Erasmus School of Economics

You Get What You Vote?

an empirical analysis on reciprocity in the Eurovision Song Contest

ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics

Abstract

Reciprocity is proposed to be an important motivation for behaviour. This paper adds to previous literature by analysing whether intergroup reciprocal relationships are not limited to the same period. Data on voting behaviour in the Eurovision Song Contest is used to analyse reciprocity in a large natural dynamic group setting. It is a great yearly music contest in which countries vote for each other's performances and it is also a pop culture example of where reciprocity is expected to influence behaviour. Panel data on voting in the Eurovision Song Contest is analysed with a model including country, recipient, and time fixed effects. Furthermore, the model controls for geographical, cultural, and musical performance influences of voting behaviour in the contest. The estimated coefficient is 0.082 points in a linear regression of the independent variable excess points received last year on the dependent variable excess points. This effect is statistically significant at a 1% significance level, which means that countries remember how many points they received from who in the previous year and reciprocate by giving excess points in the present year. A country gives 0.082 points in excess for every point they received in excess from that other country last year. So, in an intergroup context, the reward from reciprocity is about one-tenth of the initial generous behaviour. However, significant effects of the many controls show that reciprocity is a complicated concept and it can be interlinked with many other influences of behaviour such as geography and culture.

Master Thesis Behavioural Economics

Number of words: 16857
Thesis supervisors: dr. (Georg) G.D. Granić
Second reviewer: dr. (Tong) T. Wang
Date final version: 12-03-2020

Student Number	Family Name	First Name
385582	Ruijter	Daisy

“Social equilibrium and cohesion could not exist without the reciprocity of service and return service. All contacts among men rest on the scheme of giving and returning the equivalence.”

- Georg Simmel

Acknowledgments

There are many people I would like to thank greatly for their generous help, which I will not be able to competently reciprocate. Firstly, I would like to thank my thesis supervisor Georg Granić for guiding me through this tough process and his valuable comments. Importantly, I would like to thank my grandma for providing me with fresh fruit while I was writing this thesis. I would also like to thank my friends, family, and fellow students for their support and help. I would especially like to thank my friend Mapau for helping me with my struggles of dealing with the data, which was definitely the worst part. Furthermore, I would like to thank dr. Tong Wang for her patience and guiding me during the end towards the defense. Lastly, I would like to thank my mother for (involuntarily) proofreading and assuring me with kind and loving words. (I also need to thank my mother for intimidating me with threatening words so I would stop procrastinating).

Contents

I. Introduction

What you give is what you get. Reciprocity has been acclaimed to be a prevalent factor which allows harmony in our societies and economies (Kolm, 2006). Before the industrial revolution, which is not so long ago on the scale of human history, the economy involved many transactions but with very few payments. Cooperation was based on an economy of favours and local traditions until the free market introduced laws of supply and demand (Harari, 2015). I would help my neighbours pick their apples without expecting any material payments for this. Later, my neighbours would help me pick my pears. This social norm is distinct from market exchange norms because the behaviours are not conditional on each other (Ariely, 2008). So, the theory of reciprocity hypothesises that people respond to behaviour from others with an equivalent behaviour, without an external agreement (Kolm, 2006). My thesis revolves around this hypothesised relationship. Supply and demand were able to pervade in most of our lives but there are still parts they cannot reach. In situations where there is an absence of markets, there is still cooperative behaviour due to social norms like reciprocity (Fehr & Gächter, 2000). Reciprocity can sometimes solve market and organization failures, but other times it can cause them (Kolm, 2006). For example, Falk (2007) found that the frequency of donations increased when charities added presents in their letters compared to not adding presents. On the other hand, Fischbacher and Schudy (2014) found that some politicians reject reforms to remove open ballots, which could benefit society because vote trading would be avoided, because they benefit from the reciprocal vote trading.

We are social animals; throughout our lives we are part of many groups ranging from families to Eurovision Song Contest supporters to nations. This means that many of our decisions take place in a social context where they are not independent from each other. Social judgments reflect the dynamics of intergroup relationships (Sherif, 1967; Tajfel, 1981). Interestingly, this also applies to non-human animals; non-human primates display reciprocity in grooming and they also form alliances (Seyfarth & Cheney, 1988). In our lives, reciprocity can be found in for example charities and the political sector or in family, labour, and international relations (Kolm, 2006). Since most of our relationships are built over time, reciprocal relationships are also usually formed over repeated decisions. Mechanisms which are expected to determine reciprocity are complex and include individual recognition and memory (Trivers, 1971). Reciprocal behaviour is thus predicted to be influenced by social interactions as well as the historical dynamics, which leads to the research question of this thesis: ***what is the reciprocal effect of previous voting behaviour on present voting behaviour in a dynamic group setting?***

For this, I analyse voting behaviour in the Eurovision Song Contest (afterwards referred to as ESC). The ESC is the largest and longest running annual tv music contest in the world (Stockemer, Blais, Kostelka, & Chhim, 2018; European Broadcasting Union, 2019). The ESC data offers the opportunity to analyse reciprocal relationships in a naturally occurring large dynamic group setting. The amounts of points that countries give to other countries are publicly announced on television, making the votes and intergroup characteristic more salient. Points should only depend on the performance quality and music is supposed to transcend barriers of culture, language, history, and economy. But there is notoriously much speculation of voting based on politics and geography (Mantzaris, Rein, & Hopkins; 2018). Highlighting the historical voting patterns between Cyprus and Greece appears to confirm the speculations. Till 2018, Cyprus gave by far the most points to Greece with 419 points. After Greece comes Sweden as Cyprus gave them 156 points. Cyprus also received the most points from Greece.

Greece gave by far the most to Cyprus with 400 points and is followed by the United Kingdom with 157 points (European Broadcasting Union, 2019). These occurrences prompt questions about whether countries systematically give each other higher than expected points.

Though the ESC is not claimed to be the perfect representation of a dynamic group setting, it has many favourable characteristics. Reciprocity is not often examined in natural occurring situations due to difficulties of finding appropriate data. Firstly, reciprocity often involves different types of behaviour which makes it hard to compare them. Secondly, individuals or groups often vary in their ability to reward or punish (Seyfarth & Cheney, 1988). Seyfarth and Cheney (1988) explained this with an example of non-human primates: high-ranking primates can give more assistance, such as alliance formation, compared to low-ranking primates who are restricted to grooming. Researching repeated interactions over longer periods is also by nature organizationally more difficult. These difficulties can be overcome with the ESC: reciprocity would involve the same type of behaviour, namely voting, and the participants do not vary in their ability to vote. Possible drawbacks of using the ESC as a representation is that it has its own specific characteristics, which I will control for as much as possible.

The rise of experimental and behavioural economics has established that reciprocity is an important motivation for behaviour. However, reciprocity is often analysed on an individual level but not so much between groups (Doosje & Haslam, 2005). Though, many interactions are repeated with groups instead of individuals (Cason & Mui, 2019). Some previous research has extended the knowledge on intergroup reciprocal relationships. However, there is not much empirical research which acknowledges the effects that historical dynamics might have on the behaviour (Doosje & Haslam, 2005). Real-world interactions often include multiple periods with the same players; decisions in the European Union, for example, are made by the European Parliament and the Council which consists of the governments from all 28 European Union members states (European Union, 2019). To add on existing knowledge of reciprocity, I use the ESC to conduct empirical research on reciprocal relationships in a dynamic group setting. The dataset allows to analyse whether countries remember which country gave them how many points in the previous year(s) and reciprocate based on that.

Influences in the ESC could provide insights to other relationships between countries such as in the United Nations or the European Union. These groups have a real history which is hard to manipulate in experiments. If reciprocity significantly influences behaviour in dynamic group settings, it can be worthwhile to invest in intergroup relationships. Reciprocity can increase cooperation in the many real-life intergroup interactions where the market cannot or does not determine the outcome. Understanding the role of reciprocity would for example provide benefits in negotiations. Other countries could be persuaded to support a certain future goal by building and strengthening the intergroup relationships through generous acts in the present.

This paper has the following setup: the successive sections II and III include respectively a description of the ESC system and a review of related literature. Afterwards is a description of the data in section IV and the research method in section V. This is followed by section VI, in which the results of the research are presented and interpreted. Lastly, the paper is discussed and concluded in respectively section VII and section VIII.

II. The ESC System

Each participating country is represented by a broadcaster which puts forward a musical performance for the annual competition. Since 1956 the ESC grew from seven countries to a maximum participation of 44 countries per year. Along with this development came substantial changes in the voting system and regulations (Mantzaris, Rein, & Hopkins, 2018). The current voting system consists of every country being obliged to deliver two[1] sets of points. The first is from jury voting and the second by televoting. Jury voting is done by five national music experts while televoting is done by the national audiences (European Broadcasting Union, 2019). The contest also includes semi-finals since 2004. During the semi-final, the ten performances with the most votes will qualify for the grand final. Six countries skip the semi-finals and are automatically qualified, which are the following: the host (country which won the previous year), France, Germany, Italy[2], Spain, and the United Kingdom (European Broadcasting Union, 2019). During the semi-final, the countries participating in that edition and three pre-qualified countries are obliged to vote. While in the grand final, even the countries which did not qualify have to vote (European Broadcasting Union, 2019).

Voting is done by ranking all the performances. It is not permitted to rank your own country as well as giving the same rank to two different performances. The ranking determines the allotment of points, which is based on the following system: the performance with the highest rank receives 12 points, the second-best receives 10, and the third-best receives 8. Afterwards 7, 6, 5, 4, 3, and 2 points are given accordingly to the performance with the ensuring rank. Finally, the performance with the tenth-best rank receives 1 point while the rest of the performances receive 0 points (European Broadcasting Union, 2019). The points allocated by the juries and televoters are announced separately on live television, which are then combined in a 50/50 ratio. The performance with the highest total of points wins (European Broadcasting Union, 2019).

The ESC has many rules to support their commitment of a fair and valid outcome. For example, national jury members must be professionals from the music industry, and they cannot have been a member of the national jury in the preceding two years (European Broadcasting Union, 2019). Juries are required to vote without bias to for example nationality or gender. Furthermore, the decision regarding which of the two semi-finals countries participate in, is issued by a system of draws which take past voting patterns into consideration. Countries are divided into groups which make it less likely that countries which traditionally give each other high points are placed together (European Broadcasting Union, 2019).

[1] Until 2016 every country gave one set of points.
[2] Before 2011 Italy was excluded from this list

III. Reciprocity and the ESC

Reciprocity

The theory of reciprocity hypothesises that people respond to behaviour from others with equivalent behaviour, without receiving material benefits. Its relevancy goes all the way back to the birth of modern economics where Adam Smith (1759) wrote about reciprocity as a motivation for behaviour. Recognition of importance lingered but the conventional image during the 19th and 20th centuries was that of the economic man (Kolm, 2006). Most economists such as Arrow (1980) took greed, which leads to self-interest, to be the main motivation of behaviour (Fehr & Gächter, 2000). Due to observing trade-offs which are incompatible with the economic man, behavioural economists returned to social norms such as reciprocity. For example, Fehr and Gächter (2000) concluded that individuals behave according to theories of reciprocity in experiments: people often punish free riders when they are given the option to sanction each other, even when it is costly. The rewards and punishments are often bigger than predicted by theories of self-interest.

Previous research has been able to confirm theories of reciprocity on an individual level, using a wide array of methods such as experimental games.[3] With these games, the researchers focused on either reward[4] or punishment[5]. For example, reciprocity was found in ultimatum games as the rejection of low offers, despite leading to economic loss (Güth, Schmittberger, & Schwarze, 1982; Thaler, 1988; Roth, 1995). Originally, reciprocity was assigned to individuals but Gouldner (1960) also applied it to groups. People are brought together in many different combinations, producing different types of relationships. Gouldner (1960) attributed reciprocity to all social systems, for example reciprocity in marriage and reciprocity between political parties (Gouldner, 1960). Intergroup relationships occur in interactions between two or more groups, which are assembles of two or more individuals. However, it is somewhat unclear whether groups empirically display reciprocal behaviour. Whereas Song (2006) concluded from trust games that reciprocal dynamics between individuals cannot be readily applied to intergroup reciprocal dynamics; there is significantly less reciprocal behaviour between groups compared to individuals. This brings about the interest in the ESC, which is a dynamic group setting in which reciprocal relationships might take place.

Reciprocity in the ESC

Fischbacher and Schudy (2014) from laboratory experiments with individuals that open ballots allow for vote trading through reciprocity. Trust and reciprocity were an essential basis for vote trading among legislators, especially because formal commitment mechanisms are often not possible. Similarly, the characteristics of the ESC ballot structure also encourage reciprocity as it consists of open ballots and formal commitment mechanisms are not allowed. Previous research on whether countries in the ESC systematically give each other higher than expected points also suspected and found reciprocity. For

[3] These games include investment games (Berg, Dickhaut, & McCabe 1995), the prisoner's dilemma (Falk, Fehr, & Fischbacher, 2005), the gift-exchange game (Fehr, Kirchsteiger, & Riedl, 1993), and the ultimatum game (Güth, Schmittberger, & Schwarze, 1982; Thaler, 1988; Roth, 1995).

