lobo & Lewis-Beck, 2012; Nezi, 2012). Votes are dichotomized to the options 0, vote for party currently in opposition, and 1, vote for party currently in government (Costa Lobo & Lewis-Beck, 2012; Nannestad & Paldam, 1997).⁸ The vote cast in the last election is controlled for. Following popular convention (Headrick & Lanoue, 1991; Nannestad & Paldam, 1997), baseline models will be used to estimate general effects of micro and macro variables before moving on to testing the asymmetry.

First, to separate the effects of punishment from the effects of reward, the main independent variables of economic change will be split into two variables. This split means that both objective and subjective economic effects will be divided in a positive change (for example, decrease in unemployment) and a negative change (increase in unemployment)(Nannestad & Paldam, 1997). Unlike the threshold models discussed in section 3A, a clear and non-arbitrary cut-off point exists for punishment and reward, at the value of zero. In these models, separate variables for positive and negative changes will included, enabling comparisons to their strength and significance (Lau, 1985). The original economic item, Sociotropic evaluation, will then be divided into:

- Positive evaluation, which is 1 for positive scores, else it is 0
- Negative evaluation, which is 1 for negative scores, else it is 0
- Neutral evaluation, which contains neutral evaluations and will be the reference category

If the function were symmetric, the punishment and reward models would have comparable explanatory power (expressed in R^2 , pseudo- R^2 or log likelihood) and mirror each other in terms of coefficient size, direction and significance (Clagett, 1986).

⁸ Alternatively, models could be conceptualized to compare voting for the main government party vis-à-vis the main opposition party. In the literature, this choice is understandably less popular (Costa Lobo & Lewis-Beck, 2012): it assumes that the main voting choice depends on only two parties, which is at best partly true in part of the countries, and discards a substantial share of the data. Worse, can bias the results, as the senior coalition partner often cannibalizes the junior partner(s): in elections, they gain seats at the expense of the smaller incumbent parties. Focusing on the main government party only can suggest growth where in total there is none.

Second, hypotheses H3 and H4abc argue that the effect of economic evaluations is moderated by salience, news consumption, political interest and educational attainment. In linear regression models, such moderations between variables X1 and X2 can be modeled by including the product of the variables, $X1 \cdot X2$. Although this multiplicative method is also regularly used in logistic regression models, it wrongly relies on the assumption that the way variable X1 is moderated by X2 is the same for all scores of X1 and X2. However, in logistic regression, logged odds coefficients are conditional on the values other effects have. It is still possible to find interactions that are both very significant and very strong, for all ranges of values, yet interactions of moderate strength are unlikely to be found. Thus, this method runs a strong risk of finding false negatives (Norton et al., 2004). Two solutions are used to properly test the interactions: separate regression analyses for different levels of salience and political affinity, and estimating marginal instead of multiplicative effects. For separate regression analyses, the moderator variables are dichotomized based on the median value,⁹ after which the dataset was split in half, and identical models were performed for both the high and low category. By comparing the (differences in) direction and significance of the positive and negative sociotropic evaluations, the presence or absence of an interaction can be discovered.¹⁰ For marginal effects, the multiplicative interaction is added to the model.¹¹ Then, this model is used to estimate the differences in marginal logged odds for both the high and the low category of the moderator.¹² Although it is not possible to test whether the differences in marginal effects are significant, these results can be used to corroborate the findings of the separate regression models: is the pattern comparable?

4B. Data and operationalization

For the analyses, the European Election Study dataset will be used. The European Election study consists of several separate studies; for this study, the voter study will be used. EES is used by many voting studies (Costa Lobo & Lewis-Beck, 2012; Van der Brug et al., 2007) and contains questions on social, political and economic attitudes as well as demographic background characteristics. The 2004 wave is most suitable, as it contains all variables needed to test the hypotheses (unlike older waves) and was held under 'normal' economic circumstances, whereas the more recent 2009 data concerns economic voting under exceptional economic conditions. The EES surveys are conducted every five year when voters across the EU can elect a new European Parliament. Earlier research has shown that most voters treat these elections as 'second order' elections: a moment to reflect on their national governments and an instrument to convey this opinion. Concerns about European affairs may have some influence, but national effects are so powerful that they resemble the effects present at national first order elections (Van der Brug et al., 2007). Additionally, EES data is not troubled by time period effects: unlike national

⁹ Alternative specifications were also performed; these will be discussed in the results chapter.

¹⁰ While theoretically uninteresting, this methodology also 'interacts' the other variables in the models by estimating different effects for each separate model. This should not be problematic, since (a) we have no theoretical reason to expect strong interactions between these variables, (b) we can empirically discover whether these interactions exist and (c) if anything, these only put an additional strain on the effect of (for example) political interest, which would lead to an underestimation of its interaction with economic evaluations rather than to an overestimation. Finally, if the effect of, for example, age, varies strongly between these models, the models will actually more closely resemble reality through it.

¹¹ Conform expectations, this interaction was insignificant in all models.

¹² This is done using the `-lincom-` command (Buis, 2010).

elections, all respondents were interviewed in the same time frame. Moreover, the unique peculiarities of national elections will be absent from this data (Van der Brug et al., 2007). EES waves are aimed at all EU countries, except for tiny Malta which could not field a research team in 2004 (EES, 2004). This leaves us with 24 countries, as Bulgaria, Croatia and Romania were not yet EU member in 2004.¹³

Earlier research on asymmetric effects consists primarily of case studies. Hence, they were often criticized for not accounting for country-specific idiosyncratic events and developments that may or may not have trumped economic effects (Headrick & Lanoue, 1991). Since it is not known which countries are representative for the population of countries in terms of asymmetric voting, using single-N or small-N research risks drawing conclusions based on exceptional cases. The use of a substantial group of countries entails a major step in dealing with such criticism. On the other hand, the EU countries are relatively homogenous in terms of wealth, level of democracy and quality of elections, which filters out countries so different that economic voting may work differently there.

MACRO LEVEL VARIABLES

There is a relatively broad consensus on the question of which economic indicators are most important: economic growth, inflation and unemployment (Nezi, 2012; Van der Brug et al., 2007). On top of maximizing the comparability of the results with earlier studies, these indicators are preferable over other economic statistics for several reasons. First and most important, these indicators are most often mentioned by media and voters, supporting the idea that they are indeed central in electoral decisions (Nezi, 2012). Second, these three aspects jointly capture economic trends sufficiently (Van der Brug et al., 2007). Third, on the empirical level, most indicators are measured differently across countries, yet these three are characterized by a high degree of standardization. This greatly increases comparability of the scores (Van der Brug et al., 2007). Data on these three economic indicators are drawn from Eurostat. Some have argued that it is not the absolute level of economic conditions, but rather the trend, that concerns the question voters ask themselves: is the change compared to the previous year positive, absent or negative (Headrick & Lanoue, 1991)? For all three indicators, this is measured as change in percent points compared to the previous year (Nezi, 2012).

On the macro level, two characteristics of the governments are added as control variables. Firstly, duration of government: regardless of the closeness of election outcomes, voters tend to give their new leaders the benefit of the doubt; a period commonly referred to as a 'honeymoon effect' (Bouvet & King, 2013; Headrick & Lanoue, 1991). The number of months in office is added as control variable. Secondly, incumbents' ideology may matter: right-wing parties are punished more severely for high inflation; left-wing parties for rising unemployment (Bouvet & King, 2013; Van der Brug et al., 2007). Others found that ideology does not influence anti-incumbency voting (Bartels, 2011). To adjust for this potential distortion, government ideology is controlled for by using a 5-point ranging from a strongly

¹³ In ESS 2004, Great Britain and Northern Ireland were surveyed separately. Northern Irish elections for the devolved parliament should be viewed as subnational rather than as national elections. Moreover, due to political instability, London has suspended the regional Assembly for the period 2002-2007 (Northern Ireland Assembly, 2014). Therefore, Northern Ireland will be excluded from the analyses.

right-wing government (1) to a strongly left-wing government (5).¹⁴ All political macro data are obtained from the Comparative Political Data Set III (Armingeon, Gerber, Leimgruber, Beyeler, & Menegale, 2012).

MICRO LEVEL VARIABLES

Next, the micro level variables will be discussed, roughly in order of importance. The exact phrasing of the questions that were read to respondents can be found in Appendix 8B. For all variables discussed in this section, basic descriptive statistics are listed in Appendix 8C. Unfortunately, many respondents had missing scores on theoretically important variables. The dispersion of respondents that were henceforth not included in (some of) the analyses can be found in Appendix 8C as well.¹⁵

Several ways of operationalizing the expression of voter satisfaction have been used, but the most obvious distinction is between incumbent popularity (Headrick & Lanoue, 1991) and actual vote casts (Bloom & Price, 1975; Nannestad & Paldam, 1997). Since actual votes have greater political relevance, more strongly incentivize citizens to reflect on their choice and enforce a stronger distinction between different parties, it is often considered to be a better conceptualization of how the economy is translated in citizens' behavior.

First, actual voting behavior is the variable that is most relevant in terms of political consequences (Lewis-Beck & Stegmaier, 2000). It could be argued that popularity measurements are uncontaminated by effects of strategic voting, and hence, may be a better reflection of how economic conditions influence voter attitudes. In the end, however, it is actual voting behavior and not attitudes that determines the leaders of democratic countries. Second, in part due to this greater importance, voters will take the question whom to vote for more seriously than a popularity score. Since they will be more inclined to contemplate what is important to them and how well the government performs on these issues, actual votes more accurately capture the mechanism. Third, since votes can only be cast on a single party, citizens have to differentiate clearly between their first choice and the rest. On the other hand, popularity functions allow for identical scores between candidates, and a ranking of parties on their popularity may not coincide with the distinction between government and opposition parties, blurring the effect. For these three reasons, voting and not popularity is chosen as dependent variable.

In the EES 2004, respondents were asked for the party they had voted for in the most recent general

¹⁴ This Schmidt Index is based on the share of total cabinet posts held by left-wing parties versus the share held by centre and right-wing parties: 1 = left hegemony (100% of all posts held by left parties), 2 = left dominance (66%-99% of the posts), 3 = balance (33-66%), 4 = right dominance (1-33%), 5 = right hegemony (no posts).

¹⁵ To maximize the number of respondents for each analysis, no listwise deletion was applied. This way, the loss of external validity is minimized. A drawback may be that differences between the models are conflated due to compositional effects, i.e. that the models are different by virtue of being estimated for different groups of respondents. Robustness checks show no signs of substantial compositional effects. Moreover, compositional effects complicate model comparisons, yet the aim of the models is not to test relative strength of variables but whether the effect of evaluations is moderated or not.

elections, and which party they would vote for if there were general elections tomorrow. These answers were combined with data on governing parties (Slomp, 2011; supplemented with various governmental websites) to divide respondents between those that voted for party currently in opposition (0) from those voting for a party currently in government (1).¹⁶

For respondent's judgment on economic performance, this could be operationalized as either absolute economic performance (for example, the level of unemployment) or relative economic performance (for example, whether unemployment is higher or lower than last year). The latter option is preferable, as it more closely matches the theoretical notion of judging incumbents on the change they bring rather than a situation that also (perhaps mainly) was generated by previous governments. Also, citizen estimates of exact percentages of inflation, unemployment or growth are far less reliable than estimates of the general trend, and would probably generate great biases, great numbers of unanswered survey questions, or both. Therefore, it is common to use survey questions asking citizens whether they believe the economy has improved, stayed the same or deteriorated over the past year (Anderson, 2006; Bellucci, 2012). The time period should be no longer than the past year, as this is most important for voters (Bloom & Price, 1975). In the EES, this question is asked on a scale which ranges from 1 (much worse) to 5 (much better). This is recoded into a variable consisting of three categories (worse, stayed the same, improved) to increase comparability with the models in which positive and negative evaluations are modeled separately.¹⁷

In the EES, respondents were asked what they considered to be their country's most important problems (MIP) as open-ended question. MIP items are part and parcel of current day political datasets and voter studies (Anderson & Hecht, 2012), but are not uncontroversial.¹⁸ Wlezien warns against using MIP questions for measuring salience, concluding for several reasons that it is potentially problematic: (1) the MIP item mixes up the respondent's assessment of how relevant and how critical a certain issue is, (2) respondents can interpret importance in multiple ways; (3) the number of people considering P to be the most important problem also depends on shifts in importance attached to problems Q, R and S and (4) using MIP discards all information about issues that are considerably important, yet not the most important issue (Wlezien, 2005).

The remarks of Wlezien should not be ignored. Nevertheless, MIP items are considered valid and useful for the purposes of the present study. The reasoning for this conclusion is outlined below.

The first argument, as presented, is that for something to be considered the MIP, this includes both

¹⁶ For past vote, the number of opposition voters exceeds the number of government voters. This is likely the result of minority governments in seven countries (Denmark, Latvia, Poland, Romania, Slovakia, Spain and Sweden) and electoral systems that tend to grant large parties a greater share of seats than proportional to their share of votes (Armingeon et al., 2012).

¹⁷ An additional advantage is that this operationalization enables the use of Swedish data. In Sweden, only these three answer categories were asked, which can hereby still be used. In Spain, the question referred to the previous six months instead of the previous twelve months. Omitting Spain from the estimation does not substantially influence the models.

