EVOLUTION & VOTING
How Nature Makes us Public Spirited

John P.Conley

Ali Toossi

And

Myrna Wooders

No 601

WARWICK ECONOMIC RESEARCH PAPERS

THE·UNIVERSITY·OF
WARWICK

DEPARTMENT OF ECONOMICS

Evolution & Voting:

How Nature Makes us Public Spirited

(Preliminary)

John P. Conley[1]
University of Illinois
jpconley@uiuc.edu

Ali Toossi
University of Illinois
toossi@students.uiuc.edu

Myrna Wooders
University of Warwick
ecsdz@warwick.ac.uk

First Version: January 2000
Revised: February 2001

Abstract

We reconsider the classic puzzle of why election turnouts are persistently so high when formal analysis strongly suggests that rational agents should not vote. If we assume that voters are not making systematic mistakes, the most plausible explanation seems to be agents receive benefits from the act of voting itself. This is very close to assuming the answer, however, and immediately begs the question of why agents feel a warm glow from participating in the electoral process. In this paper, we approach this question from an evolutionary standpoint. We show for a range of situations that public-spirited agents have an evolutionary advantage over those who are not as public-spirited. We also explore when this kind of altruistic behavior is disadvantageous to agents. The details depend on the costs of voting, the degree to which agents have different preferences over public policies and the ratio of various preference types in the population, but we conclude that evolution may often be a force that causes agents to internalize the benefits their actions confer on others.

1. Introduction

Any rational voter should realize that the probability that his vote will have an effect on the outcome of an election is negligible. Many classical writers in voting theory, Downs (1957) and Tullock (1968) for example, have argued that it simply does not pay to a citizen to show up at the polls. Even if a voter cares passionately about the outcome, the odds that his vote will be pivotal are so small, that the expected benefit of casting a ballot would always be offset by even minor costs of voting. This is hard to reconcile with the fact that more than one hundred million Americans voted in the most recent presidential election.

Not surprisingly, there have been many attempts to provide a theory of voting that comports with actual observations. Ferejohn and Fiorina (1974), for example, suggested that voters might not be fully informed, and so may not be able to calculate the probability that their vote would make a difference. They noted that this precludes voters from making an expected utility calculation and proposed instead that voters might be using a minimax strategy. Since having voted when an agent ends up not being pivotal involves only a small regret (the cost of voting) but not having voted when an agent would have been pivotal involves very large regret, minimax agents will generally choose to participate in elections.

Ferejohn and Fiorina's argument has the virtue that it does provide a foundation for rational voting, however, it is open to criticism on at least two grounds. Most obviously, it calls for agents to choose strategies in an extremely conservative, and probably unrealistic way. For example, a minimax agent should never cross a street because it is possible that a car might hit him. More fundamentally, Ferejohn and Fiorina ignore the fact that the benefit of voting to any given agent depends on the actions of all of the other agents. While the expected utility approach can also be criticized for taking the probability a voter will be pivotal as exogenous and not depending on strategic interaction among voters, Ferejohn and Fiorina go one step

further. In suggesting that agents follow a minimax strategy, they are asserting that voters give no consideration at all to the strategic choices of others. It may be possible to justify this as an approximation for large economies, but it is at least a bit troubling to build a theory of voting on foundation of strategic myopia.

More recently, several authors have reformulated the problem to allow for strategic interaction between voters. For example, Palfrey and Rosenthal (1983) consider a model in which agents are completely informed about costs of voting and preferences of other voters. These agents play a noncooperative game in which they can either vote or abstain. They show that in some equilibria, there are substantial turnouts even for large economies. Unfortunately, these high turnout equilibria seem to be fragile, and, as they point out, the assumption of complete information is rather strong for large populations. The work of Palfrey and Rosenthal is partly based on the pioneering work of Ledyard (1981). There, and in a subsequent paper (Ledyard 1984), this author first explores the idea of strategic interaction among voters. In contrast to Palfrey and Rosenthal, Ledyard considers the case of voters who have incomplete information about the voting costs and preferences of their fellow citizens. Ledyard's key result is to prove that equilibria exist with positive turnouts. Unfortunately, Palfrey and Rosenthal (1985) were able to show that as the electorate got large, the cost of voting would again be the dominant factor for rational voters in Ledyard's model and so turnouts would be low.

To summarize, although the game theoretic approach taken by these author do suggest that turnouts will be positive in many cases, they still do not explain what we actually observe. What is missing is a model in which agents have incomplete information and at the same time exhibit robust equilibria in which the turnouts are substantial even for large populations.

Riker and Ordeshook (1968) propose quite a different explanation for why it could be rational to vote. They suggest that agents might actually get utility from the act of voting itself. They show that if agents feel a sense of civic duty that

2

is satisfied by going to polls, then large positive turnouts are not at all surprising regardless of the size of the electorate. This seems quite plausible, and the recent literature provides strong empirical and experimental evidence that agents do indeed feel a "warm glow" from public-spirited activity. See Andreoni 1995, and references therein. Despite the intuitive appeal of the civic duty explanation, it is somewhat disappointing from a theoretical standpoint. Saying that agents vote because they like to vote is essentially assuming the answer. As Andreoni (1990) points out in a somewhat different context, making such an assumption tends to rob the theory of its predictive power.

This provides the starting point for the current paper. Our main objective is to address the questions of why agent might have such a sense of civic duty. Is there some sense in which public-spiritedness in the context of voting is beneficial to agents? If so, what degree of altruism is optimal? Fundamentally, we are asking how agents might come to have preferences that incorporate the welfare of others. We approach this question from an evolutionary standpoint. We show for a range of situations in a voting game that public-spirited agents have an evolutionary advantage over those who are not as public-spirited. We also explore when this kind of altruistic behavior is disadvantageous to agents. In general, we find that agents who like to vote will have an evolutionary advantage when voting is not too costly compared to the benefits of winning elections, and when the population of like-minded voters is large enough that winning an election is a realistic possibility. The broader message is that evolution may be a force that causes agents to internalize the benefits their actions confer on others, at least to the extent that they all share a common set of preferences.

The plan of this paper is the following. In section 2 we describe the model. In section 3, we explore how the cost of voting, the size of the opposition, and the degree to which preferences over public policies affect the evolutionary benefits of voting. In section 4, we connect these results to the literature on evolution and

altruism more generally and discuss possible extensions. Section 5 concludes.

2. The Model

We consider a dynamic economy with continuum of agents uniformly distributed on the interval [0,1]. Agent are divided into two types which we will designate H and L for "high" and "low" voters, respectively. Two things distinguish these types: their preference over public policies and their propensity to vote. We denote the *share* of each type in the population by S_j. Since the population is divided between these two types we have

$$S_j \in [0, 1] \text{ for } j = H, L, \text{ and } S_H = 1 - S_L.$$

Each period, agents vote on a randomly generated public proposal.[2] These proposals produce a cost or benefit for each type of agent that is uniformly distributed on the interval $[-1, 1]$. We assume that all agents of the same type have the same preferences over proposals, but that preferences between the types differ. Formally, we denote the *benefit* that agents of type H receive from a proposal in any given period as a random variable B_H where:

$$B_H \sim \mathcal{U}(-1, 1).$$

We wish to allow the preferences of the two types of agents to be positively or negatively correlated. Thus, we denote the benefit that agents of type L receive from a proposal in any given period as a random variable B_L where:

$$B_L \in \{B_L^p, B_L^n\},$$

---

[2] Much of this model is a modification of Conley and Temimi (2001).

and

$$B_L^n = -\alpha B_H + (1 - \alpha)U_I$$

$$B_L^p = (1 - \alpha)B_H + \alpha U_I,$$

where U_I is an independent uniform distribution on the interval $[-1, 1]$, and $\alpha \in [0, 1]$ is the *preference correlation parameter*. Note that this implies that the correlation coefficients between B_H and each of B_L^n and B_L^p are:

$$Corr(B_H, B_L^n) = \frac{-\alpha}{\sqrt{1 - 2\alpha + 2\alpha^2}}$$

$$Corr(B_H, B_L^p) = \frac{1 - \alpha}{\sqrt{1 - 2\alpha + 2\alpha^2}}.$$

We denote the *propensity to vote* of the two types by V_H and V_L where

$$V_j \in (0, 1) \text{ for } j = H, L.$$

Also define the *relative public-spiritedness* of the two types as:

$$\beta = \frac{V_H}{V_L},$$

where $\beta \geq 1$. We shall assume that the likelihood that an agent chooses to vote for a proposal depends both on their innate propensity to vote (V_j) and the net costs or benefits that passage of a given realization the proposal will produce for them. More formally, we shall assume for any realization of the public proposal b_j, that $v_j \, |b_j|$ is the probability that a voter of type j will cast a ballot. This implies that the net turnout of voters of type j in any given election is a random variable given by:

$$TO_j = S_j V_j B_j.$$

Note that this number can be positive or negative. We will use the convention that a negative turnout measures the number of "No" votes while a positive one

measures the number of "Yes" votes. Putting both types together implies that the total net turnout is a random variable given by:

$$TO_{net} = TO_H + TO_L = S_H V_H B_H + S_L V_L B_L,$$

We denote the *cost of casting a ballot* by C and assume is the same for everybody. Since the voters show up at the polls with probabilities less than one, the realized voting cost to a voter of type j in any given election is also a random variable:

$$C_j = V_j \mid B_j \mid C.$$

Note that it is the probability of voting that affects the expected cost and not whether the vote was positive or negative. This is why we consider the absolute value of the proposals in the expression for the costs.