[4] Fehr et al. (1993) and Berg et al. (1995)

[5] Güth et al. (1982), Thaler (1988), Roth (1995), Fehr and Gächter (2000), and Falk et al. (2005)

example, Doosje and Haslam (2005) concluded that countries gave significantly better scores in 1996 to countries they received good scores from in 1991 to 1995 and vice versa.

A common assumption was that the effects of reciprocity remain stable as time goes on. However, a recent research by Chuan, Kessler, and Milkman (2018) on charitable giving revealed that reciprocity declines over time. Mantzaris, Rein, and Hopkins (2018) also analysed the time aspect and stated that conclusions based on consistent voting behaviour over multiple years are more reliable. They created graphs with networks of both one-sided and two-sided collusion in the ESC whereby the two-sided collusions were described as reciprocity. Tendency for two-sided collusion steadily increased over the years among one-sided collusions, which suggests that these countries see benefits in reciprocal relationships and are therefore increasingly cooperative. Hence, for their research on the ESC, Ginsburgh and Noury (2008), Spierdijk and Vellekoop (2009), and Budzinksi and Pannicke (2017) opted for panel data analysis. The results of their research will be described separately later. They all applied fixed effects models, which allowed the intercept and slope to vary per voting country. This approach is endorsed by the conclusion of Spierdijk and Vellekoop (2009) that the way factors influence voting behaviour substantially differs across countries.

A major concern in the ESC is that voting is influenced by factors other than the quality of performance. Ginsburgh and Noury (2008) analysed this by creating a proxy for quality. They searched for vote trading based on political issues, which resembles reciprocity, as an influence on voting behaviour. In their equation the number of points given and received between two countries, within a year, is set off against each other. Countries seem to partake in vote trading. However, despite that performances from countries with similar cultures and languages are generally preferred, quality prevails instead of the suspected political influences (Ginsburgh & Noury, 2008). Spierdijk and Vellekoop (2009) extended on their research and created a concept of bias using the quality proxy. This concept of bias was adopted by Budzinksi and Pannicke (2017) and together they found many non-reciprocal influences of voting bias, which will be further highlighted in the "other influences of ESC voting" section.

So, previous panel data research on the ESC already analysed whether countries assign each other higher than expected points, so-called excess points, within the same year. My research will extend on this by adding another time aspect with the lagged effect of reciprocity. As social interactions often are not limited to one time period, they can involve abilities such as memory and recognition. Therefore, I will research whether countries remember the points they have received in excess[6] last year and reciprocate with their own vote. In hypothesising the relationship of voting behaviour between countries, I drew support from the economic literature on theories of reciprocity (e.g. Falk and Fischbacher, 2006). The theories of reciprocity simply hypothesise that an action is followed by a similar action. Therefore, the following hypothesis representing reciprocity has been formulated with simply a positive linear relationship:

Receiving excess points last year has a positive effect on giving excess points this year in the ESC.
(hypothesis 1)

[6] In order to measure reciprocity, I will create a variable to represent excess points. The way this concept is created is described in the data section.

Jury vs Televoting

Another aspect of voting in the ESC is the distinction between professional juries and televoting done by the general public. One can expect that juries are less likely to be influenced by reciprocity. Firstly, they are not allowed to be a member of the national jury in the preceding two years. This should already disrupt reciprocal relationships. Furthermore, they should vote based on quality and without bias with their professional skills and experience. In addition to this, the points awarded by jury members are announced separately. Furthermore, the name as well as the country of jury members are published online (Groot, 2019). Ginsburgh and Noury (2008) argued, using the experimenter demand effect, that publicising voting behaviour will push the professional juries to be more objective. So, juries are pushed to be less likely influenced by reciprocity. However, previous research on the ESC does not have conclusive results on the difference between jury and televoting. By estimating whether the votes were similar in the period before the introduction of televoting and the period after, Ginsburgh and Noury (2008) concluded that there is no significant difference between votes from juries and televoters. Yet, Spierdijk and Vellekoop (2009) used the same approach and found that juries are less likely influenced by religious similarities and patriotic voting. Furthermore, Haan et al. (2005) concluded that juries are better judges of quality when compared to the public; the judgements of the juries are influenced less by the performance order. To gain more insight on this ambiguous difference, the following hypothesis was formulated based on the experimenter demand effect:

Juries display less reciprocal behaviour compared to the people who vote through televoting.

(hypothesis 2)

Other influences of ESC voting

Existing literature on the ESC has found many influences of voting behaviour besides reciprocity, which need to be recognized and controlled for. These can be divided into the following three categories: geographical influences, cultural influences, and musical performance influences.

Firstly, geographical influences play a role in ESC voting. Many researchers have confirmed significant collusion and cliques of countries based on geographical factors[7]. The compositions of these cliques depend on the time period as they are susceptible to change. The most comprehensive research, encompassing 1957 till 2017, was done by Mantzaris, Rein, and Hopkins (2018). The cliques found to be statistically significant from 2007 till 2017 are the most relevant for this thesis and are visualised in Appendix 1. Another type of geographical influence was confirmed by Spierdijk and Vellekoop (2009) and Budzinksi and Pannicke (2017); countries have a significant positive bias for their neighbours.

Secondly, culture also influences voting behaviour in the ESC. Cultural preferences predict that people prefer their own culture. Therefore, people favour countries with whom they share similarities. Many

[7] Yair (1995), Yair & Maman (1996), Doosje & Haslam (2005), Gatherer (2004 & 2006), Fenn et al. (2006), Spierdijk & Vellekoop (2009), Mantzaris, Rein, & Hopkins (2018)

researchers have found significant positive effects of similarity in language and culture[8]. As culture is complicated, it has many different dimensions. Hofstede (1980) created dimensions which could be measured and used for analysis. These dimensions are useful and have been used by other researchers such as Doosje and Haslam (2005), who concluded that collectivistic countries display more reciprocity than individualistic countries. However, the Hofstede dimensions do not encompass the entire complexity of culture. For example, Doosje and Haslam (2005) also concluded that low-power countries display more reciprocity than high economic power countries, which is not a Hofstede dimension. Furthermore, other researchers concluded that countries prefer religious closeness (Spierdijk & Vellekoop, 2009; Budzinksi & Pannicke, 2017)[9]. Spierdijk and Vellekoop (2009) also analysed ethnic closeness by using Turkish migrants and found a significant bias to country of origin, which they called patriotic voting.

Thirdly, characteristics of the musical performance are also expected to influence voting behaviour in the ESC. Concerning this, Haan, Dijkstra, and Dijkstra (2005) and de Bruin (2005) concluded that having an earlier musical performance is significantly related to a lower chance of receiving high points. An exception was found by Haan, Dijkstra, and Dijkstra (2005): the opening performance has a higher chance of receiving high points. However, these results have been challenged by other rearchers such as Ginsburgh and Noury (2008), Spierdijk and Vellekoop (2009), and Budzinksi and Pannicke (2017). English performances, ones performed by groups, and performances by the host country have a higher percentage of wins than the percentage of participation[10] (Spierdijk & Vellekoop, 2009). Yet, characteristics of the musical performance (gender, language, order, solo, duet, or group, and whether it was the host country performing) did not have statistically significant effects on voting behaviour (Spierdijk & Vellekoop, 2009; Budzinksi & Pannicke, 2017).

All things considered; many factors could influence voting behaviour in the ESC. While Ginsburgh and Noury (2008) and Spierdijk and Vellekoop (2009) were able to explain the voting behaviour of some countries, they concluded there is still “unexplainable noise”. Furthermore, recent research has found comparable voting biases in a national music contest similar to the ESC, despite more homogenous cultural backgrounds within countries (Budzinski & Pannicke, 2017). This suggests that, in addition to geography and culture, there are other factors influencing voting behaviour. A possible factor would be reciprocity.

[8] Ginsburgh and Noury (2008), Spierdijk and Vellekoop (2009), Blangiardo and Baio (2014), Budzinksi and Pannicke (2017), and Mantzaris, Rein, and Hopkins (2018)

[9] Budzinksi and Pannicke (2017) confirmed countries prefer performances from countries with the same religion and Spierdijk and Vellekoop (2009) confirmed it for Cyprus, Iceland, Ireland, Latvia, and the United Kingdom.

[10] 50% of the winning performances are sung in English while only 24% of all performances are sung in English and 41% of the winning performances are by groups account while this is only 25% for all performances. Also, more than 10% of the winning performances are by the host country while only less 5% of all performances is by the host country.

IV. Data

For the analysis, I constructed a dataset of 49 countries covering the period of 2004 till 2018. This added up to 31753 observations. This panel data is characterised by three dimensions: voting country i (individual unit dimension), receiving country j (receiving unit dimension), and year t (time dimension). Thus, it is a multi-dimensional panel data (MDPD). Data spanning over several years is especially important for this thesis because of the potential time aspect of reciprocal relationships. Also, the findings are more reliable if the reciprocity is consistent over a longer period (Mantzaris, Rein, & Hopkins, 2018). Due to the changing voting schemes, the analysis does not include the entire period of the ESC. This specific time period is chosen because the ESC voting scheme hardly changed during that time (Mantzaris, Rein, & Hopkins, 2018). Also, the most recent years are chosen because it is new and still relatively unexplored.

The data gives an unbalanced panel as not all countries participate or perform every year in every edition. Countries are selected for certain semi-finals or have to be qualified for the final. Furthermore, not all countries participate yearly; some countries have stopped participating in the ESC and others joined later. For example, Slovakia participated four times over the period of 2004 till 2018, while the Netherlands participated fourteen times. The data is truncated because most Eastern European countries debuted at the ESC in the 1990s or later, whereas there is data for Western European countries since 1975. Therefore, reducing the considered time period to 2004 till 2018 also somewhat limits the overrepresentation of Western European countries. This creates a more balanced dataset. All included countries participated at least three times from 2004 till 2018[11], which is a requirement inspired by Spierdijk and Vellekoop (2009) and Budzinksi and Pannicke (2017). A table of all the countries is provided in Appendix 2.

Dependent variable: excess points

The dependent variable is the continuous variable excess points. It is the excess points given on top of what would be expected based on the average. To measure reciprocity, variables were needed which could be rewarded or punished with an equivalent behaviour. In the ESC this would involve generous or stingy points. Voting generously is regarded as positive behaviour because there are no contractual obligations to do so. However, some performances receive higher points simply because they are better than others. Therefore, the variables need to account for quality because it is not regarded as behaviour worth rewarding if high points are given to a high-quality performance; it is not generous to give something which was deserved. However, the quality of art such as music performances is hard to operationalise as it is often based on unmeasurable factors. The following proxy, which was created by Ginsburgh and Noury (2008), is used for the quality of a performance:

$$quality_{jt,-i} = \frac{1}{n_{t,-ij}} \sum_{k \neq i} points_{kjt} \qquad \text{formula (1)}$$

[11] Serbia and Montenegro performed only two years as one country between this period but was not excluded from the dataset because the country did participate in the final voting for the winner in a year they withdrew from performance (European Broadcasting Union, 2019).

$points_{kjt}$ is the amount of points given to country j in year t.
$n_{t,-ij}$ is the amount of countries participating in the ESC in year t but subtracting country i and country j. So, quality of the performance by country j in year t is the average points given to country j by all countries other than country i. As can be derived from the concept, quality of the performance by country j (receiving country) differs per country i (voting country). Using quality, the dependent variable was created, which is inspired by Spierdijk and Vellekoop (2009) and Budzinksi and Pannicke (2017)[12]:

$$excess\ points_{ijt} = points_{ijt} - quality_{jt,-i} \qquad \text{formula (2)}$$

$points_{ijt}$ is the amount of points given to country j from country i in year t.
So, $excess\ points_{ijt}$ is the difference between the points from country i to country j and the average of the points all other countries gave to country j in year t. This variable is used in order to measure reciprocity because it can represent a reward or punishment. If the amount of excess points is for instance two, it means country i gave country j two points more than country j had received on average from all other countries in this period. These extra two points can be regarded as generous behaviour, which can be used as a reward.

This concept of excess points has several advantages. Issues regarding endogeneity and circularity are avoided by excluding the points of country i to country j (Budzinksi & Pannicke, 2017). In addition to this, it is a continuous variable and it controls for quality. However, it is not undisputable that the correct or best measurement of quality is how a performance is assessed by other participants. Ideally there would be a variable perfectly operationalising quality. As quality is almost impossible to operationalise, this concept satisfies since it is likely that high-quality performances receive higher points. It also provides a measurement of overvaluation (or undervaluation), which can be regarded as positive (or negative) behaviour that can be used to analyse reciprocity.

