Málaga Economic Theory Research Center Working Papers

Conformity, information and truthful voting

Bernardo Moreno, María del Pino Ramos-Sosa and Ismael Rodríguez-Lara

WP 2016-1
February 2016

Departamento de Teoría e Historia Económica
Facultad de Ciencias Económicas y Empresariales
Universidad de Málaga
ISSN 1989-6908

Conformity, information and truthful voting

Bernardo Moreno*
Universidad de Málaga

María del Pino Ramos-Sosa†
Universidad de Málaga

Ismael Rodriguez-Lara‡
Middlesex University London

Abstract

We induce conformity in a binary-decision voting game by assuming that agents may derive some utility by voting the same option that others. Theoretically, we show that truthful voting is the unique equilibrium without conformity. Introducing conformity enlarges the set of equilibria, which includes voting profiles in which agents do not necessarily vote for their preferred option. If agents are informed that others will vote truthfully, truthful voting is more pervasive in equilibrium. In our setting, the effects of conformity and information depend on the voting rule and the preferred option of each agent. We provide empirical support for our theoretical predictions by means of a laboratory experiment.

Keywords: truthful voting, conformity, information, experimental evidence.

JEL classification numbers: C91, C92, D71, D72.
Acknowledgments: We acknowledge financial support from Junta de Andalucía (SEJ-5980 and P09-SEJ-4941REC) and the Spanish Ministry of Science and Technology (ECO2014-53767-P and ECO2014-58297-R).

*Departamento de Teoría e Historia Económica, Facultad de Ciencias Económicas y Empresariales, Campus de El Ejido, 29071 Málaga, Spain. E-mail: bernardo@uma.es.

†Corresponding author: Departamento de Teoría e Historia Económica, Facultad de Ciencias Económicas y Empresariales, Campus de El Ejido, 29071 Málaga, Spain. Phone number: 0034952131298. Fax number: 0034952131299. E-mail: mapi.ramos.sosa@gmail.com.

‡Middlesex University London. Business School. Hendon Campus, The Burroughs, London NW4 4BT. E-mail: I.Rodriguez-Lara@mdx.ac.uk. Website: https://sites.google.com/site/ismaelrl/

1 Introduction

People frequently face binary decisions that require their opinion: board members choosing whether to accept or to reject a proposal (e.g., projects and budgets), senators and congressmen voting whether or not to pass a bill, or citizens voting in a referendum. In these cases, agents might have a clear and strong opinion about what is the best decision to be taken. If they do not pay attention to anything else, one can expect for them to vote truthfully. Truthful behavior is indeed a well rooted concept in social choice theory (see Barberà, 2011 for a complete survey on strategy-proofness social choice functions). Understanding whether agents vote truthfully is of first order importance. This paper proposes an explanation why, under some circumstances, agents may decide not to vote for their preferred option. Their motivation would be the desire to vote for the same option as some other agents. This phenomenon is known as *conformity*, which is defined as the tendency of agents to align attitudes, beliefs and behaviors with some other agents (Myers, 2012). Indeed, our model could be applied to explain current behavior of some members in the US Supreme Court when voting over sensible issues (see Dorff and Brenner (1992) for further empirical evidence on conformity in the US Supreme Court).

On June 26, 2015 the US Supreme Court composed of 9 justices, 5 conservatives and 4 liberals, voted over the fundamental right to marry to same-sex couples nationwide. In the end, the same-sex marriage proposal passed in a 5-4 decision with the support of the liberals and one conservative justice. For this 5-4 configuration of votes, some justices were pivotal as they could have changed the outcome by changing their vote orientation. Oral arguments heard on April 28, 2015 gave clues about the positions of some of the justices. Although conservative Chief Justice John Roberts seemed to be in favor of the law, he voted against it. In the 5-4 decision Chief Justice Roberts was not pivotal. As we argue below (see Proposition 2 in Section 2) if Chief Justice Roberts also paid attention to the votes cast by the other conservatives, voting against the law was not a dominated strategy. That is, when the decision does not depend on his vote, Chief Justice Roberts' vote can be guided by his desire to coincide with the vote of the conservatives.

Our goal in this paper is to study, both theoretically and empirically, how conformity and information affect agents' tendency to vote truthfully. On a theoretical level, we build a model of complete information in which a group of five agents face a binary decision, with one of the options requiring certain support (majority, supermajority or unanimity) to be elected. Agents have a preferred option among two possible ones, A and B. There is a group of three agents that prefer option B and a group of two agents preferring option A. In what follows, we refer to agents preferring option A (B) as type-A agents (type-B agents). We distinguish three scenarios. In the first scenario, agents only care about the adopted decision. Therefore, voting truthfully is the unique undominated Nash equilibrium. In the second scenario, all agents are conformists. They care about whether their preferred option is chosen and whether they vote the same that other agents. Conformist agents' preferences are lexicographical (Moreno and Ramos-Rosa, 2015) in the sense that they always prefer their option to be elected and would conform if they are not pivotal. In this case, agents voting truthfully is an undominated Nash equilibrium and there are also several equilibria in which some group of agents do not vote truthfully. In the third scenario, it is common information that two type-B agents will vote for their preferred option. This resembles the case of external members in a committee that are assumed to vote truthfully.[1] We show that truthful voting remains being an undominated Nash equilibrium and the set of equilibria in which agents do not vote truthfully shrinks.

To empirically investigate these predictions, we conduct a laboratory voting game. Following, among others, Gerber et al. (1998), Morton and Williams (1999), Battaglini et al. (2010) or Bassi et al. (2011), we employ monetary incentives to induce subjects to have preferences over alternatives that correspond to our theory. In our experiment, we consider two different types of subjects. Each type receives the highest possible payoff if the adopted decision coincides with their type. In our baseline treatment, subjects do also receive an small additional payoff, regardless of what they vote. In our treatment with conformity, the additional payoff is received only if the subjects' decision coincides with the decision of any other agent in their group. In a third treatment, a group of two type-B agents is forced to vote truthfully and this is known by all the agents before they vote.[2]

Consistent with our theoretical predictions, we find that agents are more likely to vote truthfully in the baseline treatment, compared with the treatment in which conformity is induced. Our data support also the importance of knowing that other agents are voting truthfully before casting the vote as truthful voting becomes more pervasive. We also provide evidence on the effects of the different voting rules (majority, supermajority and unanimity) on the likelihood of each type of agent voting truthfully. We show that the voting rule does not have any effect on the likelihood of voting truthfully when there is no conformity, as predicted by our theoretical model. In the presence

[1] Rivas and Rodriguez Alvarez (2014) refer to them as objective agents, Buechel et al. (2015) as honest agents and Battaglini et al. (2010) and Moreno and Ramos-Sosa (2015) as independent agents

[2] Note that each of the treatments corresponds to a theoretical scenario.

of conformity, the effect of the voting rule seems to depend on the type of the agent (e.g., we find that type-B agents are less likely to vote truthfully when the voting rule requires more votes for their preferred option to be elected). Finally, our data shed light into the effects of conformity and information into social efficiency. We find that conformity (information) decreases (increases) the likelihood of obtaining the maximum possible total payoff.

Next section describes a revision of the existent literature in conformity. In Section 3, we present our theoretical model and the results for this particular setup. In Section 4, we present the experimental design. We summarize the testable hypotheses in Section 5. Results are discussed in Section 6. Section 7 concludes. Additional material such as the experimental instructions and the non-parametric analysis of our data are relegated to the appendix.

2 Literature Review

Social influence started to be studied with Sherif (1937), Asch (1955) or Milgram (1965, 1974). Asch (1955) asked participants to match a line of a specific length to one of three lines. Although the correct answer was obvious, participants tended to answer wrongly upon observing that others were giving incorrect answers. *Conformity* may occur because subjects want to state an opinion that fits with the group choice (normative influence) or because subjects believe that the group is better informed than they are (informational influence) (see Myers, 2012). In this paper, we consider a voting game in which agents have a clear preference about their preferred option, despite they may want to vote for the same option that others. We then rely our study on the idea of normative influence.

In the literature on voting, Hung and Plott (2001) study *information cascades* which, according to Anderson and Holt (1997), occurs "when initial decisions coincide in a way that it is optimal for each of the subsequent individuals to ignore his or her private signals and follow the established pattern."[3] In Hung and Plott (2001), agents decide in sequence after receiving an informative signal about the correct decision. In one of the treatments, the authors focus on informational influence

[3]This is also known as herd behavior. There is an ample literature studying information cascades and/or herd behavior. For a theoretical study see Banerjee (1992), Bikhchandani et al. (1992) or Dekel and Piccione (2000), among others. See Coleman, (2004), Morton et al. (2012), and Morton et al. (2015) for empirical studies and Hung and Plott (2001), and Morton and Williams (1999, 2000), Goeree and Yariv (2015), for experimental evidence. For other settings in which subjects follow others to comply with a social norm see, for example, Henrich et al. (2001), Bardsley and Sausgruber (2005) or Cooper and Dutcher (2011).

by assuming that agents gain utility if they announce the correct decision. As for the effects of normative influence, Hung and Plott (2001) consider a treatment in which agents receive a larger payoff if their individual decision coincides with the group decision, which is determined by majority rule. In another treatment, the payoff for voting the group decision is larger than the payoff obtained from voting for the correct option (see also Goeree and Yariv (2015) for experimental evidence on the effects of normative and informational influence).

