Econometrica, Vol. 72, No. 5 (September, 2004), 1565–1581

NOTES AND COMMENTS

SOCIAL INDETERMINACY

BY GIL KALAI[1]

An extension of Condorcet's paradox by McGarvey (1953) asserts that for every asymmetric relation R on a finite set of candidates there is a strict-preferences voter profile that has the relation R as its strict simple majority relation. We prove that McGarvey's theorem can be extended to arbitrary neutral monotone social welfare functions that can be described by a strong simple game G if the voting power of each individual, measured by the *Shapley–Shubik power index*, is sufficiently small.

Our proof is based on an extension to another classic result concerning the majority rule. Condorcet studied an election between two candidates in which the voters' choices are random and independent and the probability of a voter choosing the first candidate is $p > 1/2$. Condorcet's jury theorem asserts that if the number of voters tends to infinity then the probability that the first candidate will be elected tends to one. We prove that this assertion extends to a sequence of arbitrary monotone strong simple games if and only if the maximum voting power for all individuals tends to zero.

KEYWORDS: Social choice, information aggregation, Arrow's theorem, simple games, the Shapley–Shubik power index, threshold phenomena.

1. INTRODUCTION

IN THIS PAPER we extend to general voting schemes two basic results concerning the majority rule. The first is McGarvey's theorem, an extension of Condorcet's paradox, and the second is Condorcet's jury theorem. Our extension of McGarvey's theorem in fact requires the extension of Condorcet's jury theorem, thus revealing an unexpected link between the two basic phenomena demonstrated by Condorcet.

We begin with an informal description of our results. Condorcet's famous "paradox" demonstrates that given three candidates A, B, and C, the majority rule may result in the society preferring A to B, B to C, and C to A. Arrow's impossibility theorem is an extension of Condorcet's paradox that asserts that under certain general conditions, such nontransitive social preferences cannot be avoided under *any* nondictatorial voting method.

McGarvey (1953) proved another extension of Condorcet's paradox: An asymmetric relation R on a finite set X is a binary relation such that every pair of elements $x, y \in X$ is ascribed one and only one of the relations xRy and yRx. McGarvey's theorem asserts that for every asymmetric relation R on a finite set of candidates, there is a group of individuals, each with a preference order relation on the candidates such that R coincides with the outcome of simple majority voting between every pair of candidates.

McGarvey's theorem is an early and simple manifestation of the phenomenon that choice aggregated over many individuals may lead to arbitrary outcomes (or, in other words, will not have any testable implications). We will refer to this phenomenon as

[1]I wish to thank Ehud Friedgut, Bezalel Peleg, Yuval Peres, and the referees and editor for fruitful discussions, and Yaakov Bergman, Ron Holzman, Motty Perry, Ariel Rubinstein, Ran Spiegler, Edna Ullman-Margalit, and Leeat Yariv for helpful suggestions. Supported by an ISF bikura grant.

"complete social indeterminacy." Another example of this phenomenon is the well-known Sonnenschein–Debreu–Mantel theorem concerning demand functions (see, for example, Sonnenschein (1972)).

Our extension of McGarvey's theorem is similar in spirit to Arrow's theorem (Arrow (1950)). Arrow's theorem asserts that the only voting method that can force transitive (or rational) social preferences concentrates all power in one individual. Our theorem demonstrates that the only way to impose *any restriction* on social preferences is by granting one individual a "substantial" amount of power, i.e. an amount of power that is bounded away from zero regardless of the size of the society. The measure of power that we will use is the Shapley–Shubik power index.

Condorcet's jury theorem asserts that in an election between two candidates, say Alice and Bob, if every voter votes for Alice with probability $p > 1/2$ and for Bob with probability $1 - p$ and if these probabilities are independent, then as the number of voters tends to infinity, the probability that Alice will gain the majority of votes tends to one (see Young (1988)). Condorcet's jury theorem can be interpreted as saying that even if agents receive very poor (independent) signals indicating which decision is correct, the majority rule will nevertheless result in the correct decision being taken with a high probability if there are enough agents and each agent votes according to the signal he receives. This phenomenon is referred to as "asymptotically complete aggregation of information." Our extension of Condorcet's jury theorem asserts that its conclusion remains valid when we replace simple majority with a general voting method as long as the power of every voter is sufficiently small.

Both results require some natural restrictions on the general voting method. One important restriction is neutrality, namely that the voting method is not a priori in favor of one of the alternatives. Our theorems also require strict social preferences. We note that McGarvey's theorem allows also to prescribe equality in the election's outcome between pairs of candidates. Similarly, strict social preferences are not required in Arrow's theorem. At a later stage we will examine biased voting methods and comment on the situation in which social indifference is a possibility.

In order to describe our results formally we need to define the notions of a simple game, the Shapley–Shubik power index, and a social welfare function. A simple game (or a voting game) defined on a set N of players (voters) is described by a function v that assigns to every subset (coalition) S of players the value '1' or '0.' We assume that $v(\emptyset) = 0$ and $v(N) = 1$. A candidate is elected if the set S of voters that voted for him is a winning coalition in G, i.e., if $v(S) = 1$. Recall that a simple game is *proper* if $v(S) + v(N \setminus S) \leq 1$ for every coalition S, i.e. if the complement of a winning coalition is a losing one. A simple game G is *strong* if $v(S) + v(N \setminus S) = 1$ for every coalition S, i.e. if it is proper and the complement of a losing coalition is a winning one. We will further assume that the game G is monotone, i.e. the addition of an individual to a winning coalition does not change it into a losing one.

The Shapley–Shubik power index assigns a real number between 0 and 1 to every player in a simple game. This index measures the power of the player in the game. A quick way to define the Shapley–Shubik power index (which differs from the original axiomatic definition) is as follows: Suppose that there are n voters. We say that a voter i is *pivotal* with respect to a set S of voters if $v(S \cup \{i\}) = 1$ and $v(S \setminus \{i\}) = 0$. In other words, player i can make a difference. For a probability distribution $\mathbf{P}$ on all subsets of voters, the probability that a voter is pivotal is called the *influence* of the voter (with respect to $\mathbf{P}$). The Shapley–Shubik power index of a voter measures his influence under the following distribution: First, randomly choose a real number t uniformly between 0 and 1 and then let each player i belong to S with probability t (independently of other players). The Shapley–Shubik power indices of all players sum up to one. For a simple game G we denote by $\bar{\phi}(G)$ the maximum value of the Shapley–Shubik power indices of all players in G.

DEFINITION 1.1: A sequence $(G_k)_{k=1,2,\dots}$ of monotone strong simple games has *diminishing individual power* if

$$\lim_{k\to\infty} \bar{\phi}(G_k) = 0.$$

Consider an election rule described by a monotone strong simple game G for an election between two candidates, Alice and Bob, and suppose that each voter votes for Alice with probability p and for Bob with probability $1 - p$ and that the probabilities are independent. Let $\mathbf{P}_p(G)$ denote the probability that Alice will be elected.

DEFINITION 1.2: A sequence $(G_k)_{k=1,2,\dots}$ of strong simple games has *asymptotically complete aggregation of information* if for every $p > 1/2$,

$$\lim_{k\to\infty} \mathbf{P}_p(G_k) = 1$$

(namely, the assertion of Condorcet's jury theorem holds for this sequence).