¹⁸ See, for a list of examples, (Wlezien, 2005).

whether an issue is important or not, and whether the situation on this particular issue is good or bad. War will not be MIP in times of peace even when physical safety would rank very high on the list of priorities (Wlezien, 2005). While it is beyond doubt that MIP captures elements of both, this need not be problematic. Theoretically, a positive correlation between graveness and importance of a problem is plausible, and empirically, actual economic conditions are controlled for, so MIP will primarily capture an importance effect.

Second, Wlezien argues that importance could be conceptualized in different ways: respondents could focus on short-term or on long-term importance of an issue, and could differ in whether something is important for their personal lives or important for the state of the country (Wlezien, 2005). The confusion for this question depends on how the question was phrased. In the case of the European Election Study, the questions explicitly and very specifically ask about what the respondents considers to be the MIP: “*What do you think are the most important problems in [country] at present?*” and “*Of those you have mentioned what would you say is the single most important problem?*” (EES, 2004). Therefore, it can be assumed that in this particular questionnaire, the vast majority of the respondents conceptualized MIP in the same, correct way. Additionally, I could find no argument suggesting that the number of erroneous conceptualizations varies over time or across countries.

Third, Wlezien correctly points out that a respondent may trade their perceived MIP not only due to fluctuations in importance of this particular problem, but also due to fluctuations in importance of other problems (Wlezien, 2005). When respondents have to single out one problem, naturally this is influenced by their opinions on both this particular problem and on other problems. Although presented as problematic, it simply reflects that voters prioritize. A voter could believe inflation is the MIP presently in country C due to a recent drop in crime rate or due to an increase in inflation, but either way, inflation will be crucial in shaping her vote.

Fourth, asking for the MIP creates a winner-takes-it-all distribution of answers. Only a few issues have been considered to be most important by a substantial number of people. Problems that are considered to be important but not vital will be eclipsed by the vital problems, and hardly distinguishable from issues that are deemed unimportant at all. Here, Wlezien uses the clear example of the number of Americans considering ‘welfare’ to be the MIP: this is fairly low, whereas the conclusion that it is not deemed important is strongly rejected by other data (Wlezien, 2005). Admittedly, the importance attached to second-tier issues will be strongly understated by the MIP item. However, this study involves first-tier economic issues, which in earlier case studies are fairly often ranked as MIP (Anderson & Hecht, 2012). Therefore, in the particular case of economic issues, MIP measurements will be far less troubled by the winner-takes-it-all structure distribution of the answers. Concluding, the MIP item is a useful measurement for salience.

Political affinity is conceptualized in three different ways: political interest, newspaper readership and level of education. Political interest is measured on a four-point scale ranging from 1 (not at all interested) to 4 (very interested). Newspaper readership is measured in days per week, ranging from 0

to 7.¹⁹ Education is measured as the age at which respondents have stopped full-time education. These variables are used for analyses that require splitting the data; hence, the number of categories needs to be reduced. Separating the analysis over too many categories means that there will be too few respondents in each of the cells in each country, and that there is a substantial risk of finding false negatives. Therefore, newspaper readership and education are recoded into quartiles of roughly the same size, and all are dichotomized based on the median value. Obviously, the latter operationalization is more successful in reducing the risks of too small groups per cell, and will be the main operationalization, whereas the former will be used as a robustness check. Theoretically, an argument can be made that there is a more interesting cutoff point for newspaper readership: between those who do and those who do not read newspapers at all. Therefore, the models will be repeated using this alternative dichotomization.

The analyses control for gender, age, immigration status, which is dichotomously coded (0 = native, 1 = foreign-born), a ten-point scale for left-right self-placement (1 = left, 10 = right), a ten-point scale showing the ideological difference between respondent and incumbent government,²⁰ trade union membership,²¹ a five-point scale self-placement of social class (Palmer & Whitten, 2011),²² urbanization of the respondent's place of living (1 = rural area or village, 2 = small or middle sized town, 3 = large town),²³ income which is coded in quintiles, and religiosity (0 = not religious, 1 = religious)²⁴.

4C. Endogeneity issues: the danger of reverse causality

The multilevel logistic models sketched in paragraph 5.1 may have the virtue of parsimony, yet they do not adequately deal with the risk of endogeneity. These models rely on correlational evidence, but multiple studies have argued that there may be reverse causality, and that this biases the correlations. Below, this problem will be discussed and two different solutions will be evaluated: the use of panel data and the use of exogenous instrumental variables. For multiple reasons, exogenous instrumental variables will be used in a two stage probit analysis.

¹⁹ 87 respondents answered '8 days'; these were recoded to '7'. Alternatively, this group of respondents were excluded from the analysis, which made no meaningful difference. In Sweden, a binary yes / no answer to reading newspapers 'sometimes' could be answered, which is only used for the analyses in which newspaper consumption is dichotomized.

²⁰ Here, left-right self placement is recoded in to a five-point scale and absolute difference with government ideology is calculated by $\sqrt{(\text{respondent ideology} - \text{government ideology})^2}$.

²¹ Alternatively, a variable measuring whether someone in the respondent's household is a trade union member was used. Neither had a significant effect in any model and the effect on other variables was virtually zero.

²² Since the highest category, Upper class, only contains 391 respondents, the fourth and fifth classes are combined. This only marginally changes regression coefficients but is more supportive of modeling interactions.

²³ In two countries, different answering categories had been used, which were recoded. For Poland, villages were coded 1, towns up to 99,999 were coded 2, and cities >100,000 were coded 3. For the Netherlands, not urbanized at all was coded 1, little / somewhat urbanized were coded 2, and strongly / very strongly urbanized were coded 3.

²⁴ Alternatively, a 5-point scale measuring religious attendance, in which non-religious people score 1 and 5 means attending several services per week.

Several authors warn against the endogeneity of economic perceptions vis-à-vis political preference (Evans & Andersen, 2006; Evans & Pickup, 2010; Palmer & Whitten, 2011): voting intentions or partisan preferences may influence the way the economy is assessed, i.e. Republicans will judge President Obama more harshly than their Democratic counterparts, simply because they are Republicans. The greater the ideological distance is towards the governing party or parties, the more negative the economic judgment will be. If this is the case, the correlational evidence provided by the statistical models may be correct, but has to be interpreted reversely (Anderson & Hecht, 2012; Evans & Andersen, 2006). The mechanism through which preference influences economic assessment is selective perception: people are more inclined to acquire information

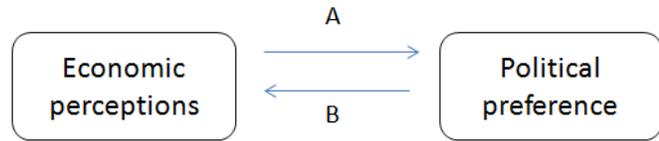


Figure 9. The problem of exogeneity

(and to continue reading) that matches their predispositions than information that argues otherwise (Hansford & Gomez, 2011).²⁵ As schematically presented in figure 9, this would mean that the strength and significance of the main theoretical mechanism, as presented by arrow A, may be biased due to the (not modeled but in reality actually existing) effect of arrow B. Drawing valid conclusions on the voting effects of the economy becomes highly problematic.

Two kinds of solution could be used to deal with endogeneity problems: (1) using panel data and (2) using exogenous instrumental variables.

The first way to deal with this issue is to use panel data: repeatedly interviewing the same respondents over time. Proponents argue that the economic perception of a respondent at time point T-1 cannot possibly be caused by the political preference and voting intentions of the same respondent at time point T (Anderson & Hecht, 2012). While this is

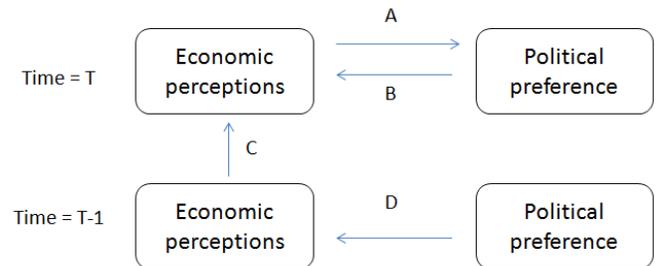


Figure 10. Panel data as a solution for endogeneity

technically true, and an improvement compared to analyses that do not attempt to combat the endogeneity problem, it is still unsatisfying. If reverse causality really is at work here, then the economic perception at T-1 is still caused by political preferences, only those held by the respondent at T-1. Schematically speaking (see Figure 10), the risk of endogeneity via arrow B is replaced by the risk of endogeneity via arrows C and D. Since political preferences are strongly autocorrelated over time, both economic assessment and political preference at time point T are influenced by political preference at time point T-1. Effectively building a time lag into the equation may have its merits, but will not be the

²⁵ A British case study of 2004–2009 has argued that the effect of economic satisfaction was stronger for citizens with high incomes. Since Labour governed the country for this period, and income is negatively correlated to voting for Labour, the authors conclude that this supports the endogeneity thesis: for low income voters, economic malaise is less often a reason to reject the Labour government because they are more inclined to vote Labour in general (Palmer & Whitten, 2011). However, this argument depends strongly on the assumption that income can be used as a valid proxy for party preference, an assumption the authors correctly admit is dubious.

solution for endogeneity issues.

A second, more promising approach is to use exogenous instrumental variables in a two-stage regression analysis (see figure #11). In the first stage, the independent variable of interest (economic perceptions) is regressed on one or multiple variables (economic situation) known to be exogenous to the dependent variable (political preference), to generate predicted scores on the variable of interest (Hansford & Gomez, 2011). In the second stage, the dependent variable is not regressed on actual scores of the independent variable, but on the predicted scores generated in the first stage. This way, the error terms of economic perceptions will not be influenced by the scores on political preference.

This method relies heavily on the selection of appropriate instruments, which need to be a theoretically valid and empirically reliable way to predict one’s economic perceptions. Exogenous instruments need to meet two theoretical criteria: they may not be influenced by the dependent variable, and they should influence the dependent variable *indirectly* via economic evaluations rather than directly influence the dependent variable. Theoretically, we

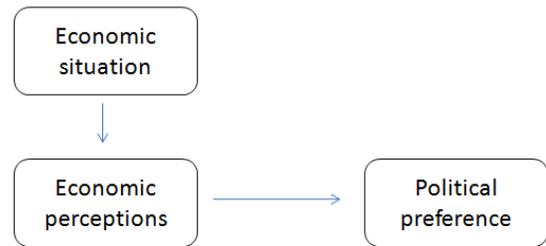


Figure 11. Exogenous instruments as a solution

should recall the debate on egotropic en sociotropic voting presented in the second chapter. Sociotropic evaluations are derived from various information sources, but the two most important are media, which inform voters about actual economic conditions, and personal experiences, which are in part related to demographic characteristics.²⁶ This means that sociotropic evaluations can be (to some extent) predicted by a combination of objective economic indicators and individual-level characteristics. Obviously, economic performance will be the same for each inhabitant of a country; hence, individual level predictors will be necessary for proper estimation. For both types of instruments, the risk of reverse causation is substantially smaller than for sociotropic evaluations: neither objective economic performance nor demographic characteristics are influenced by a respondent's ideological distance to the government.²⁷ Here, three instruments are used: social class, left-right self-placement and the level of inflation. Theoretically, all would appear to be exogenous to economic evaluations. In terms of causal mechanisms between instruments and economic satisfaction, all three make sense: people from different social classes have plausibly experienced economic fluctuations over the past year in different ways, leftist people care more about some economic indicators, rightist people care more about others, and inflation (by definition implying a longitudinal comparison) is part of the economy that respondents form an opinion about. Since governments over Europe vary strongly in types of parties in power, social class and left-right stance should have no direct effect on incumbency voting, and the relation between objective and subjective indicators has been analyzed in section 2C. To summarize: theoretically, these variables are valid instruments.

²⁶ A more direct way to model this might be an item measuring egotropic evaluations, but this was absent in the EES 2004 data.

²⁷ In fact, the risk of endogeneity biases may even be lower for macro-economic performance and demographic variables than for egotropic evaluations (Hansford & Gomez, 2011).

Empirically, three additional checks can be done to discover whether the instruments are reliable predictors of the endogenous variable. Unlike many theoretically equally plausible candidates, social class, left-right stance and inflation have sufficient scores on all three tests. Since it is impossible to perform these tests on instrumental variable probit regression models,²⁸ instrumental variable two stage least squares models are performed using the same specifications. Because the first stage regression estimates are identical for the probit and linear regression models, the first stage regression results for two stage least squares models can be used for the purpose of testing the probit models (Davison, 2012).²⁹ First, instruments need to be strong enough predictors of economic evaluations: the instruments should explain enough variance of economic evaluations. Otherwise, the standard errors estimated in the second stage will be too small and there is a risk of finding false positives. For this purpose, F-tests are used. Second, to empirically confirm that the dependent variable does not influence the instruments, an endogeneity test is performed. Finally, since multiple exogenous variables are used to estimate a single predictor, Sargan and Basman Chi² tests³⁰ can discover whether the instrumented variable is over identified by the instruments (Fisher, 2010). Over identification concerns the efficiency of the models, but over identification can be acceptable if there is a trade-off with instrument strength: adding a third instrument makes the model less efficient, but the increase in strength also makes it less biased (Söderbom, 2009). Performing these tests for models including general economic evaluations, positive evaluations and negative evaluations shows that all models score sufficiently for all tests, except for the over identification tests which show that the negative evaluations variable is over identified by the instruments (albeit only at the $p < .05$ level, yet not at the $p < .01$ level). As mentioned, the price of combating this inefficiency is that instruments would become weaker. Especially for sufficiently large datasets such as the EES, instrument strength is preferred. It should be noted that the instruments are sufficiently strong to be used, but not that much stronger than the threshold level. All test statistics can be found in Appendix 8D.