Main independent variable: excess points received last year

The main independent variable is the continuous variable excess points received last year. This concerns the points received from a country last year on top of what would have been expected based on the average. This is past behaviour from country j that country i would reciprocate with a reward (or punishment) of excess point (or negative excess points). The construction of this variable is the same as the previously described excess points, only country j is switched with country i and then the variable is lagged. Country j is usually the receiving country but for the main independent variable it is the country giving points to country i, which is usually the voting country. This means quality is defined as:

$$quality_{it,-j} = \frac{1}{n_{t,-ji}} \sum_{k \neq j} points_{kit} \qquad \text{formula (3)}$$

$points_{kit}$ is the amount of points given to country i in year t.
$n_{t,-ji}$ is the amount of countries participating in the ESC in year t but subtracting country j and country i.

[12] Spierdijk and Vellekoop (2009) created this variable and named it “bias” but in this thesis it is referred to as excess points because it does not necessarily need to be irrational.

So, quality of the performance by country i in year t is the average points given to country i by all countries other than country j. As previously explained, quality is used to create the following:

$$excess\ points\ received_{jit} = points_{jit} - quality_{it,-j} \qquad \text{formula (4)}$$

$points_{jit}$ is the amount of points given to country i from country j in year t.
So, $excess\ points\ received_{jit}$ is the difference between the points country i received from country j and the average number of points country i received from all countries other than country j in year t. Country j is hereby the country that country i gives points to in the dependent variable excess points. In order to measure the time aspect of reciprocal relationships, the variable is lagged one year. Thus, excess points received last year measures the amount of excess points a country received in the previous year. This variable is used in order to measure reciprocity because it can represent positive (or negative) behaviour that a country may want to reward (or punish) with an equivalent behaviour. If the amount of excess points received last year is for instance three, this means that last year country j gave country i three points more than country i had received on average from all other countries. These three points in excess can be regarded as generous behaviour which is worth rewarding in the present year.

This concept has the same advantages and shortcoming as excess points, but it has an extra shortcoming: there are many missing values, especially in the finals. This is because all countries can vote even when they do not perform. So, countries can vote for a certain country while not receiving points from that country simply because they do not have a performance to receive points on. The dataset assigned zero points to countries which voted but did not perform. These values were removed when creating excess points received because it would seem like these countries received low points, while in truth these countries did not perform.

Some basic descriptive statistics are presented in Table 1. The decrease in number of observations exhibit the problem of missing variables which was explained earlier; countries do not consistently perform in the ESC, which means that not all generous (or stingy) voting can be reciprocated by returning the favour. Another version of the main independent variable was created as a recognition of the missing variables problem. This version, called excess points received last time, has fewer missing variables than the original and is further highlighted in the additional analysis section of the results.

All countries	*Number of countries: 49*				
	Observations	Mean	Standard Deviation	Min	Max
Excess points	31,753	0	3.071	-9.55	12
Excess points received	21,373	-0.006	3.135	-9.542	11.957
Excess points received last year	9,349	-0.025	3.08	-9.542	11.957
Excess points received last time	14,012	-0.02	3.116	-9.542	11.957

Table 1: Summary statistics of the variables excess points, excess points received, excess points received last year, and excess points received last time for all countries; see formulas (1) and (2) for the construction of excess points and formulas (3) and (4) for the construction of excess points received last year; excess points received last year is lagged for 1 year and excess points received last time is lagged from the previous data point, regardless if it was 1 year ago or earlier.

A histogram of the dependent variable excess points, in Figure 1, shows that excess points has a skewed distribution. The histogram illustrates the fact that voting countries cannot give points to all participating countries; voting countries can only give points to ten countries, while there is a maximum participation of 28 countries. So, countries mostly receive negative excess points, which makes the positive excess points more exclusive and therefore more valuable for reciprocal relationships.

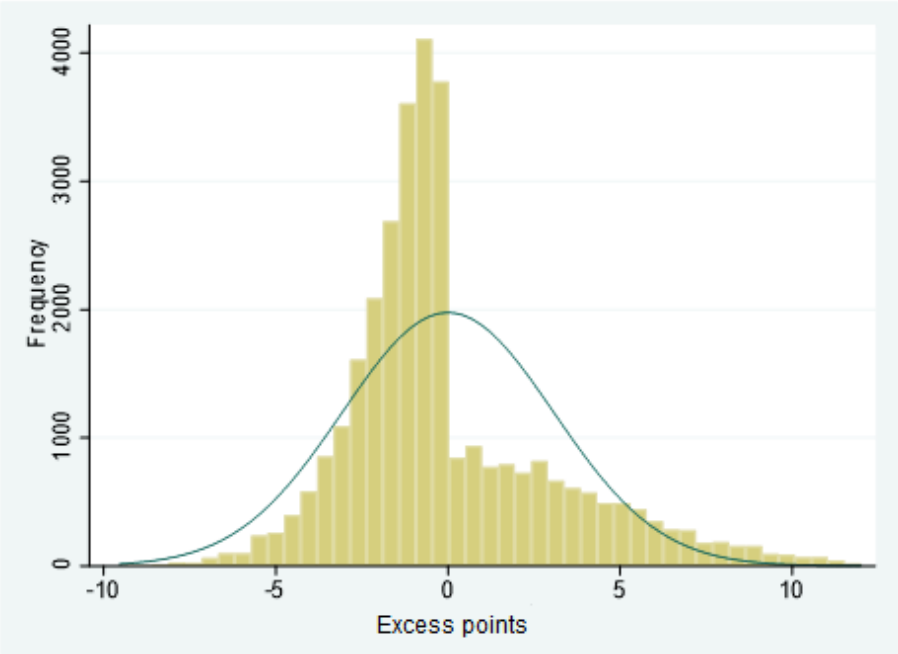


Figure 1: Histogram of the dependent variable excess points for all countries; see formulas (1) and (2) for the construction of this variable.

Scatterplots of excess points received last year on excess points seem to provide preliminary evidence of reciprocal relationships. A scatterplot of all countries is given in Figure 2 and to illustrate, a scatterplot of one country is given in Figure 3. Greece is hereby used as an example and both scatterplots can be found on the next page. A linear fitted line shows a positive relationship between excess points received last year and excess points given. This means that as excess points received last year increases, so does excess points. Also, Figure 2 shows that there are no large outliers. Scatterplots with nonlinear fitted lines are also probed and can be found in Appendix 3 for all countries and Appendix 4 for Greece. These scatterplots display a quadratic relationship of excess points received last year on excess points. The line increases slowly until excess points of zero and thereafter increases exponentially. However, the theories of reciprocity only hypothesize about a positive relationship and does not mention the shape. Therefore, this thesis will simply analyse a positive linear relationship.

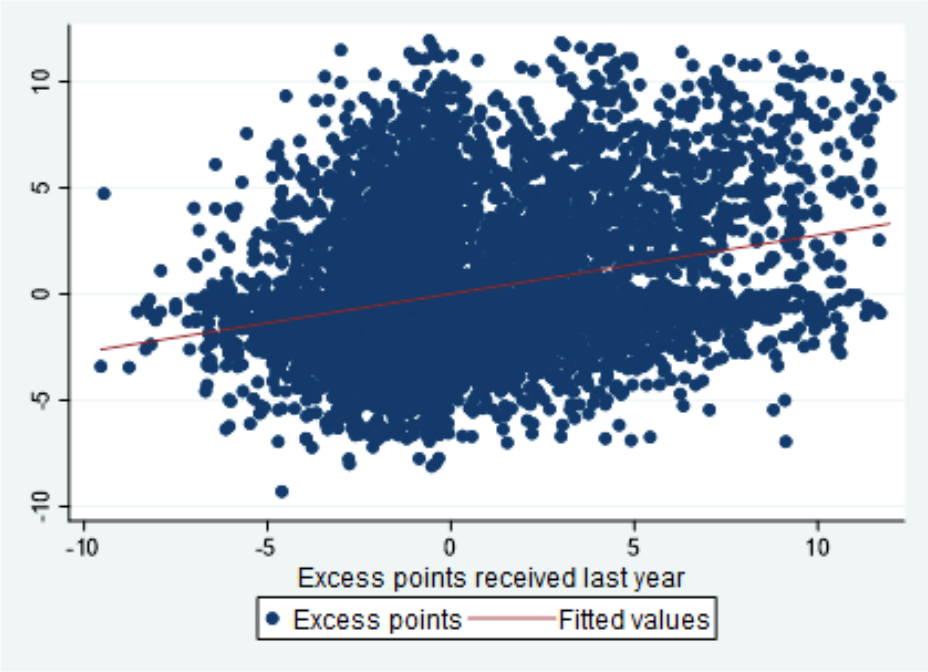


Figure 2: Scatterplot from the main independent variable excess points received last year on the dependent variable excess points for all countries with linear fitted line; see formulas (1) and (2) for the construction of excess points and formulas (3) and (4) for the construction of excess points received last year.

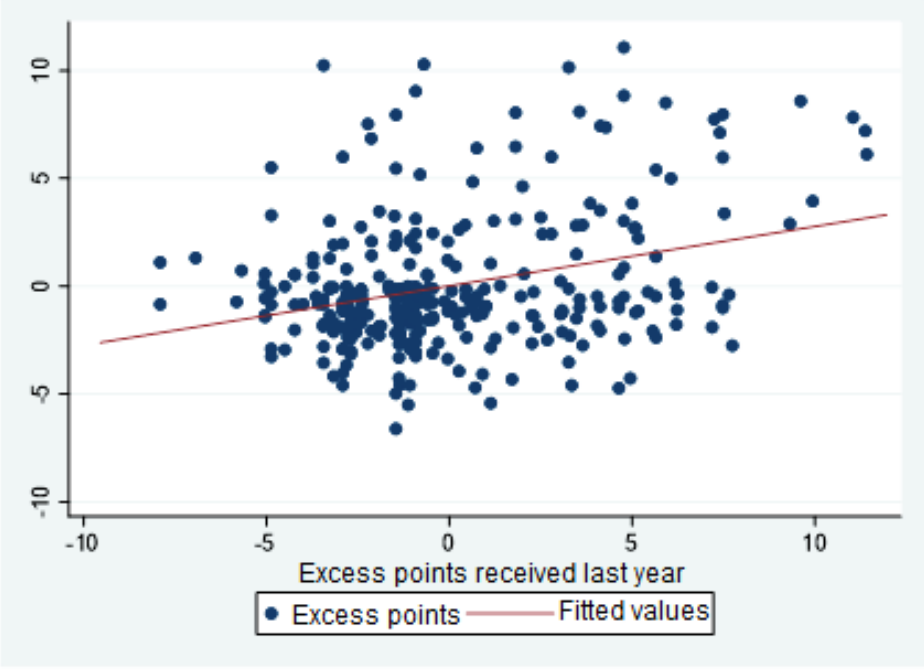


Figure 3: Scatterplot from the main independent variable excess points received last year on the dependent variable excess points for Greece with linear fitted line; see formulas (1) and (2) for the construction of excess points and formulas (3) and (4) for the construction of excess points received last year.

The dependent variable and main independent variable are constructed from country voting behaviour in the ESC. The voting data is obtained from *https://data.world/datagraver/eurovision-song-contest-scores-1975-2019*. This dataset with aggregated voting behaviour is made publicly available by *Datagraver.com*, which is an initiative from the IT service provider Itude. Itude has a respectable reputation and has worked for clients such as Greenpeace and the Dutch government (Datagraver, 2019). The dataset includes the points countries gave to each other through ranking. It distinguishes between points given during semi-finals and finals as well as points given by juries and televoters.

Jury

So, the voting data also provides for the variable which is necessary for testing hypothesis 2. The hypothesis states that juries display less reciprocal behaviour compared to people who vote through televoting. Jury is a dummy variable which takes on the value 1 if the points from country i (voting country) are from juries and 0 when the points are from televoting. Including the jury variable follows the footsteps of Haan (2005), Ginsburgh and Noury (2008), and Spierdijk and Vellekoop (2009). The variable measures the difference between excess points from juries compared to televoters. The dummy is a simple and straightforward way to measure the difference, but it is important to recognise that the data on televoting is available only since 2016. This is a rather short timespan for analysing dynamic relationships.