In our model, agents vote simultaneously and do not have any private signal about the correct option. Instead, they have a preferred option and receive the largest payoff only if the group decision coincides with their preferred option. Our agents' preferences are in line with Dutta and Sen (2012), Gerber et al. (1998), Bassi et al. (2011) and Morton et al., (2012), in the sense that agents have preferences defined not only on the outcomes but also on the messages of the rest of the agents. Although agents in our model have also incentives to conform to somebody else, we do not consider that subjects want to vote the same option as the majority. This, in turn, makes our paper divert also from bandwagon behavior, which refers to the desire of voting for the predicted winner in an election (Morton et al. 2015, Morton and Ou 2015).[4] In this literature, agents' preferences are frequently modelled by assuming that agents get a benefit or cost depending on whether they vote with the majority or not (Luzzati, 1999; Hung and Plott, 2001; Michaeli and Spiro, 2015). We relax the assumption in at least two ways. First, we consider that agents are willing to conform only if they cannot change the outcome of the election. Second, we consider that agents may not consider the vote of the majority as a reference point, but simply care about voting the same option than any other member of the group.

Our model is also very well related to the one of Bassi et al. (2011). In this model, agents are told their types, as a way to create them an identity. As in our model, there are some agents that are part of the majority group and others of the minority group. The main difference rest on the fact that Bassi et al. (2011) give incentives to agents to deviate from their assigned types and be on the winning side of an election (bandwagoning voting), while we only penalize agents that vote alone.

Last but not least, one important aspect of our paper is that we investigate the effects of different voting rules (e.g., majority, supermajority or unanimity) on truthful voting. This, in turn, is related to the work of Feddersen and Pesendorfer (1998) or Guarnaschelli et al. (2000), who study the effect

[4] Dorff and Brenner (1992) find evidence for bandwagon voting in the US Supreme Court using data from 1946 to 1975. Morton and Ou (2015) provides an excellent revision on the literature on bandwagon voting.

of different rules on voting behavior in a Condorcet-winner setting.

3 Model

Consider five agents $N = \{1, 2, 3, 4, 5\}$ that have to vote between option A and option B. Agents have their preferences defined over two possible options. Agents $1, 2$ and 3 are *type-B agents* because they prefer option B to option A. Agents 4 and 5 are *type-A agents*, because they prefer option A to option B. The list of types $t = (t_1, t_2, t_3, t_4, t_5) = (B, B, B, A, A)$ is common information, where $t_i \in \{A, B\}$ stands for the type of agent $i \in N$.

Agents vote simultaneously for one of the two options (abstention is not allowed). Let $M_i = \{A, B\}$ be the set of messages for agent $i \in N$, where $m_i = A$ ($m_i = B$) stands for agent i voting for option A (option B), respectively.[5] Let $M = \times_{i \in N} M_i$ be the set of messages and $m_{-i} \in M_{-i} = \times_{j \in N \setminus \{i\}} M_j$ be the messages of all agents except i. The set of all agents except i is denoted by $N_{-i} = N \setminus \{i\}$. We denote the profile of messages by $m \in M$.

The voting rule is such that for option B to be adopted, at least q agents have to vote for it. We refer to $q \in \{3, 4, 5\}$ as the quota. A quota $q = 3$ ($q = 4$) [$q = 5$] implies that option B requires simple majority (supermajority) [unanimity] to be adopted, respectively. We define $q(m)$ as a mapping from M to $\{A, B\}$, where $q(m) = B$ if q or more agents vote for option B. Otherwise, $q(m) = A$.

Agents' preferences are defined over alternatives that consists of the adopted decision and the profile of messages announced by the agents. Agents' preferences can then be written as follows:

$$u(t_i, q, m) = v(t_i, q(m)) + c(m)$$

where

$$v(t_i, q(m)) = \begin{cases} v & \text{if } q(m) = t_i \\ 0 & \text{if } q(m) \neq t_i \end{cases}$$

$$c(m) = \begin{cases} c & \text{if } m_i = m_j \text{, for any } j \in N_{-i} \\ 0 & \text{if } m_i \neq m_j \text{, for any } j \in N_{-i} \end{cases} .$$

[5] In our proofs, we also use m_i^A (m_i^B) to denote that agent i vote for option A (B), respectively.

The first term $v(t_i, q(m))$ indicates what agents receive if the adopted decision coincides with their type. Thus, if option B receives the required support, agents $1, 2$ and 3 receive a payoff of v whereas agents 4 and 5 receive a payoff of 0. Agents may also receive an additional payoff $c(m) = c$ if their message coincide with the one of any other agent $j \in N_{-i}$. If $m_i = m_j$ we say that agent i conforms to agent $j \in N_{-i}$. Otherwise, $c(m) = 0$. We say that there is conformity if $c(m)$ is as described above.

We assume that $v > c > 0$ to catch the idea that agents do not place their will to follow the rest of the agents before obtaining their preferred option. That is, agent i will conform only if her message is irrelevant to the adopted decision; i.e., preferences to conform are lexicographical.[6]

The aim of this paper is to investigate the effects of introducing conformity on the voting behavior of the agents. In particular, we want to study to what extent truthful behavior is affected. We hereafter refer to agent $i \in N$ voting truthfully if she votes for her own type, that is $m_i = t_i$. We also want to study voting behavior when it is common information that some agents will always vote for their type.

Next, we define the equilibrium concept that we use through our paper.

Definition. For any $q \in \{3, 4, 5\}$, $m \in M$ is a *weakly undominated Nash equilibrium* (WUNE) if for any $i \in N$,
(1) m_i is not weakly dominated and
(2) for any $m'_i \in M_i$, $u_i(t_i, q, m_i, m_{-i}) \geq u_i(t_i, q, m'_i, m_{-i})$.

In our baseline scenario, we assume that all agents receive $c(m) = c$ regardless of the profile of messages $m \in M$. Then, there is no conformity.

Proposition 1. In the baseline scenario, all agents voting truthfully is the unique weakly undominated Nash equilibrium, that is,
i) type-A agents vote for A for any $q = \{3, 4, 5\}$
ii) type-B agents vote for B for any $q = \{3, 4, 5\}$.

Proof. If there is no conformity, $u_i(t_i, q(m)) = v(t_i, q(m)) + c$. For any $m_{-i} \in M_{-i} = \times_{j \in N \setminus \{i\}} M_j$ such that the vote of agent $i \in N$ does not determine the adopted decision, $u_i(t_i, q, m_i, m_{-i}) =$

[6] Admittedly, this is not the unique way of defining conformity. We could have opted for requiring an agent to conform to some subset of N_{-i}. See Moreno and Ramos-Sosa (2015) for a discussion.

$u_i(t_i, q, m_i', m_{-i})$ where $m_i = t_i$ and $m_i' \neq t_i$. For any $m_{-i} \in M_{-i} = \times_{j \in N \setminus \{i\}} M_j$ such that the vote of agent $i \in N$ determines the adopted decision, $v + c = u_i(t_i, q, m_i, m_{-i}) > u_i(t_i, q, m_i', m_{-i}) = c$ where $m_i = t_i$ and $m_i' \neq t_i$. $\square$

Proposition 1 shows that all agents voting truthfully is the unique WUNE for any quota $q \in \{3, 4, 5\}$.

In our conformity scenario, we assume that $c(m) = c$ when $m_i = m_j$ for some $j \in N_{-i}$. Then, agents' payoffs depend on the adopted decision and how their messages relate to other agents' messages.

Proposition 2. In the conformity scenario, there are weakly undominated Nash equilibria in which:
i) type-A agents vote for A and for B for any $q \in \{3, 4\}$,
ii) type-B agents vote for A and for B for any $q = \{3, 4, 5\}$.

Proof. We show that there are WUNE in which agents do not vote truthfully. Take $q = 3$ or $q = 4$ and consider the cases in which all agents vote unanimously for one of the options. In these cases, agents get $v + c$ or c, depending on whether their preferred option is being elected or not. Any agent switching her message and voting alone would be worst off, because the adopted decision will not change and $c(m) = 0$ (i.e., the agent that switches gets v or 0, depending on whether her preferred option is being elected). When $q = 5$, these arguments apply for any type-B agent as well. However, all agents voting for B will not be an equilibrium profile as any type-A agent (say agent 5) can deviate and vote truthfully to change the elected decision, getting $v > c$. Indeed, $u_5(A, q, m_5^A, m_{-5}) > u_5(A, q, m_5^B, m_{-5})$ for any $m_{-5} \in M_{-5}$, where $u_5(A, q, m_5^A, m_{-5}) = v$ if $m_j = B$ for any $j \in N \setminus \{5\}$, and $u_5(A, q, m_5^A, m_{-5}) = v + c$ otherwise. $\square$

Proposition 2 shows that in the conformity scenario, voting truthfully remains as an equilibrium strategy for all the agents. Besides, there are also other equilibria where agents vote for an option that does not coincide with their own type.[7] This is true for all agents and quotas, except for type-A agents when $q = 5$. Type-A agents vote truthfully in that case because their decision determines the adopted decision. Interestingly, for quota $q = 3$, all agents voting for option B (option A)

[7]Note that the arguments in Proposition 1 can be applied to show that voting truthfully is also a WUNE in this setting. There are equilibrium situations in which the adopted decision does not coincide with the decision obtained if all agents vote truthfully. See Moreno and Ramos-Sosa, (2015) for a discussion.

is an equilibrium. Since no agent is pivotal, type-A (type-B) agents voting truthfully would be penalized.[8] Hence, type-A (type-B) agents are conforming to type-B (type-A) agents.