The expected payoff that members of each type receive in each period is rather complicated from an algebraic standpoint. It requires calculating the net turnout for any given realization of a proposal, and then integrating over all the proposals that pass, while subtracting the expected voting cost in each case. The net turnout, in turn, depends on the share of each type of agent in the population, the relative public-spiritedness of the types, whether the two types of agents have positively or negatively correlated preferences, and the preference correlation parameter. We relegate both the expression and the derivation to the appendix, however, we denote the *expected payoff* to agents of type j by:

$$\bar{\pi}_j, \text{ for } j = H, L$$

To model the evolution of the shares of each voter type over time we use replicator dynamics. According to this dynamic, the growth rate of the proportion of each type in the population is determined by the difference between its expected

6

payoff and the population average payoff. Any type whose expected payoff is more than average increases its share of population. Formally, the average payoff is:

$$\bar{\pi} = S_H \bar{\pi}_H + S_L \bar{\pi}_L.$$

In the interest of simplifying the model, we will treat the dynamics of this model as taking place in continuous time. Since we will mainly be interested in showing how the parameters of the models and initial conditions of the economy determine which steady state the system converges to, this is innocuous. On the other hand, if we wanted to calculate the actual dynamic path we would have to explicitly take into account the fact the proposals are distinct and arrive at discrete points in time. For the present, however, we shall assume that population shares evolve according to the following dynamic:

$$\dot{S}_j = S_j(\bar{\pi}_j - \bar{\pi})$$

where $\dot{S}_j$ is derivative of S_j with respect to time. The state of system at time t is given by the currently population shares:

$$S^t = (S_H^t, S_L^t).$$

We close with a remark. In the introduction, we told a story about evolution taking place over preferences and nature selecting for agents who felt a "warm glow" from altruistic actions. In the model above, however, altruistic actions appear to programmed into behavior and evidently does not relate preferences at all. It would have been possible to derive the behavioral voting parameter (V_j) indirectly from an altruism parameter in preferences. However, since these two would be completely correlated, nothing but additional algebraic complexity would be gained from looking at the microfoundations in this case. We therefore consider a reduced form in which V_j serves as being a proxy for altruism in preferences. We do not believe any loss of generality results.

3. Results

In this section, we focus on the steady states of the economy we describe above. We will be most interested in showing how the parameters of voting game determine the population shares in the steady state to which the system converges.

The literature on evolution in economics is most concerned with the evolutionarily stable states (ESS). Testing for the stability of a steady state requires that the strategies agents play be shown to survive the introduction of small proportions of "mutant" strategies in the sense that they yield higher average payoffs. We will take up the question of how the presence of mutant players affects our equilibria in section 5. In this section, however, we will concentrate on finding the steady states themselves and will also study the likelihood a particular steady state will emerge as the outcome of the dynamic process. With this in mind, we shall say that the *type j wins the evolutionary game* if the parameters and initial conditions are such that the economy converges over time to a stable steady state in which type j makes up the entire populations ($S_j = 1$). We do this to simply our discussion, however, and not to assert that this necessarily is a compelling definition of evolutionary success.

We begin by the showing that steady states will always exist, and that there are three distinct possible dynamic situations for the economy.

Theorem 1. *Depending on values of parameters α, β and C there are three possible outcomes for the system:*[3]

1. **High type wins:** *The system has two steady states $S_H = 0$ and $S_H = 1$ where $S_H = 1$ is globally stable and $S_H = 0$ is unstable.*

---

[3] Let $F_t(S_0)$ be the value assumed by the state variable at time t when the initial condition at time 0 is S_0. A steady state S^* is stable if for every neighborhood U of S^* there is a neighborhood U_1 of S^* in U such that if $S_0 \in U_1$, $F_t(S_0) \in U_1$, $t > 0$. A steady state is asymptotically stable if it is stable and in addition if $S_0 \in U_1$, then $\lim_{t \to \infty} F_t(S_0) = S^*$. The basin of attraction of an asymptotically stable steady state is the set of all points S_0 such that $\lim_{t \to \infty} F_t(S_0) = S^*$. If there is a unique steady state with basin equal to the entire state space it is called globally stable.

2. **Large population wins:** *The system has three steady states, $S_H = 0$, $S_H = 1$ and $S_H = S_H^* \in (0,1)$ where $S_H = 0$ and $S_H = 1$ are asymptotically stable and their basin of attraction are $[0, S_H^*)$ and $(S_H^*, 1]$ respectively, and $S_H = S_H^*$ is unstable.*

3. **Low type wins:** *The system has two steady states $S_H = 0$ and $S_H = 1$ where $S_H = 0$ is globally stable and $S_H = 1$ is unstable.*

<u>Proof/</u>

See appendix.

Figure 1 illustrates the three cases given in Theorem 1. What this result says is sometimes, the high voter types will increase their share of the population until they make up the entire society regardless of how small their numbers to begin with. This case is shown in Figure 1a. For other parameters, the low types will come to dominate the population regardless of their initial share. Figure 1c illustrates this. Both of these situations, however, are just limiting cases of what we think of as the more typical case in which the initial population shares matter. In general, there will be two stable steady states and one unstable steady state that divides the basins of attraction. Figure 1b illustrates this. We will call this unstable steady state the *tipping point* and denote it S_H^*.

We now turn to the question of when being public-spirited is more likely to lead to evolutionary success. More specifically, we will explore how the parameters of the model affect the relative size of the basins of attraction for the various steady states. Theorem 1 already shows that for some parameter configurations, either the high or low voter types will always win the evolutionary game regardless of their initial populations shares. Since these are degenerate case of the more general "large population wins" situation, we will focus our attention on Case 2 of Theorem 1 for the remainder of the section. From a formal standpoint we will, be asking how the tipping point moves in response to changes in the cost of casting a ballot, the
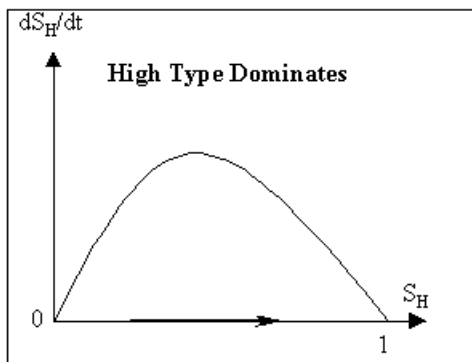
Figure 1a.
No matter what the initial population share, the high voting type will eventually make up the entire population. Thus, $S_H = 1$ is a globally absorbing state

Figure 1b.
The larger its initial population share, the more likely a type is to win the evolutionary game. The $S_H = S^*$ is an unstable steady state that divides the basins of attraction for the two stable steady states $SH = 1$ and $S_H = 0$.

Figure 1c.
No matter what the initial population share, the high voting type will eventually make up the entire population. Thus, $S_H = 1$ is a globally absorbing state
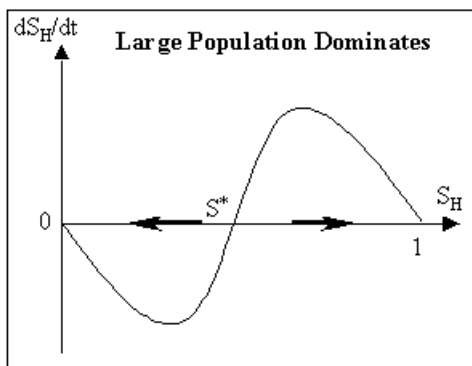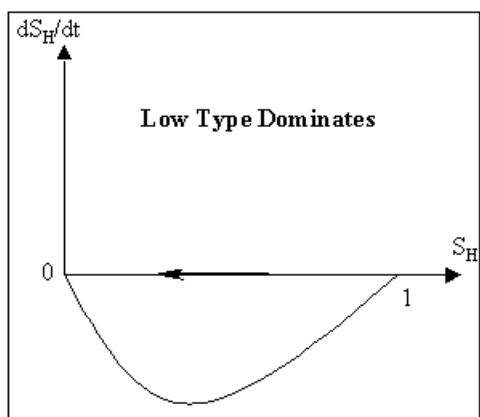
parameter of relative public-spiritedness and the parameter of preference correlation. We will interpret an increase in the size of a basin of attraction for a steady state in which a given type makes up the entire population as meaning it is easier for that type to win the evolutionary game. We begin by consider what happens as the cost of voting increases.