Controls

Aside from the two main variables, I include 23 control variables which had been argued by previous researchers to influence voting behaviour. Countries vote blindly which makes it redundant to control for points received within the same year. The controls are divided into the following three categories: geographical influences, cultural influences, and musical performance influences. Geographical influences are operationalised with the variable neighbours[13] and three country cliques (north, east, and southeast) [14]. Previous research already concluded geography is an important explanation for collusion. Furthermore, culture was also concluded to affect voting behaviour. The cultural influences are operationalised with Hofstede's six[15] cultural dimensions (power distance, individualism, masculinity, uncertainty avoidance, long-term orientation, indulgence)[16], language[17], religion[18], Turkish population[19], and economic power[20]. Cultural preferences and patriotic voting were suggested to incite bias towards

[13] Data for neighbours was collected by using Google maps.
[14] The included country cliques are based on networks of collusion found by Mantzaris, Rein, and Hopkins (2018) which were statistically significant for voting behaviour during 2007 till 2017. A visualization is displayed in Appendix 1. Country cliques which only consisted of two countries were not included in this control variable because these could be reciprocal relationships this thesis is primarily interested in.
[15] Previous research had four cultural dimensions, but two new dimensions were added by Hofstede after the previous papers were published.
[16] Data on Hofstede's cultural dimensions was collected from geerthofstede.com and hofstede-insights.com.
[17] Data on language was collected from the CIA World Factbook.
[18] Data on religion was collected from Pew Research Center, which is a nonpartisan think tank.
[19] Data on Turkish immigrant population was collected from the 2004 final Report of the Independent Commission on Turkey by the Turkish Studies Center.
[20] Data on GDP per capita was obtained from the World Development Indicator database of the World Bank Group.

countries with similar cultures or countries of origin. Ideally there is a control for minorities, but only Turkish migrants are considered due to the difficulty of finding reliable data on migration of other minorities. Economic power was also proven to be an important predictor of behaviour. Lastly, previous research also underlined possible biases towards certain characteristics of the musical performance. For example, all countries are free to choose in which language the act will be performed but the show is presented in English and French (European Broadcasting Union, 2019). Performances with English or French songs might make a different impact. Therefore, musical performance influences[21] are operationalised by the following variables inspired by previous research: automatically qualified, host country, gender of performer, English performance, French performance, duet, group, order of performance, and opening performance. All the control variables and their expected effect are described in more detail in Appendix 5 whereas summary statistics are included in Appendix 6.

[21] All variables on performance were collected from the official ESC website eurovision.tv and Wikipedia.

V. Research Method

The following model is estimated, and the variables are described in the subsequent Table 2:

$$excess\ points_{ijt} = \boldsymbol{\beta_1 excess\ points\ received\ last\ year_{jit}} + \beta_n geographical\ influences_{ij} + \beta_n cultural\ influences_{ij} + \beta_n performance\ influences_{jt} + \alpha_i + \tau_j + \lambda_t + u_{ijt}$$

Table 2: Description of the variables included in the model

$Excess\ points_{ijt}$	Excess points from country i (voting country) to country j (receiving country) in year t; excess points is the difference between the points given from country i to country j and the average points given to country j by all countries other than country i in year t; see formulas (1) and (2)
$Excess\ points\ received\ last\ year_{jit}$	Excess points country i received last year from country j in year t; excess points received last year is the difference between points country i received from country j and the average of points country i received from all countries other than country j in year t-1; see formulas (3) and (4)
$Geographical\ influences_{ij}$	Collection of control variables operationalising geographical influences[22] of country i (voting country) with country j (receiving country)
$Cultural\ influences_{ij}$	Collection of control variables operationalising cultural influences[23] of country i (voting country) with country j (receiving country)
$Musical\ performance\ influences_{jt}$	Collection of control variables operationalising musical performance influences[24] of country j (receiving country) in year t
i	The voting countries 1, 2, …, 49
j	The receiving countries 1, 2, …, 49

[22] Geographical influences include the following control variables: neighbours and three cliques (north, east, and southeast).

[23] Cultural influences include the following control variables: six cultural dimensions of Hofstede (power distance, individualism-collectivism, masculinity-femininity, uncertainty avoidance, long-term orientation, and indulgence), language, religion, Turkish population, and economic power. The cultural influences except economic power do not depend on time because they are expected to remain close to stagnant over time (Geert Hofstede, 2019). Also, economic power is therefore the only cultural influence which has different dimensions, which are the following: country i (voting country) and year t.

[24] Performance influences include automatically qualified, host, gender, English song, French song, duet, group, order, and opening. Order and opening have different dimensions than the other musical performance influences, which are the following: country j (receiving country) in edition x of year t.

Table 2 continued: Description of the variables included in the model

t	The number of years 1, 2, …, 15
α_i	The country-specific intercept
τ_j	The recipient-specific intercept
λ_t	The time specific intercept
u_{ijt}	Idiosyncratic error from country i (voting country) with country j (receiving country) at year t

The linear specification includes excess points, excess points received last year, 49 voting country binary variables (country fixed effects), 49 recipient country binary variables[25] (recipient fixed effects), 15 single-year binary variables (time fixed effects), and 23 control variables. Excess points is the dependent variable and excess points received last year is the main independent variable. β_1, which is the slope coefficient of the regression, is the effect of excess points received last year on excess points after controlling for influences of geography, culture, and musical performance. β_1 is of primary interest because it depicts the reciprocal relationship described in hypothesis 1: receiving excess points last year has a positive effect on giving excess points this year in the ESC. So, β_1 is hypothesized to be positive.

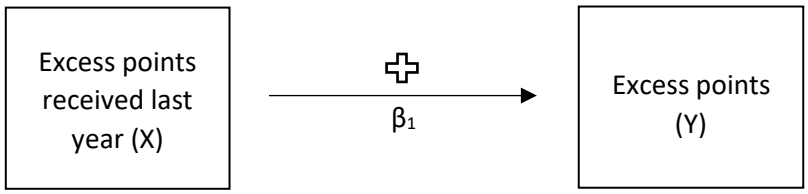


Figure 4: Representation of reciprocal relationship in the ESC

The chosen approach is a generalised least squares regression with country fixed effects, recipient fixed effects, and time fixed effects. The panel data model controls for biases which cannot be fully explained with the control variables. For example, it is expected that countries have structural biases due to historical ties which do not involve reciprocity. People from different countries vary in culture and many other aspects which could not be captured by one of the controls; each country has certain institutions which cannot be operationalised but might affect the excess points given. Therefore, country fixed effects and recipient fixed effects are included to account for unobserved country level heterogeneity bias. Including country fixed effects follows previous research on ESC voting behaviour (Ginsburgh & Noury, 2008; Spierdijk & Vellekoop, 2009; Budzinksi & Pannicke, 2017). Furthermore, time fixed effects are included to absorb shocks which are specific to the period. For example, the increased convenience of televoting with innovations such as the ESC official app. Approvingly, reciprocal relationships are expected to change over time and thus allows the use of a fixed effects model. Every year countries decide how many excess points they will give. This is not set in stone and can depend on many factors

[25] The amount of voting country binary variables and recipient country binary variables depend on the year and edition. The amount mentioned here is the amount of all included countries over the entire period.

which alternate over time; relationships fluctuate. Lastly, clustered standard errors are used to correct for autocorrelated and heteroscedastic errors. It also solves some concerns about lack of normality. This is necessary since variables are typically autocorrelated in panel data.

The four assumptions of the fixed effects regression need to be reviewed in order to conclude that the estimators are consistent and normally distributed (Stock & Watson, 2015). Firstly, the variables need to have distributions which are identical but independent across countries. This can be assumed since sampling is done by including all countries which participated at least three times in the ESC during the chosen period. Secondly, it is necessary to assume large outliers are unlikely. Evidence for this approval is given in Figure 2, which was displayed in the previous data section. Thirdly, the model requires the assumption of no perfect multicollinearity. It can be assumed that there are no issues of collinearity as these would have been automatically removed by the statistical software package STATA. Fourthly, the regression requires the assumption of strict exogeneity. This is however a difficult assumption to accept. The model accounts for many variables such as all observed and unobserved time invariant characteristics. But there might still be omitted variable bias because it is impossible to know which other unobserved variables bias the results. For example, culture is a complex concept which is probably not fully controlled for with Hofstede's dimensions or the other included cultural influences.

In order to test hypothesis 2, an interaction term of the variables excess points received last year and jury is included. This leads to the following model:

$$
\begin{aligned}
excess\ points_{ijt} &= \beta_1 excess\ points\ received\ last\ year_{jit} + \beta_2 jury_{it} \\
&+ \boldsymbol{\beta_3 excess\ points\ received\ last\ year_{jit} * jury_{it}} \\
&+ \beta_n geographical\ influences_{ij} + \beta_n cultural\ influences_{ij} \\
&+ \beta_n performance\ influences_{jt} + \alpha_i + \tau_j + \lambda_t + u_{ijt}
\end{aligned}
$$

Hypothesis 2 states that juries display less reciprocal behaviour compared to the people who vote through televoting. This is reasoned with the experimenter demand effect in combination with the facts that the voting behaviour of professional juries is exposed, and they are hired to vote objectively. In accordance with the experimenter demand effect, the interaction term is expected to display a negative effect of being a jury on reciprocal behaviour. Therefore, β_3 is hypothesized to be negative.

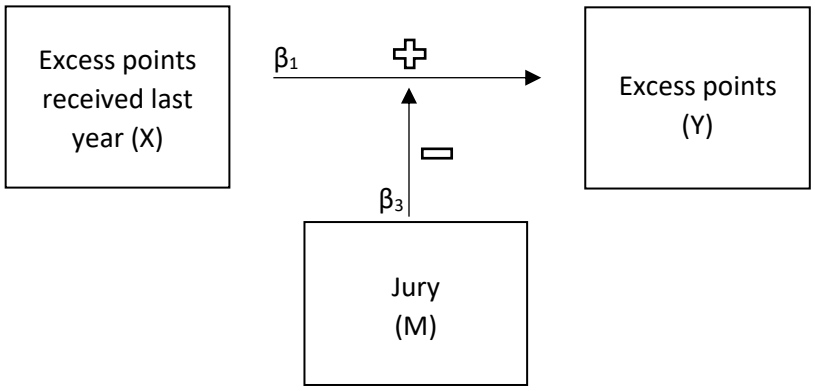


Figure 5: Representation of the effect from jury on reciprocal relationship in the ESC

To reduce potential bias, the model is controlled for 23 variables which are expected to influence voting behaviour and therefore excess points from country i (voting country) to country j (receiving country). The control variables are sorted into the following three categories: geographical influences, cultural influences, and musical performance influences. These control variables help differentiate between reciprocity and other influences. The control variables and their expected effect are described in Appendix 5.

VI. Results

Reciprocal relationship (hypothesis 1)

The results in Table 3 of the following page show that the coefficients of excess points received last year on excess points are positive and significant at a 1% significance level throughout all the models. This confirms hypothesis 1 that receiving excess points last year has a positive effect on giving excess points this year in the ESC. As mentioned before, a positive β_1 represents a reciprocal relationship. The coefficient is 0.275 points in model 1, 0.201 points in model 2, and 0.082 points in model 3, which means that the reciprocal effect diminishes after controlling for geographical influences but diminishes the most by controlling for cultural influences. This suggests that influences seemingly from reciprocity can be explained mostly with culture and partly by geography. The estimate is robust to including musical performance influences as it remains at 0.082 points in the most complete model 4. This coefficient has the following interpretation: this period a country gives 0.082 points in excess, which means on top of what one would expect on average, for every point received in excess from that other country last year. This interpretation and all following interpretations are given under the condition that all other factors remain constant. The return is about one-tenth of the initial generosity. Perceiving as an individual, the return does not seem high since individuals usually expect to get something equivalent to what they gave. In this case it would be a return of one point in excess for every point in excess, especially since every country has the same ability. Also, considering that the amount of points a country can give is scarce, it does not seem like a high return since excess points are valuable. This is conforming to the conclusion from Song (2006) that there is significantly less reciprocal behaviour between groups compared to individuals.

The results are also conforming to what was expected from theories of reciprocity and some of the previous research[26] that reciprocity significantly influences voting behaviour. Besides that, reciprocity is shown to be not only limited to behaviours within one period as it also affects the relationship of behaviour from two separate years. In some way, the results confirm the findings of Ginsburgh and Noury (2008) as well. They concluded apparent vote trading, which looks like reciprocity, is actually driven by quality, language similarities, and cultural similarities. Whereas my results also suggest that seemingly reciprocal relationships are actually based on geography and culture. Regarding the precision of the estimates, it is less precise in model 2 compared to model 1 but becomes more precise again in model 3 and retains the same precision in model 4[27].