Our extension of Condorcet's jury theorem as presented in Section 3 asserts that diminishing individual power implies aggregation of information. In fact, these two properties are equivalent:

THEOREM 1.3: *A sequence (G_k) of monotone strong simple games has an asymptotically complete aggregation of information if and only if it has the property of diminishing individual power.*

A social welfare function F is a map that associates an asymmetric relation R on the alternatives to every profile of individual preferences. The social preference relation R is not assumed to be a transitive relation. We require the condition of independence of irrelevant alternatives (IIA), which states that the social preference between two alternatives a and b is determined by the individual preferences between a and b. We also require the Pareto condition (P) that if every individual in a society prefers a to b, then so will the society. A social welfare function is called neutral if it is invariant under permutations of the alternatives. A neutral social welfare function that satisfies conditions (IIA) and (P) can be described using a strong simple game G defined on the set of individuals. Thus, given the order relations $R_1, R_2, \dots, R_n$, which represent the individual preferences, the social preference relation R is defined by the following rule: For two alternatives a and b, aRb if the set of individuals for which aR_ib is a winning coalition in G. For a social welfare function F define the *image* of F to be the set of asymmetric relations that derive from F for some profile of individual preferences.

DEFINITION 1.4: A sequence (G_k) of strong simple games leads to *complete social indeterminacy* if for every number m of alternatives and every asymmetric relation R

on a set of m alternatives, there is $k(m)$ so that for every $k \geq k(m)$, R is in the image of the neutral social welfare function based on G_k (namely, the assertion of McGarvey's theorem holds for this sequence).

In Section 2 we prove the following theorem.

THEOREM 1.5: *A sequence* (G_k) *of strong simple games with asymptotically complete aggregation of information leads to complete social indeterminacy*.

Theorems 1.5 and 1.3 lead to our main result:

THEOREM 1.6: *A sequence* (G_k) *of monotone strong simple games with diminishing individual power leads to complete social indeterminacy*.

Theorem 1.6 can be reformulated as follows: For every number m of alternatives, there is a real number $\delta = \delta(m) > 0$ such that if a neutral social welfare function F is described by a monotone strong simple game G in which the Shapley–Shubik power index of each individual is at most δ, then the image of F is the set of all asymmetric preference relations on m alternatives.

One common way to prevent chaotic majority decisions is to introduce some bias towards the status quo. For example, requiring that a law pass in both houses of parliament or requiring a supermajority (a proportion larger than half of all votes) or requiring (as is the case of amendments to the U.S. Constitution) a majority of legislatures in three fourths of the states. In Section 4 we consider social welfare functions that are a priori biased towards a default order relation and show that unless the bias is overwhelming our extension of McGarvey's theorem still holds. The setting we consider is as follows: Let R_0 be an order relation on a set of alternatives that represents the default order relation. For a monotone proper simple game G, consider a social welfare function F defined as follows: For two alternatives a and b such that aR_0b, we have bRa if and only if the set of voters that prefer b to a forms a winning coalition in G. F is called the social welfare function based on G with a default order relation.

DEFINITION 1.7: A sequence $(G_k)_{k=1,2,\ldots}$ of monotone proper simple games has *overwhelming bias* if for every $p < 1$

$$\liminf_{k\to\infty} \mathbf{P}_p(G_k) = 0.$$

The sequence (G_k) has *diminishing bias* if for every $p \neq 1/2$

$$\lim_{k\to\infty} \big(1 - \mathbf{P}_p(G_k) + \mathbf{P}_{1-p}(G_k)\big) \to 0.$$

Definition 1.4 of complete social indeterminacy extends unchanged to social welfare functions that are biased towards a default order relation.

THEOREM 1.8: *In the model of social welfare functions biased toward a default order relation*, *a sequence* (G_k) *of monotone proper simple games with diminishing individual power that does not have overwhelming bias leads to complete social indeterminacy*.

In Section 4 we also comment on another way to avoid indeterminacy, namely by allowing social indifference. Given a monotone proper simple game G, a neutral social welfare function based on G on a set of m alternatives is defined as follows: For two alternatives a and b we have aRb if and only if the set of voters that prefer a to b forms a winning coalition in G. If G is not a strong simple game it may be the case that neither aRb nor bRa in which case we say that the society is indifferent between a and b. When we extend the definition of complete social indeterminacy (Definition 1.4) unchanged to this model we obtain the following theorem:

THEOREM 1.9: *In the model of neutral social welfare functions with the possibility of social indifference, a sequence* (G_k) *of monotone proper simple games with diminishing individual power and diminishing bias leads to complete social indeterminacy* (*for strict social preference relations*).

This is the case, for example, when G_k is a simple majority on k voters. Our proofs of Theorems 1.5 and 1.3 extend unchanged to the case of diminishing bias. (For a sequence (G_k) of proper simple games to have asymptotically complete aggregation of information, it is necessary that the bias be diminishing.) Note that our definition of complete social indeterminacy deals only with asymmetric relations without the possibility of indifference. In contrast to Theorem 1.9, only a few asymmetric relations are in the image of neutral social welfare functions based on α-supermajority, for every fixed $\alpha > 1/2$ (see Alon (2002) and Salant (2003), and Section 4).

The proofs of our results are of independent interest. Theorem 1.5 is proved using an elementary probabilistic argument that can be described as: "reduction to the case of majority using sampling." The proof of Theorem 1.3 uses recent results in probability theory and combinatorics concerning threshold phenomena. Threshold phenomena refer to situations in which the probability of an event (in our case $\mathbf{P}_p(G)$) changes rapidly as some underlying parameter (in our case p) varies within some interval.

2. FROM INFORMATION AGGREGATION TO INDETERMINACY

2.1. *Social Welfare Functions*

We consider a social welfare function that, given a profile $\mathcal{R}$ of n order relations R_i, $i = 1, 2, \ldots, n$, on m alternatives, yields an asymmetric relation R for the society. Thus, $R = F(R_1, R_2, \ldots, R_n)$ where F is the social welfare function. aR_ib states that the ith individual prefers alternative a over alternative b. aRb indicates that the society prefers alternative a over alternative b. The social preferences are not assumed to be transitive.

The condition of independence of irrelevant alternatives (IIA), states that for every two alternatives a and b the individual preferences between a and b determine the social preference between a and b. Formally, the set $\{i : aR_ib\}$ determines whether aRb. The social preference between a and b can thus be described by a strong simple game $G_{a,b}$ as follows: Let S be the set of individuals that prefer alternative a over alternative b (i.e., $S = \{i : aR_ib\}$). S is a winning coalition for the game $G_{a,b}$ if aRb.

The Pareto condition is another standard assumption that asserts that if all individuals in a society prefer alternative a over b, then so will the society. This means that in the game $G_{a,b}$ for every two alternatives a and b, the set of all voters is a winning coalition and the empty set of voters is a losing one.

We will also assume that the social welfare function is monotone, which means that if an individual who prefers alternative a over alternative b changes his preferences, this will not result in the opposite change in the society's preferences.