Theorem 2. *All else equal, the higher the cost of voting C, the less likely the high voter types will win the evolutionary game.*

Proof/

    See appendix.

To be slightly more formal, Theorem 2 says that if the parameters of the system are such that the system has three steady states, then all else equal, as C increases S_H^* approaches one and the basin of attraction of $S_H = 0$ expands. This means that the as the cost of voting increases, the high voter type has to have a larger initial population share to prevent themselves from being squeezed out by the low voter type. Of course, this makes sense. If voting is costly, then the act of voting conveys that much less net increase in payoff to the high voter types. If voting is extremely costly, voting is a net loss, even to the group collectively. In this case, it is better to have a low voting parameter and we end in up in case 3 with the only stable steady state being $S_H = 0$ and the tipping point forced all the way up to $S_H = 1$.

Proof/

    See appendix.

Next we consider how the parameter of relative public-spiritedness β affects the evolutionary advantage of voting. The question is: when the high types become more public-spirited or the low types become less public-spirited, does it become more or less likely that the high voting types will win the evolutionary game?

11

Theorem 3. *All else equal, more public-spirited the high voter types are compared to the low voter types, the more likely the high voter types will win the evolutionary game.*

Proof/

    See appendix.

    More formally, assume the parameters of the system are such that the system has three steady states. Then all else equal, as β increases, S_H^* approaches zero and the basin of attraction of $S_H = 1$ expands. In the limit, low types don't vote at all, and the high types always win regardless of their initial population share. In general, this implies that high public-spiritedness can substitute for lower population. Thus, political activism may be especially important to minority groups.

    Last, we turn to the effects of preference correlation α. The question is: is public-spiritedness more or less of an advantage for a group when they are have preferences that are similar to the remaining population?

Theorem 4. *All else equal, the more positively correlated the preferences of the two types of voters, the less likely the high voter types will win the evolutionary game.*

Proof/

    See appendix.

    More formally, assume the parameters of the system are such that the system has three steady states. Then all else equal, as the preferences of two types become more positively correlated S_H^* approaches one, and the attraction of $S_H = 0$ expands. Again, this makes intuitive sense. If preferences are more negatively correlated, the payoff from voting is higher since not only do you get more benefits for your own group, but you also damage the competing group. However, if preferences are more positively correlated, there is a kind of "free-rider" effect operating. The low voter types benefit from the proposals that the high voter types work to pass

12

through the electoral process. The low types pay less of the voting cost, but receive benefits that are close to those of the high voter type. In the extreme case of perfect correlation, the high types are always supplanted by the low types in the steady state regardless or the initial population shares (case 3). This will turn out to have significant implications for the interpretation of our steady states as Evolutionary Stable equilibria. We explore this more in the next section.

As an aside, this result is a bit more complex than it first appears. Recall that we consider two cases: one in which the preferences are positively correlated and $\alpha = 0$ means perfect correlation while $\alpha = 1$ means no correlation, and another where preferences are negatively correlated and $\alpha = 0$ means no correlation while $\alpha = 1$ means perfect negative correlation. Fortunately, these cases dovetail and in both cases, a lower alpha means that the preferences are more similar. In the appendix, we show that despite these being separate mathematical situations, the qualitative results are the same.

4. Evolution and Altruism.

The literature on evolution in economics is very large, and it is not our intention to survey it here. Instead, we shall concentrate on a discussion of how the model we present agrees with and differs from the existing literature.

Evolutionary game theory is typically used to explain who how agents might choose strategies in an arational way. Thus, evolution takes place over strategic choices. See Taylor and Jonker 1978, Friedman 1991, or more recently Lagunoff 2000, among many others. In contrast, we propose that evolution takes place over the underlying preferences of agents and those in turn determine their strategic

13

choices.[4] In this, we follow such authors as Becker (1976), Hirshleifer (1978) more recently Bergstrom and Stark (1993) and Robson (1996). (See Robson 2000 for a more complete survey.)

This raises an interesting question regarding whether or not our story can be reconciled with the traditional view in economics which seems to take evolution as metaphor for learning or imitation in strategic situations. See Kandori, Mailath and Rob (1993) or Fundenberg and Levine (1998) chapter 3, for example. We take a somewhat neutral view on this. Whether preferences come from nature (no learning) or nurture (passive learning) does not really matter for the results in our model. In either case, the actions of the parents are passed on through preferences to the children. What our model does not allow is a kind of active learning in which agents might somehow choose to undertake actions to shape their preferences as in Reiter (2001), for example. All in all, the major difference that evolving over preferences rather than strategies makes in interpretation is that the agents in our model are fully rational and behave in a strictly optimal way at all points. The literature most closely related to the current paper relates to the evolutionary viability of altruism. In their seminal piece, Bergstrom and Stark (1993) consider a number of models, but focus on one in which benefits of altruistic actions are felt amongst groups of siblings. Selfish siblings are at an advantage over altruistic ones in the same family, but pass on their selfish genes to their children. Since groups of altruistic siblings are at an advantage over groups of selfish siblings, the momentary benefit or exploiting one's own altruistic sibling is far out weighed by the evolutionary disadvantage of having a completely selfish set of children. The altruistic genes end up being successful. Eshel, Samuelson and Shaked (1998) pick up on another model described in Bergstrom and Stark in which agents are arranged

---

[4] Recall that he voting propensity parameter (V_j) is a behavioral expression that reflects optimal altruistic actions of public-spirited agents. Thus, V_j is not a strategy, but rather a consequence of optimal voter choice. We treat the reduced form of the model, however.

in a circle and experience positive externalities when their direct neighbors choose to undertake costly altruistic actions. Agents choose a strategy each period by adopting the highest yielding action that they can directly observe. They show that for a correctly parameterized model, altruistic behavior survives and is stable against the introduction of mutations. Bester and Guth (1999) propose a model of externality producing duopolists. They show that if the production of one duopolist lowers the marginal cost of production for the other firm, then production choices are strategic complements. This means that when an altruistic firm chooses a higher than privately optimal production level, the other firm responses with its own higher production level, and this in turn benefits the first firm. Clearly, it is better to be selfish when paired with an altruist. Altruists, however, do much better when they happen to be paired with other altruists while egoists do much worse when they are paired with other egoists. As a result, altruists do better on the average, and are more successful from an evolutionary standpoint. (See also the comments of Bolle 2000 and Possajennikov 2000.)

There is common thread in all of these papers: local interaction. Eshel, Samuelson and Shaked's externalities extend only to adjacent neighbors, Bergstrom and Stark's only to groups of siblings, and Bester and Guth's only to pairs of duopolists. It is doubtful that any of these results could be generalized to more widespread externalities. What allows altruism to survive is the that the altruist gene is able to recapture some part of the external benefit of its behavior.[5] In Eshel, Samuelson and Shaked's case, it is through teaching one's neighbors to be altruist, for Bergstrom and Stark it is by producing kids who have an evolutionary advantage, and for Bester and Guth it is though the strategic complementarity. It may appear that the model we describe breaks with this and does indeed have widespread

---

[5] To be a bit more precise, recapturing benefits of altruism only needs to take place in a relative sense. For example, recapture happens if egoists are less likely to benefit from the acts of a-rational than other a-rational.

externalities. This is only partly true. Our two groups of voters each consist of a continuum of agents, and when a proposal passes, the costs and benefits that result are purely public in nature. In this sense, the externalities are widespread. Notice, however, that preferences and voting propensity are completely linked by construction in our model. Thus, while benefits of proposals that pass are spread across many individuals, they are in a sense localized within a given genotype. We conclude that the gene recaptures much of the externality even though the individuals themselves do not get an advantage from voting.

Although the mechanism that allows altruism to survive in our group selection setting is similar the one at work for local interaction models described above, there remains the key question regarding the robustness of the steady states to the occurrence of mutations. Unless the steady states we find can survive strategic experimentation and random genetic drift, there is little reason to believe that we would ever observe them as the outcome of any evolutionary process.