Jury vs Televoting (hypothesis 2)

The results of the models with reciprocity, jury, and an interaction are also displayed in Table 3 of the following page. The interaction of jury and excess points received last year signifies the effect of being a jury on reciprocal behaviour. A negative β_3 implies juries are influenced less by reciprocity. The estimates of the interaction term are negative in all models, but they are all not statistically different

[26] For example, Doosje & Haslam (2005) and Mantzaris, Rein & Hopkins (2018) found signs of reciprocal behaviour.
[27] The standard error is 0.019 in model 1, increases to 0.021 in model 2, decreases to 0.016 in model 3 and stays there in model 4.

from zero at a 10% significance level. Therefore, the estimates cannot confirm hypothesis 2 that juries display less reciprocal behaviour compared to the people who vote through televoting. Interestingly though, all the coefficients of excess points received last year increase slightly compared to the models without juries. As the effect of reciprocity is larger when controlling for juries, it suggests that the general public is more influenced by reciprocity. Though, the estimates are less precise as the standard errors are higher. Also, the estimates of jury on excess points are all not statistically significant at a 10% significance level. Thus, together with fact that the interaction term coefficients are not statistically significant, I cannot conclude a difference between voting behaviour from professional juries and the general public. This is neither consistent nor inconsistent with previous research because the conclusions are contradicting (Haan et al., 2005; Ginsburgh & Noury, 2008; Spierdijk & Vellekoop, 2009).

Table 3: *Fixed effects regression estimates of excess points received last year on excess points*

Dependent variable: Excess points				
	(1)	**(2)**	**(3)**	**(4)**
		Controlling for geographical influences	**Controlling for geographical and cultural influences**	**Controlling for geographical, cultural, and musical performance influences**
Models with reciprocity				
Excess points received last year	0.275***	0.201***	0.082***	0.082***
	(0.019)	(0.021)	(0.016)	(0.016)
Models with reciprocity, jury, and interaction				
Excess points received last year	0.321***	0.256***	0.103**	0.104**
	(0.042)	(0.041)	(0.044)	(0.043)
Jury	-0.057	-0.059	0.119	0.117
	(0.067)	(0.067)	(0.121)	(0.121)
Jury*Excess points received last year	-0.050	-0.061	-0.023	-0.023
	(0.043)	(0.040)	(0.043)	(0.042)
Observations	9,349	9,349	4,967	4,967
Control variables				
Geographical influences		X	X	X
Cultural influences			X	X
Musical performance influences				X
Country FE	X	X	X	X
Recipient FE	X	X	X	X
Year FE	X	X	X	X

Notes: Results of linear fixed effects regression for all 49 countries; robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; unbalanced panel between the years 2004 and 2018; the coefficients of the geographical influences, cultural influences, musical performance influences, country specific dummies, recipient specific dummies, year specific dummies, and the constant are left out for clarity; see formulas (1) and (2) for the construction of excess points and formulas (3) and (4) for the construction of excess points received last year; both excess points and excess points received last year are in points; jury*excess points received last year is an interaction term of the dummy variable jury and the continuous variable excess points received last year.

Control variables

Tables 4, 5, and 6 from the following pages display the results from respectively geographical, cultural, and musical performance influences. The statistically significant results are deemed interesting to interpret separately in the ensuing sections.

Geographical influences

Table 4 of the next page displays the estimates of the geographical influences[28]. Firstly, it shows that the positive relationship between neighbours and excess points is statistically significant at a 1% significance level in all the models. The effect decreases slightly as more control variables are included but it is still larger compared to reciprocity. For a certain country, according to the most complete model 4, the excess points increase with 1.291 points if the countries are neighbours compared to when the countries are not neighbours. Thus, the estimates are consistent with the expectation that being a neighbour has a positive effect. Also, the results conform to what was predicted by cultural preferences and conclusions of previous research (Spierdijk & Vellekoop, 2009; Budzinksi & Pannicke, 2017).

Furthermore, Table 4 displays that the positive coefficients of the north clique and the east clique are statistically significant at a respectively 1% and 5% significance level in the most complete model 4. The interpretations of the model 4 estimates are the following: for a certain country, the excess points increase with 2.773 points when the two countries are part of the north clique compared to not being in the north clique together. For a certain country, the excess points increase with 2.468 points when countries are part of the east clique compared to not being in the east clique together. These estimates confirm the expectation that being in the country clique north or country clique east has a positive relationship with excess points, which is also conforming to the earlier research by Mantzaris, Rein, and Hopkins (2018). These effects are also quite economically significant because they are about one fifth on the scale of 12 points. This is higher than the influence of reciprocity. The same cannot be concluded for the southeast clique. The estimate begins statistically significant at a 1% significance level in model 2 but loses the significance when cultural influences are included; in model 3 the effect is not statistically significant at a 10% significance level.

[28] The geographical influences were tested for joint significance and are confirmed to be so at a 5% significance level

Table 4: *Fixed effects regression estimates of geographical influences on excess points*

Dependent variable: Excess points				
	(1)	**(2)**	**(3)**	**(4)**
		Controlling for geographical influences	**Controlling for geographical and cultural influences**	**Controlling for geographical, cultural, and musical performance influences**
Excess points received last year	0.275***	0.201***	0.082***	0.082***
	(0.019)	(0.021)	(0.016)	(0.016)
Neighbours		1.673***	1.274***	1.291***
		(0.312)	(0.301)	(0.301)
North clique		3.182***	2.774***	2.773***
		(0.311)	(0.408)	(0.403)
East clique		2.176***	2.434**	2.468**
		(0.682)	(0.985)	(1.001)
Southeast clique		1.799***	0.984	0.927
		(0.391)	(0.718)	(0.723)
Observations	9,349	9,349	4,967	4,967
Control variables				
Geographical influences		X	X	X
Cultural influences			X	X
Musical performance influences				X
Country FE	X	X	X	X
Recipient FE	X	X	X	X
Year FE	X	X	X	X

Notes: See notes table 3; robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; the estimates for the cultural influences, musical performance influences, country specific dummies, recipient specific dummies, year specific dummies, and the constant are left out for clarity.

Cultural influences

The results in the ensuing Table 5 show that the coefficients of uncertainty avoidance, language, religion, Turkish population, and economic power are statistically significant at a 1% or 10% significance level. Model 4 is the most complete model and is therefore used for the interpretations.

Remarkably, distance in uncertainty avoidance is the only cultural dimension from Hofstede[29] which is statistically significant at either a 1%, 5%, or 10% significance level. This suggests that, concerning the cultural dimensions, excess points are only influenced by whether countries have a similar tolerance for unpredictability. For a certain country, the excess points decrease with 0.022 points as the distance in uncertainty avoidance increases with one on Hofstede's cultural scale. It is statistically significant at a 1% significance level. Furthermore, the excess points increase with 0.825 points for a certain country when the countries share an official language compared to not sharing an official language. This is statistically significant at a 10% significance level. Religion has an estimated coefficient with a larger economic significance. For a certain country, the excess points increase with 2.639 points when the countries share a major religion compared to when the countries do not share a major religion. It is statistically significant at a 1% significance level. The estimated coefficient of Turkish population has an even larger economic significance. For a certain country, the excess points increase with 7.027 points when the countries share a substantial Turkish population compared to when the countries do not share a substantial Turkish population. The economic significance is much larger than the effect of reciprocity (0.082 points in excess) and it is statistically significant at a 1% significance level. All these estimates confirm earlier research[30] and the expectation that sharing similar cultures and patriotism has a positive effect on excess points. Lastly, economic power has a negative relationship with excess points, which is statistically significant at a 10% significance level. Excess points decrease by 0.047 points when the GDP per capita increases by one percent. This confirms the expectation that countries with more economic power are less likely to give excess points, which was based on earlier research (Fiske, 1993; Doosje & Haslam, 2005).

[29] The variables of Hofstede's cultural dimensions were tested for joint significance and are confirmed to be so at a 5% significance level.

[30] The estimate of uncertainty avoidance confirms Spierdijk and Vellekoop (2009) and Budzinksi and Pannicke (2017). The estimate of language confirms Budzinksi and Pannicke (2017). The estimate of religion confirms Spierdijk and Vellekoop (2009) and Budzinksi and Pannicke (2017). The estimate of Turkish population confirms Spierdijk and Vellekoop (2009).

Table 5: *Fixed effects regression estimates of cultural influences on excess points*

Dependent variable: Excess points				
	(1)	**(2)**	**(3)**	**(4)**
		Controlling for geographical influences	**Controlling for geographical and cultural influences**	**Controlling for geographical, cultural, and musical performance influences**
Excess points received last year	0.275***	0.201***	0.082***	0.082***
	(0.019)	(0.021)	(0.016)	(0.016)
Power distance			0.007	0.007
			(0.005)	(0.005)
Individualism-collectivism			-0.006	-0.006
			(0.005)	(0.005)
Masculinity-femininity			-0.005	-0.006
			(0.005)	(0.005)
Uncertainty avoidance			-0.022***	-0.022***
			(0.004)	(0.004)
Long-term orientation			-0.002	-0.001
			(0.007)	(0.007)
Indulgence			0.002	0.002
			(0.005)	(0.005)
Language			0.822*	0.825*
			(0.453)	(0.454)
Religion			2.618***	2.639***
			(0.257)	(0.270)
Turkish population			7.059***	7.027***
			(1.417)	(1.415)
Economic power			-0.497*	-0.468*
			(0.270)	(0.273)
Observations	9,349	9,349	4,967	4,967
Control variables				
Geographical influences		X	X	X
Cultural influences			X	X
Musical performance influences				X
Country FE	X	X	X	X
Recipient FE	X	X	X	X
Year FE	X	X	X	X

Notes: See notes table 3; robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; the estimates for the geographical influences, musical performance influences, country specific dummies, recipient specific dummies, year specific dummies, and the constant are left out for clarity.

Musical performance influences

The musical performance influences, which are displayed in Table 6, do not have any statistically significant coefficients. Therefore, it was regarded uninteresting to interpret separately. Thus, there is no significant voting bias towards certain musical performance characteristics, which corresponds to some earlier research (Spierdijk & Vellekoop, 2009; Budzinksi & Pannicke, 2017).

Table 6: *Fixed effects regression estimates of musical performance influences on excess points*

Dependent variable: Excess points				
	(1)	**(2)**	**(3)**	**(4)**
		Controlling for geographical influences	**Controlling for geographical and cultural influences**	**Controlling for geographical, cultural, and musical performance influences**
Excess points received last year	0.275***	0.201***	0.082***	0.082***
	(0.019)	(0.021)	(0.016)	(0.016)
Automatically qualified				0.258
				(0.706)
Host country				-0.309
				(0.648)
Gender of performer				0.106
				(0.116)
English song				-0.008
				(0.139)
French song				-0.250
				(0.282)
Duet				0.083
				(0.125)
Group				0.068
				(0.127)
Order				0.005
				(0.003)
Opening				-0.225
				(0.190)
Observations	9,349	9,349	4,967	4,967
Control variables				
Geographical influences		X	X	X
Cultural influences			X	X
Musical performance influences				X
Country FE	X	X	X	X
Recipient FE	X	X	X	X
Year FE	X	X	X	X

Notes: See notes table 3; robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; the estimates for the geographical influences, cultural influences, country specific dummies, recipient specific dummies, year specific dummies, and the constant are left out for clarity.

Additional Analysis

Additional analysis is conducted to increase confidence in the results. Firstly, the model is tested with another version of the main independent variable excess points received last year. This checks whether the estimates are sensitive to another time dynamic. The original version created lagged variables of excess points exclusively from the preceding year. The less "restricting" version of the lagged excess points includes lags from the previous year there is data. So, it includes the previous time country i received points from country j regardless if it was last year. This lag is therefore not exclusively from the preceding year but could also be a two-year lag or more. The estimates of this analysis are included in Appendix 7. It appears there are no major differences in the primary interest coefficient β_1. Model 4 from the regression with excess points received last time has a coefficient of 0.094 points which is statistically significant at a 1% significance level. In comparison, model 4 from the regression with the original excess points received last year has a coefficient of 0.082 points which is also statistically significant at a 1% significance level. These two models both include all control variables.

Furthermore, a random-effects model is also tested to check for sensitivity. These results are included in Appendix 8. Model 4 from the random effects regression with excess points received last year has a coefficient of 0.095 points which is also statistically significant at a 1% significance level. According to the Hausman test, the coefficients are not significantly different. This means that the Hausman test results favour random effects. But it was decided to stick with fixed effects because the fixed effects model is still more convincing for the analysis of this data. Not being able to reject does not mean the random effects assumption can be accepted. It is very improbable that, for example, all cultural aspects are captured by Hofstede and that these country-specific effects are uncorrelated with excess points given and received.

Lastly, the main models with reciprocity from Table 3 are also tested for sensitivity to differential effects of professional jury voting. To test this, regressions of the models are analysed using a sample with only jury voting. The coefficients of reciprocity do not change much from the original model[31] and they also all remain significant at the same 1% significance level.

[31] The coefficient of excess points received last year on excess points was 0.271 in model 1, 0.197 in model 2, 0.0777 in model 3, and 0.0775 in model 4. The coefficients of the original results were 0.275 in model 1, 0.201 in model 2, 0.082 in model 3, and 0.082 in model 4.