Finally, we assume that the social welfare function is *neutral*, namely that it is invariant to permutations of the *alternatives*. Assuming neutrality is equivalent to the assertion that all simple games $G_{a,b}$ are strong and identical. Therefore, a neutral social welfare function can be described in terms of a single strong simple game.

A convenient way to think about the function is as a rule for elections between two candidates. There is a pool of several candidates and every individual has an order relation on all candidates. We wish to understand the possible outcomes of two-candidate elections between each pair of candidates within the pool.

2.2. *The Proof of Theorem* 1.5

The basic idea of the proof of Theorem 1.5 is simple: We have a sequence of strong simple games (G_k) such that for every $p > 1/2$, $\lim_{k\to\infty} \mathbf{P}_p(G_k) = 1$. Suppose that G_k is a simple game with $n(k)$ players. We wish to realize an asymmetric relation R on m alternatives with a social welfare function based on G_k for large enough k. We begin with a society (of n_0 voters), whose existence is guaranteed by McGarvey's theorem, that realizes R as its strict majority preference relation. Next, we sample $n(k)$ voters (with repetitions) from this society. Consider two alternatives a and b such that aRb. The probability p that a voter will prefer alternative a over alternative b is bounded away from 1/2. In fact, $p \geq 1/2 + 1/2n_0$. However, $\mathbf{P}_p(G_k)$ tends to 1. Therefore, as k tends to infinity, the society will prefer a over b with a probability that approaches 1. Since this is true for every pair of alternatives, we deduce that when k is large enough, the social preference relation will be R with a high probability.

To present the formal proof of Theorem 1.5 we require the following definition: Let $\mathcal{R} = (R_1, R_2, \dots, R_{n_0})$ be a profile of order relations on a fixed set X of m alternatives and let R be an asymmetric relation on X. We say that the profile $\mathcal{R}$ realizes R with *quality* t if for every two alternatives a and b in X such that aRb we have

$$|\{i : aR_ib\}| \geq (1/2 + t)n_0.$$

For example, the profile consisting of the three order relations $a > b > c$, $b > c > a$, and $c > a > b$ realizes a cyclic relation on the three alternatives a, b, and c with quality 1/6.

Using the notion of the quality of a representation of an asymmetric relation, the statement and proofs of representations with arbitrary strong simple games become quite straightforward:

THEOREM 2.1: *Let R be an asymmetric order relation on m alternatives and suppose that R can be realized by some profile of order relations with quality t. Let G be a strong simple game that satisfies*

$$\mathbf{P}_p(G) > 1 - 1\Big/\binom{m}{2}, \quad \textit{for every}\, p, \quad p \geq 1/2 + t. \tag{2.1}$$

Then R is in the image of a social welfare function on m alternatives based on G.

Let $t(m)$ be the largest real number such that every asymmetric relation R on m alternatives can be realized by some profile with quality $t(m)$.

COROLLARY 2.2: *Let G be a strong simple game that satisfies*

$$(2.2) \qquad \mathbf{P}_p(G) > 1 - 1\bigg/\binom{m}{2}, \quad \textit{for every } p, \quad p \geq 1/2 + t(m).$$

Then every asymmetric relation on m alternatives is in the image of a social welfare function on m alternatives based on G.

REMARK: Theorems 2.1 and 1.5 do not require monotonicity. If G is monotone we can simply replace relation (2.1) by $\mathbf{P}_{1/2+t}(G) > 1 - 1/\binom{m}{2}$.

DERIVATION OF THEOREM 1.5: McGarvey's theorem implies that $t(m) > 0$. Therefore combined with McGarvey's original result, Corollary 2.2 implies Theorem 1.5. Indeed, in order to derive Theorem 1.5 note that a sequence (G_k) with asymptotically complete aggregation of information satisfies $\lim_{k\to\infty} \mathbf{P}_p(G_k) = 1$ for every $p > 1/2$ and therefore relation (2.2) is satisfied when k is sufficiently large.

PROOF OF THEOREM 2.1: Let G be a strong simple game with n players that satisfies relation (2.1) and let F be the corresponding social welfare function when there are m alternatives. Let $\mathcal{R}_0 = (R_1, R_2, \dots, R_{n_0})$ be a profile of n_0 order relations that realizes the asymmetric relation R with quality t. Consider a random voter profile $\mathcal{R}$ for n individuals for which the order preference relation for every individual is equal to R_j with probability $1/n_0$, $j = 1, 2, \dots, n_0$, and these probabilities are independent. For two alternatives a and b, if aRb then the probability p that the ith player will prefer a to b is at least $1/2 + t$. The probability that the society will prefer a to b is $\mathbf{P}_p(G)$. Relation (2.1) asserts that if $p \geq 1/2 + t$, then $\mathbf{P}_p(G) > 1 - 1/\binom{m}{2}$. We conclude that the probability that the social preferences will coincide with R is larger than $1 - \binom{m}{2}/\binom{m}{2} = 0$. *Q.E.D.*

REMARKS: 1. McGarvey's proof of his theorem (see Section 4) implies that $t(m) \geq 1/m(m-1)$ and this bound can be increased to $c \log m/m$ (for some constant c) using a subsequent result by Erdős and Moser (1964). Alon (2002) recently showed that $c_1/\sqrt{m} \leq t(m) \leq c_2/\sqrt{m}$, where $c_2 > c_1 > 0$ are constants. (See Section 4.)

2. The proof of Theorem 1.5 applies unchanged to neutral social welfare functions based on proper simple games. Of course, in order to have asymptotically complete aggregation of information it is necessary to have diminishing bias.

3. DIMINISHING POWER AND AGGREGATION OF INFORMATION

Consider an election between two candidates, Alice and Bob, where the election rule is given by a monotone strong simple game $G = \langle N, v\rangle$. The game G describes the way in which the individual preferences between the two candidates aggregate. Now we will discuss how the voters are going to vote.

We suppose that the ith voter receives a signal s_i where $s_i = 1$ with probability $p > 1/2$, $s_i = 0$ with probability $1 - p$ and the signals are independent. The signal $s_i = 1$ means "vote for Alice" and we assume that voters act according to the signals they receive. Therefore, the set S of voters who vote for Alice is given by $S = \{i : s_i = 1\}$. The set S is a random set of players such that for each player i, $i \in S$ with probability p, $i \notin S$ with probability $1 - p$ and the events "$i \in S$" are independent for $i = 1, 2, \dots, n$.

For a specific set $S \subset N$, the probability that the set of Alice's voters is precisely S is denoted by $\mathbf{P}_p(S)$ and is equal to $p^{|S|}(1-p)^{n-|S|}$. Denote by $\mathbf{P}_p(G)$ the probability that the random set S of Alice's voters is a winning coalition, i.e. the probability that Alice wins the election is:

$$\mathbf{P}_p(G) = \sum \{\mathbf{P}_p(S) : v(S) = 1\}.$$

REMARK: It is often useful to regard the payoff function v as a function defined on the vector of signals $(s_1, s_2, \ldots, s_n)$ as follows: $v(s_1, s_2, \ldots, s_n) = v(S)$, where $S = \{i : s_i = 1\}$. Functions from the set of 0–1 vectors of length n to the set $\{0, 1\}$ are called Boolean functions. Most of the results we use in this section originally appeared in the literature in the language of Boolean functions rather than simple games.