As it turns out, the steady states in which the high voter types prevail are robust to the introduction almost all types of mutants. To see this, suppose we are in a steady state in which the high voter type makes up the entire population. Now introduce a small fraction of mutants with tastes that differ from the dominant type. Because the mutants make up such a small fraction of the economy, they have a negligible effect on elections and the proposals most favored by the dominant type will continue to pass. Thus, *provided that the tastes of the mutant are sufficiently different* from the dominant type, they will get a systematically smaller payoff than the dominant type regardless of their propensity to vote and will not upset the steady state. On the other hand, if the mutants have the same (or at least very similar) tastes for public proposals as the dominant type, then they can successfully free-ride on their voting efforts. Thus, a mutant with the same tastes, but a lower voting parameter can upset the steady state and will eventually supplant the original dominant type. Observe, however, that the free riding mutants are in turn

16

vulnerable to even less public-spirited mutants who otherwise share their tastes.

At first glance, this may seem like bad news. This analysis suggests that no steady state with agents who have any positive voting propensity is an ESS. There is a kind of Gresham's Law at work in which bad citizens force out good ones. We believe that the news is not so bad, however, and there are at least two possible ways to address the fact that the steady states we discover in our model are not ESS.

First, notice that the mutants we are worried about have to have the same tastes but different voting propensity as the dominant type. There are reasonable arguments for why this may be an unlikely scenario in the real world. To the extent that preferences are literally based on genes, for example, it might be impossible to inherit a love of high levels of public spending without also having the public-spiritedness to vote. Both may be driven by the same "empathy" or "responsibility" gene, for example. To the extent that preferences are learned from the environment, the same argument might apply. Parents may teach their children to be empathetic and socially responsibly and this would inform both their voting behavior and preferences over public proposals. If a child rejects his parent's teaching or gets a truly mutant gene he would necessarily find himself equipped both with preferences over public proposals and voting propensities that differ from those of his parents. Even if such mixed mutations were possible, it might be that there are social sanctions that keep it from taking over. In other words, suppose a free rider arises. If this is detectable, the dominant group may protect itself by refusing to provide a mate for this mutant. After all, who wants a child to marry a selfish person? Or it may be that social sanctions imposed by the dominant group more than offset the gain the free rider receives from not voting.[6] Thus, even though having the high voter types

---

[6] See Harbaugh (1996) for some interesting evidence that social sanctions and rewards do play a role in getting people to vote, and that people even try to lie about their voting behavior to receive these rewards.

win the evolutionary game is not an ESS in a strict sense, there are still reasons to believe that this steady state my arise in real world settings.

Second, let us put aside the arguments just given, and suppose that free riding mutants with high type tastes can arise. The remarkable thing is that, provided that mutation happens slowly enough, this actually improves the welfare of the dominant type and in a sense does not threaten its evolutionary success. Consider the following dynamic story: Initially we have two types of agents, high and low, and a small leavening of all possible types of mutants. Suppose that we are in a situation that converges to the high types dominating the population. If the mutants are small enough in number, their presence is not enough to prevent the high type for forcing the low voter types close to extinction. The only agent type that manages to increase its population proportion is the free riding mutant with high type tastes. Eventually, of course, enough time passes that the free riding mutant replaces the high type. This mutant in turn is eventually replaced by an even less public-spirited mutant with the same taste as the original high type and so on until public-spiritedness converges to zero. Thus, tastes of the original high types are evolutionary stable, but the altruism is not in the long run.

Notice, however, that in the initial state, there was a compelling reason for the high type to vote. At the beginning of time, there are a significant fraction of low voter types with different tastes over public policies, and voting is needed to assure that the public proposals favored by the high voter type pass. As the low voter types begin to disappear, however, the high voters could win the elections even if they were less public-spirited since there are fewer of the low types to oppose them. Thus, in the steady state, *continuing to vote is socially wasteful because all of the opposition has been vanquished*. At this point, not only the individuals, but also the species itself benefits from having a lower voting parameter. In this modified environment, free riders can thrive without threatening the survival of their types.

We think of this as a kind of decline and fall of the Roman empire story. Ini-

tially, for Rome to thrive, its citizens must be vigilant and willing to make sacrifices for the common good. If the neighboring cities contain less public-spirited citizens, they will be conquered and added to the empire. Eventually, however, Rome will have vanquished all of its enemies, and then it is better for everyone to spend public money on bread and circuses instead of a large standing army. Public-spirited sacrifice ceases to serve a useful purpose and it is time for Romans to rest on their laurels. The key, however, is to make sure that all of Rome's enemies have been destroyed before this decline into decadence. If the decline happens before all the Gauls have been pacified, decline turns into fall.

5. Conclusion

One feature of our model which may be open to criticism is that we find that only one type of agent can survive in the steady state. In reality, however, we seldom observe a completely homogeneous society. An interesting extension of our model might be to assume that agents experience diminishing marginal utility of projects. In this case, the benefits accruing whichever type of voter makes up the winning coalition would decline, while the prospective benefits of winning an election to the opposition group would remain high. This would suppress the winning coalition's turnout, make it more likely the opposition would begin to win election, and slow the winning coalition's grow rate even if they continued to win. It might be possible to find a stable interior solution in which both types of agents persist for such a model. Another interesting generalization would be allow more than two types. Simulation results suggest that if the groups are equally numerous and preferences over public policies are uncorrelated, then the type with the highest voting propensity will prevail in the evolutionary game. It is harder to prove theorems about this case,

however, as the initial conditions can vary widely, and it is not immediately clear which are the most compelling benchmark cases.

Our work is motivated by our interpretation of the literature which suggests that it is difficult to explain observed voting behavior on the basis of rational choice unless one assumes that agents get utility from the act of voting itself. In this paper we have attempted to provide foundation for the warm glow associated with behaving in a public-spirited manner using evolutionary game theory. The basic result is that being public-spirited can confer an evolutionary advantage. Having a high propensity to vote is more advantageous when voting is less costly, when your group's preferences over public project differ sharply from those of competing groups, and when the competing group is less public-spirited or less numerous. We conclude that Evolutionary forces may indeed play a role in causing agents to internalize the benefits their actions confer on there fellow agents, at least to the degree that they share a common set of preferences.

Appendix

Derivation of Payoff Functions

The preferences could be positively or negatively correlated. We first give the payoff functions in each case and then illustrate the procedure for computing the functions when preferences are negatively correlated. The procedure for computing the payoff functions when the preferences are positively correlated is very similar.

In the case where preferences are negatively correlated we have $B_L = B_L^n$ where $0 \leq \alpha \leq 1$.

Define $\theta = \frac{(1-\alpha)(1-S_H)}{\alpha-(\alpha+\beta)S_H}$.

Then the expected payoff for each group as S_H varies is as follows:

- For $0 \leq \alpha \leq \frac{1}{2}$:

$$\bar{\pi}_H = \begin{cases} -\frac{1}{6\theta} - \frac{V_H C}{2} & \text{for } 0 < S_H \leq \frac{1}{1+\beta} \\ .25 - \frac{\theta^2}{12} - \frac{V_H C}{2} & \text{for } \frac{1}{1+\beta} < S_H \leq 1 \end{cases}$$

$$\bar{\pi}_L = \begin{cases} \frac{\alpha}{6\theta} - \frac{1-\alpha}{12\theta^2} + \frac{1-\alpha}{4} - \frac{V_L C(4\alpha^2 - 6\alpha + 3)}{6(1-\alpha)} & \text{for } 0 \leq S_H \leq \frac{1}{1+\beta} \\ \frac{\alpha\theta^2}{12} - \frac{(1-\alpha)\theta}{6} - \frac{\alpha}{4} - \frac{V_L C(4\alpha^2 - 6\alpha + 3)}{6(1-\alpha)} & \text{for } \frac{1}{1+\beta} < S_H < 1 \end{cases}$$

- For $\frac{1}{2} < \alpha < 1$:

$$\bar{\pi}_H = \begin{cases} -.25 + \frac{\theta^2}{12} - \frac{V_H C}{2} & \text{for } 0 < S_H < \frac{2\alpha-1}{(2\alpha-1)+\beta} \\ -\frac{1}{6\theta} - \frac{V_H C}{2} & \text{for } \frac{2\alpha-1}{(2\alpha-1)+\beta} \leq S_H \leq \frac{1}{1+\beta} \\ .25 - \frac{\theta^2}{12} - \frac{V_H C}{2} & \text{for } \frac{1}{1+\beta} < S_H \leq 1 \end{cases}$$

$$\bar{\pi}_L = \begin{cases} -\frac{\alpha\theta^2}{12} + \frac{(1-\alpha)\theta}{6} + \frac{\alpha}{4} - \frac{V_L C(4\alpha^2 - 2\alpha + 1)}{6\alpha} & \text{for } 0 \leq S_H < \frac{2\alpha-1}{(2\alpha-1)+\beta} \\ \frac{\alpha}{6\theta} - \frac{1-\alpha}{12\theta^2} + \frac{1-\alpha}{4} - \frac{V_L C(4\alpha^2 - 2\alpha + 1)}{6\alpha} & \text{for } \frac{2\alpha-1}{(2\alpha-1)+\beta} \leq S_H \leq \frac{1}{1+\beta} \\ \frac{\alpha\theta^2}{12} - \frac{(1-\alpha)\theta}{6} - \frac{\alpha}{4} - \frac{V_L C(4\alpha^2 - 2\alpha + 1)}{6\alpha} & \text{for } \frac{1}{1+\beta} < S_H < 1 \end{cases}$$