Finally, we study voting behavior when it is common information that some agents will always vote for their type and conformity is still induced in the preferences of the agents. This will be our informational scenario in which agents 1 and 2 vote for their preferred option. We then show that for the type-B agent voting truthfully is a weakly dominant strategy but type-A agents' may not vote truthfully.[9]

Proposition 3. In the informational scenario, there are weakly undominated Nash equilibria in which:

i) type-A agents vote for A and B for any $q \in \{3, 4\}$ and

ii) type-B agents vote for B for any $q \in \{3, 4, 5\}$.

Proof. We show that there are WUNE in which type-A agents do not vote truthfully.[10] Note that all agents voting unanimously for A and only one agent voting for B is dismissed as agents 1 and 2 are forced to vote for B.

Take $q = 3$ or $q = 4$ and consider the case in which all agents vote unanimously for B. In this case, this is a WUNE since any type-A agent (say agent 4) voting truthfully is worst off as $u_4(A, q, m_4^B, m_{-4}^B) = c > 0 = u_4(A, q, m_4^A, m_{-4}^B)$. Note that, in this case, type-A agents vote for B. The remaining type-B agent (agent 3) in this case votes truthfully as $u_3(B, q, m_3^B, m_{-3}^B) = v + c > v = u_3(B, q, m_3^A, m_{-3}^B)$. Besides, when $q = 3$, the profile of messages where only two agents vote for B is not an equilibrium due to agent 3 voting truthfully can make the decision be B when voting for B. When $q = 4$, agent 3 also votes truthfully for any $m_{-3} \in M_{-3} = \times_{j \in N \setminus \{3\}} M_j$ since $u_3(B, q, m_3^B, m_{-3}) \geq u_3(B, q, m_3^A, m_{-3})$. Therefore, two agents voting for B not being 1 and 2 are

[8]This is much in line with the comment made by Charles Plott to the authors in the sense that "the votes under majority-rule institutions are typically overwhelming majorities because the minority, when anticipating a loss on the vote, just go along with the majority."

[9]Although the theoretical result also holds when only one type-B agent votes truthfully, we force two agents for experimental purposes. By forcing two agents to vote truthfully, the payoff for conformity $c > 0$ is always guaranteed. Moreno and Ramos-Sosa (2015) provides the number of forced agents (that they refer as independent agents) within the voting group guaranteeing that the adopted decision always coincides with the decision obtained when all agents vote truthfully. Their results hold for any quota $q \in \{1, n\}$, any number of agents n and any list of agents' types t. In our setting, two forced agents guarantee that the adopted decision will coincide with the decision under truthful voting for $q = 3$ and $q = 5$.

[10]Following Proposition 1, voting truthfully is a WUNE for each possible $q \in \{3, 4, 5\}$.

profiles of strategies that are not possible and agent 3 always votes truthfully.

Take $q = 5$. Consider the case in which all agents vote unanimously for B. In this case, this is not a WUNE since type-A agent (say agent 4) voting truthfully makes the decision be A. That is, she is strictly better off as $u_4(A, q, m_4^A, m_{-4}^B) = v > c = u_4(A, q, m_4^B, m_{-4}^B)$. Besides, agent 3 voting for A is not a WUNE since $u_3(B, q, m_3^B, m_{-3}) \geq u_3(B, q, m_3^A, m_{-3})$ for any $m_{-3} \in M_{-3} = \times_{j \in N \setminus \{3\}} M_j$. Therefore, for $q = 5$, the unique WUNE is all agents voting truthfully. □

Compared with the conformity scenario, in the informational scenario the equilibrium strategies reduce to only two: all agents voting truthfully and all agents unanimously voting for B (except for $q = 5$). Type-A agents, however, are not affected in that voting for any option is an equilibrium strategy when $q = 3$ or $q = 4$, but voting truthfully is a dominant strategy when $q = 5$.

In the next section we present an experiment that is aimed to test our theoretical predictions.

4 Experimental design

A total of 390 subjects were recruited to participate in our computerized sessions (Fischbacher, 2007). Subjects were Economics or Business students from the undergraduate population of the Universidad de Valencia, with no previous experience in similar experiments. Subjects were invited to participate in our experiment using the recruitment system of the laboratory (LINEEX).

We run a total of 13 sessions with 30 subjects each. At the beginning of each session, subjects were randomly assigned a type (Player A or Player B), which was kept constant through the session. Subjects were told that they were in a group of 5 subjects. It was common information that each group consisted of 2 Players A and 3 Players B.

When subjects were informed about their types, they were asked to vote between option A and option B across three different rounds. In each of the rounds, option B required a different number of votes $q = \{3, 4, 5\}$ to be the adopted decision. It was common information that if option B did not get at least q votes then option A would be the adopted decision. In our experiment, q was announced to subjects at the beginning of each round. The order of q was randomly selected and we balance the number of observations across sequence of decisions; e.g., we have the same number of observations for sequence $3, 4, 5$ and $5, 3, 4$. As for the feedback, subjects voted in each q receiving no information whatsoever about previous decisions in their groups. Thus, subjects voted without knowing what other subjects in their group had voted for or what was the adopted decision in previous rounds. This, in turn, implies that our method for eliciting the relevant behavior is free of historical contagion and learning effects.

At the end of the experiment, one round was randomly selected for payment. Payoffs for each subject depended on whether or not *i*) the adopted decision in that round coincided with their own type, and *ii*) they voted for the same option that any other subject in their group.

Our experiment relies on a between-subject design (i.e., subjects do only participate in one of the three possible treatments). We summarize these treatments below:

- **Baseline** (BL, 120 subjects, 24 groups). Subjects received 75 ECUs if the adopted decision coincided with their own type. Subjects received an additional amount of 25 ECUs, regardless of the option they decided to vote for.

- **Conformity** (CON, 120 subjects, 24 groups). Subjects received 75 ECUs if the adopted decision coincided with their own type. They received the additional payoff of 25 ECUs only if their vote coincided with the one of any other subject in their group.

- **Informational** (INF, 150 subjects, 30 groups). Payoffs were as in the conformity treatment but it was common information that 2 subjects in the role of Player B would be given no option but to vote for option B. The subjects forced to vote for B were the same through the whole session. It was common information that they had to vote for option B in each round, as this was the only option that appeared in their computer screen.[11]

Instructions in each treatment were read aloud by the session monitor and subjects were allowed to ask any question in private before starting the treatment. We minimized the probability of subjects missing how payoffs were generated with a pre-experimental quiz, in which subjects were asked to compute the payoffs of randomly generated examples.

Each session lasted around 1 hour and subjects received approximately 7.5 Euros for participating (10 ECUs = 1 Euro). The experiment included an additional phase in which individual characteristics were elicited. Our questionnaire, together with a translated version of our original instructions, is presented in Appendix A. We note that the questionnaire includes gender, age, cog-

[11]We decided to have two subjects instead of the computer in the role of forced agents to avoid any concern about social preferences; e.g., type-A agents can vote differently depending on whether they impose an externality on the computer or on other human subject. For the effects of social preferences on bandwagon voting see Morton and Ou (2015) or Corazzini and Greiner (2007).

nitive abilities, risk aversion, social preferences, trust, happiness, satisfaction and inequality. We shall use these variables as controls in our econometric analysis in Section 6.

5 Testable Hypotheses

Our first treatment (BL) resembles our theoretical model for the case in which $c(m) = c$ for any $m \in M$, therefore subjects are expected to vote truthfully in this treatment (Proposition 1). We induce conformity in our second treatment (CON) by paying subjects the additional amount of money only if they vote for the same option that any other member in their group. Given our theoretical results (Proposition 2), we predict for truthful voting to be less pervasive in this treatment. The first hypothesis that we want to reject is then as follows:

Hypothesis 1. *The presence of conformity will not affect the likelihood of voting truthfully, which will be the same in BL and CON.*

We expect that conformity will affect the likelihood of truthful voting for both type of agents, except when option B requires unanimity. Proposition 2 highlights that type-A agents vote truthfully in that case, therefore we do not expect any effect of conformity if $q = 5$.

In our third treatment (INF), it is common knowledge that two type-B agents are forced to vote for B. Our theoretical results display that voting truthfully is the unique equilibrium strategy for the other type-B voter (Proposition 3). As a result, we expect for type-B agents to vote more truthfully in INF, compared with CON. The hypothesis that we want to reject is then as follows:

Hypothesis 2. *In a conformity setting, the presence of two type-B agents agents voting for their types will not affect the likelihood of truthful voting of the other type-B agent, which will be the same in CON and INF.*

Proposition 2 shows that type-A agents can vote for A or B when $q = \{3, 4\}$, whereas voting truthfully is the unique WUNE for them when $q = 5$. The same is true when forced agents are included, therefore we do not expect any positive effect of knowing the presence of forced agents on the likelihood of type-A agents on voting truthfully.