Earlier studies have empirically shown that on the one hand, economic perceptions are indeed influenced by political attitudes. On the other hand, when economic perceptions are modeled in such a way that they are (made) exogenous, economic perceptions still have a powerful impact on voting decisions (Stevenson & Duch, 2013). This section has shown that jointly, the exogenous variables are theoretically plausible and empirically reliable instruments for economic evaluations. Therefore, they can be used for hypothesis testing in section 5.

²⁸ Although it should theoretically be possible, STATA 12 does not allow these estimations. Fortunately, the results can be obtained nevertheless in the way explained in the text.

²⁹ Clearly, the second stage regression results in which incumbency voting is treated as a continuous variable are meaningless, and can be ignored.

³⁰ Unfortunately, these tests cannot be done on cluster-robust standard errors. Hence, these tests are performed on single-level rather than multilevel models with otherwise identical specifications.

5. EMPIRICAL RESULTS

5A. Testing of asymmetry between punishment and reward

The first step in the analyses is to establish whether the predicted asymmetry between punishment and reward is indeed present in the EES 2004 data.³¹ In its most basic form, the relation between economic developments and incumbency voting is modeled in Table 1.

Table 1. Incumbency voting regressed on economic indicators, multilevel logistic regression

	General	Positive	Negative	Split
Economic evaluation	.452 ***			
Positive evaluation		.723 ***		.356 ***
Negative evaluation [#]			.876 ***	.727 ***
Economic growth	.089 *	.085 ***	.089 ***	.009 *
Unemployment change	.234	.203 ***	.203 ***	.220
Inflation change	.014	.017	.027 ***	.020
Constant	-1.688 ***	-.993 ***	-.497 ***	-.971 ***
Pseudo R ²	.043	.025	.038	.042
Log pseudo likelihood	-11,290	-11,515	-11,360	-11,320
Micro level N	17,536	17,536	17,536	17,536
Country level N	22	22	22	22

***= p<0.001; **=p<0.01; *=p<0.05. Reported are logged odds.

[#]: this variable is reversely coded to enable testing equality

As shown in the general model, retrospective sociotropic economic evaluations have a significant effect on incumbency voting. Conform the expectations, this effect is strong, positive, and far stronger than the objective macro indicators. In the split model, the variable on sociotropic evaluations is separated in two dummies. Since negative evaluations were reversely coded, the coefficients can be compared vis-à-vis each other.³² The difference in strength is clearly observable in Table 1 and significant at p<.05.³³ In terms of model strength, adding the variable for negative sociotropic evaluations leads to a greater increase in Pseudo R² (and to a greater decrease in log likelihood) than adding the positive sociotropic

³¹ All analyses are performed in STATA 12.

³² Since in logistic regression, coefficient sizes are estimated conditionally to other effects, the reversely coded variable is no identical mirror counterpart of the original. Nevertheless, it approximates the true variable and only marginally influences estimates of the other variables.

³³ Chi-square test is done using the –test positive = negative- command after model specification.

evaluations.³⁴ On the whole, Table 1 suggests that economic effects on incumbency voting exist, and that the reward for economic improvements is smaller than the punishment for economic decline. Jointly, these comparisons provide support for hypothesis 1: punishment is indeed greater than reward.

On the other hand, the difference between punishment and reward is far from uniform across countries. As shown in Appendix 8E, the punishment effect exceeds the reward in only 10 of 22 countries, whereas reward exceeds punishment in 10 other countries, and neither achieves significance in Finland and Poland. At this point, it appears that on average punishment for economic malaise is greater than reward for economic improvements, but also that this effect is strongly differential. On the one hand, such cross-country differences could be caused by country-level variation in, most probably, various aspects of economic performance. On the other hand, it could be caused by compositional effects, as countries contain different populations (i.e. country A has more high educated citizens than country B). The former effects will be tested first in section 5B by relaxing the assumption that voters treat ‘the economy’ as a single concept. The latter will be analyzed in sections 5C and 5D on issue salience and political affinity.

5B. Splitting up ‘the economy’

In section 3C, it is argued that using different indicators for economic performance may not be a question of finding the optimal measuring stick. Instead, different indicators could simply have different effects on the voting calculus: it is hypothesized that the asymmetry is greater for inflation than for economic growth and unemployment. If this is true, then countries scoring comparably on economic satisfaction may still differ in terms of punishment and reward, if the economic performance is based on different indicators. After all, there is a significant correlation between the three main economic indicators (GDP per capita growth, unemployment and inflation), but this relation is far from perfect and most often rather weak.³⁵

Tables 2, 3 and 4 present the results for the various economic indicators. Each table consist of a model containing only the main effect (adjusted for the vote in the previous election), a macro model containing all economic indicators, a sociotropic model also accounting for general economic evaluation and finally a full model with all demographic controls. Each model in Table 2 shows the same pattern: the contribution of economic growth to the model is insignificant. This is rather remarkable, since so many earlier studies have found the economic growth matters. The most plausible explanation is that there is too little variance at the country level: in contrast to the large N at the micro level, only 22 countries are included. Therefore, the lack of variance leads to insignificant coefficients for economic growth.

³⁴ Wald and Likelihood ratio tests on these non-nested models show that the contribution of negative evaluations is greater, but cannot test whether the difference is greater.

³⁵ Pearson correlation coefficients are, for changes in GDP per capita and unemployment: -.234; for GDP per capita and inflation: -.070, for inflation and unemployment: .467, all significant at $p < .001$ level.

Table 2. Incumbency voting regressed on economic indicators, multilevel logistic regression

	Single		Macro		Sociotropic		Full	
Vote in last election	4.667	***	4.693	***	4.609	***	4.542	***
Evaluation					0.344	***	.380	***
Economic growth increase	-.090		-.062		-0.056		-.068	
Economic growth decrease	-.046		-.005		-0.005		-.121	
Unemployment change			.434		0.500		.167	
Inflation change			-.049		-0.061		-.044	
Ideological distance							.001	
Age							-.442	***
Social class							-.107	
Gender							.013	**
Immigrant status							.038	
Union membership							.141	
Urbanization							.084	
Religiosity							-.069	
Income quintile							.176	
Time since election							-.025	
Constant	-3.009	***	-3.112	***	-3.731	***	-3.928	***
Pseudo R ²	.538		.540		.541		.572	
Log pseudo likelihood	-4,973		-4,953		-4,675		-2,914	
Micro level N	15,900		15,900		15,040		10,101	
Country level N	22		22		22		21	

***= p<0.001; **=p<0.01; *=p<0.05. Reported are logged odds.

A comparable situation is shown in Table 3 for unemployment, which, too, contributes little when to explaining economic voting. In most models, both an increase and a decrease in unemployment yield insignificant coefficients. The sociotropic model appears to be an exception, yet this estimate is in the wrong direction and lacks robustness, as the effect disappears when control variables are added or removed. The conclusions based on these results are the same as in the models for economic growth: there are simply too few country-level cases for proper testing of this variable; hence, effects are insignificant when controlled for basic demographic and attitudinal characteristics.

Lastly, for inflation, results are presented in Table 4. For an increase in inflation, the same interpretation can be used as in the previous tables: its insignificance is likely caused by the small country-level variance.³⁶ For a decrease in inflation, an effect is robust enough to be found in the full model that controls for all demographic and attitudinal characteristics. However, the effect is negative: a decrease in inflation is punished by voters. The hypothesis predicted that the reward effect of decreasing inflation would be smaller than the punishment for increasing inflation, or potentially absent. A negative effect

³⁶ This problem is absent in the longitudinal studies that tend to be the habitat of objective economic indicators (Alesina et al., 1993; Bloom & Price, 1975; Clagett, 1986; Headrick & Lanoue, 1991)

for a decrease in inflation was not expected.

Table 3. Incumbency voting regressed on economic indicators, multilevel logistic regression

	Single		Macro		Sociotropic		Full	
Vote in last election	4.653	***	4.680	***	4.591	***	4.567	***
Evaluation					.369	***	.381	***
Unemployment increase	.597		.676		.806	*	.573	
Unemployment decrease	.115		.391		.278		.972	
Economic growth			.091		.094		.053	
Inflation change			-.062		-.073		-.086	
Ideological distance							-.006	
Age							-.391	***
Social class							-.150	*
Gender							.013	**
Immigrant status							-.056	
Union membership							.166	
Urbanization							.127	*
Religiosity							-.056	
Income quintile							.212	
Time since election							-.031	
Constant	-3.293	***	-3.653	***	-4.367	***	-4.195	***
Pseudo R ²	.539		0.540		.543		.569	
Log pseudo likelihood	-4,964		-4,935		-4,653		-2,935	
Micro level N	15,900		15,900		15,040		10,101	
Country level N	22		22		22		21	

***= p<0.001; **=p<0.01; *=p<0.05. Reported are logged odds.

A potential explanation is that absolute levels of inflation in the analyzed countries are relatively low: for all but five, it is lower than four percent. The arguments presented earlier on why low inflation may be preferable over zero inflation may be the reason why voters do not intend to punish incumbents for a (minor) increase in inflation. The idea that low inflation is preferred over both high inflation and very low inflation corroborates the findings of a Greek case study: inflation has a positive effect on government evaluations, up to a tipping point after which a further increase in inflation is judged strongly negatively (Nezi, 2012). This mechanism, too, supports splitting up effects of ‘the economy’ into several different effects that are all part of ‘the economy’, but are treated differently by voters. These analyses merit further studies that no longer treat economic performance as a single construct, but explore the differential effects of inflation, economic growth and unemployment. For now, the hypothesis that the difference between punishment and reward is greater for inflation than for economic growth and unemployment has to be rejected: none of the indicators performed as hypothesized.

Table 4. Incumbency voting regressed on economic indicators, multilevel logistic regression

	Single		Macro		Sociotropic		Full	
Vote in last election	4.642	***	4.707	***	4.621	***	4.612	***
Evaluation					.341	***	.366	***
Inflation increase	.248		.136		.119		.182	
Inflation decrease	-.402	*	-.385		-.376		-.492	**
Economic growth			.075		.078		.043	
Unemployment change			.492		.531		.688	
Ideological distance							.004	
Age							-.421	***
Social class							-.163	*
Gender							.014	**
Immigrant status							.030	
Union membership							.086	
Urbanization							.120	*
Religiosity							-.044	
Income quintile							.242	
Time since election							-.025	
Constant	-3.549	***	-3.908	***	-4.502	***	-4.971	***
Pseudo R ²	.514		.546		.547		.575	
Log pseudo likelihood	-4,939		-4,893		-4,619		-2,893	
Micro level N	15,900		15,900		15,040		10,101	
Country level N	22		22		22		21	

***= p<0.001; **=p<0.01; *=p<0.05. Reported are logged odds.

At this point, this study has shown that economic voting is on average asymmetrical, and that this asymmetry varies strongly between different countries, but splitting different macro indicators has not contributed much to explaining these findings. The next step will be to analyze cross-level interactions between issue salience on the one hand, and general economic performance and the three macro indicators on the other hand.

5C. Issue salience

The third hypothesis argues that the more salient an issue is, the more value citizens attach to the performance on the issue, and the stronger its effect on voting behavior will be. Also, the more salient an issue is, the greater the asymmetry between punishment and reward will be, because negative effects will trigger voters on a more emotional level. On average, voting populations tend to care most about economic issues, but this varies between and within countries. Therefore, in Table 5, the population is divided between those who viewed any non-economic issue as the most important problem in the country, in the low salience models, and respondents who selected any economic issue as the most salient issue, in the high salience models. In Appendix 8f1, identical models with economic evaluations as a single variable is presented. This procedure will be repeated in section 5D for political affinity models. Table 5 shows that negative perceptions are important in voting behavior regardless of whether respondents believe the economy is the most important issue or not. Positive evaluations, on the other hand, are only significant for respondents who did not view the status of the economy as the most important problem facing the country. Respondents who feel that their country faces severe economic issues, even though they believe that the situation is better than a year before, will not reward their government for it. This confirms the expected mechanisms.

It should be noted that ideological distance towards the government is a vital control variable: models that did not control for ideological distance lead to different results. Coefficients for negative evaluations are unscathed (they actually increase in size) but positive evaluations are no longer significant at the $p < .05$ level ($z = 1.80$, $p = .072$). Ideological distance is an important and robust predictor of the inclination to vote for the government and should be included for this theoretical reason as well. Nevertheless, it casts doubt on the strength of this finding. Naturally, this issue is also present in the instrumental variables probit analysis in Table 6: since left-right stance is one of the instruments, and left-right stance is used to calculate ideological distance, ideological distance has to be omitted as a control variable. A final remark on general economic salience can be made using the marginal effects³⁷ presented in Figure 12. These do not confirm the trend of the regression models: negative evaluations have stronger effects in both groups, but the difference is (marginally) smaller for the high salience group.