In the case where preferences are positively correlated we have $B_L = B_L^p$ where $0 \leq \alpha \leq 1$. Let's denote $\theta = -\frac{\alpha(1-S_H)}{(1-\alpha)+S_H(\beta+\alpha-1)}$. Then the expected payoff for each group as S_H varies is as follows:

- For $\alpha = 0$

$$\bar{\pi}_H = .25 - \frac{V_H C}{2} \text{ for } 0 < S_H \leq 1$$

$$\bar{\pi}_L = .25 - \frac{V_L C}{2} \text{ for } 0 \leq S_H < 1$$

- For $0 < \alpha \leq \frac{1}{2}$

$$\bar{\pi}_H = .25 - \frac{\theta^2}{12} - \frac{V_H C}{2} \text{ for } 0 < S_H \leq 1$$

$$\bar{\pi}_L = .25 - \frac{(1-\alpha)\theta^2}{12} - \frac{\alpha\theta}{6} - \frac{\alpha}{4} - D_1 V_L C \text{ for } 0 \leq S_H < 1$$

- For $\frac{1}{2} < \alpha \leq 1$

$$\bar{\pi}_H = \begin{cases} -\frac{1}{6\theta} - \frac{V_H C}{2} & \text{for } 0 < S_H \leq \frac{(2\alpha-1)}{(2\alpha-1)+\beta} \\ .25 - \frac{\theta^2}{12} - \frac{V_H C}{2} & \text{for } \frac{(2\alpha-1)}{(2\alpha-1)+\beta} < S_H \leq 1 \end{cases}$$

$$\bar{\pi}_L = \begin{cases} -\frac{(1-\alpha)}{6\theta} - \frac{\alpha}{12\theta^2} + \frac{\alpha}{4} - D_2 V_L C & \text{for } 0 \leq S_H < \frac{(2\alpha-1)}{(2\alpha-1)+\beta} \\ .25 - \frac{(1-\alpha)\theta^2}{12} - \frac{\alpha\theta}{6} - \frac{\alpha}{4} - D_2 V_L C & \text{for } \frac{(2\alpha-1)}{(2\alpha-1)+\beta} \leq S_H < 1 \end{cases}$$

21

Note that $\bar{\pi}_H$ and $\bar{\pi}_L$ in both cases are continuous and well behaved functions of S_H, the proportion of high voting parameter type.

Here is an outline of how the payoffs for each agent is calculated when $B_L = B_L^n$. The calculations when preferences are positively correlated are very similar. Let's denote the probability that a proposal passes by P. By the definition of P we have:

$$P = Prob(S_H V_H B_H + S_L V_L B_L^n > 0 \mid B_H = b_H)$$
$$= Prob(S_H V_H b_H + (1 - S_H) V_L (-\alpha b_H + (1 - \alpha) U_I > 0)$$
$$= Prob(U_I > \frac{\alpha(1 - S_H) - \beta S_H}{(1 - \alpha)(1 - S_H)} b_H)$$
$$= Prob(U_I > \frac{b_H}{\theta}) = 1 - Prob(U_I < \frac{b_H}{\theta})$$
$$= \begin{cases} 1 & \text{for } \frac{b_H}{\theta} \leq -1 \\ \frac{1}{2} - \frac{b_H}{2\theta} & \text{for } -1 < \frac{b_H}{\theta} < 1 \\ 0 & \text{for } \frac{b_H}{\theta} \geq 1 \end{cases}$$

Thus the probability that a given proposal passes is a function of S_H. The value of P for all $-1 \leq b_H \leq 1$ (unless otherwise stated) as S_H varies over 0 and 1 is as follows:

- If $0 \leq \alpha \leq \frac{1}{2}$:

  ∘ For $0 \leq S_H \leq \frac{1}{1+\beta}$ we have

  $$P = \frac{1}{2} - \frac{b_H}{2\theta}.$$

  ∘ For $\frac{1}{1+\beta} < S_H < 1$ we have:

  $$P = \begin{cases} 0 & \text{for } -1 \leq b_H \leq \theta \\ \frac{1}{2} - \frac{b_H}{2\theta} & \text{for } \theta < b_H < -\theta \\ 1 & \text{for } -\theta \leq b_H \leq 1 \end{cases}$$

  ∘ For $S_H = 1$ we have:

  $$P = \begin{cases} 1 & \text{for } 0 < b_H \leq 1 \\ 0 & \text{for } -1 \leq b_H \leq 0 \end{cases}$$

- If $\frac{1}{2} < \alpha \leq 1$:

  ∘ For $0 \leq S_H < \frac{2\alpha-1}{(2\alpha-1)+\beta}$ we have:

  $$P = \begin{cases} 1 & \text{for } -1 \leq b_H \leq -\theta \\ \frac{1}{2} - \frac{b_H}{2\theta} & \text{for } -\theta < b_H < \theta \\ 0 & \text{for } \theta \leq b_H \leq 1 \end{cases} .$$

22

○ For $\frac{2\alpha-1}{(2\alpha-1)+\beta} \leq S_H \leq \frac{1}{1+\beta}$,

$$P = \frac{1}{2} - \frac{b_H}{2\theta}.$$

○ For $\frac{1}{1+\beta} < S_H < 1$ we have:

$$P = \begin{cases} 1 & \text{for } -\theta \leq b_H \leq 1 \\ \frac{1}{2} - \frac{b_H}{2\theta} & \text{for } \theta < b_H < -\theta \\ 0 & \text{for } -1 \leq b_H \leq \theta \end{cases}$$

○ For $S_H = 1$ we have:

$$P = \begin{cases} 1 & \text{for } 0 < b_H \leq 1 \\ 0 & \text{for } -1 \leq b_H \leq 0 \end{cases}$$

The payoff that a high voting parameter agent can expect for a given proposal is:

$$
\begin{aligned}
E(\pi_H \mid B_H = b_H) &= P(b_H - C_H) + (1-P)(-C_H) \\
&= Pb_H - C_H
\end{aligned}
$$

Therefore the average payoff of a high voting type agent over all possible values of b_H is:

$$
\begin{aligned}
\bar{\pi}_H &= E_{b_H}[E(\pi_H \mid B_H = b_H)] \\
&= \int_{-1}^{1} \frac{(Pb_H - C_H)}{2} db_H
\end{aligned}
$$

Here we show the results only in the case for $0 < \alpha \leq \frac{1}{2}$ and when $\frac{1}{1+\beta} < S_H < 1$. The procedure in other cases would be the same. As it was shown above in this case we have:

$$P = \begin{cases} 1 & \text{for } -\theta \leq b_H \leq 1 \\ \frac{1}{2} - \frac{b_H}{2\theta} & \text{for } \theta < b_H < -\theta \\ 0 & \text{for } -1 \leq b_H \leq \theta \end{cases}$$

Substitution gives:

$$\bar{\pi}_H = \int_{b_H=-1}^{b_H=\theta} \frac{-V_H C \mid b_H \mid}{2} db_H + \int_{b_H=\theta}^{b_H=-\theta} \left(\frac{b_H}{2}(\frac{1}{2} - \frac{b_H}{2\theta}) - \frac{V_H C \mid b_H \mid}{2} \right) db_H +$$

$$\int_{b_H=-\theta}^{b_H=1} \left(\frac{b_H}{2} - \frac{V_H C \mid b_H \mid}{2} \right) db_H$$

$$= .25 - \frac{\theta^2}{12} - \frac{V_H C}{2}$$

which is the desired result. Note that

$$\int_{b_H=-1}^{b_H=1} \mid b_H \mid dP_H = 1$$

For the calculation of the low voting parameter agent payoff we take a different route. The payoff a low voting type can expect for a given proposal, $B_H = b_H$ is:

$$E(\pi_L \mid B_H = b_H) = \int_{-1}^{\frac{b_H}{\theta}} \frac{-C_L}{2} du_I + \int_{\frac{b_H}{\theta}}^{1} \frac{(b_L^n - C_L)}{2} du_I$$