VII. Discussion

The results have provided two main novel insights regarding reciprocity in a dynamic intergroup context. Firstly, the results provide evidence that reciprocal relationships exist between behaviours from two different years. This suggests that cooperation in repeated intergroup interactions is not only based on what happens within the same period. Therefore, reciprocal relationships involve memory and recognition. This insight could be applied to the many other situations where voting takes place in dynamic group settings such as the United Nations or the European Union. The United Nations for example was also found to have relatively stable voting blocs (Holcombe & Sobel, 1996). In the Security Council or Council of the EU, historical dynamics of voting behaviour should be acknowledged as an influence of cooperation between countries. Secondly, since the return is smaller than would have been expected in reciprocal relationships between individuals, the results confirm that knowledge from interindividual contexts cannot be readily applied to intergroup contexts. Individuals form many relationships including relationships as part of a group. However, they need to be attentive when applying norms of relationships between individuals to their intergroup relationships. A possible explanation could be something comparable to the diffusion of responsibility phenomenon, which in this case hypothesises that individuals reciprocate less because they are aware that others in their country could reciprocate by giving the excess points (Choo, Grimm, Horváth, & Nitta, 2019). However, further research is needed to conclude anything about the reasons.

Furthermore, the result show that reciprocity plays a statistically significant role but there are many other influences in the ESC. Many seemingly reciprocal relationships could be explained with similarities in geography and culture. Also, the effect of reciprocity is smaller than some of the other significant influences such as religion[32]. Excess points can be regarded as a form of affinity between countries. Therefore, the other influences could also provide insights about affinities between countries in situations like the European Union. As distance in uncertainty avoidance is found to be the only statistically significant cultural dimension from Hofstede, it appears that having a similar tolerance for unpredictability is an important cultural factor for affinity. However, care must be taken in assuming external validity since the generalisability of relationships in the ESC is unclear. The groups in the ESC might not representative for groups outside this music contest. A music competition between European countries is a specific situation which has its own specific characteristics. Budzinksi and Pannicke (2017) found biases similar to those from the ESC in a national music contest while Holcombe and Sobel (1996) found stable voting blocs in the UN. But this is not enough to assume the ESC has external validity.

Lastly, the results provide insights for how to create a fairer system in the ESC. It would be ironic and harmful if the relationships between participating countries would worsen due to alliances or biases. As mentioned earlier, the ESC uses a system of draws to decide which countries participate in which semi-final. Historical voting patterns are taken into consideration in order to guarantee fair outcomes. Countries which traditionally give each other high points have a smaller chance of being drawn into the same semi-final. This means that there is less opportunity for reciprocal behaviour, which might explain the relatively small effect of reciprocity. However, no causal claims can be made since the analysis does

[32] The estimated coefficient of religion is 2.639 points in excess while the estimated coefficient of reciprocity is 0.082 points in excess.

not test the effect of this draw system. Though, the results provide inspiration for tests to improve the system. For example, the ESC organizers can test whether considering similarities of religion in the draws of semi-finals decreases structural affinities.

Regarding policy, this analysis does not make any claims about causation of the relationships and therefore restrains from making recommendations. As the control variables show, there are many influences of voting behaviour in the ESC. It is difficult to prove causation by reciprocity because it could be interlinked with many other influences. For example, it is unclear whether the geographical cliques are caused by reciprocity or the other way around. Countries might start reciprocal relationships which leads to a geographical clique or countries give excess points to their neighbours which eventually lead to reciprocal relationships. The results solely provide interesting insights to the relationships. Also, the motivation of behaviour needs to be understood in order to develop good policies. The motivation of reciprocal relationships is complicated and not understood well enough.

Limitations

Despite my efforts there are still reasons why the estimates might be biased and inconsistent. Firstly, there might be limitations in the methodology. The fixed effects model requires the relationship to be the same for all the countries. However, assuming homogeneity in the voting behaviour could be problematic. For example, there are many possible motivations for reciprocity such as a preference for balance. The desire or preference for balance can be universally found but it also depends on characteristics such as culture, personality, and education (Kolm, 2006). Lack of universal attitudes to reciprocity means certain countries might display stronger reciprocal behaviour than others. Therefore, panel data estimations could suffer from heterogeneity bias. Time-series estimates would be more appropriate. Another problem of the fixed effects model is that it might not work well if reciprocal relationships change slowly over time within countries (Stock & Watson, 2015). It is also uncertain whether the strict exogeneity assumption is met. Even after accounting for some unobserved heterogeneity there are still variables which could cause bias. Country, recipient, and time fixed effects cannot control for all omitted variables because it does not control for variables which change over time as well as across countries (Stock & Watson, 2015). Cultural or psychological influences, for example, can change over time and across countries. It is probable that these cultural influences are not captured by Hofstede since the imagined dimensions are not a flawless or complete measurement of culture. Culture is very complex and there are many other ways to describe it (Geert Hofstede, 2019). For example, affection for other countries changes over time and is also different across countries. Adding more time periods does not reduce the inconsistency (Stock & Watson, 2015).

Furthermore, there might be limitations in the data. Firstly, there are many missing values for excess points received last year. This mostly has to do with the way this variable was created but it means that the data could be sensitive to variations. Also, there is for example a bias towards reciprocity if countries which have steady reciprocal relationships have a higher chance of participating in the finals and therefore have more data. Countries might also be discouraged to participate by receiving low points. The lowest amount of participation is three times, but some countries participate in all 15 years. This would cause an upwards bias if countries drop out because they are not able to form reciprocal relationships. Secondly, the quality of a performance is operationalised with a proxy. It is not professed that average assessment is the correct measurement for quality. Controlling for the true quality would better isolate the effect of reciprocity. Thirdly, the dummies for geographical and linguistic influences

might be too simplistic. For example, Budzinksi and Pannicke (2017) also gathered the distance between the country capitals while Spierdijk and Vellekoop (2009) included the lexicostatistical distance between countries. Fourthly, there is another measure for culture called the World Values Survey which could have been used to provide robustness checks. Lastly, it might be problematic to differentiate between professional jury voting and televoting since televoting is only available since 2016.

Recommendations

Limitations offer opportunities for further research. For example, as mentioned before, there are many possible motivations for reciprocity. The mechanisms which are expected to determine reciprocity are complex (Trivers, 1971). This thesis does not analyse how these mechanisms work, but they make interesting extensions. Kolm (2006) mentions three possible motivations, which can be mixed together: liking, balance, and continuation. Liking reciprocity is motivated by the receiver liking the generous giver. Balance reciprocity is the aspiration of balance and sometimes a preference for fairness. Continuation reciprocity is led by the expectation that the behaviour will be rewarded (Kolm, 2006). For example, balance reciprocity is an interesting extension. Kahneman, Knetsch, and Thaler (1986) confirmed through experiments that people are commonly prepared to enforce fairness. Falk and Fischbacher (2006) illustrated that people also take the intention into account when deciding whether behaviour should be rewarded or punished. Falk, Fehr, and Fischbacher (2008) concluded intentions of fairness are important for reciprocal behaviour. Future research could analyse whether reciprocal behaviour between groups is also dependent on intentions. (Kolm, 2006). Further research could also try to analyse whether reciprocity caused the relationships. Gouldner (1960) explains that reciprocity is one of the starting mechanisms of cooperative behaviour because it provides confidence that they will be compensated. Song (2006) suggests that reciprocity is intertwined with trusting behaviour. Whether reciprocity caused these relationships as it increased trust between countries. This is especially interesting as distance in uncertainty avoidance is the only cultural dimension from Hofstede which is a statistically significant. Having a similar tolerance for unpredictability increases the willingness to give excess points, which could relate to trust.

Furthermore, a more complete empirical examination of the relationship can be undertaken. Firstly, the scatterplots of the relationship in Appendix 3 and Appendix 4 display a quadratic reciprocal relationship; the reciprocal relationship appears to increase slowly until excess points of zero and increases quicker as the excess points increase. Further research can be done on the shape of the relationship with a quadratic function. It appears like countries react differently to stinginess and they reward higher excess points more than lower amounts. Lasty, further analysis can be done to make sure the model is measuring reciprocity. The variable Turkish population is significant and has a relatively large effect. This suggest ethnicity plays a big role in voting behaviour of the ESC. Further research can try to isolate the effect of reciprocity from patriotic voting by for example running the model for several pairs of countries with no or little shared ethnicity but still give many excess points to each other.

VIII. Conclusion

In conclusion, the results provide evidence that there is a reciprocal relationship between previous voting behaviour and present voting behaviour in a dynamic group setting. This thesis contributes to literature by concluding that intergroup reciprocal relationships transcend years; countries remember which country gave them how many points in the previous year(s) and they reciprocate this in the present year. Since behaviours from two succeeding years can have a reciprocal relationship, it indicates the involvement of memory and recognition in reciprocity. In a linear regression of excess points received last year on excess points, the estimated coefficient is 0.082 points, which is statistically significant at a 1% significance level. The main hypothesis is accepted, which states that there is a positive relationship between excess points received in the past and excess points given. The positive relationship represents a reciprocal relationship. This means that in this period a country gives 0.082 points in excess[33] for every point they received in excess from that other country last year. The effect of reciprocity is smaller than would be expected by individuals since individuals expect rewards or punishments of equivalent behaviours, especially when they have the same abilities. Reciprocity has been hailed to be an essential influence of voting behaviour and was expected to be so in the ESC. However, in the intergroup context of the ESC, the reward is about one-tenth of the initial generosity. The second hypothesis that professional juries display less reciprocal behaviour compared to people who vote through televoting cannot be accepted; the interaction term of jury and excess points received last year is not statistically significant from zero. Furthermore, reciprocity is complicated and interlinked with many other influences. Most of seemingly reciprocal relationships can be explained with cultural influences and partly by geographical influences. The many possible influences also make it difficult to prove causation. Reciprocity could for example be interlinked with geography in a way that is not clear whether for example reciprocity lead to the northern clique or the other way around. Also, when comparing to the other statistically significant influences, some based on geography and culture have a larger effect on excess points than reciprocity. For example, patriotic voting had the most economically significant effect of 7.027 points on the excess points. It takes more than being generous to win over the voters.

[33] Excess points are points given on top of what would be expected based on the average points given to a musical performance.

IX. References

Ariely, D. (2008). *Predictably Irrational.* New York: Harper Collins.

Arrow, K. (1980). Discrimination in the Labour Market. In J. E. King, Readings in Labour Economics. Oxford: Oxford University Press.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, Reciprocity, and Social History. *Games and Economic Behaviour*, 122-142.

Blangiardo, M., & Baio, G. (2014). Evidence of bias in the Eurovision song contest: modelling the votes using Bayesian hierarchical models. *Journal of Applied Statistics*, 2312-2322.

Bruine de Bruin, W. (2005). Save the last dance for me: unwanted serial position effects in jury evaluations. *Acta Psychologica*, 245-260.

Budzinski, O., & Pannicke, J. (2017). Culturally biased voting in the Eurovision Song Contest: Do national contests differ? *Journal of Cultural Economics*, 343-378.

Cason, T. N., & Mui, V.-L. (2019). Individual versus group choices of repeated game strategies: A strategy method approach. *Games and Economic Behaviour*, 128-145.

Chuan, A., Kessler, J. B., & Milkman, K. L. (2018). Field study of charitable giving reveals that reciprocity decays over time. *Proceedings of the National Academy of Sciences of the United States of America*, 1766-1771.

Choo, L., Grimm, V., Horváth, G., & Nitta, K. (2019). Whistleblowing and diffusion of responsibility: An experiment. *European Economic Review*, 287-301.

Coupe, T., Gergaud, O., & Noury, A. (2018). Biases and Strategic Behaviour in Performance Evaluation: The Case of the FIFA's best soccer player award. *Oxford Bulletin of Economics and Statistics*, 358-379.

Datagraver. (2019, June 26). *Over Datagraver*. Retrieved from Datagraver: https://datagraver.com

Doosje, B., & Haslam, S. A. (2005). What Have They Done for Us Lately? The Dynamics of Reciprocity in Intergroup Contexts. *Journal of Applied Social Psychology*, 508.

European Broadcasting Union. (2019, April 3). *Facts & Figures*. Retrieved from Eurovision Song Contest: https://eurovision.tv

European Broadcasting Union. (2019, May 22). *Rules*. Retrieved from Eurovision Song Contest: https://eurovision.tv

European Union. (2019, March 26). *How EU decisions are made*. Retrieved from European Union: https://europa.eu

Falk, A. (2007). Gift Exchange in the Field. *Econometrica* , 1501-1512.

Falk, A., & Fischbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior*, 293-315.

Falk, A., Fehr, E., & Fischbacher, U. (2005). Driving Forces Behind Informal Sanctions. *Econometrica*, 2017-2030.

