Kirstein, Roland

Working Paper
The Condorcet Jury-Theorem with Two Independent Error-Probabilities

CSLE Discussion Paper, No. 2006-03

Provided in Cooperation with:
Saarland University, CSLE - Center for the Study of Law and Economics

Suggested Citation: Kirstein, Roland (2006) : The Condorcet Jury-Theorem with Two Independent Error-Probabilities, CSLE Discussion Paper, No. 2006-03

This Version is available at:
http://hdl.handle.net/10419/23081

Standard-Nutzungsbedingungen:
Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden. Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen. Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes. You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
The Condorcet Jury-Theorem with Two Independent Error-Probabilities

Roland Kirstein*

Center for the Study of Law and Economics
Discussion Paper 2006-03, version (04) of April 2006

Abstract

The Condorcet jury-theorem is derived from the implicit assumption that jury members may only commit one type of error. In binary decision situations however, two error types may occur, the probability of which is independent of each other. Taking this into account leads to a generalization of the theorem.

JEL-Code: D71, K40, L22
Encyclopedia of Law and Economics: 0610, 7100
Keywords: Group decisions, judicial, imperfect decision-making.

*Corresponding author: Center for the Study of Law and Economics, Universität des Saarlandes, Bldg. 31, Room 3.13, D-66041 Saarbrücken, Tel. ++49-681-302-3582, email: rol@mx.uni-saarland.de, homepage: http://rolandkirstein.de. I am grateful to Sven Berg, Giuseppe Dari-Mattiacci, Veronika Nolte, and Dieter Schmidtchen for helpful comments on earlier versions. However, remaining errors are entirely mine.
1 Introduction

The Condorcet jury theorem (henceforth: CJT) states the conditions under which a jury that decides with absolute majority is less likely to commit an error than each single member.\(^1\) The CJT focuses on binary decisions.\(^2\) Such decision situations are governed by two error types, the probabilities of which are independent from each other. This fact is not taken into account in the original CJT. It is the aim of this paper to derive a modified CJT with two error probabilities, which I coin “Condorcet-Heiner-Theorem” (CHT). Two decades ago, Ronald Heiner has set up an influential theory of rule-governed behavior based on binary decision-making. The central parameters of his theory are the two independent error probabilities.\(^3\)

The theorem has been used in Schofield (2002), (2005) and Congleton (2005) to evaluate the merits of representative democracy. Other possible areas of application are public choice, e.g., the analysis of federalism in Mueller (2001), or business administration. Here are possible interpretations of the theoretical paper by Boland (1989), which examines whether it is better to split a jury of, say, nine members into three subcommittees, let each of these subcommittees vote on the issue, and then aggregate the three votes to one decision. Boland comes to the result that such an indirect majority system does not improve the quality of a group decision.

Another application to organizational theory has been provided by Ladha (1992), while Berg/Marañon (2001) and Koh (2005) have analyzed hierarchies. Moreover, the CJT may help to theoretically determine the decision quality of collegial courts compared to that of single judges. An empirical study by Karotkin (1994) has demonstrated that chambers composed of three judges do not come to better judgements in private law cases. In penal law cases, however, the opposite is true. The criterion for decision quality

\(^{1}\)For an overview of Condorcet’s contributions to mathematical economics see Crépeau/Rieucau (2005) and Rothschild (2005). Many real-world examples can be found in Surowiecki (2004) who, however, fails to even mention the name Condorcet.

\(^{2}\)The case of more than two options has been dealt with in List/Goodin (2001).

\(^{3}\)See Heiner (1983, 1985a, 1985b, 1986a, 1986b, 1990a, 1990b). Note that Heiner’s theory is one of boundedly rational behavior. In this paper, however, it is assumed that decision-makers decide rationally.
was the rejection rate in appeal courts.

The CJT may provide valuable insights for the design of court systems. Society wishes courts to avoid errors. If the theorem is true, then society faces a trade-off between decision quality (demanding larger chambers or juries) and the cost of running the court system, as collegial courts are more cost-intensive. Moreover, the duration of a court case might be increased if more judges are involved, as Tullock (1994) has argued. Juries of peers are also costly, as ordinary citizens may face enormous opportunity costs when serving in a jury. These cost aspects are assumed away in the CJT.

The rest of this paper is organized as follows: section 2 briefly repeats the CJT, which only serves to introduce my notation. Section 3 presents the modified theorem: a modified criterion for decision quality is introduced in 3.1, the CHT is proven for the case of three jury members in 3.2, and 3.3 completes the presentation by deriving the results for juries with a higher (and odd) number of members. Section 4 briefly discusses the derived insights.

2 The Condorcet-Jury-Theorem

Assume that a decision body is composed of an odd number of members, denoted as \( k = 2h - 1 \) with \( h \in \mathbb{N} \) and \( h \geq 2 \), and that each of these members decides independently of the others. The collective decision is made with absolute majority, while abstention is neglected and prior communication is excluded. Juries with members who do not decide independently from each other are analyzed by Berg (1993) and Ladha (1995). The qualified majority rule was analyzed by Nitzan/Paroush (1984) and Ben-Yasar/Nitzan (1997). The reliability of jury decisions under alternative majority rules has been compared by Klausner/Pollak (2001).