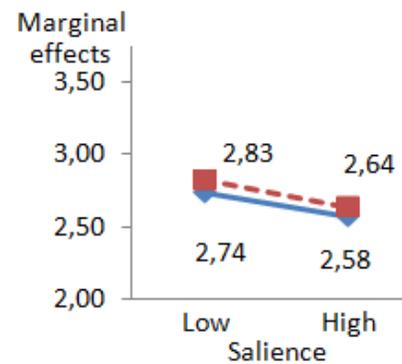


Figure 12. Marginal effects of positive (solid line) and negative (dashed) evaluations by salience

The separate regression models were generated for interactions between unemployment score and salience of unemployment, and in the same fashion, for inflation and economic growth. At no point did this yield significant results, not even for negative economic

³⁷ In Appendix 8f2, marginal effects for all four moderations tested in this and the next paragraph are also summarized in the form of a table.

evaluations. Several reasons could be ground for this. First, the very small number of people choosing inflation or the economy as the most important issue makes estimates less reliable and creates a power problem. For the latter issue, this seems to be related to the specificness of alternative economic items available, which differed per country (an issue absent from the analyses in Table 5). Second, for unemployment, the effect was not significant in other models either, likely (as mentioned earlier) due to the small number of level 2 units. In short, the results are mixed: Everything considered, the hypothesis on differential effects for different economic issues has not yielded a conclusive answer. Therefore, the final set of hypotheses, on political affinity, will be tested on the way these variables condition the asymmetry between punishment and reward.

Table 5. Incumbency voting regressed on economic indicators, multilevel logistic regression

Economic salience	Low		High		Low		High	
Vote in last election	4.395	***	4.460	***	4.495	***	4.680	***
Positive evaluation	.280	*	.212		.264	*	.271	
Negative evaluation [#]	.473	***	.364	*	.553	***	.471	***
Ideological distance	-.454	***	-.411	***	-.430	***	-.363	***
Age					.012	**	.015	*
Social class					.129	*	.076	
Gender					-.031		-.306	*
Immigrant status					.059		-.176	
Union membership					.209		.085	
Urbanization					-.050		-.054	
Religiosity					.154		.262	
Income quintile					-.024		-.035	
Economic growth	.020		.156		-.004		.138	
Unemployment change	.473		.872	*	.426		.974	*
Inflation change	.0139		-.150		.002		1.175	
Time since last election	-.005		-.004		-.007		-.002	
Constant	-2.918	***	-2.878	***	-3.055	***	-3.929	***
Pseudo R ²	.562		.570		.562		.583	
Log pseudo likelihood	-2,409		-1,764		-1,672		-1,242	
Micro level N	8,060		6,125		5,596		4,505	
Country level N	22		21		21		20	

***= p<0.001; **=p<0.01; *=p<0.05. Reported are logged odds.

[#]: this variable is reversely coded to enable testing equality

Table 6. Incumbency voting regressed on economic indicators, instrumental probit analysis

Economic salience	Low		High		Low		High	
Vote in last election	1.725		1.870	*	1.079		1.851	
Positive evaluation	1.692	*	1.658					
Negative evaluation [#]					1,988	***	1.377	
Constant	-1.639	***	-1.617	***	-.052		-.665	
Log pseudo likelihood	-7,216		-4,857		-7,586		-5,817	
Micro level N	7,831		5,893		7,831		5,893	
Country level N	22		21		22		21	

***= p<0.001; **=p<0.01; *=p<0.05. Reported are probit coefficients.

[#]: this variable is reversely coded to enable testing equality

Instruments: inflation, social class, left-right self placement

Positive models: F = 13.82 Partial R² .015 Adjusted R² = .041

Negative models: F = 10.44 Partial R² .009 Adjusted R² = .051

5D. Political affinity

Finally, political sophistication was hypothesized to moderate the difference between punishment and reward: highly sophisticated citizens are more likely to both reward and punish, where lowly sophisticated citizens are expected to primarily punish bad economic trends, but hardly rewarding positive economic trends. For all three aspects of political affinity (news consumption, political interest and education), interactions with sociotropic evaluations have been modeled. In all models, no interaction achieved significance at the $p < .05$ level. As noted in section 4A, this absence of results is probably caused by the fact that logistic regressions violate the linearity assumption: for linear regressions, variables can be multiplied to estimate their interaction effects under the assumption that for each value of variable X1, the moderation of variable X2 is the same. The interaction effects that are found, then, are not interaction effects for actual citizens but rather the average. In the case of particularly strong interaction effects, significant results may still be found, but more often, false negatives are found (i.e. a Type II error) (Buis, 2010; Norton et al., 2004). Indeed, alternative methods of estimating the interaction between the asymmetry of economic incumbency voting and political sophistication suggest that sophistication indeed conditions the effect sociotropic economic evaluations. The clearest method to solve this is to estimate separate regression models for different levels of political sophistication. For each variable, the respondents are split into a low-category and a high-category based on the median respondent. This is a loss of information compared to the original larger number of categories, but this data reduction is needed to guarantee every category in every country has sufficient respondents for reliable model estimation.

In this chapter only the regression models modeling positive and negative evaluations separately are presented; models containing the uniform effects (sociotropic evaluations estimated as a single variable) can be found in Appendix 8g1, 8g2 and 8g3. In all models, the general effect is a strong and significant predictor of voting behavior, and the models in the Appendix are comparable to those presented here. The first aspect of sophistication is news consumption, measured in days per week reading newspapers. Table 7 presents models with and without demographic controls. As in all other models in this section, the vote in the previous election and ideological distance towards the government are strong and robust predictors, with signs in the predicted direction: those who have voted for government parties before are far more likely to do so again than opposition voters, and the greater the ideological distance is towards the government, the less likely the respondent is to vote for a government party.

Table 7 confirms that economic voting is asymmetric: for all levels of newspaper consumption, negative effects are greater and more significant than positive effects. Nevertheless, reading newspapers matters. For high newspaper consumption, differences are small, but for low newspaper consumption, here is no reward effect. Various model specifications yield comparable results. Adding demographic controls also removes a third of the sample, but results are not meaningfully different. Additionally, in Appendix 8g4 and 8g5, these models are repeated using a different cut-off point: instead of the above-median versus below-median, the groups consist of those reading newspapers versus those not reading any newspapers at all. Here, too, both groups punish weak economic performance, but reward effects only exist for the group reading newspapers. Furthermore, instrumental variables probit analyses are

presented.

Table 8 confirms this effect. Although coefficient sizes are (predictably) not identical between logit and probit models, the interaction is comparably visible in Table 8: positive evaluations are only significant for those scoring high on news consumption, negative evaluations are significant for all respondents. Tests that use the alternative dichotomization (between those who do and those who do not read newspapers) show even greater differences between the low and the high group, further substantiating the theoretical argument. Finally, the marginal effects shown in Figure 13 confirm the trend that has been noted: marginal effects of negative evaluations are greater than their positive counterparts for low news consumption, but are smaller for high news consumption. The crossing lines indicate an interaction between news consumption and economic evaluations (although they cannot test it). This change in direction of the difference is a strong indicator of an interaction in the predicted direction. In general, all the empirical evidence supports hypothesis H4a: the asymmetry exists at all levels, but is greater for citizens who consume less news.³⁸ Since news consumption is related to the other dimensions of political affinity, this is a good omen for the other two dimensions. Nevertheless, as argued before, the dimensions are both theoretically and empirically distinct, and will have to be judged on their own empirical contributions.

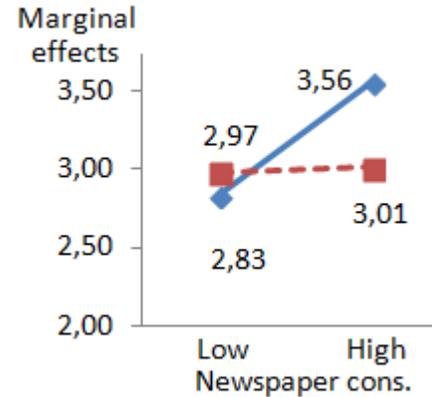


Figure 13. Marginal effects of positive (solid line) and negative (dashed) evaluations by news

The second dimension of political sophistication is political interest. As argued in the theoretical chapter, learning about news and developments does not necessitate forming a mental connection between the economic trend and the incumbents politically responsible for them. Citizens who are hardly interested in politics will be less likely to link societal trends to the government, which will reduce the strength of their economic incumbency vote. This effect is expected to moderate the asymmetry between punishment and reward: if disinterested citizens vote economically, they will mentally connect only a small amount of economic information to their governments, and as discussed before, smaller amounts will be more skewed than greater amounts. Political interest scores are based on respondents' self-assessments into four categories, then dichotomized based on the median value.

³⁸ Additionally, the models were performed on the original variable containing eight categories ranging from 0 to 7 days per week, presented in Appendix 8g6. The trends were comparable: punishment is stronger than reward, punishment effects exist both at high and low levels of news consumption, and positive evaluations were only significant for the highest group of news consumption. However, since most groups were rather small, not all categories matched the general pattern, most likely due to a lack of statistical power. Nevertheless, this table indicates the same trend as the other models for news consumption.

Table 7. Incumbency voting regressed on economic indicators, multilevel logistic regression

News consumption	Low		High		Low		High	
Vote in last election	4.368	***	4.404	***	4.479	***	4.577	***
Positive evaluation	.198		.372	**	.234		.369	*
Negative evaluation [#]	.498	***	.385	**	.557	***	.478	***
Ideological distance	-.411	***	-.424	***	-.366	***	-.418	***
Age					.013	***	.015	*
Social class					.091		.218	***
Gender					-.130		-.149	
Immigrant status					.006		-.092	
Union membership					.219		.342	*
Urbanization					-.016		-.035	
Religiosity					.152		.228	
Income quintile					-.036		-.022	
Economic growth	.078		.076		.049		.079	
Unemployment change	.903	*	.779		.977	*	.991	
Inflation change	-.107		-.067		-.135		-.114	
Time since election	-.012		-.002		-.012		.006	
Constant	-2.248	***	-2.573	***	-3.125	***	-4.300	***
Pseudo R ²	.563		.559		.560		.572	
Log pseudo likelihood	-1,804		-2,141		-1,315		-1,387	
Micro level N	6,136		7,113		4,455		4,772	
Country level N	21		21		20		20	

***= p<0.001; **=p<0.01; *=p<0.05. Reported are logged odds.

[#]: this variable is reversely coded to enable testing equality

Table 8. Incumbency voting regressed on economic indicators, instrumental probit analysis

News consumption	Low		High		Low		High	
Vote in last election	2.289		1.016	*	1,766		.767	
Positive evaluation	1.082		2.085	***				
Negative evaluation [#]					1,478	*	2.030	***
Constant	-1.768	***	-1.291	***	-.592		.206	
Log pseudo likelihood	-5,057		-6,768		-5,388		-6,768	
Micro level N	5,922		6,894		5,922		6,894	
Country level N	21		21		21		21	

***= p<0.001; **=p<0.01; *=p<0.05. Reported are probit coefficients.

[#]: this variable is reversely coded to enable testing equality

Instruments and their testing statistics are identical to those reported in Table 6.

Table 9 shows that for all groups of respondents, negative evaluations have stronger effects than positive evaluations: the same effect as with newspaper consumption. In the same sense, it behaves comparably with respect to the interaction: positive evaluations are absent for citizens with low political interest, but are present for politically interested voters. Negative evaluations affect voting behavior for all groups, and its effect is always greater than the effect of positive evaluations. Adding demographic controls once again strongly reduces the number of respondents, but the results are quite robust. Additional models are presented in Appendix 8g7 and 8g8. Here, the original four categories of political interest are used, with roughly the same results: for all categories, negative evaluations matter more than positive evaluations. Only for the most politically interested citizens, positive evaluations significantly contribute to the model. The group of most disinterested citizens shows no significant effects for negative economic voting; this is most likely the result of overstretching the data. This bottom group only contains 729 respondents, clustered in 20 countries and split by economic evaluations, which means that cells, on average, have less than twenty respondent in them. This is too low to reliably estimate multilevel regression models. The other three levels of political interest contain more respondents and provide more reliable estimations. Additionally, Table 10 replicates the model in an instrumental variables probit regression analysis. Once more, positive evaluations are only relevant for respondents scoring high on political affinity, whereas negative evaluations are relevant for all groups of respondents. Finally, like the newspaper effects, the marginal effects shown in Figure 14 indicate that there is an interaction as well. Again, the difference in marginal effects changes direction: reward effects are greater than punishment effects in the high affinity model, and smaller than punishment effects in the low affinity model. In total, the models support the second political affinity hypothesis H4b: the asymmetry is greater for politically disinterested citizens than for politically interested citizens.