Although we focus our attention on the effects of conformity and the presence of forced agents on voting truthfully, we have some testable predictions for the effect of the quota as well. We have shown that voting truthfully is the unique WUNE in our BL treatment, therefore the quota should not affect the agents' behavior in this treatment. This argument is also valid for type-B agents in the INF treatment. We expect to observe the following behavioral pattern in our data.

Conjecture. *The quota does not have any effect on the agents' likelihood of voting truthfully in BL. The same is true for the likelihood of type-B agents voting truthfully in INF.*

We note that in the rest of the cases (e.g., in the CON treatment), voting truthfully is an equilibrium for any possible quota $q = \{3, 4, 5\}$, therefore there is no clear-cut prediction for the effect of the quota on the likelihood of voting truthfully. Arguably, voting truthfully is the unique equilibrium for type-A agents if unanimity is required, therefore voting truthfully may be more likely for them if $q = 5$, compared with $q = 3$ or $q = 4$. By the same token, type-B agents can anticipate that increasing the quota makes it more difficult for option B to be chosen. Type-B agents can then be less likely to vote truthfully when the quota increases, as a result.

6 Results

This section presents our experimental evidence. Section 6.1 focuses on the first two treatments (BL vs. CON) to show the effects of conformity on the likelihood of voting truthfully. In Section 6.2 we assess whether knowing that two agents are voting truthfully influences on agents' voting behavior in a conformity setting (CON vs. INF). To provide some insights into efficiency, we evaluate our findings in terms of total surplus in Section 6.3.

6.1 On the effects of conformity

Figure 1 shows the effect of conformity by plotting the average likelihood of voting truthfully in BL and CON, for each type of agent separately. Error bars reflect standard error of the mean. The results disaggregated by quota are presented in Table 1 below the figure.

In line with our theoretical prediction, we observe that the presence of conformity decreases the average frequency of voting truthfully for type-A (from 0.75 to 0.60) and type-B agents (from 0.87 to 0.75). As it is shown in Table 1, conformity decreases the likelihood of type-A agents voting truthfully for any possible quota, with the smallest effect in $q = 5$. As for type-B agents, conformity decreases the likelihood of truth-telling for each possible quota, except when $q = 3$.[12]

[12] A non-parametric analysis suggests that these effects are statistically significant (see Appendix B).

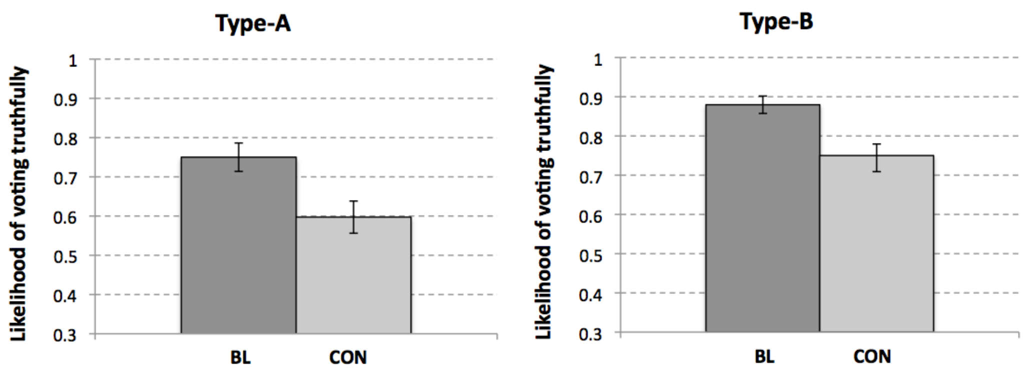

Figure 1: Effect of conformity on the likelihood of voting truthfully.

Table 1: Frequency of truthful voting in BL and CON for each possible quota.

	Type-A		**Type-B**	
	BL	**CON**	**BL**	**CON**
$q = 3$	0.71	0.46	0.90	0.94
$q = 4$	0.85	0.71	0.90	0.74
$q = 5$	0.69	0.62	0.83	0.57
N	48	48	72	72

Table 1 allows us to observe also the effect of the quota on the likelihood of voting truthfully. When there is no conformity, Proposition 1 shows that there is a unique WUNE in which all agents vote truthfully regardless of the quota. Our findings seem to support this prediction by suggesting no effect of the quota on the likelihood of voting truthfully in the BL treatment. The quota, however, seems to have a significant effect on the likelihood of voting truthfully in the CON treatment.

One final aspect that is worth mentioning is that type-B agents are more likely to vote truthfully than type-A agents. This occurs in both treatments for any possible quota. This is an interesting finding because type-B agents are in a majority group. If decisions are taken by majority rule, there may be a bias for type-A agents not to vote truthfully, as already suggested in Bassi et al. (2011). Arguably, any type-A agent can ensure option A to be elected when $q = 5$ if they vote truthfully.

In what follows, we perform an econometric analysis to study the behavior of type-A and type-B agents in more detail. For each type of agent, we estimate a logit model for the likelihood of voting truthfully. The set of independent variables includes a dummy variable that takes the value of 1 if there is conformity (d_{CON}), and two dummy variables for the value of the quotas ($d_{q4} = 1$ if $q = 4$ and $d_{q5} = 1$ if $q = 5$). Our specification includes also the interaction terms between the treatment and the quotas ($d_{CON}d_{q4}$ and $d_{CON}d_{q5}$) to capture the (possibly different) effects of conformity on the likelihood of voting truthfully depending on the quota. We control for individual heterogeneity by including the answers to the questionnaire as independent variables.

Figure 2 depicts the average marginal effect (ME) of conformity, together with the 95% confidence intervals, for each possible quota.[13] The results for type-A agents (type-B agents) are presented in the left-panel (right-panel), separately. The reported effects take into account Ai and Norton (2003) and Karaca−Mandic et al. (2012).[14]

We observe that conformity decreases the likelihood of type-A agents voting truthfully when $q = 3$ (ME $= -0.222$, $p = 0.017$) and $q = 4$ (ME $= -0.126$, $p = 0.003$). Consistent with our theoretical results, we do not observe any effect of conformity on the likelihood of type-A agents voting truthfully when $q = 5$ (ME $= -0.036$, $p = 0.703$). As for type-B agents, we cannot reject the null hypothesis of conformity having no effect on the likelihood of voting truthfully when $q = 3$

[13] Appendix B contains the estimates of our logit regressions. As we discuss in Appendix B, our findings are robust if we consider a linear probability model instead.

[14] Ai and Norton (2003) and Karaca−Mandic et al. (2012) discuss the correct way of estimating marginal effects in nonlinear models that include interaction terms. Hereafter, all the marginal effects reported in our paper take their work into account.

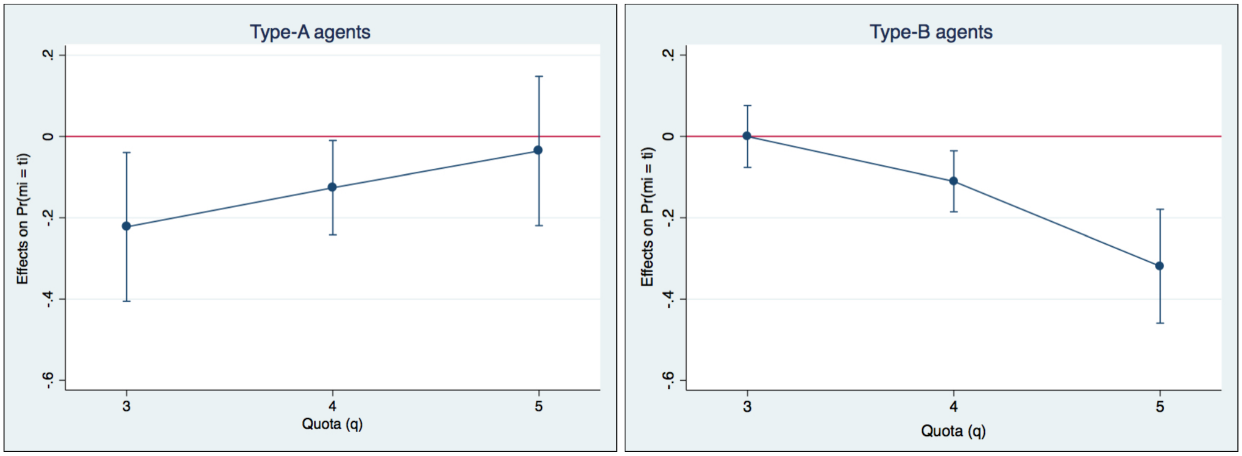

Figure 2: Marginal effects of conformity on the likelihood of voting truthfully after logit specification. The reported effects control for individual heterogeneity and take into account Ai and Norton (2003) and Karaca−Mandic et al. (2012).

(ME $= -0.0003$, $p = 0.993$). This null hypothesis is rejected at any common significance level when $q = 4$ (ME $= -0.11$, $p = 0.004$) or $q = 5$ (ME $= -0.319$, $p < 0.001$).[15] We conclude that these findings provide evidence against our first hypothesis.

Observation 1. *Conformity decreases the likelihood of voting truthfully.*

Next, we study the effects of the quota on the likelihood of agents voting truthfully in BL and CON by means of a logit analysis, where the explanatory variables are the dummies for the quotas (d_{q4} and d_{q5}). The estimated marginal effects are summarized in Table 2.