The proofs of the following two simple lemmas are presented in the Appendix.

LEMMA 3.1: *If G is a proper simple game, then for* $0 < p < 1$,

$$\mathbf{P}_p(G) \leq 1 - \mathbf{P}_{1-p}(G). \tag{3.1}$$

Equality holds if and only if G is strong.

LEMMA 3.2: *If G is a monotone simple game, then the function $\mathbf{P}_p(G)$ is a strictly monotone continuous function of p in the interval* $[0, 1]$.

Let ϵ, $0 < \epsilon < 1/2$, be a real number. Since $\mathbf{P}_p(G)$ is a strictly monotone and continuous function of p, there is a unique value of p denoted by p_1 such that $\mathbf{P}_{p_1}(G) = \epsilon$. There is also a unique value of p denoted by p_2 such that $\mathbf{P}_{p_2}(G) = 1 - \epsilon$. The interval $[p_1, p_2]$ is called a *threshold interval* and its length $p_2 - p_1$ is denoted by $T_\epsilon(G)$. The value $p_c = p_c(G)$, at which $\mathbf{P}_{p_c}(G) = 1/2$, is called the *critical probability* for G. It follows from relation (3.1) that if G is a proper simple game, then $p_c \geq 1/2$ with equality if and only if G is strong.

We now require the important notion of influence.[2] The *influence* of the kth player on G, denoted by $I_k^p(G)$, is the probability that the player is *pivotal*, i.e. the probability that for a random coalition S (according to the probability distribution $\mathbf{P}_p$) that does not contain k, S is a losing coalition and $S \cup \{k\}$ is a winning one. The influence of a player is a normalized version of the correlation between his vote and the election's outcome. The total influence $I^p(G)$ equals $\sum_{k=1}^n I_k^p(G)$.

Let $\phi_k(G)$ denote the Shapley–Shubik power index for the kth player in G. The following integral representation of $\phi_k(G)$ is due to Owen (1988):

$$\phi_k(G) = \int_0^1 I_k^p(G)\, dp.$$

[2]The mathematical study of pivotal agents and influences is fundamental not only in the context of power indices in game theory, but also in other areas of economics (see, for example, Pesendorfer and Swinkels (2000) and Al-Najjar and Smorodinsky (2000)), as well as in reliability theory, statistical physics, probability theory and statistics, distributed computing, and complexity theory.

Owen's representation of the Shapley–Shubik power index coincides with the description given in the Introduction but differs from Shapley's original axiomatic definition. (The Banzhaf index of the kth player equals $I_k^{1/2}(G)$.) Define $\bar{\phi}(G) = \max(\phi_1(G), \phi_2(G), \ldots, \phi_n(G))$.

Our first task is to prove that if the threshold interval is small, then so is the power of every individual.

THEOREM 3.3: *Let G be a monotone strong simple game. If $T_\epsilon(G) \leq \gamma$, then $\bar{\phi}(G) \leq \gamma + 3\sqrt{\epsilon}$.*

PROOF: Suppose that $T_\epsilon(G) \leq \gamma$ or, in other words, $\mathbf{P}_p(G) < \epsilon$ if $p < 1/2 - \gamma/2$. It follows from the definition that $I_k^p(G) \leq \min(1, \mathbf{P}_p(G)/p)$. Therefore,

$$\begin{aligned}\phi_k(G) &= \int_0^1 I_k^p(G) = 2\int_0^{1/2} I_k^p(G)\\ &\leq 2\int_0^{\sqrt{\epsilon}} 1 + 2\int_{\sqrt{\epsilon}}^{1/2-\gamma/2} \mathbf{P}_p(G)/\sqrt{\epsilon} + 2\int_{1/2-\gamma/2}^{1/2} 1 \leq 2\sqrt{\epsilon} + \sqrt{\epsilon} + \gamma.\end{aligned}$$

Q.E.D.

Our main task is to prove that if the power of every individual is small, then the threshold interval must also be.

THEOREM 3.4: *For every $\epsilon, \gamma > 0$ there exists $\delta > 0$ such that for every monotone strong simple game G if $\bar{\phi}(G) \leq \delta$, then $T_\epsilon(G) \leq \gamma$.*

Theorems 3.3 and 3.4 provide the two directions of the equivalence in Theorem 1.3 between diminishing individual power and asymptotically complete aggregation of information. We note that the assertion of Theorem 1.3 and the proof extend unchanged to the case of monotone proper simple games with diminishing bias. The proof of Theorem 3.4 extends to the following more general result. This generalization will be useful in Section 4 where we study social welfare functions that are biased towards a default order relation.

THEOREM 3.5: *For every $a, \epsilon, \gamma > 0$ there exists $\delta > 0$ such that for every monotone simple game G if $\bar{\phi}(G) \leq \delta$ and $a \leq p_c(G) \leq 1 - a$, then $T_\epsilon(G) \leq \gamma$.*

I will now present the mathematical concepts and results required for proving Theorem 3.4. We first need the following fundamental result:

PROPOSITION 3.6 (Russo's lemma; see Grimmett (1989)):

$$\frac{d\mathbf{P}_p(G)}{dp} = I^p(G). \tag{3.2}$$

Russo's lemma implies that if the total influence for every p in the threshold interval is large, then the threshold interval itself must be small. The rather simple proof is presented in the Appendix.

Next, we require a result that shows that for a specific value of p, if all individual influences $I_k^p(G)$ are small, then their sum $I^p(G)$ is large.

THEOREM 3.7 (Talagrand (1994)): *For some constant* $C > 0$, *and every positive real* $\delta > 0$, *if* G *is a monotone simple game and if* $I_k^p(G) \leq \delta$ *for every* $k = 1, 2, \dots, n$, *then*

$$I^p(G) \geq C \log(1/\delta) \mathbf{P}_p(G)\big(1 - \mathbf{P}_p(G)\big). \tag{3.3}$$

I will now present the outline of the proof of Theorem 3.5 leaving the full details to the Appendix. Set $x = \log(1/\delta)$. Talagrand's theorem asserts that if all influences $I_k^p(G)$ for every p are smaller than δ and if p belongs to the threshold interval where $\epsilon \leq \mathbf{P}_p(G) \leq 1 - \epsilon$, then the sum of influences $I^p(G)$ is larger than $C(\epsilon/2)x$. By Russo's lemma the sum of influences is the slope of the function $\mathbf{P}_p(G)$ (as a function of p). Therefore, if for every p, $I_k^p(G) < \delta$, the size of the threshold interval $T_\epsilon(G)$ is at most $(1/C)(2/\epsilon)(1/\log(1/\delta))$. We conclude that if ϵ is fixed, then as δ tends to zero, so does $T_\epsilon(G)$.

This is quite close to the statement of Theorem 3.5. However, one problem still remains: In order to apply Talagrand's theorem we need to assume that the individual influences are small for every p, which is stronger than the assumption we have made. When we assume that the Shapley–Shubik power index is small, this only means that the influences $I_k^p(G)$ are small when we average over p. To complete the proof we need to exclude the possibility that the threshold interval $[p_1, p_2]$ (or a large chunk of it) is the union of many small parts such that in every part there is a different player with a high degree of influence. In the proofs of Theorems 3.4 and 3.5 presented in the Appendix, we make some additional observations that will enable us to relate the influences for one value of p to the influences for the entire threshold interval.