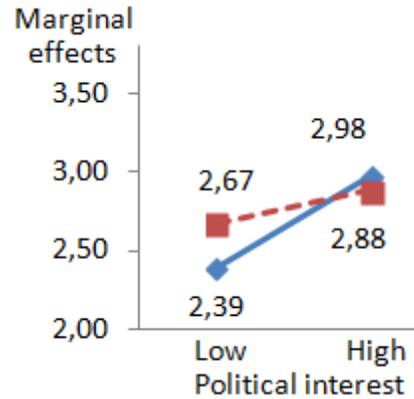


Figure 14. Marginal effects of positive (solid line) and negative (dashed) evaluations by interest

Table 9. Incumbency voting regressed on economic indicators, multilevel logistic regression

Political interest	Low		High		Low		High	
Vote in last election	4.368	***	4.404	***	4.390	***	4.791	***
Positive evaluation	.198		.372	**	.224		.307	*
Negative evaluation [#]	.498	***	.385	**	.417	***	.605	***
Ideological distance	-.411	***	-.424	***	-.335	***	-.453	***
Age					-.112		-.200	*
Social class					.017	**	.013	**
Gender					-.522	**	.343	
Immigrant status					.237		.135	
Union membership					.165	*	.111	*
Urbanization					.009		-.087	
Religiosity					.291	*	.115	
Income quintile					-.040		-.006	
Economic growth	.078		.076		.044		.065	
Unemployment change	.903	*	.779		.753		.639	
Inflation change	-.107		-.067		-.178		.032	
Time since last election	-.012		-.002		-.006		-.002	
Constant	-2.248	***	-2.573	***	-3.713	***	-3.633	***
Pseudo R ²	.563		.559		.540		.560	
Log pseudo likelihood	-1,804		-2,141		-1,323		-1,573	
Micro level N	6,136		7,113		4,221		5,871	
Country level N	21		21		21		21	

***= p<0.001; **=p<0.01; *=p<0.05. Reported are logged odds.

[#]: this variable is reversely coded to enable testing equality

Table 10. Incumbency voting regressed on economic indicators, instrumental probit analysis

Political interest	Low		High		Low		High	
Vote in last election	1.894		1.361		.805		1.208	
Positive evaluation	1.635		1.927	***				
Negative evaluation [#]					1,977	***	1.902	*
Constant	-1.561	*	-1.521	***	-.246		-.169	
Log pseudo likelihood	-4,703		-7,388		-5,623		-7,801	
Micro level N	5,577		8,131		5,577		8,131	
Country level N	22		22		22		22	

***= p<0.001; **=p<0.01; *=p<0.05. Reported are probit coefficients.

[#]: this variable is reversely coded to enable testing equality

Instruments and their testing statistics are identical to those reported in Table 6.

The first and second dimensions of political affinity have been affirmed: asymmetric information input means that negative information is more often relevant, and asymmetric transformation of information into political issues means that negative issues more often become politicized. The broader mechanism of political affinity has passed two of the three tests, but to further assess its importance in explaining the asymmetry, the third aspect will also have to be analyzed. The final dimension of political affinity is education. Here, the expected mechanism is that more highly educated voters will cast their vote based on a large number of issues they know about and have mentally connected to their political leaders, whereas more lowly educated voters will cast their vote based on a few issues, which generates a negativity bias. Education is recoded based on years of full-time education, dichotomized based on the median value.

The results in Table 11 show the same pattern as the other political affinity variables: negative evaluations are relevant for all groups of respondents, and positive evaluations only for higher educated voters. Remarkably, for more highly educated citizens, coefficients are slightly larger for positive evaluations than for negative evaluations, although the differences are rather small. In Appendix 8g9 and 8g10, the model is repeated with education recoded into quartiles of roughly comparable size. This model shows that for very low educated citizens, only the negative effect is present; for below-average and above-average, both effects exist but negative effects are stronger; and for very high educated citizens, positive effects are even (slightly) larger than negative effects. Next, Table 12 shows the instrumental variables probit counterpart. Here, both positive and negative evaluations are only significant for highly educated citizens. Thus, the instrumental variables analysis shows no interaction effect. On the other hand, the marginal effects in Figure 15 show the same confirmation as for the other two dimensions of political affinity: the effect of negative evaluations is greater for low educated citizens, the effect of positive evaluations is greater for high educated citizens. Once again, the general picture confirms the interaction of the political affinity hypothesis H4c.

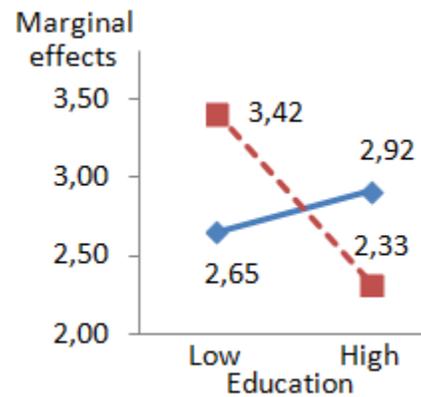


Figure 15. Marginal effects of positive (solid line) and negative (dashed) evaluations by education

Table 11. Incumbency voting regressed on economic indicators, multilevel logistic regression

Education	Low		High		Low		High	
Vote in last election	4.537	***	4.295	***	4.624	***	4.506	***
Positive evaluation	.188		.367	**	.126		.445	**
Negative evaluation [#]	.522	***	.325	***	.625	***	.347	**
Ideological distance	-.425	***	-.427	**	-.386	***	-.405	**
Age					-.121		-.228	**
Social class					.013	*	.011	
Gender					-.187		.004	
Immigrant status					.009		.329	
Union membership					.140		.105	*
Urbanization					-.014		-.089	
Religiosity					.155		.325	
Income quintile					-.002		-.029	
Economic growth	.083		.063		.026		.057	
Unemployment change	.951	*	.553		.908	*	.617	
Inflation change	-.134		.001		-.161		-.037	
Time since last election	-.007		-.005		-.012		.005	
Constant	-2.487	***	-2.481	***	-3.233	***	-3.718	***
Pseudo R ²	.579		.547		.579		.560	
Log pseudo likelihood	-1,886		-2,079		-1,389		-1,414	
Micro level N	6,589		6,800		4,870		4,782	
Country level N	22		22		21		21	

***= p<0.001; **=p<0.01; *=p<0.05. Reported are logged odds.

[#]: this variable is reversely coded to enable testing equality

Table 12. Incumbency voting regressed on economic indicators, instrumental probit analysis

Education	Low		High		Low		High	
Vote in last election	2.204	*	1.059	*	1,957		.767	
Positive evaluation	1.337		2.028	***				
Negative evaluation [#]					1,395		2.030	***
Constant	-1.731	***	-1.351	***	-.693		.206	
Log pseudo likelihood	-5,280		-6,178		-6,319		-6,428	
Micro level N	5,922		6,894		5,922		6,894	
Country level N	22		22		22		22	

***= p<0.001; **=p<0.01; *=p<0.05. Reported are probit coefficients.

[#]: this variable is reversely coded to enable testing equality

Instruments and their testing statistics are identical to those reported in Table 6.

The empirical results can be summarized more or less in three main points. The first is that the asymmetry between punishment and reward is indeed present. On average, the punishment incumbents receive for economic dissatisfaction is substantially stronger than the reward they can earn for economic satisfaction. However, this asymmetry varies strongly across countries, and is even reversed in a subset of the countries studied. Therefore, this paper has set out to identify causal mechanisms explaining the variation in the strength of asymmetric economic voting. In the second point, spanning hypotheses H2 and H3, this has been done by putting the economic variables under closer scrutiny. Differences in asymmetry caused by different economic indicators seems theoretically plausible: a decrease in unemployment or increase in economic growth will always be welcomed by voters, but for low levels of inflation it appears that this will actually not be appreciated. Although hints of the latter mechanism can be found in the data, in general, there were too few countries in the analysis for proper testing of these hypotheses. Many earlier studies have found that unemployment and economic growth matter; their insignificance is likely the result of a lack of macro level variance. Nevertheless, in this study with EES 2004 data, the second hypothesis will have to be rejected. The results for economic salience were mixed: some results confirmed the expected effect of higher salience increasing the asymmetry. Other techniques found no significant differences between both groups of citizens. Also, nothing was found for salience of individual economic indicators. The results for this hypothesis are inconclusive.

With regards to the third point, the EES data were more tailored to suit the needs of testing micro-level interactions, as has been done with the hypotheses on political affinity H4a, H4b and H4c. All three mechanisms of political affinity are confirmed by the empirical results. Together, this provides empirical support for the broader mechanisms underlying economic voting. On average, people tend to more often punish governments for economic malaise than reward them for economic prosperity. This asymmetry exists for all groups of voters, but is especially great for voters with little political affinity. Voters with greater political affinity are more likely to exhibit both pro and anti economic incumbency voting. Multilevel logistic regression, models, instrumental variable probit analyses and marginal effects estimations, using multiple operationalizations, all support the hypotheses of political affinity, which is the most important result of the empirical analyses.

Table 13. Main findings on asymmetry between punishment and reward		
#	Hypothesis tested	Result
H1	Punishment effects are stronger than rewards	Confirmed
H2	Asymmetry is larger for inflation than for unemployment and growth	<i>Rejected</i>
H3	Salience increases the asymmetry	<i>Inconclusive</i>
H4a	News consumption reduces asymmetry	Confirmed
H4b	Political interest reduces asymmetry	Confirmed
H4c	Education reduces asymmetry	Confirmed

6. CONCLUSION AND DISCUSSION

Decades of research economic voting have resulted in a large increase in our knowledge of the circumstances under which economic incumbency voting is present (that is, in virtually all countries, at all times). In most cases, economic voting is not only one of the factors, but the most important of issues on which voters judge their incumbent governments. Various factors have been found to moderate the strength of economic voting, which can be related to the economy (such as the clarity of responsibility of the government for economic affairs) or to other issues replacing economic concerns as the top priority (such as political violence). In response to criticism of ecological and individualistic fallacies, macro assumptions have been tested at the micro level, and micro assumptions have been tested at the macro level. Hardly tested, on the other hand, is the important (yet often implicit) assumption of the vast majority of studies that economic voting is symmetric. However, both theoretical arguments and empirical case studies hint that the electoral punishment for economic downturn by far surpasses the reward for economic growth. This study has set out not only to show whether there is an asymmetry between punishment and reward, but, if so, also test micro and macro level explanations for this asymmetry. After summarizing and discussing the approach and findings of this study, they will be situated against the broader study of economic voting. To reiterate, the central research question was:

Research Question: *To what extent do differential macro-economic conditions and micro-level political attitudes account for the asymmetry between reward and punishment of economic voting?*

Using data from 2004 on 22 European countries, this study confirms that, on average, punishment effects are far stronger than reward effects. On the other hand, the degree to which they are asymmetrical varies substantially across countries, and in some countries the reward effects were even stronger than the punishment effects. Loss aversion plausibly contributes to the asymmetry, but biological factors are unhelpful in explaining variation between comparable countries and over short time spans. To explore this variation, three theoretical explanations at both macro and micro level are analyzed. The first explanation for variation in economic voting is that voters do not treat ‘the economy’ as a single issue in their voting considerations, but rather as an umbrella of related but essentially different concepts. Theoretically, it can be argued that unemployment and GDP per capita growth are relatively symmetrical, linear functions: voters will want less unemployment and more economic growth regardless of the actual levels these indicators currently hold. Governments can both win and lose popularity on these two issues. For inflation, the situation is different: citizens do not like very high inflation, but do not mind low to intermediate levels of inflation. In fact, inflation levels of a few percentage points may be preferable over zero or near-zero inflation, as the European Central Bank recently confirmed. Increasing inflation, then, may lead to a punishment if the levels are too high, but reducing inflation may not yield rewards. The empirical models find no effects for unemployment or economic growth, but increasing inflation has a positive effect on incumbency voting. This is likely the result of the fact that absolute levels of inflation were rather low in Europe in 2004: in most cases below five percent, and only in Greece over ten percent. In the 22 countries in the analysis, the economic trends of 2004 did not fluctuate very strongly, which made variation on the macro level rather small.

This makes it likely that a larger, more varied case set would lead to even stronger results. For now, however, the results indicate that different economic indicators indeed have differential effects, and we should not treat 'the economy' as a single issue.

The second mechanism is that of issue salience. Voters cannot possibly form an opinion of governmental performance on every issue that the government is involved, and have to pick a select number of issues which are important to them. Issues become more salient when they are more problematic: salience consists of both the value attached to an item and the actual performance on this item. Since bad performance is to increase salience, issue salience was expected to contribute to the asymmetry between reward and punishment. This was done both for the three indicators of the economy separately and for economic performance as a whole. Likely due to data issues, no effects were found for separate indicators. On the other hand, the far larger dataset on economic performance as a whole did enable more rigorous testing, and confirmed that economic salience indeed moderates the asymmetry: the difference between punishment and reward is greater when 'the economy' was the most salient issue.

Lastly, political affinity was analyzed. Political affinity was conceptualized as three related, but different steps in the causal chain from economic developments to vote choice, each of which can contribute to a more asymmetric vote function. The first step is for citizens to gain information on the economy. Bad news is most prominently and readily available, meaning that those who consume little news will only pick up the bad news, whereas the more news one consumes, the more positive news will also slip through. Empirical results confirm this mechanism: the effects for groups with little media consumption are far more different than those who consume much media. The second step is political interest: not all issues are politicized in the minds of voters. The more interested a voter is, the more issues will be mentally linked to the incumbent government's performance. Comparably, negative issues will be stronger than positive, so voters voting on a single or small number of issues will have a greater negativity bias (since only the strongest issues matter) than those who are more interested in politics and judge governments on a broader range of issues. Empirical results on this topic are mixed, but may improve if superior data were to be used. The third and final step is education: people may link multiple issues to the government, but this generates a complex picture to base the vote on. Low educated people may decide to simplify the picture by omitting several issues and still voting primarily on a small number of issues, where the same negativity bias is present. Like the first step in political affinity, education is also found to contribute to asymmetry: the lower one is educated, the greater the asymmetry will be.

In the past decades, several case studies have made an empirical argument that economic voting should be modeled asymmetrically, separating punishment from reward effects. A few case studies have argued that the effect is symmetric, whereas most do not pore upon this issue and implicitly assume asymmetry. This study has confirmed on a larger scale that economic voting is indeed asymmetric, but also that the asymmetry varies over countries and may in some cases be absent or even reversed. This is a compelling argument for cross-sectional designs: studies that select either Estonia or Latvia may argue that the countries are seemingly comparable, but still end up with opposite results. In addition to the

empirical argument for modeling the effects separate, several theoretical arguments have been made to not only show that it exists, but also explain why and under what conditions. However, we have only scratched the surface of explaining the asymmetry between punishment and reward, and future studies analyzing other aspects are needed to truly understand why the asymmetry is so present.