After a change of variable and taking expectation over all possible values of b_H we will have:

$$\bar{\pi}_L = E_{b_H}[E(\pi_L \mid B_H = b_H)]$$
$$= \frac{1}{4(1-\alpha)} \left(\int_{-1}^{1} \int_{\frac{b_H(1-\alpha-\theta\alpha)}{\theta}}^{-\alpha b_H+(1-\alpha)} b_L^n db_L^n db_H - \int_{-1}^{1} \int_{-\alpha b_H-(1-\alpha)}^{-\alpha b_H+(1-\alpha)} C_L db_L^n db_H \right)$$

Now let's calculate $\bar{\pi}_L$ in the case where $0 < \alpha \le \frac{1}{2}$ and $\frac{1}{1+\beta} < S_H < 1$. First the expected payoff for the low voting parameter for a given proposal $B_H = b_H$ is:

bullet for $-1 \le b_H \le \theta$, we have $\frac{b_H}{\theta} > 1$, thus:

$$E[\pi_L \mid B_H = b_H] = 0 - \int_{-1}^{1} \frac{V_L C \mid b_L^n \mid}{2} du_I$$

24

Here we make a change of variable. Let's denote $a_2 = -\alpha b_H - (1 - \alpha)$ and $a_3 = -\alpha b_H + (1 - \alpha)$. At $u_I = 1$, $B_L^n = a_3$ and at $u_I = -1$, $B_L^n = a_2$ and $du_I = \frac{db_L^n}{1-\alpha}$. Hence:

$$E[\pi_L \mid B_H = b_H] = -\int_{a_2}^{a_3} \frac{V_L C \mid b_L^n \mid}{2(1 - \alpha)} db_L^n$$

• for $\theta < b_H < -\theta$, we have $-1 < \frac{b_H}{\theta} < 1$, thus:

$$E[\pi_L \mid B_H = b_H] = \int_{\frac{b_H}{\theta}}^{1} \frac{b_L^n}{2} du_I - \int_{-1}^{1} \frac{V_L C \mid b_L^n \mid}{2} du_I$$

By the same change of variable and denoting $a_1 = \frac{1-\alpha-\alpha\theta}{\theta} b_H$ we will have:

$$E[\pi_L \mid B_H = b_H] = \frac{1}{2(1 - \alpha)} \left(\int_{a_1}^{a_2} b_L^n db_L^n - \int_{a_2}^{a_3} V_L C \mid b_L^n \mid db_L^n \right)$$

• for $-\theta \leq b_H \leq 1$, we have $\frac{b_H}{\theta} \leq -1$, thus:

$$E[\pi_L \mid B_H = b_H] = \int_{-1}^{1} \frac{b_L^n}{2} du_I - \int_{-1}^{1} \frac{V_L C \mid b_L^n \mid}{2} du_I$$

after change of variables we will have:

$$E[\pi_L \mid B_H = b_H] = \frac{1}{2(1 - \alpha)} \left(\int_{a_2}^{a_3} b_L^n db_L^n - \int_{a_2}^{a_3} V_L C \mid b_L^n \mid db_L^n \right)$$

Now we find $\bar{\pi}_L$, the expected payoff for a low voting parameter agent by taking expectation over all possible values of b_H.

$$\bar{\pi}_L = \int_{b_H=\theta}^{b_H=-\theta} \int_{b_L^n=a_1}^{b_L^n=a_3} \frac{b_L^n}{4(1 - \alpha)} db_L^n db_H +$$

$$\int_{b_H=-\theta}^{b_H=1} \int_{b_L^n=a_2}^{b_L^n=a_3} \frac{b_L^n}{4(1 - \alpha)} db_L^n db_H - \frac{V_L C}{4(1 - \alpha)} \left(\int_{b_H=-1}^{b_H=1} \int_{b_L^n=a_2}^{b_L^n=a_3} \mid b_L^n \mid db_L^n db_H \right)$$

$$= \frac{\alpha\theta^2}{12} - \frac{(1 - \alpha)\theta}{6} - \frac{\alpha}{4} - \frac{V_L C(4\alpha^2 - 6\alpha + 3)}{6(1 - \alpha)}$$

which is the desired result. Note that:

$$\int\limits_{b_H=-1}^{b_H=1} \int\limits_{b_L^n=a_2}^{b_L^n=a_3} \mid b_L^n \mid db_L^n db_H = \frac{8\alpha^2 - 12\alpha + 6}{3}$$

As mentioned before the other cases are computed in a very similar manner.

Proofs of Theorems

Theorem 1. *Depending on values of parameters α, β and C there are three possible outcomes for the system:*

1. **High type wins:** *The system has two steady states $S_H = 0$ and $S_H = 1$ where $S_H = 1$ is globally stable and $S_H = 0$ is unstable.*
2. **Large population wins:** *The system has three steady states, $S_H = 0$, $S_H = 1$ and $S_H = S_H^* \in (0,1)$ where $S_H = 0$ and $S_H = 1$ are asymptotically stable and their basin of attraction are $[0, S_H^*)$ and $(S_H^*, 1]$ respectively, and $S_H = S_H^*$ is unstable.*

3. **Low type wins:** *The system has two steady states $S_H = 0$ and $S_H = 1$ where $S_H = 0$ is globally stable and $S_H = 1$ is unstable.*

Proof/

The steady states are solution to $\dot{S}_H = 0$. The replicator dynamics can be written as follows:

$$\dot{S}_H = S_H(\bar{\pi}_H - \bar{\bar{\pi}}) = S_H(1 - S_H)(\bar{\pi}_H - \bar{\pi}_L)$$

Thus $S_H = 0$ and $S_H = 1$ are always steady states. The other steady state, if it exists, is the solution to $\bar{\pi}_H - \bar{\pi}_L = 0$. The calculations are a little tedious, but the results are straightforward to verify. We show the calculations in detail for one case when large population dominates. For the case where high type (low type) wins refer to the proof of next theorems.

Let's consider the case where the preferences are negatively correlated, i. e. $B_L = B_L^N$ and $0 < \alpha \leq \frac{1}{2}$. Substituting the values of payoff functions in $\bar{\pi}_H - \bar{\pi}_L = 0$ when $0 < S_H \leq \frac{1}{1+\beta}$, We get:

$$F_2\theta^2 + 2(1 + \alpha)\theta - (1 - \alpha) = 0$$

Where: $F_2 = 3(1 - \alpha) + 6V_L C(\beta - \frac{(4\alpha^2 - 6\alpha + 3)}{3(1-\alpha)}) > 0$

Solution to the above equation gives us θ and Given the definition of θ we will get: $S_H^* = \frac{1-\alpha-\alpha\theta}{1-\alpha-(\alpha+\beta)\theta}$ Solving the equation we get two solutions for θ:

$$\theta_1 = -\frac{(1+\alpha) + \sqrt{(1+\alpha)^2 + F_2(1-\alpha)}}{F_2}$$

$$\theta_2 = -\frac{(1+\alpha) - \sqrt{(1+\alpha)^2 + F_2(1-\alpha)}}{F_2}$$

Of these $0 < \theta_2 < 1$ which considering the definition of θ means that $S_H < 0$. So this is not permissable as a solution. It is easy to show that if $C < C_1^*$ where $C_1^* = \frac{2\alpha}{3V_H - \frac{V_L(4\alpha^2 - 6\alpha + 3)}{1-\alpha}}$ then $\theta_1 < -1$ which guarantees that $\frac{\alpha}{\alpha+\beta} < S_H^* \leq \frac{1}{1+\beta}$

For finding the stability properties of steady states we basically find the sign of $\dot{S}_H$ which is equivalent to finding the sign of $\bar{\pi}_H - \bar{\pi}_L$ for different values of S_H. In this case we have:

$$\bar{\pi}_H - \bar{\pi}_L = -(F_2\theta^2 + 2(1+\alpha)\theta - (1-\alpha))$$

The function $F_2\theta^2 + 2(1+\alpha)\theta - (1-\alpha)$ is positive for all values of $\theta < \theta_1$ and $\theta > \theta_2$. It is negative for all the values of $\theta_1 < \theta < \theta_2$. Thus $\dot{S}_H$ is negative for all $0 < S_H < S_H^*$ and positive for $S_H^* < S_H \leq 1$.