We support for our conjecture suggesting no effect of the quota on the likelihood of voting truthfully in the BL treatment. As for the CON treatment, there is no-clear cut prediction regarding how subjects should vote for each quota. In that case, our findings suggest that type-A agents are more likely to vote truthfully as the quota increases, whereas type-B agents seem to vote less truthfully as the quota increases.

Observation 2. *i) The quota has no effect on the likelihood of voting truthfully when there is no conformity. ii) Under conformity, the quota has an effect on the likelihood of voting truthfully. The effect seems to depend on the type of the agent.*

[15] Our results indicate that conformity reduces the likelihood of type-A agents voting truthfully by roughly 13%. The estimates for type-B agents suggest a reduction in the likelihood of voting truthfully by 14%. Statistically, both effects are significant at the 1% level.

Table 2: Marginal effects of the quota on the likelihood of voting truthfully in BL and CON.

	Type-A		Type-B	
	BL	CON	BL	CON
($d_{q4} = 1$ if $q = 4$)	0.149	0.25***	0.000	-0.208***
	(0.09)	(0.10)	(0.05)	(0.06)
($d_{q5} = 1$ if $q = 5$)	-0.021	0.167	-0.069	-0.375***
	(0.10)	(0.10)	(0.06)	(0.06)
Controls	Yes	Yes	Yes	Yes
Wald χ^2-test	25.73**	19.55*	19.43	29.19***
Pseudo- R^2	0.09	0.08	0.08	0.17
Observations	144	144	216	216

Overall, these findings highlight the importance of conformity on the likelihood of voting truthfully. Next, we investigate the influence in voting truthfully when agents know that there are two forced agents voting for their preferred option.

6.2 On the effects of information

We replicate our previous analysis and show the effect on voting when agents know that some agents are voting truthfully. Figure 3 and Table 3 report in aggregate levels the likelihood of voting truthfully in INF for each type of agent. For the sake of comparisons, we include the results for CON. Importantly, two type-B agents were forced to vote for their preferred in our INF treatment. The reported data for the INF treatment do not consider these agents; i.e., it only focusses on the type-B agents who were not forced to vote truthfully but allowed to vote for any of the two options. Error bars reflect standard errors of the mean.

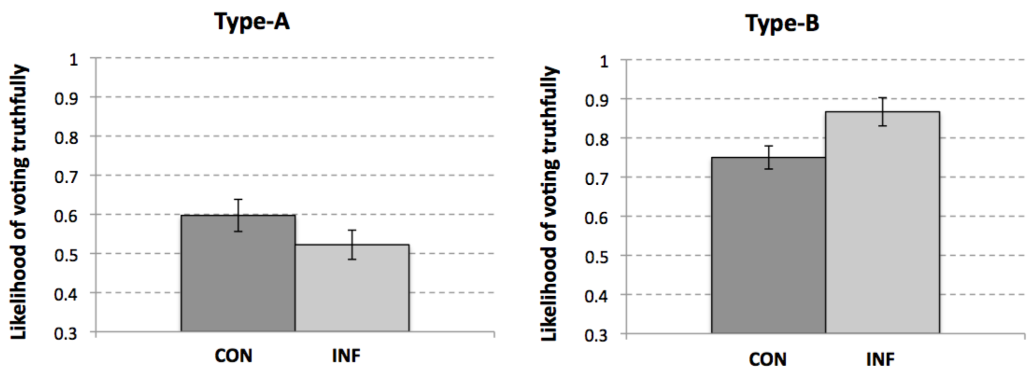

Figure 3: Effect of information on the likelihood of voting truthfully.

Table 3: Frequency of truthful behavior in CON and INF for each possible quota.

	Type-A		**Type-B**	
	CON	**INF**	**CON**	**INF**
$q = 3$	0.46	0.35	0.94	0.97
$q = 4$	0.71	0.55	0.74	0.97
$q = 5$	0.62	0.67	0.57	0.67
N	48	30	72	60

We observe that the presence of forced agents increases the likelihood of type-B agents on voting truthfully (from 0.75 to 0.86), which is in line with our theoretical prediction. The results seem to be consistent for each possible quota. As for type-A agents, the presence of forced agents seems to decrease their likelihood of voting truthfully (from 0.60 to 0.52). This negative effect occurs for each possible quota except when $q = 5$, which is the only quota in which voting truthfully is the unique WUNE strategy for type-A agents in both treatments.[16]

Next, we perform our econometric analysis. We estimate a logit specification for the probability of each type of agent on voting truthfully. We control for individual heterogeneity and include a dummy variable that takes the value of 1 if there are forced agents (d_{INF}), and two dummies for the quotas (d_{q4} and d_{q5}). We allow for different effects of the presence of forced agents depending on the quotas by including the interaction terms. The logit estimates, together with the results of

[16]See Appendix B for the results of the non-parametric analysis.

a linear probability model, are presented in Appendix B. The estimated marginal effects and the corresponding 95% confidence interval are summarized in Figure 4.

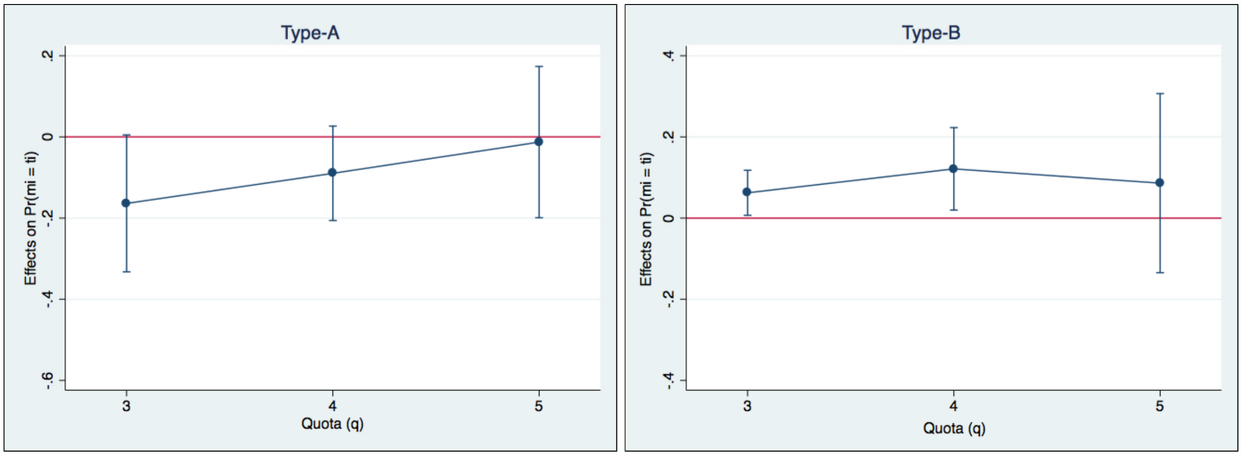

Figure 4: Marginal effects of information on the likelihood of voting truthfully after logit specification. The reported effects control for individual heterogeneity and take into account Ai and Norton (2003) and Karaca−Mandic et al. (2012).

The marginal effects indicate that the presence of forced agents decreases the likelihood of type-A agents on voting truthfully. The effect is significant when $q = 3$ (ME $= -0.164$, $p = 0.057$). However, we cannot reject the null hypothesis of no effect of the presence of forced agents on the behavior of type-A agents when $q = 4$ and $q = 5$ (ME $= -0.090$, $p = 0.131$ and ME $= -0.013$, $p = 0.894$, respectively).

Consistent with our theoretical prediction, we do observe a positive effect of the presence of forced agents on the likelihood of type-B agents voting truthfully when $q = 3$ (ME $= 0.062$, $p = 0.028$) and $q = 4$ (ME $= 0.121$, $p = 0.052$). The effect is not significant when $q = 5$ (ME $= 0.086$, $p = 0.444$).

When we estimate the average effect of the presence of forced agents on the likelihood of type-A agents voting truthfully we find that this is not significant at any common level (ME $= -0.089$, $p = 0.12$). The overall effect for type-B agents is, however, significant (ME $= 0.090$, $p = 0.043$).

Observation 3. *When agents know that two type-B agents are forced to vote truthfully, the likelihood of voting truthfully increases for the remaining type-B agents.*

Finally, we can look at the effect of the quota on the likelihood of voting truthfully in INF. Our results suggest that type-A agents are more likely to vote truthfully when $q = 4$ and $q = 5$ (ME $= 0.20$, $p = 0.029$ and ME $= 0.317$, $p = 0.001$, respectively). As a result, we conclude that the quota has a positive effect on the likelihood of type-A agents voting truthfully. As for type-B agents, there is no effect of the quota when we compare the likelihood of voting truthfully in $q = 3$ and $q = 4$. Surprisingly enough, type-B agents are less likely to vote truthfully in the INF treatment when $q = 5$ (ME = 0.310, $p < 0.001$). This occurs despite voting truthfully being the unique WUNE strategy for type-B agents in all possible quotas. Although we find that the presence of forced agents increase truthful behavior, our findings suggest that type-B agents are less likely to vote truthfully as we increase the quota, probably as a response to type A's agents voting more truthfully when the quota increases.