Methodologically this study has aimed to move beyond earlier studies by using a highly ambitious design. The core of the analyses consisted of multilevel logistic regression models. Through testing multiple operationalizations, it was possible to gain more understanding of political affinity. As an alternative for modeling interactions by multiplying variables, splitting regressions by different variables has proven itself as an insightful alternative. Moreover, the instrumental variables probit analyses behaved the way they were predicted to behave: modeling economic evaluations so that they are exogenous to voting intention did change the coefficients, but also confirmed the results found by the more parsimonious models. These methods should be a fruitful point of departure for future studies into (asymmetric) economic incumbency voting. Future studies could improve on the research design in several ways.

In terms of data, future studies could improve on several points. The European Election Survey wave of 2004 has been useful for testing micro-level interactions, but contained too few country-level observations for proper testing the macro effects of economic indicators. This issue can be dealt with by including more countries, more diverse countries, and / or taking a longitudinal approach in order to improve macro level variance. Second, for many countries, the dataset was incomplete. Dichotomizing certain variables was an effective way to ensure each cell contained enough respondents, but some other datasets are less hindered by this problem. Third, to procure better instruments for instrumental variable analyses, improvements could be made at both micro and macro level. At the micro level, including more socioeconomic items would help capturing individual level variation in economic experiences. At the macro level, within-country differences could be captured by also including regional data such as NUTS levels 1, 2 and or 3. After all, especially in larger countries, there can be great economic differences between (for example) the capital region and the periphery. Fourth, using a more diverse set of countries could show to what degree the asymmetry persists in non-EU countries. Since cross-country differences in asymmetry are substantial, analyzing an even more diverse set of countries can test how broadly the findings can be generalized.

In terms of methodology, future research using multi-method designs could subject some of the assumptions made in this paper to closer scrutiny. For example, this study has made the assumption that all newspapers have on average the same effect on citizens. However, there may be differences in the way newspapers present and frame financial and economic news, whether they focus on business news or general news, whether they are associated with a certain ideological position or not. Combining data on newspaper readership with content analysis could more specifically test this mechanism and differentiate between various newspapers. Alternatively, qualitative interviews could be held with a smaller number of voters to explore the extent to which they are aware of asymmetries in economic voting behavior, and what reasons they believe drive them to vote asymmetrically (or more subtle ways of exploring the same mechanism). Finally, as mentioned in 3B, economists and psychologists have used

experiments to test the mechanism of loss aversion. In a comparable fashion, political scientists could use experiments or experimental surveys to observe behavior in controlled scenarios tailored to test, for example, differential effects of various economic indicators. In terms of theory development, modeling interactions with political affinity has led to interesting results. It may be worthwhile to investigate the effects of other, related micro-level characteristics on the asymmetry, such as political efficacy or social capital. Additionally, this study has treated parties as entities able to govern, which are in varying degrees successful in this task and can be placed on a left-right axis. In reality, parties actively engage in strategic political behavior to influence their popularity and the election results. Investigating how party behavior influences the asymmetry will lead to better understanding of the subject.

This study is part of the debate on economic voting. The results confirm majority positions on some of the most debated issues within economic voting: retrospective, sociotropic voting is indeed a powerful predictor, and incumbency voting was identifiable despite incumbents consisting of coalitions of up to six different parties. All innovative findings in this study were based on subjective economic indicators rather than objective indicators. General opinions on the economy were far more powerful than objective macro indicators in explaining voting behavior. However, if the theoretical arguments made can be supported by a dataset containing more country level variance than the EES 2004 wave, even the general sociotropic judgment may have to be replaced or complemented by perceptions on different economic indicators instead of perceptions on the economy as a whole. The robust findings on how political affinity moderates the asymmetry contributes to the debate on economic voting in multiple ways. On the one hand, it emphasizes the value of micro-macro approaches to test how different groups structurally use economic information differently. This provides a stepping stone for analyzing how many other findings on economic voting, such as the moderation by external economic events or clarity of governmental responsibility, may also vary over high and low affinity voters. On the other hand, the asymmetry present in judgments could well be complemented by asymmetries in other parts of economic voting, such as party behavior or institutional effects.

Finally, economic voting is one of the most important elements of valence politics, the increasing role of voters judging incumbents on their performance rather than on ideology, identity or cleavages. Voting behavior may also be asymmetric for voters who cast their votes primarily based on other issues such as health, safety or environment. In the same sense, voting on these and other issues may be moderated by political affinity in a comparable way. Elections are central in representative democracies, and vital in shaping the way countries are governed. To better understand why citizens vote the way they do remains the holy grail of a large stream within comparative political science. Although it may never be reached, the field may move ever closer towards this goal. If anything, this paper has shown that the path to the goal is often not linear but asymmetrical.

7. REFERENCES

- Abdellaoui, M. (2007). Loss aversion under prospect theory: A parameter-free measurement. *Management Science*, 53(10), 1659–1674.
- Alesina, A., Londregan, J., & Rosenthal, H. (1993). A model of the political economy of the United States. *The American Political Science Review*, 87(1), 12–33.
- Althaus, S. (2002). American news consumption during times of national crisis. *Political Science & Politics*, 35(3), 517–521.
- Anderson, C. (2000). Economic voting and political context: a comparative perspective. *Electoral Studies*, 19, 151–170.
- Anderson, C. (2006). Economic voting and multilevel governance: a comparative individual-level analysis. *American Journal of Political Science*, 50(2), 449–463.
- Anderson, C., & Hecht, J. D. (2012). Voting when the economy goes bad, everyone is in charge, and no one is to blame: The case of the 2009 German election. *Electoral Studies*, 31(1), 5–19.
- Armingeon, K., Gerber, M., Leimgruber, P., Beyeler, M., & Menegale, S. (2012). Comparative Political Data Set III 1990-2011. *Institute of Political Science, University of Berne*.
- Bartels, L. (2011). Ideology and retrospection in electoral responses to the Great Recession. *Conference on Popular Reactions to the Great Recession*, (June), 1–55. Retrieved from <http://www.princeton.edu/~bartels/stimulus.pdf>
- Bellucci, P. (2012). Government accountability and voting choice in Italy, 1990–2008. *Electoral Studies*, 31(3), 491–497.
- Bellucci, P., Costa Lobo, M., & Lewis-Beck, M. S. (2012). Economic crisis and elections: The European periphery. *Electoral Studies*, 31(3), 469–471.
- Benton, A. (2005). Dissatisfied democrats or retrospective voters?: Economic hardship, political institutions, and voting behavior in Latin America. *Comparative Political Studies*, 38(4), 417–442.
- Bloom, H. S., & Price, H. D. (1975). Voter response to short-run economic conditions: the asymmetric effect of prosperity and recession. *American Political Science Review*, 69(4), 1240–1254.
- Bouvet, F., & King, S. (2013). Income inequality and election outcomes in OECD Countries: New evidence following the economic crisis of 2008-2009. Retrieved from http://www.euce.org/eusa/2013/papers/8a_bouvet.docx, on October 2, 2013
- Buis, M. (2010). Stata tip 87: Interpretation of interactions in non-linear models. *The Stata Journal*, 10(2), 305–308.

- Clagett, W. (1986). A reexamination of the asymmetry hypothesis: economic expansions, contractions and congressional elections. *The Western Political Quarterly*, 39(4), 623–633.
- Clarke, H. D., Scotto, T. J., & Kornberg, A. (2011). Valence politics and economic crisis: Electoral choice in Canada 2008. *Electoral Studies*, 30(3), 438–449.
- Clarke, H. D., & Stewart, M. (1994). Prospections, retrospections, and rationality: The “bankers” model of presidential approval reconsidered. *American Journal of Political Science*, 38(4), 1104–1123.
- Costa Lobo, M., & Lewis-Beck, M. S. (2012). The integration hypothesis: How the European Union shapes economic voting. *Electoral Studies*, 31(3), 522–528.
- Davison, R. (2012). Re: st: IV Probit. Retrieved April 29, 2014, from <http://www.stata.com/statalist/archive/2012-03/msg00672.html>
- De Martino, B., Camerer, C. F., & Adolphs, R. (2010). Amygdala damage eliminates monetary loss aversion. *Proceedings of the National Academy of Sciences of the United States of America*, 107(8), 3788–3792.
- Duch, R., & Stevenson, R. T. (2005). Context and the economic vote: A multilevel analysis. *Political Analysis*, 1–23.
- Duch, R., & Stevenson, R. T. (2008). *The economic vote. How political and economic institutions condition election results*. Cambridge: Cambridge University Press.
- Duch, R., & Stevenson, R. T. (2010a). The Global Economy, Competency, and the Economic Vote. *The Journal of Politics*, 72(01), 105–123.
- Duch, R., & Stevenson, R. T. (2010b). Context and economic expectations: When do voters get it right? *British Journal of Political Science*, 41(01), 1–31.
- EES. (2004). European Election Study, Codebook. Design, data description and documentation.
- Evans, G., & Andersen, R. (2006). The political conditioning of economic perceptions. *Journal of Politics*, 68(1), 194–207.
- Evans, G., & Pickup, M. (2010). Reversing the Causal Arrow: The Political Conditioning of Economic Perceptions in the 2000–2004 U.S. Presidential Election Cycle. *The Journal of Politics*, 72(04), 1236–1251.
- Feddersen, T. (2004). Rational choice theory and the paradox of not voting. *Journal of Economic Perspectives*, 18(1), 99–112.
- Fisher, H. (2010). Introduction to Stata – Lecture 4 : Instrumental variables, (March), 1–11.
- Fraile, M., & Lewis-Beck, M. S. (2012). Economic and elections in Spain (1982–2008): Cross-measures, cross-time. *Electoral Studies*, 31(3), 485–490.

- Freire, A., & Santana-Pereira, J. (2012). Economic voting in Portugal, 2002–2009. *Electoral Studies*, 31(3), 506–512.
- Hammerstein, P., & Hagen, E. H. (2005). The second wave of evolutionary economics in biology. *Proceedings of the National Academy of Sciences of the United States of America*, 20(11), 604–609.
- Hansford, T. G., & Gomez, B. T. (2010). Estimating the electoral effects of voter turnout. *American Political Science Review*, 104(02), 268–288.
- Hansford, T. G., & Gomez, B. T. (2011). Reevaluating the sociotropic economic voting hypothesis. Retrieved from [http://faculty.ucmerced.edu/thansford/Working Papers/Reevaluating the Sociotropic Economic Voting Hypothesis.pdf](http://faculty.ucmerced.edu/thansford/Working%20Papers/Reevaluating%20the%20Sociotropic%20Economic%20Voting%20Hypothesis.pdf), on June 7, 2013
- Headrick, B., & Lanoue, D. (1991). Attention, asymmetry, and government popularity in Britain. *The Western Political Quarterly*, 44(1), 67–86.
- Hellwig, T. (2008). Globalization, policy constraints, and vote choice. *The Journal of Politics*, 70(4), 1128–1141.
- Jonung, L., & Wadensjö, E. (1979). The effect of unemployment, inflation and real income growth on government popularity in Sweden. *The Scandinavian Journal of Economics*, 81(2), 343–353.
- Keele, L., & Morgan, J. (2013). Stronger Instruments By Design. Available at SSRN 2280347, 1–37. Retrieved from <https://pages.shanti.virginia.edu/PolMeth/files/2013/07/KeeleMorgan.pdf>
- Kinder, D., & Kiewiet, D. (1981). Sociotropic politics: the American case. *British Journal of Political Science*, 11(2), 129–161.
- Krause, G. (1997). Voters, information heterogeneity, and the dynamics of aggregate economic expectations. *American Journal of Political Science*, 41(4), 1170–1200.
- Lau, R. R. (1985). Two explanations for negativity effects in political behavior. *American Journal of Political Science*, 29(1), 119–138.
- Lewis-Beck, M. S., & Stegmaier, M. (2000). Economic determinants of electoral outcomes. *Annual Review of Political Science*, 3, 183–219.
- Lupia, A. (1994). Shortcuts versus encyclopedias: information and voting behavior in California insurance reform elections. *American Political Science Review*, 88(1), 63–76.
- Mair, P. (2006). Ruling the void?, 25–51.
- Marsh, M., & Mikhaylov, S. (2012). Economic voting in a crisis: The Irish election of 2011. *Electoral Studies*, 31(3), 478–484. doi:10.1016/j.electstud.2012.02.010
- Mueller, J. (1970). Presidential popularity from Truman to Johnson. *The American Political Science Review*, 64(1), 18–34.