∎

Theorem 2. *All else equal, the higher the cost of voting C, the less likely the high voter types will win the evolutionary game.*

Proof/

We show the steady states and their stability properties for different levels of cost of voting. The calculation are similar to ones illustrated in the previous theorem. first we consider the case where preferences are negatively correlated. Let's denote $D_1 = \frac{4\alpha^2 - 6\alpha + 3}{6(1-\alpha)}$, $D_2 = \frac{4\alpha^2 - 2\alpha + 1}{6\alpha}$, $F_1 = 3(1+\alpha) - 6V_L c(\beta - 2D_1)$, $F_2 = 3(1-\alpha) + 6V_L c(\beta - 2D_1)$, $F_3 = 3(1-\alpha) + 6V_L c(\beta - 2D_2)$ and $F_4 = 3(1+\alpha) - 6V_L c(\beta - 2D_2)$. Let also define $C_1^* = \frac{2\alpha}{3V_L(\beta - 2D_1)}$, $C_2^* = \frac{1+\alpha}{2V_L(\beta - 2D_1)}$, $C_3^* = \frac{2\alpha}{3V_L(\beta - 2D_2)}$ and $C_4^* = \frac{1+\alpha}{2V_L(\beta - 2D_2)}$. It can be verified that $C_1^* < C_2^*$ and $C_3^* < C_4^*$. It can be shown that:

- for $0 \leq \alpha \leq \frac{1}{2}$:
  - If $0 \leq C \leq C_1^*$, then there are three steady states $S_H = 0$, $S_H = 1$ and $S_H = S_H^* \leq \frac{1}{1+\beta}$, where $S_H^* = \frac{1-\alpha-\alpha\lambda_1}{1-\alpha-(\alpha+\beta)\lambda_1}$ and $\lambda_1 = -\frac{(1+\alpha) + \sqrt{(1+\alpha)^2 + F_2(1-\alpha)}}{F_2}$. Of these $S_H = 0$ and $S_H = 1$ are asymptotically stable and their basin of attraction are $[0, S_H^*)$ and $(S_H^*, 1]$ respectively. $S_H = S_H^*$ is unstable.
  - If $C_1^* < C < C_2^*$, then there are three steady states $S_H = 0$, $S_H = 1$ and $\frac{1}{1+\beta} < S_H = S_H^* < 1$, where $S_H^* = \frac{1-\alpha-\alpha\lambda_2}{1-\alpha-(\alpha+\beta)\lambda_2}$ and $\lambda_2 = \frac{(1-\alpha) - \sqrt{(1-\alpha)^2 + F_1(1+\alpha)}}{(1+\alpha)}$. Of these $S_H = 0$ and $S_H = 1$ are asymptotically

27

stable and their basin of attraction are $[0, S_H^*)$ and $(S_H^*, 1]$ respectively. $S_H = S_H^*$ is unstable.

o If $C \geq C_2^*$, then there are two steady states $S_H = 0$ and $S_H = 1$. Of these $S_H = 0$ is globally stable and $S_H = 1$ is unstable.

- for $\frac{1}{2} < \alpha \leq 1$:

o If $0 \leq C \leq C_3^*$, then there are three steady states $S_H = 0$, $S_H = 1$ and $S_H = S_H^*$, Where $\frac{2\alpha-1}{(2\alpha-1)+\beta} < S_H^* = \frac{1-\alpha-\alpha\omega_1}{1-\alpha-(\alpha+\beta)\omega_1} \leq \frac{1}{1+\beta}$ and $\omega_1 = -\frac{(1+\alpha)+\sqrt{(1+\alpha)^2+F_3(1-\alpha)}}{F_3}$. Of these $S_H = 0$ and $S_H = 1$ are asymptotically stable and their basin of attraction are $[0, S_H^*)$ and $(S_H^*, 1]$ respectively. $S_H = S_H^*$ is unstable.

o If $C_3^* < C < C_4^*$, then there are three steady states $S_H = 0$, $S_H = 1$ and $S_H = S_H^*$, Where $\frac{1}{1+\beta} < S_H^* = \frac{1-\alpha-\alpha\omega_2}{1-\alpha-(\alpha+\beta)\omega_2} \leq 1$ and $\omega_2 = \frac{(1-\alpha)-\sqrt{(1-\alpha)^2+F_4(1+\alpha)}}{(1+\alpha)}$. Of these $S_H = 0$ and $S_H = 1$ are asymptotically stable and their basin of attraction are $[0, S_H^*)$ and $(S_H^*, 1]$ respectively. $S_H = S_H^*$ is unstable.

o If $C \geq C_4^*$, then there are two steady states $S_H = 0$ and $S_H = 1$. Of these $S_H = 0$ is globally stable and $S_H = 1$ is unstable.

Now consider the case when the preferences are positively correlated. i. e. the preferences of high and low voting parameter types are B_H and B_L^p respectively. In what follows we denote $\theta = -\frac{\alpha(1-S_H)}{(1-\alpha)+S_H(\beta+\alpha-1)}$, $k_1 = 6V_Lc(\beta - 2d) - 3\alpha$, $k_2 = 6V_Lc(\beta - 2e) - 3\alpha$, $\delta_1 = \frac{\alpha-\sqrt{\alpha^2-\alpha k_1}}{\alpha}$, $\delta_2 = \frac{\alpha-\sqrt{\alpha^2-\alpha k_2}}{\alpha}$ and as defined above $D_1 = \frac{4\alpha^2-6\alpha+3}{6(1-\alpha)}$, $D_2 = \frac{4\alpha^2-2\alpha+1}{6\alpha}$. Let also denote $\hat{C}_1 = \frac{4\alpha^3-8\alpha^2+3\alpha}{6V_L(1-\alpha)^2(\beta-2d)}$, $\hat{C}_2 = \frac{\alpha}{2V_L(\beta-2d)}$ and $\tilde{C} = \frac{\alpha}{2V_L(\beta-2e)}$. The steady states and their stability properties are also found in a similar way and are as follows:

- For $\alpha = 0$

o there are two steady states $S_H = 0$ and $S_H = 1$. Of these $S_H = 0$ is globally stable and $S_H = 1$ is unstable.

- For $0 < \alpha \leq \frac{1}{2}$

o If $C \leq \hat{C}_1$, then there are two steady states $S_H = 0$ and $S_H = 1$. Of these $S_H = 0$ is unstable and $S_H = 1$ is globally stable.

o If $\hat{C}_1 < C < \hat{C}_2$, then there are three steady states, $S_H = 0$, $S_H = 1$ and $S_H = S_H^*$, where $S_H^* = \frac{\alpha\delta_1-\delta_1-\alpha}{\delta_1(\alpha+\beta)-\delta_1-\alpha}$. Of these $S_H = 0$ and $S_H = 1$ are asymptotically stable and $S_H = S_H^*$ is unstable. The basin of attraction for $S_H = 0$ and $S_H = 1$ are $[0, S_H^*)$ and $(S_H^*, 1]$ respectively.

o If $C \geq \hat{C}_2$, then there are two steady states $S_H = 0$ and $S_H = 1$. Of these $S_H = 1$ is unstable and $S_H = 0$ is globally stable.

- For $\frac{1}{2} < \alpha \leq 1$

o If $0 \leq C < \tilde{C}$, then there are three steady states, $S_H = 0$, $S_H = 1$ and

28

$S_H = S_H^*$, where $\bar{S}_H = \frac{\alpha\delta_2 - \delta_2 - \alpha}{\delta_2(\alpha+\beta) - \delta_2 - \alpha}$. Of these $S_H = 0$ and $S_H = 1$ are asymptotically stable and $S_H = S_H^*$ is unstable. The basin of attraction for $S_H = 0$ and $S_H = 1$ are $[0, S_H^*)$ and $(S_H^*, 1]$ respectively.

o If $C \geq \tilde{C}$, then there are two steady states $S_H = 0$ and $S_H = 1$. Of these $S_H = 1$ is unstable and $S_H = 0$ is globally stable.

∎

Theorem 3. *All else equal, more public spirited the high voter types are compared to the low voter types, the more likely the high voter types will win the evolutionary game.*