Moreover, the jury members are assumed to be homogenous: each comes to the correct decision with probability \( q \in [0, 1] \). For larger juries, however, in 1970, the US Supreme Court ruled that state juries need not consist of twelve members (No. 399-U.S. 78, Williams vs. Florida). This decision which has provoked research activities regarding the impact of jury size on the probability of conviction; see Gelfand/Solomon (1973).

See Feddersen/Pesendorfer (1998) who ask whether the decision quality of a jury increases if it switches from a majority to an unanimity rule.
this assumption is hardly satisfied; Berg (1996, 231) derives results for heterogenous juries (see also Paroush (1998)). Finally, it is assumed that the members do not face incentive problems when making their decisions.\footnote{Strategic voting is analyzed in Feddersen/Pesendorfer (1998).}

Now denote as $Q(j, q)$ the probability that $j$ members come to the correct decision, and as $Q_k(q)$ the probability that more than $h$ decide correctly. Then

\[
Q(j, q) = \binom{k}{j} q^j (1 - q)^{k-j}, j \leq k \tag{1}
\]

and

\[
Q_k(q) = \sum_{j=h}^{k} Q(j, q). \tag{2}
\]

Using these definitions, the main claim of the CJT is $\forall h \geq 2 : Q_k(q) > q$ if $q > 1/2$. Moreover, $\partial Q_k(q)/\partial k > 0$ and $\lim_{k \to \infty} Q_k(q) = 1$ can be shown. A proof which does not only derive sufficient, but also necessary conditions is given in Berend/Paroush (1998).

Thus, the CJT states that, with $q > 1/2$, a majority decision of the body is always better than a decision of a single member, and that the probability of a correct decision is strictly increasing in the size of the body. For a body of infinite size, this probability converges to certainty. For $q < 1/2$, the opposite claims are true, in particular $\forall h \geq 2 : Q_k(q) < q$.

3 The Condorcet-Heiner-Theorem

3.1 Imperfect Decision-Making

Courts or juries often face a binary decision and, thus, may commit two types of errors.\footnote{Tullock (1994), Kirstein/Schmidtchen (1997).} E.g., a judge may convict a suspect even though the latter is in fact innocent. Or the judge may acquit a guilty suspect. There is no reason why these two errors should occur with identical probabilities. In general, these error probabilities are independent of each other. However, the CJT completely neglects that a second error probability exists which may differ from the first one.
With two independent error probabilities, however, defining judicial decision quality is not as trivial as in the one-probability case. Assume that two options exist, A and B. Denote as \( r \) the conditional probability that option A is chosen when it is the better one. \( w \) represents the probability that option A is chosen when in fact B is better. A judge or a jury member is said to have “positive detection skill” if \( r > w \). With \( r = w \), the judge would decide the case independently of what the suspect actually did. Moreover, with \( r < 1 \) and \( w > 0 \), the detection skill of the judge is called “imperfect”. This terminology regarding courts has been introduced in Kirstein/Schmidtchen (1997), inspired by the theory of rule-governed behavior by Heiner (1983), (1986), (1990). The case \( r < w \) can safely be neglected.

Now we compare the decision-making abilities of two judges in binary decision-situations. As \( r > w \) denotes positive detection skill, it appears to be sensible to measure the decision quality of a decision-maker by the ratio \( r/w \), which was called “reliability ratio” in Heiner (1983, 566).\(^8\) The reliability ratio is greater than one if the decision-maker has positive detection skill: the probability of a correct decision for option A exceeds the probability of a wrong decision for A. A higher reliability ratio represents higher decision-quality.

The ratio \( r/w \), however, only reflects the decision for option A (which can be right or wrong). We also have to state a quality criterion with regard to the possible decision for option B, which is left out of focus as long as one looks at the ratio \( r/w \) only. Positive detection skill \( r > w \) implies that \( 1 - w > 1 - r \): the probability of correctly choosing B is greater than the probability of wrongly choosing it. For some notational reason, I will use the reciprocal value: better decision quality, with regard to option B, is therefore characterized by a a smaller ratio \( (1 - r)/(1 - w) \).

Now we have completed the criterion for decision-quality, which rests on two conditions. Let a jury member \( i \) be characterized by his decision parameters \( r_i \) and \( w_i \). Then, \( i = 1 \) decides with higher quality than \( i = 2 \) if the

\(^8\)See also Swets (1988), (1998).
following condition holds:

\[
\left[ \frac{r_1}{w_1} > \frac{r_2}{w_2} \right] \land \left[ \frac{1 - r_1}{1 - w_1} < \frac{1 - r_2}{1 - w_2} \right] \quad (3)
\]

If both parts of the condition are not fulfilled, then \( i = 2 \) exhibits higher
decision quality than \( i = 1 \).\(^9\) E.g., consider two judges with \((r_1, w_1) = (0.8, 0.2)\) and \((r_2, w_2) = (0.9, 0.7)\). A direct comparison of the respective error
probabilities \((1 - r_i) \) and \( w_i \) does not allow for a ranking of these two judges’
decision qualities, as \( i = 2 \) performs better with regard to \( r_i \), while judge \( i = 1 \)
shows the better \( w_i \) value. The above mentioned criterion, however, reveals
that the first judge is the better one: \( r_1/w_1 = 4 \) is greater than \( r_2/w_2 = 9/7 \),
and \((1 - r_1)/(1 - w_1) = 1/4 \) is smaller than \((1 - r_2)/(1 - w_2) = 1/3 \).