We conclude this section with a simple example of a sequential voting procedure that demonstrates the economic relevance of Theorem 1.3 on its own. Theorem 1.3 seems especially useful when simple games are used to model not just the voting method but more involved situations of aggregation. The intuition that in order to reduce the probability of a mistake we should reduce the individual influences is consistent with various real-life procedures for collective decisions.

Consider a committee (with an odd number n of members) set up to choose between Alice and Bob in which the members openly vote according to some order and the decision is made by simple majority. Suppose that each member of the committee receives an independent signal s_i such that $s_i = 1$ with probability p. However, the vote of a committee member depends also on earlier votes. For example, suppose that for every committee member i there is a set S_i of other members whom he respects and who voted previously and that he will vote against his signal if more than two thirds of S_i voted the opposite way. Although the voting method is simple majority, Condorcet's jury theorem does not apply since the probabilities of individuals voting for Alice are not independent. The outcome of the voting process as a function of the *original* signals (rather than the actual votes) can be described by a (complicated) strong simple game and since the original signals are independent our theorem applies.

4. BIAS AND INDIFFERENCE

In this section we study social welfare functions based on proper simple games. We first study social welfare functions that are biased towards a default order relation and later comment on neutral social welfare functions with the possibility of indifference.

We return now to McGarvey's theorem and examine whether we can achieve some restrictions on the social preferences by introducing a bias towards a default order relation. Let G be a proper monotone simple game. Let R_0 be an order relation on a set of m alternatives that represents the default order relation. We will assume from now on (without loss of generality) that the set of m alternatives is $A = \{a_1, a_2, \dots, a_m\}$ and that R_0 is defined by $a_i R_0 b_j$ if and only if $i < j$. We consider a social welfare function biased towards the default order relation R_0. Recall that F is defined as follows: For two alternatives a and b such that aR_0b, we have bRa if the set of voters that prefer b to a forms a winning coalition in G.

Extreme bias can reduce indeterminacy. An example of a biased social welfare function that implies some restrictions on the social preferences is the case where in order to reverse the default relation between two alternatives we require a unanimous vote for the opposite relation. Here, no matter how large the society, if for the default preference relation R_0, aR_0b, and bR_0c, while for the social preference relation R, bRa, and cRb, then it must be the case that cRa. In other words, the social preferences that violate the default order relation must obey transitivity. Therefore, for social welfare functions based on unanimous voting not every asymmetric relation can be realized.

Let α, $1/2 < \alpha < 1$, be a real number. Define an α-supermajority game as a game with a set N of players in which the winning coalitions are subsets S of N that satisfy $|S| > \alpha|N|$. McGarvey's theorem extends to α-supermajority games. In fact, the following more general statement is true: Let (d_n) be a sequence of positive integers satisfying $d_n < n/2$. Consider the sequence (G_n) of games where G_n has n players and the winning coalitions in G_n are subsets S of players such that $|S| > n - d_n$. In this case the assertion of McGarvey's theorem will apply if $\lim_{n\to\infty} d_n = \infty$. To realize an asymmetric relation R on m alternatives we can simply apply McGarvey's theorem to a simple majority social welfare function with $2d_n + 1$ voters (for large enough n) and add $n - 2d_n - 1$ voters whose preference relation is the opposite linear order to R_0.

To understand the source of the difference between supermajority and the unanimous rule we will need the following definition:

DEFINITION 4.1: A sequence (G_k) of proper monotone simple games has the *sharp threshold property* if for every $\epsilon > 0$

$$T_\epsilon(G_k) = o\big(1 - p_c(G_k)\big). \tag{4.1}$$

This definition can be motivated by quoting a (rather simple) theorem due to Bollobas and Thomason (1987) that asserts that for every monotone simple game G, $T_\epsilon(G) = O(\min(p_c(G), 1 - p_c(G)))$.

If G_n is the α-supermajority game with n players, $1/2 < \alpha < 1$, then the sequence (G_n) has the sharp threshold property. (This follows from the law of large numbers: The weak law of large numbers asserts that $\lim_{n\to\infty} \mathbf{P}_p(G_n) = 0$ for $p < \alpha$ and $\lim_{n\to\infty} \mathbf{P}_p(G_n) = 1$ for $p > \alpha$. Therefore, $\lim_{n\to\infty} p_c(G_n) = \alpha$ and for every fixed $\epsilon > 0$, $\lim_{n\to\infty} T_\epsilon(G_n) = 0$.) On the other hand, if G_n represents the unanimous rule on n players, then the sequence (G_n) does not have a sharp threshold. ($\mathbf{P}_p(G_n) = 1/2$ for $p = 1 - s/n$, where $s \approx \log_e 2$, and $\mathbf{P}_{1-\lambda s/n}(G_n) \approx (1/2)^\lambda$ is bounded away from zero.)[3]

Theorem 1.5 extends to the following result:

[3]The discontinuity we observe at $\alpha = 1$ for α-supermajority appears to be related to similar discontinuities described in the literature and the connection to the sharp threshold property may extend to other such cases as well. For example, Dasgupta and Maskin (1997) showed that the domain on which an α-supermajority rule leads to transitive social preferences gets discontinuously

THEOREM 4.2: *If a sequence* (G_k) *of monotone proper simple games has the sharp threshold property, then it leads to complete social indeterminacy.*

PROOF OF THEOREM 1.8: If the sequence G_k does not have overwhelming bias, then for some $q < 1$, $\liminf_{k\to\infty} \mathbf{P}_q(G_k) > 0$. It follows from the theorem of Bollobas and Thomason that $p_c(G_k)$ is bounded away from one. However, when $p_c(G_k)$ is bounded away from one, then having a sharp threshold simply means that $\lim_{n\to\infty} T_\epsilon(G_k) = 0$. Theorem 3.5 asserts that if $p_c(G_k)$ is bounded away from one, the property of diminishing individual power implies the sharp threshold property. Theorem 1.8 now follows from Theorem 4.2. *Q.E.D.*

Theorem 1.8 can be described as follows: The only two ways to constrain the social preferences when the number of individuals is large is either to introduce strong bias among the individuals, namely to give one individual a substantial amount of power or to introduce an overwhelming bias among the alternatives.

In order to prove Theorem 4.2 we present a quantitative version that extends Theorem 2.1. Let $1 > \alpha \geq 1/2$ and set $\beta = 1 - \alpha$. We say that an asymmetric relation R on A is realized by α-supermajority with quality q if there is a collection of voters and a voter profile $\mathcal{R}$ with the following property: For any two alternatives a and b such that bR_0a, if aRb then more than a fraction of $\alpha + q$ of the voters prefer a to b, and if bRa then less than a fraction of $\alpha - q$ of them prefer a to b. Recall that $t(m)$ was defined as the largest real number such that every asymmetric relation R on m alternatives can be realized by simple majority with quality $t(m)$. By our simple observation concerning supermajority, it follows that every asymmetric relation on m alternatives can be realized by α-supermajority with quality $\beta t(m)$.