6.3 Maximizing total surplus

We look at the issue of efficiency losses in this section. If agents care about the welfare of the group (Coate and Conlin, 2004; Feddersen and Sandroni, 2006) they will vote to maximise total payoffs. In our setting this occurs when option B is elected, regardless of all agents are voting for B or not. Our previous findings suggest that conformity decreases truthful voting, whereas receiving information foster it. How do these results translate into efficiency gains or losses in terms of total surplus?

We use the behavioral data in each treatment to form all possible configurations of groups consisting in three type-B agents and two type-A agents.[17] We then compute the average expected payoff and the likelihood of getting the maximum possible payoff on each treatment for each possible quota. These results are reported in Table 4.

As we increase the quota, we expect for the likelihood of electing option B to decrease. As a result, we observe that the average expected payoff and the likelihood of getting the highest possible payoff decreases with the quota. When we compare across treatments, Table 4 shows that CON results in a smaller average expected payoff and a smaller likelihood of getting the highest possible payoff for any possible quota, compared with the BL treatment. Efficiency improves in INF, as the expected payoff and the likelihood of maximizing the total payoff increases for each possible quota.

Observation 4. *The likelihood of maximizing total payoffs is higher when there is no conformity. With the presence of forced agents, the likelihood of maximizing total payoffs increases.*

[17]Given our data, we can form a total of 961,056 different groups in BL and CON. We can form at least 53,100 groups in INF (when we do not combine the behavior of different subjects in the role of forced type-B voters).

Table 4: Efficiency and total surplus

	BL	CON	INF
A. Average expected payoff			
$q = 3$	338.9	335.4	338.4
$q = 4$	290.6	285.4	314.0
$q = 5$	279.1	269.3	272.2
	302.9	296.7	308.2
B. Likelihood of getting the highest possible payoff			
$q = 3$	0.85	0.50	0.54
$q = 4$	0.21	0.03	0.20
$q = 5$	0.06	0.03	0.07
	0.37	0.19	0.27

7 Concluding Remarks

This paper studies the effects of conformity and information in a binary-decision voting game in which agents are heterogeneous with regard to their preferred outcome. We show that conformity reduces the likelihood of truthful voting for all agents, whereas some agents are more likely to vote truthfully when they know that others will vote for their preferred option. Our experimental data support these predictions and suggest the importance of the voting rule in driving the results.

Very much attention has been paid to the incentives for agents to vote truthfully in the literature on Social Choice. There is a bunch of studies characterizing social choice functions that satisfy strategy-proofness (i.e. truth-telling is a weakly dominant strategy in the direct mechanism). Arguably, there are social choice functions satisfying desirable properties that fail to be strategy-proof. If we insist on the desirability of those social choice functions we should measure their manipulability. Although we lack from a unanimously way of doing it theoretically, laboratory experiments provide an excellent manner of addressing this question empirically. This paper is part of a more ambitious project in which we want to test several well-known social choice functions in their appropriate domains of definition. The current paper can be understood as a first step in this direction.

References

[1] Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. Economics letters, 80(1), 123-129.

[2] Anderson, L. R., & Holt, C. A. (1997). Information cascades in the laboratory. The American economic review 87(5), 847-862.

[3] Asch, S. E. (1955). Opinions and social pressure. Readings about the social animal, 193, 17-26.

[4] Banerjee, A. V. (1992). A simple model of herd behavior. The Quarterly Journal of Economics, 797-817.

[5] Battaglini, M., Morton, R. B., & Palfrey, T. R. (2010). The swing voter's curse in the laboratory. The Review of Economic Studies, 77(1), 61-89.

[6] Barberà, S., (2011). Strategy-proof social choice. In: Arrow, K. J., Sen, A., and Suzumura, K. (Eds.). Handbook of Social Choice and Welfare (Vol. 2), 731-831. Elsevier.

[7] Bardsley, N., & Sausgruber, R. (2005). Conformity and reciprocity in public good provision. Journal of Economic Psychology, 26(5), 664-681.

[8] Bassi, A., Morton, R. B., & Williams, K. C. (2011). The effects of identities, incentives, and information on voting. The Journal of Politics, 73(02), 558-571.

[9] Bikhchandani, S., Hirshleifer, D., & Welch, I., (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. Journal of political Economy, 992-1026.

[10] Buechel, B., Hellmann, T., & Klößner, S. (2015). Opinion dynamics and wisdom under conformity. Journal of Economic Dynamics and Control, 52, 240-257.

[11] Coate, S., & Conlin, M. (2004). A group rule: Utilitarian approach to voter turnout: Theory and evidence. American Economic Review, 1476-1504.

[12] Coleman, S. (2004). The effect of social conformity on collective voting behavior. Political Analysis, 12(1), 76-96.

[13] Cooper, D. J., & Dutcher, E. G. (2011). The dynamics of responder behavior in ultimatum games: a meta-study. Experimental Economics, 14(4), 519-546.

[14] Corazzini, L., & Greiner, B. (2007). Herding, social preferences and (non-) conformity. Economics letters, 97(1), 74-80.

[15] Dekel, E., & Piccione, M. (2000). Sequential voting procedures in symmetric binary elections. Journal of Political Economy, 108(1), 34-55.

[16] Dorff, R. H., & Brenner, S. (1992). Conformity Voting on the United States Supreme Court. The Journal of Politics, 54(03), 762-775.

[17] Dutta, B., & Sen, A. (2012). Nash implementation with partially honest ividuals. Games and Economic Behavior, 74(1), 154-169.

[18] Feddersen, T., & Pesendorfer, W. (1998). Convicting the innocent: The inferiority of unanimous jury verdicts under strategic voting. American Political science review, 92(01), 23-35.

[19] Feddersen, T., & Sandroni, A. (2006). A theory of participation in elections. The American Economic Review, 1271-1282.

[20] Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. Experimental economics, 10(2), 171-178.

[21] Gerber, E. R., Morton, R. B., & Rietz, T. A. (1998). Minority representation in multimember districts. American Political Science Review, 92(01), 127-144.

[22] Goeree, J. K., & Yariv, L. (2015). Conformity in the Lab. Journal of the Economic Science Association, 1-14.

[23] Guarnaschelli, S., McKelvey, R. D., & Palfrey, T. R. (2000). An experimental study of jury decision rules. American Political Science Review, 94(02), 407-423.

[24] Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). In search of homo economicus: behavioral experiments in 15 small-scale societies. American Economic Review, 73-78.

[25] Hung, A. A., & Plott, C. R. (2001). Information cascades: Replication and an extension to majority rule and conformity-rewarding institutions. American Economic Review, 1508-1520.

[26] Karaca-Mandic, P., Norton, E. C., & Dowd, B. (2012). Interaction terms in nonlinear models. Health services research, 47(1), 255-274.

[27] Luzzati, T. (1999). Economic theory and conformism (Vol. 68). Edward Elgar Publishing, Cheltenham Glos, UK.

[28] Michaeli, N., & Spiro, D., (2015). Norm Conformity across Societies. Journal of Public Economics, forthcoming.

[29] Milgram, S. (1965). Some conditions of obedience and disobedience to authority. Human relations, 18(1), 57-76.

[30] Milgram, S. (1974). Obedience To Authority: An Experimental View. London.

[31] Moreno, B. & M.P. Ramos-Sosa (2015) Voting by Conformity. Universidad de Málaga, mimeo.

[32] Morton, R. B., Muller, D., Page, L., & Torgler, B. (2015). Exit polls, turnout, and bandwagon voting: Evidence from a natural experiment. European Economic Review, 77, 65-81.

[33] Morton, R. B., & Ou, K. (2015). What motivates bandwagon voting behavior: Altruism or a desire to win?. European Journal of Political Economy, forthcoming.

[34] Morton, R. B., Piovesan, M., & Tyran, J. R. (2012). The dark side of the vote: Biased voters, social information, and information aggregation through majority voting. Harvard Business School Research Paper, (13-017).

[35] Morton, R. B., & Williams, K. C. (1999). Information asymmetries and simultaneous versus sequential voting. American Political Science Review, 93(01), 51-67.

[36] Morton, R. B., & Williams, K. C. (2000). Learning by voting: Sequential choices in presidential primaries and other elections. University of Michigan Press.

[37] Myers, D. G. (2012). Social psychology (11th ed.). Boston, MA: McGraw-Hill.

[38] Rivas, J., & Rodriguez Alvarez, C. (2014). Deliberation, Leadership and Information Aggregation. Documentos de Trabajo (ICAE), (4), 1-34.

[39] Sherif, M. (1937). An experimental approach to the study of attitudes. Sociometry, 1(1/2), 90-98.

Appendix A: Translated instructions (originally in Spanish)

Welcome to the experiment!

This is an experiment to study decision-making. Instructions are simple and if you follow them carefully you will be paid an amount of money at the end of the experiment. Your earnings in this experiment may depend on your decisions and the decisions of other participants. Your earnings will be given confidentially so neither the other participants in the lab nor the instructors will know your payoffs. Similarly, your identity will be anonymous throughout the whole experiment. Thus, nobody will know the decisions you have taken during the experiment.

Please, from no on, do not communicate with other participants during the experiment. If you have any questions please raise your hand. Out of this type of questions, any kind of communication among the participants of the experiment is forbidden and will be subject to immediate exclusion from the experiment.

What's the experiment about?