- Nannestad, P., & Paldam, M. (1997). The grievance asymmetry revisited: A micro study of economic voting in Denmark, 1986–1992. *European Journal of Political Economy*, 13(1), 81–99.
- Nezi, R. (2012). Economic voting under the economic crisis: Evidence from Greece. *Electoral Studies*, 31(3), 498–505.
- NorthernIrelandAssembly. (2014). History of the Assembly. Retrieved May 7, 2014, from <http://www.niassembly.gov.uk/ABOUT-THE-ASSEMBLY/General-Information/History-of-the-Assembly/>
- Norton, E., Wang, H., & Ai, C. (2004). Computing interaction effects and standard errors in logit and probit models. *Stata Journal*, 4(2), 154–167.
- Palmer, H., & Whitten, G. D. (2011). Through thick and thin? The dynamics of government support across income groups during economic crises. *Electoral Studies*, 30(3), 427–437.
- Scotto, T. J. (2012). Conclusion: Thinking about models of economic voting in hard times. *Electoral Studies*, 31(3), 529–531.
- Singer, M. M. (2011). Who says “It’s the economy”? Cross-national and cross-individual variation in the salience of economic performance. *Comparative Political Studies*, 44(3), 284–312.
- Slomp, H. (2011). *Europe, a political profile. An American companion to European politics (II)*. Santa Barbara: ABC-CLIO.
- Söderbom, M. (2009). Applied Econometrics Lecture 2 : Instrumental Variables , 2SLS and GMM. Retrieved June 30, 2014, from www.soderbom.net/lec2n_final.pdf
- Sovey, A. J., & Green, D. P. (2011). Instrumental Variables estimation in Political Science: A readers’ guide. *American Journal of Political Science*, 55(1), 188–200.
- Stevenson, R. T., & Duch, R. (2013). The meaning and use of subjective perceptions in studies of economic voting. *Electoral Studies*, 32(2), 305–320. doi:10.1016/j.electstud.2013.02.002
- Strömbäck, J., & Shehata, A. (2010). Media malaise or a virtuous circle? Exploring the causal relationships between news media exposure, political news attention and political interest. *European Journal of Political Research*, 49(5), 575–597.
- Sunshine Hillygus, D. (2005). The MISSING LINK: Exploring the Relationship Between Higher Education and Political Engagement. *Political Behavior*, 27(1), 25–47.
- Van der Brug, W., Van der Eijk, C., & Franklin, M. (2007). *The economy and the vote. Economic conditions and elections in fifteen countries*. Cambridge: Cambridge University Press.
- Wlezien, C. (2005). On the salience of political issues: The problem with “most important problem”. *Electoral Studies*, 24(4), 555–579.

8. APPENDIX

8A. Tentative literature overview

	Socio	Ego	Retro	Prosp	Subject	Oject	Asymm	Symmetry
Austria	O		O			O	O	
Australia	A		A		A			
Belgium	ABO		AB		AB			
Canada	A		A		A			
Denmark	ABO		ABO		AB	O	O	
Finland	O		O			O	O	
France	BDO		BO	D ³⁹	BD			
Germany	ABO	F ⁴⁰	ABFO		ABF	O	O	
Greece	BCEO		BCEO		BCE	O	EO	
Iceland	A		A		A			
Ireland	BGO		BGO		BG	O	GO	
Italy	BCHO		BCHO		BCH	O	O	
Japan	A		A		A			
Luxembourg	BO		BO		B	O	O	
Netherlands	ABO		ABO		AB	O	O	
New Zealand	A		A		A			
Norway	A		A		A			
Portugal	ACIOR		ACIOR		ACIR	O	IO	R
Spain	ABCO		ABCO		ABC	O	O	
Sweden	AJO		AJO		A	JO	O	
Switzerland	A		A		A			
UK	ABDKO		ABDKO		ABD	K ⁴¹ O	KO	
US	ALMN		ALMN		A	LMN	MN	

L	(Alesina, Londregan, & Rosenthal, 1993)	N	(Clagett, 1986)	D	(Hellwig, 2008)
F	(Anderson & Hecht, 2012)	P	(Clarke, Scotto, & Kornberg, 2011)	J	(Jonung & Wadensjö, 1979)
B	(Anderson, 2000)	C	(Costa Lobo & Lewis-Beck, 2012)	G	(Marsh & Mikhaylov, 2012)
A	(Anderson, 2006)	I	(Freire & Santana-Pereira, 2012)	E	(Nezi, 2012)
H	(Bellucci, 2012)	R	(Freire & Santana-Pereira, 2012)	O	(Van der Brug et al., 2007)
M	(Bloom & Price, 1975)	K	(Headrick & Lanoue, 1991)	K	K

³⁹ For France, prospective data were argued to be useable as proxy for retrospective evaluations.

⁴⁰ The authors explain this counter-intuitive finding by declaring these German elections to be exceptional in terms of the low attribution of responsibility towards the government

⁴¹ The authors unhelpfully use the term ‘perceived’, while they use actual objective data on inflation and the change in the number of unemployed adults.

8B. Exact wording of the items in the (English-language) questionnaire

Item	Exact wording of the question	Answering categories	Notes
V015 – V027	What do you think are the most important problems in [country] at present? Any other important problems?	<i>(open answer)</i>	
V028	Of those you have mentioned what would you say is the single most important problem?	<i>(open answer)</i>	
V069	And how many days of the week do you read a newspaper?	0 – 7 days a week	Q1 Q6
V070	For Sweden: Do you usually read one or several newspapers regularly? With regularly I mean at least one time a week?	0 = no 1 = yes	Q0
V113	Which party did you vote for at the [General Election] of [Year of Last General Election]?	<i>(open answer)</i>	
V114	And if there was a general election tomorrow, which party would you vote for?	<i>(open answer)</i>	
V134	In political matters people talk of “the left” and “the right”. What is your position? Please indicate your views using any number on a 10-point scale. On this scale, where 1 means “left” and 10 means “right”, which number best describes your position?	1 = left 10 = right	Q1 Q2 Q8
V149	What do you think about the economy? Compared to 12 months ago, do you think that the general economic situation in [country] is ...	1 = a lot better 2 = a little better 3 = stayed the same 4 = a little worse 5 = a lot worse	
V154	To what extent would you say you are interested in politics?	1 = very 2 = somewhat 3 = a little 4 = not at all	
V215	Are you yourself a member of a trade union or is anyone else in your household a member of at trade union?	1 = Yes I am 2 = Yes, someone else is 3 = Yes, both (1) and (2) 4 = No	Q1 Q2 Q3
V216	How old were you when you stopped full-time education?	<i>(open answer)</i>	Q1 Q2 Q6
V217	Are you ... [gender]	1 = male 2 = female	Q6
V218	What year were you born?	<i>(open answer)</i>	Q1 Q4 Q6
V219	In which country were you born?	0 = country of interview 1 = other country	Q1 Q6
V221	How many people live in your household including yourself, who are 18 years of age or older?	<i>(open answer)</i>	Q1
V224	If you were asked to choose one of these five names for your social class, which would you say you belong to – the working class, the lower middle class, the middle class, the	1 = working class 2 = lower middle class 3 = middle class	Q1 Q2 Q6

Jelle Lössbroek – Pocketsize Plus and Massive Minus

	upper middle class or the upper class?	4 = upper middle class 5 = upper class	
V225	Would you say you live in a rural area or village, in a small or middle size town, or in a large town?	1 = rural area or village 2 = small or middle size town 3 = large town	Q3 Q7
V226	...	1 = very strongly urbanized 2 = strongly urbanized 3 = somewhat urbanized 4 = little urbanized 5 = not at all urbanized	Q0
V227	...	1 = village 2 = town to 19,999 3 = town 20,000-49,999 4 = town 50,000-99,999 5 = city 100,000-499,999	Q0
V229	How often do you attend religious services: several times a week, once a week, a few times a year, once a year or less, or never?	1 = several times a week 2 = once a week 3 = a few times a year 4 = once a year or less 5 = never	Q1 Q2
V230	We also need some information about the income of this household to be able to analyze the survey results for different types of households. Please count the total wages and salaries per month of all members of this household; all pensions and social insurance benefits; child allowances and any other income like rents etc...Of course your answer as all other replies in this survey will be treated confidentially and referring back to you or your household will be impossible. Can you please tell me, what about is the monthly income of your household?	<i>Information made available in quintiles</i> 1 = first 2 = second 3 = third 4 = fourth 5 = fifth quintile	

- Q0 Question only asked in
V070: Sweden
V226: Netherlands
V227: Poland
- Q1 Not asked in Lithuania
- Q2 Not asked in Swedish short survey
- Q3 Not asked in Belgium
- Q4 Not asked in Latvia
- Q5 In Spain, the question compared to 6 months ago
In Sweden, the answering categories were 1 = improved, 3 = remained the same, 5 = gone worse
- Q6 For Sweden, an alternative data source was used.
V069: see v070
V216: derived from question about the level of education
V217, v218, v219: register data
V224: Constructed from subjective occupational group and subjective family class
- Q7 Asked in differently phrased questions for Netherlands (v226) and Poland (v227)
- Q8 The Swedish long survey used 0-10 scale (EES, 2004)

8C. Descriptive statistics

Below, in Appendix 8c1, basic descriptive statistics are presented for the variables included in the models. Next, in Appendix 8c2, an overview is given of how many respondents were omitted from the analysis due to missing scores on certain key variables.

Appendix 8c1. Descriptive statistics of EES 2004, excluding Northern Ireland

	N	Minimum	Maximum	Mean	Std. Dev.
<i>Core variables</i>					
Vote intention	20,188	.00	1.00	.38	.48
Vote in previous general elections	20,124	.00	1.00	.46	.50
Sociotropic evaluation	26,155	1.00	3.00	1.83	.77
Positive sociotropic evaluation	26,155	.00	1.00	.23	.42
Negative sociotropic evaluation	26,155	-1.00	.00	-.40	.49
Is economy the most important problem?	26,880	.00	1.00	.43	.49
<i>Political affinity</i>					
Newspaper (days per week)	25,389	.00	7.00	3.98	.74
Political interest	26,853	1.00	4.00	2.44	0.87
Education (age when completed)	24,783	.00	95.00	19.23	5.17
<i>Demographic control variables</i>					
Age	26,362	18.00	101.00	47.25	16.95
Gender	27,781	.00	1.00	.52	.50
Immigrant	27,592	.00	1.00	.05	.23
Income quintiles	24,259	1.00	5.00	2.85	1.45
Left-right distance to government	23,940	.00	4.50	1.77	1.17
Religiosity	24,992	.00	1.00	.74	.44
Social class	25,701	1.00	4.00	2.37	1.02
Union member in household	26,483	.00	1.00	.32	.47
Urbanization	27,758	1.00	3.00	1.98	.81

The 2004 wave of EES originally contains 28,861 respondents in 24 EU members. First, Belgium is omitted, as the question of voting intention was not asked there, and Northern Ireland is omitted as the political system is too distinct (see Data chapter for more specifications). Second, all respondents in other countries that scored a missing on voting intention were omitted. Third, the respondents that did not score on the central independent variable, sociotropic economic judgment, were dropped. Fourth, respondents not answering the question on the most powerful control variable, vote in previous election, were removed. Fifth, only respondents with valid scores on each variable, including demographic controls, remain.

Appendix 8c2. Number of valid respondents after those with missing scores on several variables are omitted. Final columns are expressed in percentages of the original N per country.

		[1	[2	[3	[4	[5	[%2	[% 3	[%4	[%5
Austria	1,010	1,010	759	746	672	557	75%	74%	67%	55
Belgium	889									
Cyprus	500	500	442	427	369	289	88%	85%	74%	58%
Czech Republic	889	889	652	641	469	388	73%	72%	53%	44%
Denmark	1,317	1,317	1,066	999	920	705	81%	76%	70%	54%
Estonia	1,606	1,606	1,076	1,041	858	588	67%	65%	53%	37%
Finland	900	900	656	646	576	427	73%	72%	64%	47%
France	1,406	1,406	1,038	1,034	873	777	74%	74%	62%	55%
Germany	596	596	387	381	339	231	65%	64%	57%	39%
Greece	500	500	385	373	347	218	77%	75%	69%	44%
Hungary	1,200	1,200	860	848	728	539	72%	71%	61%	45%
Ireland	1,154	1,154	903	892	830	657	78%	77%	72%	57%
Italy	1,553	1,553	1,165	1,151	1,019	609	75%	74%	66%	39%
Latvia	1,000	1,000	511	508	411	304	51%	51%	41%	30%
Lithuania	1,005	1,005	624				62%			
Luxembourg	1,335	1,335	1,244	1,235	1,009		93%	93%	76%	
Netherlands	1,586	1,586	1,285	1,260	1,220	1,077	81%	79%	77%	68%
Poland	960	960	563	555	382	263	59%	58%	40%	27%
Portugal	1,000	1,000	615	605	491	280	62%	61%	49%	28%
Slovakia	1,063	1,063	797	792	640	548	75%	75%	60%	52%
Slovenia	1,002	1,002	516	506	370	235	51%	50%	37%	23%
Spain	1,208	1,208	913	887	801	451	76%	73%	66%	37%
Sweden	2,100	2,100	1,599	905	830		76%	43%	40%	
UK	1,500	1,500	1,132	1,104	886	530	75%	74%	59%	35%
N. Ireland	1,582									
Total	28,861	26,390	19,188	17,536	15,040	9,673	73%	66%	57%	37%

- [1] Removal of Belgium and Lithuania
- [2] Removal of respondents with missing scores on the dependent variable
- [3] Removal of respondents with missing scores on sociotropic economic evaluations
- [4] Removal of respondents with missing scores on vote in the previous election
- [5] Removal of respondents with missing scores on any of the variables used

8D. Post-estimation test statistics of instrumental variables

All tests are performed in STATA 12 using the commands `-estat firststage-`, `-estat endogenous-` and `-estat overid-` (Fisher, 2010; Söderbom, 2009). As mentioned in section 4C, these post estimation tests have been generated on two stage least squares models, as it is not possible to do so on instrumental variables probit models (Davison, 2012). For the overidentification tests, flat models were used instead of multilevel models for the same reason.