THEOREM 4.3: *Let G be a monotone simple game. Set* $\alpha = p_c(G)$ *and suppose that* $\alpha \geq 1/2$. *Suppose that*

(i) $$\mathbf{P}_{\alpha+t}(G) > 1 - 1\Big/\binom{m}{2}, \quad \textit{and}$$

(ii) $$\mathbf{P}_{\alpha-t}(G) < 1\Big/\binom{m}{2}.$$

Let R be an asymmetric relation on m alternatives that can be realized by α*-supermajority with quality t. Then R is the image of the social welfare function based on G biased toward the default order relation* R_0.

The proof is identical to that of Theorem 2.1. To deduce Theorem 4.2 note first that there is no loss in generality in assuming that $\alpha_k = p_c(G_k) \geq 1/2$ for every k. Set $\beta_k = 1 - \alpha_k$. A sequence (G_k) with the sharp threshold property satisfies $\lim_{k\to\infty} \mathbf{P}_{\alpha_k + c\beta_k}(G) = 1$ and $\lim_{k\to\infty} \mathbf{P}_{\alpha_k - c\beta_k}(G) = 0$ for every constant $c > 0$. Therefore, if we let $c = t(m)$ and use the fact that every asymmetric relation R can be realized by α_k-supermajority with quality $t(m) \cdot \beta_k$, we realize that relations (1) and (2) are satisfied when k is sufficiently large.

larger at $\alpha = 1$. Feddersen and Pesendorfer (1998) considered discontinuity of asymptotically complete aggregation of information for strategic voting.

REMARK: When G is an α-supermajority with n voters, then the law of large numbers implies that $p_c(G) \approx \alpha$. The central limit theorem implies that

$$T_\epsilon(G) \approx C(1/\sqrt{n})\sqrt{\log(1/\epsilon)}, \tag{4.2}$$

for some constant C. It follows that for biased rules similar to those used in practice, such as, "In at least three quarters of the states the default must be voted out by at least two thirds of the legislators," the threshold behavior can be computed precisely. In such cases, where there is a bounded number of layers and in each layer supermajority is used, the computation of the threshold interval, Theorem 2.1, and Alon's estimate $t(m) \geq c_1/\sqrt{m}$ implies that every asymmetric relation R on m alternatives can be realized by no more than $Cm\log m$ voters (or legislators) where C is a constant that depends on the precise rule.

Finally, we make a few comments concerning neutral social welfare functions F with the possibility of indifference. A partially asymmetric relation R on a finite set X is a binary relation such that every pair of elements $x, y \in X$ is ascribed at most one of the relations xRy and yRx. Given a monotone proper simple game G and a set A of alternatives, the society prefers alternative a over alternative b if the set of voters that prefer a to b forms a winning coalition. The image of such a neutral social welfare function is a set of partially asymmetric relations. There are several ways to extend the notion of indeterminacy: Can we realize every asymmetric relation? Can we realize every partially asymmetric relation? Theorem 1.9 extends our result concerning asymmetric relations to the case of monotone proper simple games with diminishing bias and as we note in Sections 2 and 3 the proofs, which are probabilistic, extend to this case unchanged.

(i) We already encounter the possibility of social indifference when we consider simple majority with an even number of voters. In this case we have diminishing bias in a very strong sense since the probability of a tie $(1 - \mathbf{P}_p(G) - \mathbf{P}_{1-p}(G))$ tends to zero even for $p = 1/2$. McGarvey's original theorem also allows us to prescribe in an arbitrary way the cases of equality. McGarvey's proof relies on the following simple observation: Aggregating two voters with order relations $1 <_1 2 <_1 3 <_1 \cdots <_1 m$ and $m <_2 m-1 <_2 \cdots <_2 3 <_2 1 <_2 2$ results in a situation in which alternative '2' is preferred to alternative '1' but there is indifference between every other pair. Consider a partially asymmetric relation R on the set of m alternatives. For every pair of alternatives a and b, if aRb we can define order relations for two voters according to which both prefer a to b but have the opposite preferences on every other pair of alternatives. Therefore, for every partially asymmetric preference relation R on m alternatives, combining such pairs of voters for every two alternatives a and b such that aRb implies that R can be realized and by at most $m(m-1)$ voters.

(ii) For social welfare functions based on α-supermajority for $1/2 < \alpha < 1$, there are severe restrictions on the social preferences. For example, if $\alpha > 3/4$ then it is clearly impossible that the society will prefer a to b, b to c, and c to a. Alon (2002) proved that there are two constants $c_2 > c_1 > 0$ such that every asymmetric relation R on m alternatives can be realized by $(1/2 + c_1/\sqrt{m})$-supermajority and there is an asymmetric relation R on m alternatives that cannot be realized by $(1/2 + c_2/\sqrt{m})$-supermajority. (It is not known when all partially asymmetric relations can be realized.) For a fixed $t > 0$, the class of asymmetric relations on m alternatives that can be realized by $(1/2 + t)$-supermajority was studied by Salant (2003).

Realizing an asymmetric relation R by an $(1/2+t)$-supermajority, $t>0$, is rarely possible when t is fixed and the number of alternatives grows. Suppose that an asymmetric relation R on m alternatives is based on $(1/2+t)$-supermajority. Theorem 2.1, combined with the threshold properties of the majority function given by relation (4.2), asserts that R is realized with high probability by a simple majority of the preferences of T random voters when $T=10\log m/t^2$. However, the number of relations that can be so realized is at most the number of order preference relations for these T individuals, which is $m!^T \le \exp(10m(\log m)^2/t^2)$. For a fixed $t>0$ and large m, this quantity is much smaller than the total number of asymmetric relations, which is $2^{\binom{m}{2}}$.

(iii) Let (G_n) be a sequence of proper simple games with diminishing individual power. If $\lim_{n\to\infty} p_c(G_n)=1/2$ then (G_n) has diminishing bias and the image of a neutral social welfare function on m alternatives based on (G_n) will contain for a large n all asymmetric relations. Suppose next that $\lim_{n\to\infty} p_c(G_n)=\alpha>1/2$ and $\alpha<1$, namely, that the bias is neither diminishing nor overwhelming. Let F_n be the neutral social welfare functions based on G_n on a set $A=\{a_1,a_2,\ldots,a_m\}$ of m alternatives. Consider the following $\binom{m}{2}$ questions: Is a_iRa_j, $1\le i<j\le m$? Theorem 1.8 asserts that when n is large any sequence of length $\binom{m}{2}$ consisting of the words 'yes' and 'no' can serve as the sequence of answers for some voter profile. This implies that the image of F_n contains at least $2^{\binom{m}{2}}$ partially asymmetric relations (among all $3^{\binom{m}{2}}$ partially asymmetric relations). On the other hand, as we remarked above, for α-supermajority the image of F_n contains only a "small" number of asymmetric relations.

Institute of Mathematics and the Center for the Study of Rationality, Hebrew University, Givat Ram Jerusalem 91904, Israel, and Departments of Computer Science and Mathematics, Yale University, New Haven, CT 06520, U.S.A.; kalai@math.huji.ac.il; http://www.ma.huji.ac.il/˜kalai.

Manuscript received June, 2002; final revision received February, 2004.