<u>Proof/</u>

As an example consider the case when $B_L = B_l^n$ and $0 < \alpha \leq \frac{1}{2}$. We define $C_1^* = \frac{2\alpha}{3V_L(\beta - 2D_1)}$ and $C_2^* = \frac{1+\alpha}{2V_L(\beta - 2D_1)}$. As we saw in the proof to previous theorem if $0 \leq C \leq C_1^*$, then there are three steady states $S_H = 0$, $S_H = 1$ and $S_H = \rho_1 \leq \frac{1}{1+\beta}$, Where $\rho_1 = \frac{1-\alpha-\alpha\lambda_1}{1-\alpha-(\alpha+\beta)\lambda_1}$ and $\lambda_1 = -\frac{(1+\alpha)+\sqrt{(1+\alpha)^2 + F_2(1-\alpha)}}{F_2}$. Of these $S_H = 0$ and $S_H = 1$ are asymptotically stable and their basin of attraction are $[0, \rho_1)$ and $(\rho_1, 1]$ respectively. $S_H = \rho_1$ is unstable. If $C_1^* < C < C_2^*$, then there are three steady states $S_H = 0$, $S_H = 1$ and $\frac{1}{1+\beta} < S_H = \rho_2 < 1$, Where $\rho_2 = \frac{1-\alpha-\alpha\lambda_2}{1-\alpha-(\alpha+\beta)\lambda_2}$ and $\lambda_2 = \frac{(1-\alpha)-\sqrt{(1-\alpha)^2 + F_1(1+\alpha)}}{(1+\alpha)}$. Of these $S_H = 0$ and $S_H = 1$ are asymptotically stable and their basin of attraction are $[0, \rho_2)$ and $(\rho_2, 1]$ respectively. $S_H = \rho_2$ is unstable. And lastly If $C \geq C_2^*$, then there are two steady states $S_H = 0$ and $S_H = 1$. Of these $S_H = 0$ is globally stable and $S_H = 1$ is unstable. Now given everything else remains constant as β approaches ∞ (by keeping V_H constant and reducing V_L), C_1^* approaches $\frac{2\alpha}{3V_H}$ and C_2^* approaches $\frac{1+\alpha}{2V_H}$. Also it is easy to see both ρ_1 and ρ_2 approach zero. Now as long as C is less than $\frac{1+\alpha}{2V_H}$ then as β increases S_H^* approaches zero. In other cases the argument would be the same. For $C > \frac{1+\alpha}{2V_H}$, low voting type will dominate. In the case of positive preferences, as long as $C < \frac{\alpha}{2V_H}$, as β increases, S_H^* approaches zero.

∎

Theorem 4. *All else equal, more correlated the preferences of the two types of voters, the less likely the high voter types will win the evolutionary game.*

<u>Proof/</u>

We start with considering the case when preferences are perfectly negatively correlated. Let's define $C^* = \frac{1}{V_H - V_L}$. The steady states and their properties are as follows:

• For $C < C^*$, There are two steady states $S_H = 0$ and $S_H = 1$. Both the steady states are asymptotically stable. The basin of attraction for $S_H = 0$ is $[0, \frac{1}{1+\beta}]$ and for $S_H = 1$ it is $(\frac{1}{1+\beta}, 1]$.

- For $C = C^*$, the steady states are $S_H = 0$ and all the points $\frac{1}{1+\beta} < S_H \leq 1$. For any initial value of $S_H \leq \frac{1}{1+\beta}$ the group with low voting parameter is going to dominate. For values of $S_H > \frac{1}{1+\beta}$ the proportions of each group will not change over time.
- For $C > C^*$, there are two steady states $S_H = 0$ which is globally stable and $S_H = 1$ which is unstable.

  In this case the payoff for each group will be:

$$
\bar{\pi}_H = \begin{cases} -.25 - \frac{V_H c}{2} & \text{for } 0 \leq S_H < \frac{1}{1+\beta} \\ -\frac{V_H c}{2} & \text{for } S_H = \frac{1}{1+\beta} \\ .25 - \frac{V_H c}{2} & \text{for } \frac{1}{1+\beta} < S_H \leq 1 \end{cases}
$$

$$
\bar{\pi}_L = \begin{cases} .25 - \frac{V_L c}{2} & \text{for } 0 \leq S_H < \frac{1}{1+\beta} \\ -\frac{V_L c}{2} & \text{for } S_H = \frac{1}{1+\beta} \\ -.25 - \frac{V_L c}{2} & \text{for } \frac{1}{1+\beta} < S_H \leq 1 \end{cases}
$$

The steady states and their properties are derived with similar procedures to above theorems and lemmas. Note that in this case when α approaches 1, depending on the value of C, S_H^* might initially be above $\frac{1}{1+\beta}$, then decrease below it and then increases back and settles down at it. Now for given values of V_H and V_L if $C > C^*$, no matter the value of α, the low voting type will dominate the population. But for values of $0 < C < C^*$, as α decreases, S_H^* eventually approaches 1. In the case that preferences are positively correlated, for given values of V_H and V_L and $C > \frac{C^*}{2}$, no matter the value of α, the low voting type will dominate the population. But for values of $0 < C < \frac{C^*}{2}$, as α decreases, S_H^* eventually approaches 1. note that S_H^* initially start at a value below 1, and as α decreases it might further go down but it eventually settles down at 1.

∎

References

Aldrich, John H. (1997): "When is it rational to vote," in *Perspectives on Public Choice,* by Dennis C. Muller Cambridge University Press, chapter 17.

Andreoni, James (1990): "Impure Altruism and Donations to Public Choice: A Theory of Warm-Glow Giving," *The Economic Journal,* 100: 464-477.

Andreoni, James (1995): "warm-glow versus cold-prickle: the effects of positive and negative framing on cooperation in experiments," *The Quarterly Journal Of Economics,* pp. 1-21.

Becker, Gary (1976): "Altruism, Egoism, and Genetic Fitness: Economics and Sociobiology," *Journal of Economic Literature,* V. 14, pp. 817-26.

Bergstrom, Theodore C. and Stark Oded (1993): "How can Altruism prevail in an Evolutionary Environment," *American Economic Review,* v.83 pp. 149-155.

Bester, Helmut and Guth, Werner (1998): "Is altruism evolutionarily stable?," *Journal Of Economic Behavior & Organization,* pp. 193-209.

Bolle, Friedel (2000): "Is altruism evolutionarily stable? And envy and malevolence? Remarks on Bester and Guth," *Journal Of Economic Behavior & Organization,* pp. 131-133.

Conley, John and Akram Temimi (2000): "Endogenous Enfranchisement when Groups' Preferences Conflict," *Journal of Political Economy,* V. 109, pp. 79-102.

Downs, Anthony (1957): *An Economic Theory of Democracy.* New York: Harper and Row.

Eshel, Ilan ; Samuelson, Larry and Shaked Avner (1998): "Altruists, Egoists, and Hooligans in a Local Interaction Model," *The American Economic Review,* v.88 pp. 157-179.

Ferejohn, John A. , and Morris P. Fiorina (1974): "The Paradox of Not Voting: A Decision Theoretic Analysis," *American Political Science Review,* 68:525-36.

Friedman, Daniel (1991): "Evolutionary Games in Economics," *Econometrica,* pp. 637-666.

Fudenberg, Drew, and David K. Levine (1998): *The Theory of Learning in Games.* The MIT press.

Green, Donald P. , and Ian Shapiro (1994): *Pathologies of Rational Choice Theory: A Critique of Applications in Political Science.* Yale University Press.

Hirshleifer, Jack (1977): "Economics from a Biological Viewpoint," *Journal of Law and Economics,* V. 20, pp 1-52.

Hirshleifer, Jack (1978): "Natural Economics Versus Political Economy," *Journal of Social and Biological Structures.,* V. 20, pp 1-52.

Harbaugh, William (1996): "If people vote because they like to, then why do so many of them lie?," *Public Choice,* ???.

Lagunoff, Roger (2000): "On the Evolution of Pareto Optimal Behavior in Repeated Coordination Problems," *International Economic Review,* v. 41 pp. 273-93.

Ledyard, John O. (1981): "The Paradox of Voting and Candidate Competition: A general Equilibrium Analysis," in *Essays in Contemporary Fields of Economics,* by G. Horwich and J. Quirk West Lafayette: Purdue University Press.

Ledyard, John O. (1984): "The Pure Theory of Large Two-Candidate Elections," *Public Choice,* 44:7-41.

Muller, Dennis C. (1989): *Public Choice II: A revised Edition of Public Choice.* Cambridge University Press.

Palfrey, Thomas R. , and Howard Rosenthal (1983): "A Strategic Calculus of Voting," *Public Choice,* 41:7-53.

Palfrey, Thomas R. , and Howard Rosenthal (1985): "Voter Participation and Voter Uncertainty," *American Political Science Review,* 79:62-78.

Possajennikov Alex (2000): "On the evolutionary stability of altruistic and spiteful preferences," *Journal Of Economic Behavior & Organization,* pp. 125-129.

Riker, William H. , and Peter C. Ordeshook (1968): "A Theory of the Calculus of Voting," *American Political Science Review,* 62:25-43.

Smith, J. Maynard and Price G. R. (1973): "The logic of animal conflict," *Nature,* v. 246, pp. 15-18.

Reiter, Stanley (2001): "Interdependent Preferences and Groups of Agents," *Journal of Economic Theory,* v. 3 pp. 27-69.

Robson, A (1996): "A Biological Basis for Expected and Non-Expected Utility Theory," *Journal of Economic Theory,* V. 68, pp. 397-424.

Robson, A (2000): "The Biological Basis for Economic Behavior," Manuscript.

Taylor, Peter D. and Jonker, Leo B. (1978): "Evolutionary Stable Strategies and Game Dynamics," *Mathematical Bioscience,* pp. 145-156.

Weibull, Jörgen W. (1997): *Evolutionary Game Theory.* The MIT Press.