For instrument strength, an F-score > 10 is commonly used as threshold value to indicate that the instruments are strong enough to properly predict the endogenous variable in the first stage. There are no threshold values for partial or adjusted R²; generally, greater scores are preferable but lower scores have also been used (Hansford & Gomez, 2010, 2011; Keele & Morgan, 2013; Sovey & Green, 2011). For all variables, F scores are high enough.

For instrument exogeneity, a Robust F test is performed for which the null hypothesis is that the instrument is exogenous. For p values below .05, the exogeneity has to be rejected and the alternative hypothesis of instrument endogeneity has to be accepted. For all variables, endogeneity has to be rejected.

For over identification tests, two Chi² tests are used with virtually identical results. The null hypothesis is that the variable is not over identified, the alternative hypothesis is that the variable is over identified. For general evaluations and for positive evaluations, there is no over identification. For negative evaluations, the null hypothesis has to be rejected at the p<.01 level (but not at the p<.001 level). Following the arguments in section 4C, this is nevertheless acceptable.

		Economic evaluation	Positive evaluation	Negative evaluation
Instrument strength	F-score	18.586	13.820	10.436
	Partial R ²	.015	.015	.009
	Adjusted R ²	.060	.041	.051
Instrument exogeneity	Robust F	1.969 (p = .175)	1.587 (p = .222)	1.996 (p = .172)
Over identification test	Sargan Chi ²	3.640 (p = .162)	4.025 (p = .134)	12.570 (p = .002)
	Basmann Chi ²	3.640 (p = .162)	4.025 (p = .134)	12.576 (p = .002)

8E. Asymmetry between punishment and reward by country

Appendix 8e. Incumbency voting regressed on positive and negative evaluations of the economy, split per country, logged odds.

Country	N	Positive evaluation	Negative evaluation [#]	P>N	N>P
Austria	746	.244	.630***		1
Cyprus	427	.108	.890***		2
Czech Republic	641	.007	.754***		3
Denmark	999	.501***	1.191***		4
Estonia	1,041	.418**	.103	1	
Finland	646	.265	-.115		
France	1,034	.530*	1.103***		5
Germany	381	.673*	.416	2	
Greece	373	1.48***	.847***	3	
Hungary	848	1.132***	1.705***		6
Ireland	892	.755***	.429*	4	
Italy	1,151	.222	2.281***		7
Latvia	508	-.124	.497*		8
Luxembourg	1,235	.337*	.188	5	
Netherlands	1,260	.427**	.263	6	
Poland	555	.434	.193		
Portugal	605	2.065***	.471	7	
Slovakia	792	1.306***	.274	8	
Slovenia	506	1.031***	.056	9	
Spain	887	.680**	.486**	10	
Sweden	905	-.185	.811***		9
United Kingdom	1,104	.381*	1.016***		10

***=p<0.001; **=p<0.01; *=p<0.05

#: this variable is reversely coded to enable testing equality

8F. Additional tables macro-economy

Appendix 8f1. Incumbency voting regressed on economic indicators, multilevel logistic regression

Economy most salient	Low		High		Low		High	
Vote in last election	4.556	***	4.480	***	4.493	***	4.686	***
Economic evaluation	.362	***	.275	*	.407	***	.377	**
Ideological distance					-.432	***	-.364	**
Age					.012	**	.015	*
Social class					.129*		.077	
Gender					-.033		-.308	*
Immigrant status					.058		-.182	
Union membership					.210		.090	
Urbanization					-.051		-.056	
Religiosity					.149		.262	
Income quintile					-.024		-.036	
Economic growth					-.004		.139	
Unemployment change					.426		.968	*
Inflation change					.001		-.175	
Time since last election					-.008		-.003	
Constant	-3.737	***	-3.489	***	-3.921	***	-4.734	***
Pseudo R ²	.536		.534		.562		.583	
Log pseudo likelihood	-2,703		-2,029		1,673		-1,242	
Micro level N	8,546		6,494		5,596		4,505	
Country level N	22		22		21		20	

***= p<0.001; **=p<0.01; *=p<0.05. Reported are logged odds.

Appendix 8f2. Marginal effects of economic evaluations on incumbency voting, logged odds

	Salience		Newspaper		Interest		Education	
	Low	High	Low	High	Low	High	Low	High
Better	2.738***	2.575**	2.832**	3.560***	2.387**	2.983***	2.652***	2.917**
Worse	2.826***	2.643***	2.972***	3.010***	2.674***	2.884***	3.420***	2.326***
Difference	.088	.068	.140	-.550	.287	-.099	.768	-.591

8G. Additional tables political affinity

Appendix 8g1. Incumbency voting regressed on economic indicators, multilevel logistic regression

News consumption	Low		High		Low		High	
Vote in last election	4.372	***	4.404	***	4.486	***	4.578	***
Economic evaluation	.357	***	.379	***	.403	***	.423	***
Ideological distance	-.411	***	-.424	***	-.366	***	-.419	***
Age					.013	***	.015	*
Social class					.094		.219	***
Gender					-.127		-.151	
Immigrant status					-.007		-.091	
Union membership					.219		.342	*
Urbanization					-.020		-.035	
Religiosity					.149		.228	
Income quintile					-.035		-.023	
Economic growth	.078		.076		.048		.079	
Unemployment change	.893	*	.779		.959	*	.987	
Inflation change	-.106		-.067		-.135		-.114	
Time since last election	-.013		-.002		-.013		.006	
Constant	-3.040	***	-3.334	***	-4.008	***	-5.164	***
Pseudo R ²	.563		.559		.561		.572	
Log pseudo likelihood	-1,805		-2,141		1,314		-1,388	
Micro level N	6,136		7,113		4,455		4,772	
Country level N	21		21		20		20	

***= p<0.001; **=p<0.01; *=p<0.05. Reported are logged odds.

Appendix 8g2. Incumbency voting regressed on economic indicators, multilevel logistic regression

Political interest	Low		High		Low		High	
Vote in last election	4.372	***	4.404	***	4.391	***	4.794	***
Economic evaluation	0.357	***	0.379	***	0.330	***	0.452	***
Ideological distance	-0.411	***	-0.424	***	-0.337	***	-0.453	***
Age					-0.116		-0.201	*
Social class					0.017	**	0.012	**
Gender					-0.519	**	0.333	
Immigrant status					0.240		0.139	
Union membership					0.164	*	0.112	*
Urbanization					0.008		-0.086	
Religiosity					0.292	*	0.110	
Income quintile					-0.041		-0.006	
Economic growth	0.078		0.076		0.045		0.065	
Unemployment change	0.893	*	0.779		0.750		0.634	
Inflation change	-0.106		-0.067		-0.178		0.031	
Time since last election	-0.013		-0.002		-0.007		-0.003	
Constant	-3.040	***	-3.334	***	-4.396	***	-4.618	***
Pseudo R ²	.563		.559		.539		.560	
Log pseudo likelihood	-1,805		-2,141		-1,323		-1,574	
Micro level N	6,136		7,113		4,221		5,871	
Country level N	21		21		21		21	

***= p<0.001; **=p<0.01; *=p<0.05. Reported are logged odds.

Appendix 8g3. Incumbency voting regressed on economic indicators, multilevel logistic regression

Education	Low		High		Low		High	
Vote in last election	4.538	***	4.295	***	4.626	***	4.505	***
Economic evaluation	.371	***	.346	***	.398	***	.398	***
Ideological distance	-.430	***	-.427	**	-.393	***	-.405	**
Age					-.124		-.227	**
Social class					.013	*	.011	
Gender					-.180		.006	
Immigrant status					.009		.326	
Union membership					.142		.105	*
Urbanization					-.020		-.089	
Religiosity					.150		.326	
Income quintile					-.002		-.029	
Economic growth	.083		.063		.026		.057	
Unemployment change	.949	*	.553		.898	*	.618	
Inflation change	-.136		.001		-.163		-.037	
Time since last election	-.008		-.005		-.014		.005	
Constant	-3.308	***	-3.162	***	-4.112	***	-4.489	***
Pseudo R ²	.578		.547		.578		.560	
Log pseudo likelihood	-1,888		-2,079		-1,391		-1,414	
Micro level N	6,589		6,800		4,870		4,782	
Country level N	22		22		21		21	

***= p<0.001; **=p<0.01; *=p<0.05. Reported are logged odds.

Appendix 8g4. Incumbency voting regressed on economic indicators, multilevel logistic regression

Reading newspaper	No		Yes		No		Yes	
Vote in last election	4.647	***	4.391	***	4.985	***	4.520	***
Economic evaluation	.313	*	.363	***	.360	*	.404	***
Ideological distance	-.461	***	-.408	***	-.370	***	-.384	***
Age					-.179		-.165	*
Social class					.021	**	.012	*
Gender					.617		-.100	
Immigrant status					.078		.201	
Union membership					.248	**	.127	*
Urbanization					.165		-.079	
Religiosity					.306		.218	
Income quintile					.041		-.030	
Economic growth	.112		.064		.091		.049	
Unemployment change	.792		.688		.875		.704	
Inflation change	-.085		-.050		-.096		-.077	
Time since last election	-.023		-.002		-.025		-.001	
Constant	-2.758	***	-3.333	***	-5.309	***	-4.359	***
Pseudo R ²	.613		.556		.628		.561	
Log pseudo likelihood	-524		-3,622		-356		-2,546	
Micro level N	1,981		12,077		1,394		8,629	
Country level N	21		21		20		20	

***= p<0.001; **=p<0.01; *=p<0.05. Reported are logged odds.

Appendix 8g5. Incumbency voting regressed on economic indicators, multilevel logistic regression

Reading newspaper	No		Yes		No		Yes	
Vote in last election	4.982	***	4.519	***	4.648	***	4.390	***
Positive evaluation	.101		.301	*	.121		.295	*
Negative evaluation	.568	**	.505	***	.466	*	.428	***
Ideological distance	-.369	***	-.382	***	-.460	***	-.407	***
Age	-.182		-.163	*				
Social class	.022	**	.012	*				
Gender	.635	*	-.097					
Immigrant status	.075		.198					
Union membership	.247	**	.127	*				
Urbanization	.176		-.079					
Religiosity	.310		.220					
Income quintile	.045		-.030					
Economic growth	.091		.049		.110		.064	
Unemployment change	.905		.706		.805		.688	
Inflation change	-.096		-.077		-.086		-.050	
Time since last election	-.023		.000		-.022		-.002	
Constant	-4.546	***	-3.51	***	-2.043	***	-2.571	***
Pseudo R ²	.629		.561		.614		.556	
Log pseudo likelihood	-355		-2,545		-524		-3,621	
Micro level N	1,394		8,629		1,981		12,077	
Country level N	20		20		21		21	

***= p<0.001; **=p<0.01; *=p<0.05. Reported are logged odds.

Appendix 8g6. Incumbency voting regressed on economic indicators, multilevel logistic regression

Days reading news	0	1	2	3	4	5	6	7
Positive evaluation	.083	-.013	.573	.367	.435	.371	.383	.353**
Negative evaluation [#]	.610**	.405	.577**	.695**	.595	.472	.409	.514**

Micro and macro level control variables are the same as in the full models of other tables. Full model specifications are available at request.

Micro level N	1,350	851	886	832	536	630	1,007	3,135
---------------	-------	-----	-----	-----	-----	-----	-------	-------

***= p<0.001; **=p<0.01; *=p<0.05. Reported are logged odds.

[#]: this variable is reversely coded to maximize comparability

Appendix 8g7. Incumbency voting regressed on economic indicators, multilevel logistic regression

Political interest	Very low	Low	High	Very high
Economic evaluation	.199	.41***	.45***	.68***

Micro and macro level control variables are the same as in the full models of other tables. Full model specifications are available at request.

N	652	2,969	3,835	1,224
---	-----	-------	-------	-------

***= p<0.001; **=p<0.01; *=p<0.05. Reported are logged odds.

Appendix 8g8. Incumbency voting regressed on economic indicators, multilevel logistic regression

Political interest	Very low	Low	High	Very high
Positive evaluation	-.189	.283	.226	.626**
Negative evaluation [#]	.536	.406**	.571***	.683**

Micro and macro level control variables are the same as in the full models of other tables. Full model specifications are available at request.

N	729	3,492	4,475	1,396
---	-----	-------	-------	-------

***= p<0.001; **=p<0.01; *=p<0.05. Reported are logged odds.

[#]: this variable is reversely coded to maximize comparability

Appendix 8g9. Incumbency voting regressed on economic indicators, multilevel logistic regression

Education	Very low	Low	High	Very high
Economic evaluation	.46***	.39**	.49***	.45***

Micro and macro level control variables are the same as in the full models of other tables. Full model specifications are available at request.

N	2,159	2,278	2,210	2,033
---	-------	-------	-------	-------

***= p<0.001; **=p<0.01; *=p<0.05. Reported are logged odds.

Appendix 8g10. Incumbency voting regressed on economic indicators, multilevel logistic regression

Education	Very low	Low	High	Very high
Positive evaluation	-.073	.279**	.393***	.792***
Negative evaluation [#]	.805***	.629***	.833***	.730***

Micro and macro level control variables are the same as in the full models of other tables. Full model specifications are available at request.

N	1,641	6,229	7,354	2,281
---	-------	-------	-------	-------

***= p<0.001; **=p<0.01; *=p<0.05. Reported are logged odds.

[#]: this variable is reversely coded to maximize comparability