APPENDIX

Overview

Most of the Appendix is devoted to the proofs of Theorems 3.4 and 3.5. With the exception of Talagrand's theorem, which we stated in Section 3, and a related theorem by Friedgut that we will state below, our presentation is self-contained. The reader is encouraged to consult the literature (Talagrand (1994), Kahn, Kalai, and Linial (1988), Friedgut and Kalai (1996), and Friedgut (1998)) for a wider perspective on the topic discussed here and for the proofs of Talagrand's and Friedgut's results. The proofs of these two theorems are quite accessible and rely on harmonic analysis on the discrete n-dimensional cube. (Note that the terminology used in these papers differs from our own.)

Russo's Lemma and Other Basic Facts

PROOF OF LEMMA 3.1: Let $G=\langle N,v\rangle$ be an arbitrary simple game. Consider an election between Alice and Bob and suppose that Alice wins if the voters for Alice form a winning coalition in G. Now consider a simple game $G'=\langle N,v'\rangle$ defined by $v'(S)=1$ if and only if $v(N\setminus S)=0$. A winning coalition for G' is a coalition whose complement is losing in G. Suppose that every voter independently votes for Alice with probability p. The probability that Alice wins is $\mathbf{P}_p(G)$. The probability that Alice does not win is precisely the probability that the coalition of voters that did not vote for Alice is a winning coalition in G'. Since every voter does not vote Alice with probability $1-p$, this probability is $\mathbf{P}_{1-p}(G')$. Therefore $\mathbf{P}_p(G)+\mathbf{P}_{1-p}(G')=1$. It is easy to see that G is proper if and only if $G\subset G'$, which yields relation (3.1). The game G is strong if and only if $G=G'$ and this implies that given that G is proper, for every p, $0<p<1$, we have equality in relation (3.1) if and only if G is strong. *Q.E.D.*

PROOF OF LEMMA 3.2: Let $0<p<q<1$. Consider a random subset S of N according to the probability measure $\mathbf{P}_p$. Let R be a random subset of N according to the probability measure $\mathbf{P}_r$ where $r=(q-p)/(1-p)$. Consider $T=S\cup R$. The probability that a player i belongs to T is $1-(1-p)(1-r)=q$ and the events $i\in T$ are independent. $\mathbf{P}_p(G)$ is the expected value of $v(S)$ and $\mathbf{P}_q(G)$ is the expected value of $v(T)$. By monotonicity, $v(T)\geq v(S)$ and therefore $\mathbf{P}_q(G)\geq \mathbf{P}_p(G)$. Since there is a positive probability that $S=\emptyset$ and $T=N$, we conclude that $\mathbf{P}_q(G)>\mathbf{P}_p(G)$. (The monotonicity of $\mathbf{P}_p(G)$ also follows from Russo's lemma, but the direct argument used here will serve us again in the proof of Theorem 1.3.) *Q.E.D.*

PROOF OF RUSSO'S LEMMA: Let $G=\langle N, v\rangle$ be a fixed simple game on the set N of n players. Suppose that $N=\{1,2,\dots,n\}$ and consider a general product probability distribution on subsets S of N defined as follows:

$$\mathbf{P}^{p_1,p_2,\dots,p_n}(S)=\prod\{p_i : i\in S\}\cdot\prod\{(1-p_i): i\notin S\}.$$

Let $A(p_1,p_2,\dots,p_n)=\sum\{\mathbf{P}^{p_1,p_2,\dots,p_n}(S): v(S)=1\}$. Let $I_k(p_1,p_2,\dots,p_n)$ be the probability that the kth player is pivotal according to the probability measure $\mathbf{P}^{p_1,p_2,\dots,p_n}(S)$. Thus, $I_k^p(G)=I_k(p,p,\dots,p)$. Note that $A(p_1,p_2,\dots,p_n)$ is a linear function of p_k and

$$\partial A(p_1,p_2,\dots,p_n)/\partial p_k = I_k(p_1,p_2,\dots,p_n).$$

Without loss of generality we will demonstrate this relation for $k=n$. First note that

$$\begin{aligned}&A(p_1,p_2,\dots,p_n)\\&\quad=\sum_{S\subset\{1,2,\dots,n-1\}}\Big(\prod_{i\in S}p_i\cdot\prod_{i\in\{1,2,\dots,n-1\}\setminus S}(1-p_i)\Big)\big(v(S)(1-p_n)+v(S\cup\{n\})p_n\big).\end{aligned}$$

Next note that for those subsets S where $v(S)=v(S\cup\{n\})$ the summand does not depend on p_n. The contribution of a set S for which $v(S\cup\{n\})=1$ and $v(S)=0$ is $\prod\{p_i : i\in S\}\cdot\prod\{(1-p_i) : i\in\{1,2,\dots,n-1\}\setminus S\}\cdot p_n$ and the derivatives with respect to p_n sum up precisely to $I_n(p_1,p_2\dots,p_n)$.

Russo's lemma follows from the chain rule: Write $B(p)=(p,p,\dots,p)$ (n times) and $C(p)=\mathbf{P}_p(G)$. Here $A:[0,1]^n\to[0,1]$, $B:[0,1]\to[0,1]^n$, and $C:[0,1]\to[0,1]$. $C(p)$ is the function we want to differentiate. Note that $C(p)=A(B(p))$ and that $dB/dp=(1,1,\dots,1)$ is the all ones vector of length n. Therefore according to the chain rule the derivative of C at the point p is equal to $\sum_{k=1}^n \partial A/\partial p_k(p,p,\dots,p)$, which is equal to $\sum_{k=1}^n I_k^p(G)=I^p(G)$ as required. *Q.E.D.*

Proof of Theorem 1.3

We require the following result by Friedgut (1998) that asserts (in our terminology) that a simple game with a small influence (w.r.t. $\mathbf{P}_p$) is determined with high probability (w.r.t. $\mathbf{P}_p$) by a small set of players.

THEOREM A.1: *For every real number $z>0$, $A>1$, and $\gamma>0$, there is $C=C(\gamma,A,z)$ such that if $z\leq p\leq 1-z$, the following assertion holds: For a monotone simple game $G=\langle N,v\rangle$, if*

$I^p(G) \leq A$ then there exists a collection S of at most C players in N and a monotone simple game $H = \langle S, v_0 \rangle$ such that

$$\text{(A.1)} \qquad \mathbf{P}_p\{T \subset N : v(T) \neq v_0(T \cap S)\} < \gamma.$$

PROOF OF THEOREM 3.4: Since by Russo's lemma $I^p(G) = d\mathbf{P}_p(G)/dp$, Theorems 3.7 and A.1 give conditions for the derivative of $\mathbf{P}_p(G)$ to be large at a given point p. In order to prove that the threshold interval is small, we need to move from local information (for specific values of p) to global information for the entire threshold interval. The following crucial lemma was proved in collaboration with Ehud Friedgut:

LEMMA A.2: *Let G be a monotone simple game. Let $p < q \in [1/3, 2/3]$. Suppose that $\mathbf{P}_p(G) \geq a > 0$, $I^p(G) \leq A$, and that $\mathbf{P}_q(G) \leq b < 1$. Put $\gamma = a(1-b)/4$. Let S be the set of players guaranteed by Theorem* A.1. *Then* $\max\{I_k^q(G) : k \in S\} \geq U$ *where $U > 0$ depends only on a, A, and b.*