Falk, A., Fehr, E., & Fischbacher, U. (2008). Testing theories of fairness - Intentions matter. *Games and Economic Behavior*, 287-303.

Fehr, E., & Gächter, S. (2000). Fairness and retaliation: The economics of reciprocity. *The Journal of Economic Perspectives*, 159-181.

Fehr, E., Kirchsteiger, G., & Riedl, A. (1993). Does Fairness Prevent Market Clearing? An Experimental Investigation. *The Quarterly Journal of Economics*, 437–459.

Fenn, D., Suleman, O., Efstathiou, J., & Johnson, N. F. (2006). How does Europe Make Its Mind Up? Connections, cliques, and compatibility between countries in the Eurovision Song Contest. *Physica A: Statistical Mechanics and its Applications*, 576-598.

Fischbacher, U., & Schudy, S. (2014). Reciprocity and resistance to comprehensive reform. *Public Choice*, 411-428.

Fiske, L. (1993). structures of social life: the four elementary forms of human relations. *American ethnologist*, 193.

Gatherer, D. (2004). Birth of a Meme: The Origin and Evolution of Collusive Voting Patterns in the Eurovision Song Contest. *Journal of Memetics*, 28.

Gatherer, D. (2006). Comparison of Eurovision Song Contest Simulation with Actual Results Reveals Shifting Patterns of Collusive Voting Alliances. *Journal of Artificial Societies and Social Simulation*.

Geert Hofstede. (2019, July 30). *The 6-D model of national culture*. Retrieved from Geert Hofstede: https://geerthofstede.com

Ginsburgh, V., & Noury, A. G. (2008). The Eurovision Song Contest. Is voting political or cultural? *European Journal of Political Economy*, 41-52.

Gouldner, A. W. (1960). The Norm of Reciprocity: A Preliminary Statement. American Sociological Review, 161.

Groot, E. (2019, April 30). *Exclusive: These are the judges who will vote in Eurovision 2019!* Retrieved from Eurovision Song Contest: https://eurovision.tv

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization*, 367-388.

Haan, M. A., Dijkstra, S. G., & Dijkstra, P. T. (2005). Expert Judgment Versus Public Opinion - Evidence from the Eurovision Song Contest. *Journal of Cultural Economics*, 59.

Harari, Y. N. (2015). *Sapiens: A Brief History of Humankind.* London: Vintage.

Hofstede, G. (1980). Culture and Organizations. *International Studies of Management & Organization*, 15-41.

Holcombe, R. G., & Sobel, R. S. (1996). The Stability of International Coalitions in United Nations Voting from 1946 to 1973. *Public Choice*, 17-34.

Independent Commission on Turkey. (2004). *Turkey in Europe: More than a promise?* Brussels: The British Council.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1986). Fairness and the Assumptions of Economics. *The Journal of Business*, 285-300.

Kolm, S.-C. (2006). Introduction to the Economics of Giving, Altruism and Reciprocity. In S.-C. Kolm, & J. M. Ythier, *Handbook of the economics of giving, altruism and reciprocity* (pp. 1-122). London: Elsevier.

Mantzaris, A., Rein, S., & Hopkins, A. (2018). Examining collusion and voting biases between countries during the eurovision song contest since 1957. *Journal of Artificial Societies and Social Simulation*, 1.

Regenwetter, M., Kim, A., Kantor, A., & Ho, M.-H. R. (2007). The Unexpected Empirical Consensus Among Consensus Methods. *Psychological Science*, 629 - 635.

Roth, A. (1995). Bargaining experiments. In J. Kagel, & A. Roth, *Handbook of Experimental Economics* (pp. 253-348). Princeton: Princeton University Press.

Seyfarth, R. M., & Cheney, D. L. (1988). Empirical tests of reciprocity theory: Problems in assessment. *Ethology and Sociobiology*, 181-187.

Sherif, M. (1967). *Group conflict and co-operation: their social psychology.* London: Routledge and Kegan.

Smith, A. (1759). *The theory of moral sentiments.* New York: E. Duyckinck.

Song, F. (2006). Trust and reciprocity in inter-individual versus inter-group interactions: The effects of social influence, group dynamics, and perspective biases. *Experimental Economics*, 179–180.

Spierdijk, L., & Vellekoop, M. (2009). The structure of bias in peer voting systems: lessons from the Eurovision Song Contest. *Empirical Economics*, 403–425.

Stock, J. H., & Watson, M. W. (2015). *Introduction to Econometrics.* Edinburgh: Pearson Education Limited.

Stockemer, D., Blais, A., Kostelka, F., & Chhim, C. (2018). Voting in the Eurovision Song Contest. *Politics*, 428-442.

Tajfel, H. (1981). *Social stereotypes and social groups.* Oxford: Blackwell.

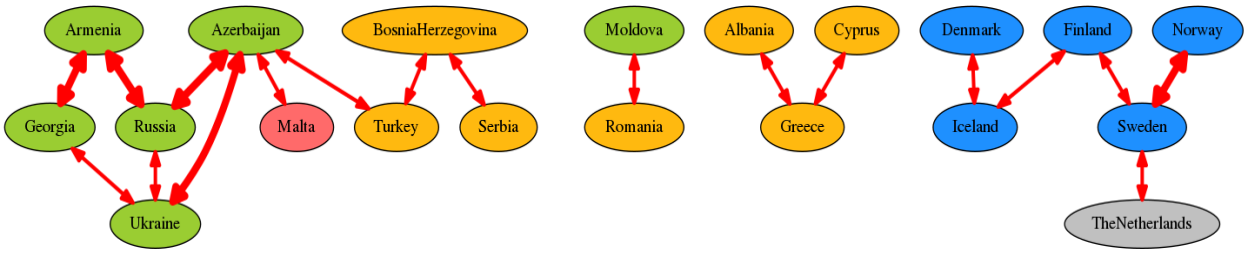
Thaler, R. H. (1988). Anomalies: The ultimatum game. *Journal of Economic Perspectives*, 195-206.

Trivers, R. L. (1971). The Evolution of Reciprocal Altruism. *The Quarterly Review of Biology*, 35-57.

Yair, G. (1995). 'Unite Unite Europe' The political and cultural structures of Europe as reflected in the Eurovision Song Contest. *Social networks*, 147.

Yair, G., & Maman, D. (1996). The Persistent Structure of Hegemony in the Eurovision Song Contest. *Acta Sociologica*, 309-325.

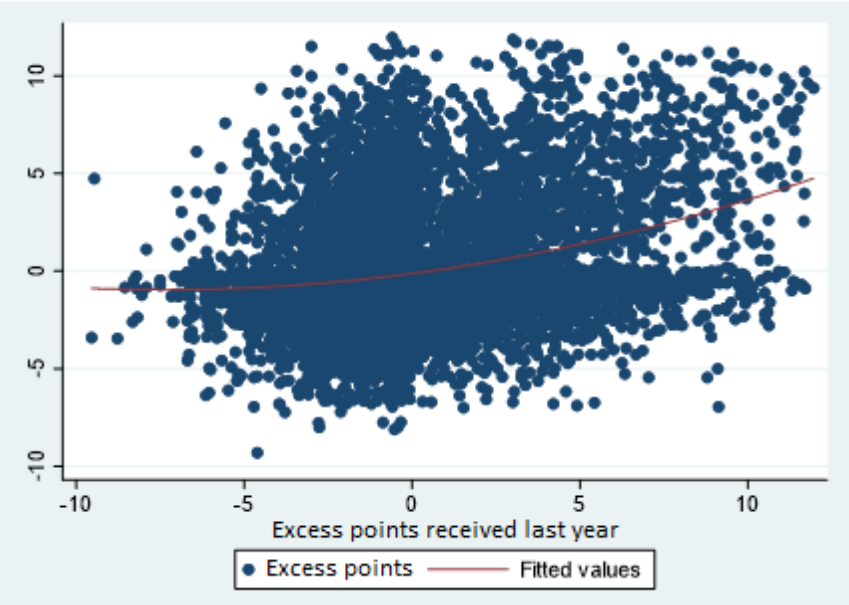
X. Appendix



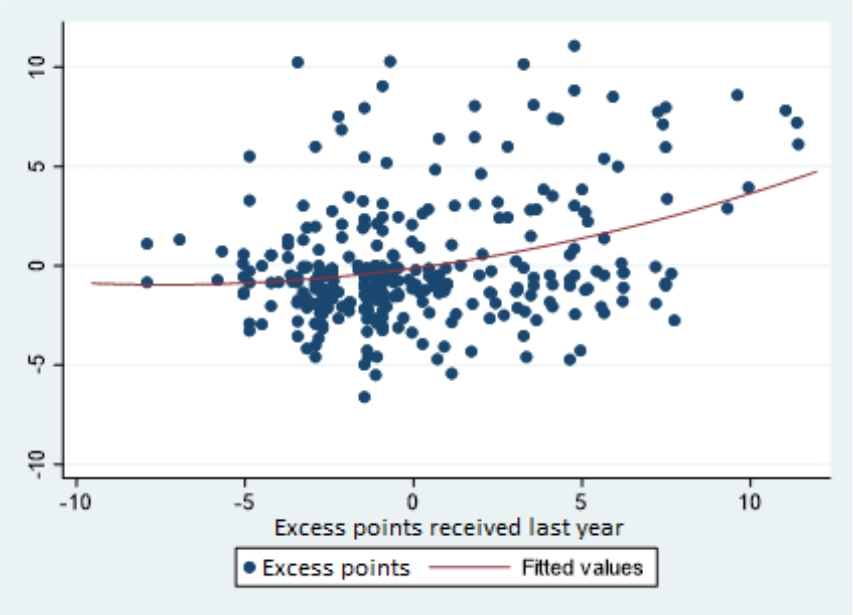
Appendix 1: An illustration of the network of country cliques in the ESC which had statistically significant collusion during the period 2007 till 2017; green signifies the clique east, red signifies the clique southwest, orange signifies the clique southeast, gray signifies central, and blue signifies north.
Source: Mantzaris, Rein, and Hopkins (2018)

Country ID	***Country name***	***Country ID***	***Country name***
1	Albania	26	Italy
2	Andorra	27	Latvia
3	Armenia	28	Lithuania
4	Australia	29	Malta
5	Austria	30	Moldova
6	Azerbaijan	31	Monaco
7	Belarus	32	Montenegro
8	Belgium	33	Norway
9	Bosnia & Herzegovina	34	Poland
10	Bulgaria	35	Portugal
11	Croatia	36	Romania
12	Cyprus	37	Russia
13	Czech Republic	38	San Marino
14	Denmark	39	Serbia
15	Estonia	40	Serbia & Montenegro
16	F.Y.R. Macedonia	41	Slovakia
17	Finland	42	Slovenia
18	France	43	Spain
19	Georgia	44	Sweden
20	Germany	45	Switzerland
21	Greece	46	The Netherlands
22	Hungary	47	Turkey
23	Iceland	48	Ukraine
24	Ireland	49	United Kingdom
25	Israel		

Appendix 2: An overview of all the included countries with their ID; Serbia & Montenegro, Serbia, and Montenegro are considered to be three separate countries.



Appendix 3: Scatterplot from the main independent variable excess points received last year on the dependent variable excess points for all countries with nonlinear fitted line; see formulas (1) and (2) for the construction of excess points and formulas (3) and (4) for the construction of excess points received last year.



Appendix 4: Scatterplot from the main independent variable excess points received last year on the dependent variable excess points for Greece with nonlinear fitted line; see formulas (1) and (2) for the construction of excess points and formulas (3) and (4) for the construction of excess points received last year.