In the experiment, there are two types of participants: A and B. Before starting the experiment, the computer will randomly decide whether you are a type-A or type-B participant. The type assigned by the computer will remain throughout the experiment.

In total, this experiment consists of three rounds. Before starting the first round, we will match you with other participants in the lab to form a group of 5 participants. Your group will remain throughout the whole experiment and consists of 2 type-A participants and 3-type B participants.

[BL and CON] In each of the rounds, you can choose between voting for *Option A* or voting for *Option B* to determine which is the ***Chosen Option*** of that round. Before you have to vote, we will announce to you (and to all the members of your group) how many votes Option B needs to be chosen in that round.

[INF] In each of the rounds, the computer force two type-B participants to vote
for Option B. If you are a type-B participant and you are forced by the computer to vote for Option B you can only choose that option in the three rounds of the experiment. The rest of the members of the group can choose in each round whether to vote for *Option A* or *Option B* to determine which is the ***Chosen Option*** of that round. Before you have to

vote, we will announce to you (and to all the members of your group) how many votes Option B needs to be chosen in that round.

Option B may need 3, 4 or 5 votes to be chosen, depending on the round.

- If Option B gets the number of votes needed, Option B will be the *Chosen Option* in that round.
- If Option B doesn't get the number of votes needed, Option A will be the *Chosen Option* in that round.

Baseline treatment:

How can I earn money in this experiment?

At the end of the experiment, we will announce which has been the *Chosen Option* in each round and what were your earnings in each round.

Your earnings in each round will depend, in part, on the Chosen option. You will also receive an additional amount just for casting your ballot. The total earnings you can receive are explained in detail as follows:

<u>Earnings for the chosen option</u>: If the Chosen Option in your group coincides with your *type* you will receive 75 ECUS. That is, if you are a type-A participant and Option A is the Chosen Option for the members of your group, you will receive 75 ECUS in that round. In the same way, if you are a type-B participant and Option B is the Chosen Option for the members of your group, you will receive 75 ECUS in that round.

<u>Earnings for casting your ballot</u>: Independently on whether you choose Option A or Option B, you will receive 25 ECUS just for casting your ballot. This gain is received independently on your type of participant and the Chosen Option in that round.

To determine your earning in one round, we sum your earnings for the chosen option and your earnings for casting the ballot. At the end of the experiment, after informing you about your total earnings in each of the three rounds of the experiment, the computer will randomly choose one round and we will pay you out depending on the earnings obtained in that round.

Your will receive the equivalent in EURO according to the exchange rate 10 ECUS = 1 Euro.

Next, we present some examples in order to show you how payoffs are calculated

(*Examples*)

Now, the computer will randomly choose whether you are a type-A or type-B participant in this experiment. Next, we present a screen with two more examples. We ask you to pay attention to these examples in order to understand correctly how earnings are calculated before starting the experiment as after that you will take a simple test to check that you have already understood everything. If after reading the examples or solving the test you have doubts about how earnings are calculated, please, raise your hand and ask the instructors. Knowing how earnings are calculated can help you to obtain more money during the experiment.

Besides, bear in mind that:

1. Your type (A or B) will be the same and in your group there always be 2 type-A participants and 3 type-B participants.

2. To determine which is the *Chosen Option* in one round, we check whether Option B has obtained a determined number of votes in that round (3 votes, 4 votes or 5 votes). This information may change from round to round and will be announced to all the members of your group before voting.

3. Your earnings will depend on whether the *Chosen Option* in one round coincides with your type of participant. Besides, you will receive an additional amount just for casting the ballot (regardless of your vote).

Conformity and informational treatment:

How can I earn money in this experiment?

At the end of the experiment, we announce which has been the Chosen Option in each round and what were your earnings in each round.

Your earnings in each round will depend, on the one hand, on the Chosen Option and on the other, *on the number of members of your group voting for the same option than you*. The total earnings you can receive are explained in detail as follows:

<u>Earnings for the chosen option</u>: If the Chosen Option in your group coincides with your *type* you will receive 75 ECUS. That is, if you are a type-A participant and Option A is the Chosen Option for the members of your group, you will receive 75 ECUS in that round. In the same way, if you are a type-B participant and Option B is the Chosen Option for the members of your group, you will receive 75 ECUS in that round.

<u>Earnings for coinciding</u>: If the option you have voted for in one round has in total 2 or more votes (that is, if somebody else has vote for the same option than you), then you will receive 25 ECUS. You will receive these earnings regardless of your type of participant and the Chosen Option in the round.

To determine your earnings in one round, we sum your earnings for the Chosen Option and your earnings for coinciding. At the end of the experiment, after informing you about your total earnings in each of the three rounds of the experiment, the computer will randomly choose one round and we will pay you out depending on the earnings obtained in that round.

Your will receive the equivalent in EURO according to the exchange rate 10 ECUS = 1 Euro.

Next, we present some examples in order to show you how payoffs are calculated

(Examples)

Now, the computer will randomly choose whether you are a type-A or type-B participant in this experiment. Next, we present a screen with two more examples. We ask you to pay attention to these examples in order to understand correctly how earnings are calculated before starting the experiment as after that, you will take a simple test to check that you have already understood everything. If after reading the examples or solving the test you have doubts about how earnings are calculated, please, raise your hand and ask the instructors. Knowing how earnings are calculated can help you to obtain more money during the experiment.

Besides, bear in mind that:

1. Your type (A or B) will be the same and in your group there always be 2 type-A participants and 3 type-B participants.

2. To determine which is the *Chosen Option* in one round, we check whether Option B\ has obtained a determined number of votes in that round (3 votes, 4 votes or 5 votes). This information may change from round to round and will be announced to all the members of your group before voting.

3. Your earnings will depend on whether the *Chosen Option* in one round coincides with your type of participant. Besides, you will receive an additional amount if someone else has voted for the same option than you.

[INF only] 4. There are 2 type-B participants in your group that are forced to vote for Option B in the three rounds of the experiment.

Questionnaire

- **Age:** What is your age?. . . years.
- **Gender:** What is your gender? (00 male, 01 female)
- **Risk aversion**: We elicited risk attitudes using the investment decision in Gneezy and Potters (1997). Each participant hypothetically received 10 Euros and was asked to choose how much of it, *x*, she wanted to invest in a risky option and how much she wished to keep. The amount invested yielded a dividend equal to 2.5*x* with 1/2 probability, being lost otherwise. The money not invested in the risky option (10-*x*) was kept by the participant. In this situation, the expected value of investing is positive and increasing in the amount invested; therefore a risk-neutral (or risk-loving) participant should invest the 10 Euros, whereas a risk-averse participant will invest less. The amount not invested in the risky asset is a natural measure of risk aversion.
- **Trust:** We use the question in the GSS to elicit attitudinal trust. Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people? (Trust = 1 if the answer is "most people can be trusted")
- **Cognitive abilities**: We use the Cognitive Reflection Test (CRT) in Frederick (2005).
- **Social preferences:** We use the answers to the social value orientation (SVO) in Van Lange et al. (1997) to classify subjects as individualistic, prosocial, competitive, or others.
- **Satisfaction**: "How do you feel in this moment with your life?" (1-7-scaled answer from 1 (very satisfied) to 7 (Not at all satisfied)
- **Happiness**: "Taking everything into consideration, would you call yourself..." (01 not very happy, 02 quite happy, 03 very happy)
- **Inequality:** "Consider the following situation: Two secretaries with the same age do exactly the same work. However, one of them earns 20 euros per week more than the other. The one that is paid more is more effcient and faster, while working. Do you believe it is fair that one earns more than the other?" ($Ineq = 1$ if Yes)

Appendix B

1. Non-parametric analysis

1.1 Effect of conformity

Table 1 in the main text shows the effect of conformity by reporting he average likelihood of voting truthfully in BL and CON, for each type of each and each possible quota separately. If we consider the Chi-2 test (one-side alternative), we find that the presence of conformity decreases the average frequency of voting truthfully from 0.83 to 0.69 ($p < 0.001$).[1] The negative effect of conformity on voting truthfully is significant for both type-A ($p < 0.003$) and type-B agents ($p < 0.001$).

When we can look at the effect of conformity for each possible quota in Table 1, we observe that conformity decreases the likelihood of type-A agents voting truthfully when the quota is 3 ($p = 0.006$), and when it is 4 ($p = 0.042$). As expected, conformity has no effect for type-1 voters when the quota is equal to 5 ($p = 0.52$). As for type-B agents, we expect that conformity decreases the likelihood of truth telling for each possible quota. Our findings suggest no significant effect when the quota is equal to 3 (p = 0.34), but when the quota equals to 4 (p = 0.004) or 5 ($p < 0.001$). Overall, all these findings are in line with our prediction.

Table 1 allows us to observe also the effect of the quota on the likelihood of voting truthfully. When there is no conformity, Proposition 1 shows that there is a unique WUNE in which all agents vote truthfully regardless of the value of the quota. Our findings seem to support this prediction by suggesting no effect of the quota on the likelihood of voting truthfully in the BL treatment. The quota, however, seems to have a significant effect in the likelihood of voting truthfully in the CON treatment. In principle, we expect for type-A agents to vote more truthfully in $q = 5$. When we do pairwise comparisons we find that the likelihood of truth telling increases in $q = 5$ with respect to $q = 3$, but the effect is weakly significant ($p = 0.10$). In fact, it seems that type-A agents are more likely to vote truthfully in the CON treatment when $q = 4$. As for type-B agents, the likelihood of voting truthfully seems to decrease in the value of the quota. We note that our theoretical result do not provide any prediction for this behavior.