PROOF: Theorem A.1 guarantees the existence of a set S of at most $C(\gamma, A)$ players and a simple game $H = \langle S, v_0 \rangle$ such that

$$\mathbf{P}_p\{T \subset N : v(T) \neq v_0(T \cap S)\} < \gamma. \qquad \textit{Q.E.D.}$$

CLAIM:

$$\mathbf{P}_p\{T : v(T \cup S) = 1\} \geq 1 - (2/a)\gamma.$$

PROOF OF THE CLAIM: Let $\mathbf{P}_p^0$ and $\mathbf{P}_p^1$ be the probability distributions induced from $\mathbf{P}_p$ on subsets of S and on subsets of $N \setminus S$, respectively. Note that whether $v(T \cup S) = 1$ depends only on $T \setminus S$ so we have to show that

$$\mathbf{P}_p^1\{T : T \cap S = \emptyset, v(T \cup S) = 1\} \geq 1 - (2/a)\gamma.$$

Now, $\mathbf{P}_p^0\{R \subset S : v_0(R) = 1\} \geq a/2$ and therefore if T is disjoint from S and $v(T \cup S) = 1$, then $\mathbf{P}_p^0\{R : R \subset S, v(T \cup R) \neq v_0(R)\} \geq a/2$. It follows that indeed

$$\mathbf{P}_p^1\{T : T \cap S = \emptyset, v(T \cup S) = 1\} \geq 1 - (2/a)\gamma.$$

We return now to the proof of the lemma. Consider the following operation: Start with a random subset R of players according to $\mathbf{P}_p$. For $j \notin R$ add j to R with probability $(q-p)/(1-p)$. Let R^* be the resulting set of players. The probability that $v(R^*) = 0$ is at least $1 - b$ and the probability that in addition $v(R^* \cup S) = 1$ is at least $1 - b - (2/a)\gamma$ (since $v(R \cup S) = 1$ implies $v(R^* \cup S) = 1$). This means that when we draw a coalition R^* at random according to $\mathbf{P}_q$, the probability that $v(R^*) = 0$ and $v(R^* \cup S) = 1$ is at least $1 - b - (2/a)\gamma$. Now we can examine the effect of adding the players in S one by one. Since $q \in [1/3, 2/3]$ we deduce that $\max\{I_k^q(G) : k \in S\} \geq C^{-1}3^{-C}(1 - b - (2/a)\gamma)$, as required. *Q.E.D.*

We return now to the proof of Theorem 3.4 We can assume that $\gamma \leq 1/10$. Suppose that $T_\epsilon(G) > \gamma$. By Russo's lemma (and the mean-value theorem) there exists p in $[1/3, 2/3]$ such that $I^p(G) \leq A$, where $A = 3/\gamma$. Since G is a strong simple game we can assume that $p \leq 1/2$. By Theorem A.1, there is a set S of players and a simple game $H = \langle S, v_0 \rangle$ such that relation (A.1) holds. The cardinality of S is bounded by a function C of A and ϵ. By our lemma, for every $q \geq p$ in the threshold interval there is a player $k \in S$ such that $I_k^q(G) \geq U$, where U depends only on ϵ and A. Since for every $q \geq p$ in the threshold interval there is a player in S whose influence is at least U, we conclude that $\bar{\phi}(G)$ is larger than $(1/2)\gamma C^{-1}U$. *Q.E.D.*

The proof of Theorem 3.5 is identical. Simply replace the interval $[1/3, 2/3]$ with an appropriate interval around the critical probability of the game.

REMARKS: 1. The proof shows that if all influences $I_k^p(G)$ are sufficiently small for some p in the threshold interval (and in particular for $p = 1/2$), then the threshold interval itself is small. In particular, diminishing individual power according to the Banzhaf power index also implies (but is not equivalent to) asymptotically complete aggregation of information.

2. A simple game is *anonymous* if the game is invariant under all permutations of the players. We define a simple game to be *weakly anonymous* if "every two players are identical." In formal terms this means that the game is invariant under a transitive group of permutations on the players. Simple majority (on an odd number of voters) is the only anonymous strong simple game although the family of weakly anonymous strong simple games is quite rich. Examples include electoral voting systems such as that in the US where all states have the same number of voters and electors. For weakly anonymous strong simple games with n players, the upper bound on the threshold interval is

$$T_\epsilon(G) \le C \log(1/\epsilon)/\log n, \tag{A.2}$$

and this bound is tight (see Friedgut and Kalai (1996)).

REFERENCES

AL-NAJJAR, N., AND R. SMORODINSKY (2000): "Pivotal Players and the Characterization of Influence," *Journal Economic Theory*, 92, 318–342.

ALON, N. (2002): "Voting Paradoxes and Digraphs Realizations," *Advances in Applied Mathematics*, 29, 126–135.

ARROW, K. (1950): "A Difficulty in the Theory of Social Welfare," *Journal of Political Economy*, 58, 328–346.

BOLLOBÁS, B., AND A. THOMASON (1987): "Threshold Functions," *Combinatorica*, 7, 35–38.

DASGUPTA, P. S., AND E. MASKIN (1997): "The Robustness of Majority Rule," Mimeo, Department of Economics, University of Cambridge.

ERDŐS, P., AND L. MOSER (1964): "On the Representation of Directed Graphs as Unions of Orderings," *Magyar Tud. Akad. Mat. Kutat. Int. Kzl.* (*Publications of the Mathematical Institute of the Hungarian Academy of Science*), 9, 125–132.

FEDDERSEN, T., AND W. PESENDORFER (1998): "Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting," *The American Political Science Review*, 92, 23–35.

FRIEDGUT, E. (1998): "Boolean Functions with Low Average Sensitivity," *Combinatorica*, 18, 27–35.

FRIEDGUT, E., AND G. KALAI (1996): "Every Monotone Graph Property Has a Sharp Threshold," *Proceedings of the American Mathematical Society*, 124, 2993–3002.

GRIMMETT, G. (1989): *Percolation*. Berlin: Springer-Verlag.

KAHN, J., G. KALAI, AND N. LINIAL (1988): "The Influence of Variables on Boolean Functions," *Proceedings of the 29th Annual Symposium on Foundations of Computer Science*, 68–80.

MCGARVEY, D. C. (1953): "A Theorem on the Construction of Voting Paradoxes," *Econometrica*, 21, 608–610.

PESENDORFER, W., AND J. M. SWINKELS (2000): "Efficiency and Information Aggregation in Auctions," *American Economic Review*, 90, 499–525.

OWEN, G. (1988): "Multilinear Extensions of Games," in *The Shapley Value*, ed. by A. E. Roth. Cambridge, U.K.: Cambridge Univ. Press, 139–151.

SALANT, Y. (2003): "Learning the Decisions of Small Committees," Discussion Paper 332, Center for the Study of Rationality, Hebrew University, Jerusalem.

SONNENSCHEIN, H. (1972): "Market Excess Demand Functions," *Econometrica*, 40, 549–563.

TALAGRAND, M. (1994): "On Russo's Approximate Zero–One Law," *The Annals of Probability*, 22, 1576–1587.

YOUNG, H. P. (1988): "Condorcet's Theory of Voting," *American Economic Review*, 82, 1231–1244.