Appendix 5: *Description of the control variables included in the model*

Variable	***Description***	***Expected effect on excess points***	***Reference***
	Geographical influences		
Neighbours$_{ij}$	Dummy variable which is 1 if country i (voting country) and country j (receiving country) are neighbours; the variable is 0 otherwise	Positive due to biases from geographical closeness or cultural preferences It is expected that countries have a stronger preference for countries which are close. Neighbouring countries also often have relatively similar cultures.	Spierdijk and Vellekoop (2009); Budzinksi and Pannicke (2017)
North$_{ij}$	Dummy variable which is 1 if country i (voting country) and country j (receiving country) are part of the north country clique; the variable is 0 otherwise The north country clique consists of Denmark, Iceland, Finland, Sweden, and Norway.	Positive due to collusion or biases from country cliques It is expected that countries have a stronger preference for countries which are a part of their group or clique.	Mantzaris, Rein, and Hopkins (2018)
East$_{ij}$	Dummy variable which is 1 if country i (voting country) and country j (receiving country) are part of the east country clique; the variable is 0 otherwise The east country clique consists of Armenia, Georgia, Russia, Ukraine, and Azerbaijan.	Positive due to collusion or biases from country cliques It is expected that countries have a stronger preference for countries which are a part of their group or clique.	Mantzaris, Rein, and Hopkins (2018)
Southeast$_{ij}$	Dummy variable which is 1 if country i (voting country) and country j (receiving country) are part of the southeast country clique; the variable is 0 otherwise The southeast country clique consists of Bosnia & Herzegovina, Turkey, Serbia, Albania, Cyprus, Greece, and Serbia & Montenegro.	Positive due to collusion or biases from country cliques It is expected that countries have a stronger preference for countries which are a part of their group or clique.	Mantzaris, Rein, and Hopkins (2018)
	Cultural influences		
Power distance$_{ij}$	Continuous variable of the difference in power distance between country i (voting country) and country j (receiving country) *See next page*	Negative due to biases for similar cultures, which is according to cultural preferences *See next page*	Spierdijk and Vellekoop (2009); Budzinksi and Pannicke (2017)

Appendix 5 continued: Description of the control variables included in the model

Power distance$_{ij}$	Power distance is one of Hofstede's cultural dimensions and roughly runs on a scale from 0 to 100. It is how much the less powerful accept and expect an unequal distribution of power (Geert Hofstede, 2019).	It is expected that countries have a stronger preference for cultures which are similar to their own.	
Individualism-collectivism$_{ij}$	Continuous variable of the difference in individualism-collectivism between country i (voting country) and country j (receiving country) Individualism-collectivism is one of Hofstede's cultural dimensions and roughly runs on a scale from 0 to 100. It represents how independent people feel as a contrast to feeling interdependent as a part of a bigger collective. Collectivistic cultures care more about the group they belong to, as opposed to people from individualistic cultures who care more about themselves (Geert Hofstede, 2019).	Negative due to biases for similar cultures, which is according to cultural preferences It is expected that countries have a stronger preference for cultures which are similar to their own.	Spierdijk and Vellekoop (2009); Budzinksi and Pannicke (2017)
Masculinity-femininity$_{ij}$	Continuous variable of the difference in masculinity-femininity between country i (voting country) and country j (receiving country) Masculinity-femininity is one of Hofstede's cultural dimensions and roughly runs on a scale from 0 to 100. It represents how much use of force is socially supported (Geert Hofstede, 2019).	Negative due to biases for similar cultures, which is according to cultural preferences It is expected that countries have a stronger preference for cultures which are similar to their own.	Spierdijk and Vellekoop (2009); Budzinksi and Pannicke (2017)
Uncertainty avoidance$_{ij}$	Continuous variable of the difference in uncertainty avoidance between country i (voting country) and country j (receiving country) Uncertainty avoidance is one of Hofstede's cultural dimensions and roughly runs on a scale from 0 to 100. It is how much the society tolerates uncertainty (Geert Hofstede, 2019).	Negative due to biases for similar cultures, which is according to cultural preferences It is expected that countries have a stronger preference for cultures which are similar to their own.	Spierdijk and Vellekoop (2009); Budzinksi and Pannicke (2017)
Long-term orientation$_{ij}$	Continuous variable of the difference in long-term orientation between country i (voting country) and country j (receiving country) *See next page*	Negative due to biases for similar cultures, which is according to cultural preferences *See next page*	Spierdijk and Vellekoop (2009); Budzinksi and Pannicke (2017)

Appendix 5 continued: Description of the control variables included in the model

Long-term orientation$_{ij}$	Long-term orientation is one of Hofstede's cultural dimensions and roughly runs on a scale from 0 to 100. It represents the extent to which a culture deals with change (Geert Hofstede, 2019).	It is expected that countries have a stronger preference for cultures which are similar to their own.	
Indulgence$_{ij}$	Continuous variable of the difference in indulgence between country i (voting country) and country j (receiving country) Indulgence is one of Hofstede's cultural dimensions and roughly runs on a scale from 0 to 100. It represents how much of a culture is about support of the good things in life (Geert Hofstede, 2019).	Negative due to biases for similar cultures, which is according to cultural preferences It is expected that countries have a stronger preference for cultures which are similar to their own.	Spierdijk and Vellekoop (2009); Budzinksi and Pannicke (2017)
Language$_{ij}$	Dummy variable which is 1 if country i (voting country) and country j (receiving country) share the same official language; the variable is 0 otherwise	Positive due to biases for similar cultures, which is according to cultural preferences It is expected that countries have a stronger preference for cultures which are similar to their own.	Budzinksi and Pannicke (2017)
Religion$_{ij}$	Dummy variable which is 1 if country i (voting country) and country j (receiving country) share at least one major religion; the variable is 0 otherwise A religion is considered to be a major religion either if it has the highest percentage of followers or it has 20% of people following it when it is the second largest religion in that country. The dataset differentiates between Christianity, Islam, Hinduism, Buddhism, Folk religions, and Judaism.	Positive due to biases for similar cultures, which is according to cultural preferences It is expected that countries have a stronger preference for cultures which are similar to their own.	Spierdijk and Vellekoop (2009); Budzinksi and Pannicke (2017)
Turkish population$_{ij}$	Dummy variable which is 1 if country i (voting country) has a substantial population with Turkish immigrant roots and country j (receiving country) is Turkey; the variable is 0 otherwise *See next page*	Positive due to bias from patriotic feelings and cultural preferences It is expected that migrants feel connected to their country of origin and therefore have a stronger preference for that country.	Haan et al. (2005); Ginsburgh and Noury (2008); Spierdijk and Vellekoop (2009)

Appendix 5 continued: Description of the control variables included in the model

Turkish population$_{ij}$	Due to data restrictions Turkish immigrants are used to represent the effect of minorities. The following nine countries are classified as countries with a substantial Turkish immigrant population: Austria, Belgium, Denmark, France, Germany, Netherlands, Sweden, Switzerland, and the United Kingdom.		
Economic power$_{it}$	Natural logarithm of GDP (gross domestic product) per capita in current U.S. dollars from country i (voting country) in year t Natural logarithm is used to increase comprehensibility of the coefficients.	Negative because people with high power are influenced less by the behaviour of others compared to people with low power It is expected that countries who have more economic power are less likely to give excess points.	Fiske (1993); Doosje and Haslam (2005)
Musical performance influences			
Automatically qualified$_{jt}$	Dummy variable which is 1 if country j (receiving country) is automatically qualified for the finals in year t; the variable is 0 otherwise Automatically qualified countries are France, Germany, Spain, the United Kingdom, and the host country until 2011. After 2011, Italy was included in this list (European Broadcasting Union, 2019).	Negative due to countries envying other countries which are automatically qualified without having to perform well	Author's own reasoning
Host$_{jt}$	Dummy variable which is 1 if country j (receiving country) is the host in year t; the variable is 0 otherwise	No hypothesized relationship due to conflicting literature or no statistically significant results	Spierdijk and Vellekoop (2009); Budzinksi and Pannicke (2017)
Gender$_{jt}$	Dummy variable which is 1 if the performer of country j (receiving country) is a male solo singer in year t; the variable is 0 otherwise	No hypothesized relationship due to conflicting literature or no statistically significant results	Spierdijk and Vellekoop (2009); Budzinksi and Pannicke (2017)
English$_{jt}$	Dummy variable which is 1 if the musical performance of country j (receiving country) contains English in year t; the variable is 0 otherwise *See next page*	No hypothesized relationship due to conflicting literature or no statistically significant results	Spierdijk and Vellekoop (2009); Budzinksi and Pannicke (2017)

Appendix 5 continued: Description of the control variables included in the model

$English_{jt}$	A musical performance is regarded as English even if it is partially sung in other languages.		
$French_{jt}$	Dummy variable which is 1 if the musical performance of country j (receiving country) contains French in year t; the variable is 0 otherwise A musical performance is regarded as French even if it is partially sung in other languages.	No hypothesized relationship due to conflicting literature or no statistically significant results	Spierdijk and Vellekoop (2009); Budzinksi and Pannicke (2017)
$Duet_{jt}$	Dummy variable which is 1 if the musical performance of country j (receiving country) is done by a duet in year t; the variable is 0 otherwise	No hypothesized relationship due to conflicting literature or no statistically significant results	Spierdijk and Vellekoop (2009); Budzinksi and Pannicke (2017)
$Group_{jt}$	Dummy variable which is 1 if the musical performance of country j (receiving country) is done by a group in year t; the variable is 0 otherwise	No hypothesized relationship due to conflicting literature or no statistically significant results	Spierdijk and Vellekoop (2009); Budzinksi and Pannicke (2017)
$Order_{jtx}$	Continuous variable for the order of the musical performance from country j (receiving country) in edition x (semi-final 1, semi-final 2 or final) of year t The order ranges from 1 till 28.	No hypothesized relationship due to conflicting literature or no statistically significant results	Haan, Dijkstra, and Dijkstra (2005); de Bruin (2005); Spierdijk and Vellekoop (2009); Budzinksi and Pannicke (2017)
$Opening_{jtx}$	Dummy variable which is 1 if the musical performance of country j (receiving country) is the opening performance in edition x (semi-final 1, semi-final 2 or final) of year t; the variable is 0 otherwise	Positive due to earlier research finding opening musical performances have a higher chance of receiving high points	Haan, Dijkstra, and Dijkstra (2005)

Appendix 6: Summary statistics of the possible explanatory variables for all countries.

All countries	*Number of countries: 49*				
	Observations	***Mean***	***Standard Deviation***	***Min***	***Max***
Geographical influences					
Neighbours	31753	0.079	0.269	0	1
North clique	31753	0.012	0.111	0	1
East clique	31753	0.01	0.102	0	1
Southeast clique	31753	0.014	0.119	0	1
Cultural influences					
Power distance	19659	27.526	19.349	0	93
Individualism-collectivism	19659	22.798	16.585	0	70
Masculinity-femininity	19659	27.33	19.181	0	105
Uncertainty avoidance	19659	26.281	18.916	0	89
Long-term orientation	28744	20.591	14.334	0	65.239
Indulgence	27751	24.204	16.966	0	64.732
Language	31753	0.035	0.184	0	1
Religion	31753	0.846	0.361	0	1
Turkish population	31753	0.003	0.058	0	1
Economic power	28051	27562.32	22953.67	831.2	133000
Musical performance influences					
Automatically qualified	31753	0.125	0.331	0	1
Host country	31753	0.023	0.148	0	1
Gender of performer	31753	0.316	0.465	0	1
English song	31753	0.793	0.405	0	1
French song	31753	0.031	0.173	0	1
Duet	31753	0.1	0.3	0	1
Group	31753	0.19	0.393	0	1
Order	31753	12.023	6.967	1	28
Opening	31753	0.045	0.207	0	1

Appendix 7: Fixed effects regression estimates of the relationship between excess points received last time and excess points

Dependent variable: Excess points				
	(1)	**(2)**	**(3)**	**(4)**
		Controlling for geographical influences	**Controlling for geographical and cultural influences**	**Controlling for geographical, cultural, and musical performance influences**
Excess points received last time	0.281***	0.209***	0.094***	0.094***
	(0.018)	(0.019)	(0.014)	(0.014)
Observations	14,012	14,012	7,039	7,039
Control Variables				
Geographical influences		X	X	X
Cultural influences			X	X
Musical performance influences				X
Country FE	X	X	X	X
Recipient FE	X	X	X	X
Year FE	X	X	X	X

Notes: Results of linear fixed effects regression for all 49 countries; robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; unbalanced panel between the years 2004 and 2018; the coefficients of geographical influences, cultural influences, musical performance influences, country specific dummies, recipient specific dummies, year specific dummies, and the constant are left out for clarity; see formulas (1) and (2) for the construction of excess points; excess points received last time is the excess points country i received from country j, of which the lag is not exclusively from the preceding year; both excess points and excess points received last time are in points.

Appendix 8: Random effects regression estimates of the relationship between excess points received last year and excess points

Dependent variable: Excess points				
	(1)	**(2)**	**(3)**	**(4)**
		Controlling for geographical influences	**Controlling for geographical and cultural influences**	**Controlling for geographical, cultural, and musical performance influences**
Excess points received last year	0.275***	0.206***	0.095***	0.095***
	(0.019)	(0.021)	(0.016)	(0.016)
Observations	9,349	9,349	4,967	4,967
Control Variables				
Geographical influences		X	X	X
Cultural influences			X	X
Musical performance influences				X
Country FE				
Recipient FE	X	X	X	X
Year FE	X	X	X	X

Notes: Results of linear random effects regression for all 49 countries; robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; unbalanced panel between the years 2004 and 2018; the coefficients of geographical influences, cultural influences, musical performance influences, recipient specific dummies, year specific dummies, and the constant are left out for clarity; see formulas (1) and (2) for the construction of excess points and formulas (3) and (4) for the construction of excess points received last year; both excess points and excess points received last year are in points.