[1] Our findings are robust to other statistical analyses (e.g., t-test, test of proportions or Mann-Whitney-Wilcoxon test).

1.2 Effect of information

The introduction of two type-B agents forced to vote truthfully does not affect the likelihood of voting truthfully for type-A agents (p = 0.18) but it does for type-B agents (p = 0.012).

When we look at the effects of forced agents for each possible quota in Table 1, we observe that knowing that two agents are forced to vote for B does not affect the likelihood of type-A agents voting truthfully when the quota is 3 (p = 0.253) or 5 (p = 0.652). The effect is, however, significant when the quota is 4 (p = 0.046). As for type-B agents, the non-parametric analysis rejects the null hypothesis that forced agents do not affect the likelihood of voting truthfully when the quota is equal to 4 (p = 0.004). This is, however, rejected when the quota equals 4 (p = 0.64) or 5 (p = 0.36).

We can also test the effect of the quota on the likelihood of voting truthfully in the INF treatment. Pairwise comparisons suggest a positive effect of the quota on the likelihood of type-A agents voting truthfully. This is because the likelihood of voting truthfully is smaller in q = 3 than in q = 4 (p =0.014$) or q = 5 (p < 0.001). If we compare q = 4 and q = 5 we find no effect of the quota p = 0.190. As for type B agents, the quota has no effect on the likelihood of voting truthfully when we compare q = 3 and q = 4 (p = 1), but type-B agents are less likely to vote truthfully when q = 5 compared with any of the other two quotas (p = 0.003).

2. Econometric analysis

2.1 Effect of conformity

The first two columns of Table C.1 present the maximum likelihood estimates of our logit regression, for each of the two types of agents separately. We complement our analysis by reporting the results of a linear probability model. In this latter case, the estimates can be interpreted as the effects of each independent variable on the probability of voting truthfully. The set of independent variables include controls for gender, risk aversion, trust, or social value orientation of the subject (see Appendix A for further details on the questionnaire).

Table C. 1. Econometric analysis for the effect of conformity

Dependent variable: Probability of voting truthfully				
	Logit model		Linear Probability model	
	Type-A	Type-B	Type-A	Type-B
dq_4 (=1 if q = 4)	0.911	-1.10 e16	0.146	-2.62e-15
	(0.57)	(0.56)	(0.09)	(0.05)
dq_5 (=1 if q = 5)	-0.103	-0.629	-0.021	-0.069
	(0.50)	(0.56)	(0.10)	(0.06)
d_{CON} (=1 if conformity)	-0.990**	0.518	-0.228**	0.027
	(0.49)	(0.68)	(0.10)	(0.05)
$d_{CON}\,dq_4$	0.206	-1.852**	0.104	-0.208**
	(0.74)	(0.84)	(0.14)	(0.080)
$d_{CON}\,dq_5$	0.822	-2.003**	0.187	-0.306***
	(0.68)	(0.81)	(0.15)	(0.09)
Age	-0.039	0.049*	-0.009	0.006**
	(0.04)	(0.03)	(0.01)	(0.003)
Women	-0.669*	0.278	-0.131*	0.041
	(0.34)	(0.31)	(0.07)	(0.04)
Risk aversion	-0.026	0.121**	-0.005	0.015**
	(0.05)	(0.06)	(0.01)	(0.01)
Trust	-0.111	0.096	-0.025	0.011
	(0.32)	(0.32)	(0.06)	(0.04)
Inequality	0.503	0.014	0.104	0.006
	(0.33)	(0.34)	(0.07)	(0.04)
CRT	-0.077	0.708	-0.019	0.094
	(0.45)	(0.50)	(0.09)	(0.06)
Individualistic	-0.274	-0.294	-0.032	-0.031
	(0.50)	(0.50)	(0.09)	(0.07)
Prosocial	-0.346	0.025	-0.046	0.009
	(0.46)	(0.49)	(0.08)	(0.06)
Competitive	-0.084	-0.609	-0.001	-0.078
	(0.80)	(0.86)	(0.14)	(0.13)
Satisfaction	-0.003	-0.167	-0.001	-0.023
	(0.12)	(0.14)	(0.02)	(0.02)
Happiness	0.058	-0.337	0.010	-0.045
	(0.28)	(0.31)	(0.06)	(0.04)
Constant	2.130	1.546	0.946***	0.810***
	(1.35)	(1.43)	(0.28)	(0.18)
Wald-test	32.38***	42.41***	2.96***	2.93***
(Pseudo) R^2	0.09	0.13	0.11	0.131
Observations	288	432	288	432

Notes: Robust standard errors are reported in parentheses. The log-pseudo likelihood in the logit regressions is -165.94 and -179.33 for type-A and type-B agents respectively. Significance *** p<0.01, ** p<0.05, * p<0.1

We see in the first column that conformity has a significant effect on the likelihood of type-A agents voting truthfully. The marginal effect reported in Figure 2 in the main text indicates that the likelihood of voting truthfully indeed decreases by 22.2 percent when q = 3. This is in line with the results of the linear probability model in column 3, which estimates a decrease of 22.8 percent. The Chi-2 test suggests that the effect of conformity is also significant when q = 4, but not when q = 5. Thus, we can observe that the estimates of conformity in the linear probability regression (-0.228) is roughly in the same magnitude (but different sign) than the estimate when we interact the treatment variable with the quota q = 5 (0.187).

Our analysis for type-B agents in columns 2 and 4 suggests that conformity has no effect on the likelihood of voting truthfully when q = 3, but in the rest of the cases. Thus, note that the dummy variable for the treatment condition is not significantly different from 0, but the interaction terms with the quotas are significant. The estimates of the linear probability model indeed suggest that the negative effect of conformity when q = 4 is around 18.1 percent (0.027-0.208). It is roughly 27.9 percent when q = 5. Recall that the marginal effects after the logit regressions reported in the main text are -0.11 and -0.32 respectively.

2. 2. Effect of knowing that others vote truthfully

The first two columns of Table C.2 present the maximum likelihood estimates of our logit regression, for each of the two types of agents separately. We complement our analysis by reporting the results of a linear probability model.

In line with our discussion in the main text, we find that information affects differently to type-A (who tend to vote less truthfully when we introduce independent voters) and type-B agents (who tend to vote more truthfully). When q = 3, we find that the effect for type-B agents is not significant according to the linear probability model, but it is when q = 4. This is the main difference between the logit model reported in the main text and the linear probability model. We also note that the reported estimates for the logit model in Table C.2 cannot be used to compute the marginal effects, according to Ai and Norton (2003). The correct marginal effects are reported in the main text.

Table C. 1. Econometric analysis for the effect of information

Dependent variable: Probability of voting truthfully				
	Logit model		Linear Probability model	
	Type-A	Type-B	Type-A	Type-B
dq_4 (=1 if q = 4)	1.077**	-1.871***	0.250**	-0.208***
	(0.45)	(0.64)	(0.10)	(0.06)
dq_5 (=1 if q = 5)	0.693	-2.665***	0.167	-0.375***
	(0.44)	(0.62)	(0.11)	(0.07)
d_{INF} (=1 if information)	-0.522	0.353	-0.123	-0.006
	(0.41)	(1.13)	(0.10)	(0.05)
$d_{INF}\,dq_4$	-0.239	1.871	-0.050	0.208**
	(0.60)	(1.63)	(0.14)	(0.08)
$d_{INF}\,dq_5$	0.647	-0.178	0.150	0.065
	(0.63)	(1.31)	(0.15)	(0.12)
Age	-0.034	0.053	-0.008	0.005
	(0.03)	(0.03)	(0.01)	(0.003)
Women	-0.425	0.313	-0.097	0.046
	(0.29)	(0.35)	(0.07)	(0.05)
Risk aversion	0.026	0.164***	0.006	0.020***
	(0.06)	(0.06)	(0.01)	(0.01)
Trust	0.047	0.113	0.011	0.013
	(0.25)	(0.35)	(0.06)	(0.05)
Inequality	-0.127	-0.152	-0.028	-0.016
	(0.31)	(0.39)	(0.07)	(0.05)
CRT	-0.439	0.664	-0.099	0.083
	(0.40)	(0.548)	(0.09)	(0.07)
Type_i	-0.746	-0.366	-0.166	-0.043
	(0.63)	(0.69)	(0.14)	(0.09)
Type_p	-0.647	-0.242	-0.143	-0.022
	(0.59)	(0.62)	(0.13)	(0.08)
Type_c	-0.522	0.172	-0.119	0.043
	(0.81)	(1.02)	(0.172)	(0.13)
Satisfaction	0.050	-0.272*	0.0114	-0.036
	(0.11)	(0.15)	(0.03)	(0.02)
Happiness	0.118	-0.374	0.027	-0.047
	(0.26)	(0.40)	(0.06)	(0.05)
Constant	1.128	2.405	0.752**	0.904***
	(1.39)	(1.72)	(0.32)	(0.23)
Wald-test	20.34	46.51***	1.60*	4.86***
(Pseudo) R^2	0.06	0.20	0.08	0.18
Observations	324	303	324	303