In the next section, the modified jury theorem for the case of three judges
is derived. The analysis is completed in the overnext section with a similar
analysis for higher odd numbers.

### 3.2 Juries of Three Homogenous Judges

Now assume that a chamber consists of three homogenous judges. Homoge-
nous means that the decision-making quality of each single judges is described
by identical parameters \( r \) (the conditional probability of a correct conviction)
and \( w \) (wrongful conviction). For convenience, define a judge as “suspect-
friendly” if \( r < 0.5 - w \), and as “suspect-hostile” if \( r > 1.5 - w \). Hence, a
judge is neither suspect-hostile nor suspect-friendly if \( 0.5 < r + w < 1.5 \).

The chamber decides with absolute majority, while abstention is prohib-
ited. The probability that \( j \) judges correctly convict the suspect is denoted
\( Q(j, r) \), while the probability that \( j \) judges wrongly do so is \( Q(j, w) \). More-
over, the probability that the chamber correctly (wrongly) convicts is denoted
\( Q_3(r) \) and \( Q_3(w) \), respectively. According to the above definition, the deci-
sion of a chamber of \( k \geq 2 \) members is better than the decision of a single
judge if

\[
\left[ \frac{Q_k(r)}{Q_k(w)} > \frac{r}{w} \right] \land \left[ \frac{1 - Q_k(r)}{1 - Q_k(w)} < \frac{1 - r}{1 - w} \right]. \quad (4)
\]

\(^9\)If only one of the two parts of this condition is fulfilled then this definition would not
allow for a ranking. This is a weakness of this quality criterion which it shares with the
Pareto-criterion. However, this problem is irrelevant for my results.
This allows to prove the following theorem (subsequently denoted as CHT) for the case of $k = 3$.

**Condorcet-Heiner-Theorem:** The decision of a chamber which consists of three homogenous judges is better than the decision of a single judge if, and only if, the judges have imperfect, but positive detection skill, and they are neither suspect-friendly nor suspect-hostile.

**Proof:** For $k = 3$, the theorem reads

$$\left[ \frac{Q_3(r)}{Q_3(w)} > \frac{r}{w} \right] \land \left[ \frac{1 - Q_3(r)}{1 - Q_3(w)} < \frac{1 - r}{1 - w} \right] \iff 0 < w < r < 1 \land \left[ \frac{3}{2} > r + w > \frac{1}{2} \right].$$

As $Q_3(q) = 3q^2(1 - q) + q^3 = 3q^2 - 2q^3, q \in \{r; w\}$, rearrangement of the inequality in the left hand part of (5) yields:

$$\frac{3r^2 - 2r^3}{3w^2 - 2w^3} > \frac{r}{w} \iff 3r - 2r^2 > 3w - 2w^3 \iff 3(r - w) > 2(r - w)(r + w).$$

Obviously, this cannot be fulfilled with $r = w$. With $r < w$, the last expression in (6) is equivalent to $r + w > 1.5$. With with $r > w$, this is equivalent to $r + w < 1.5$.

A symmetric analysis, applied to the right hand side of condition (5), yields $1.5 > (1 - r) + (1 - w) = 2 - r - w \iff r + w > 0.5$ for the case $r > w$, and $r + w < 0.5$ for $r < w$ (while this condition cannot be fulfilled with $r = w$).

The only subset of $(r, w)$ combinations which satisfies both parts of condition (5) is, therefore, characterized by $r > w$ and $0.5 < r + w < 1.5$, which completes the proof. □

Figure 1 displays the result. A perfect judge is symbolized by the bold dot in $r = 1$ and $w = 0$. If a single judge decides perfectly, three homogenous judges would commit no error either, so that the chamber is not better than the single judge. The analysis is therefore limited to imperfect judges.

The hostile judge is characterized by an $(r, w)$ combination in the upper right corner, which is labeled d). The parameters of a suspect-friendly judge are in the lower left corner of the unit box, labeled a). The area called
“chamber” comprises the \((r, w)\) combinations which fulfill the prerequisites of the CHT. Here, a chamber of three judges performs better than a single judge. In the other area (“single judge”), we have the \((r, w)\) combinations which imply that the decision of a single judge has superior quality.

The first statement of the CJT is obviously a special case of the CHT. Implicitly, the CJT either assumes \(w = 0\): there are no wrongful convictions, hence judicial error may only occur in form of a wrongful acquittal (with probability \(1 - r\)). In this case, only the left border of the r-w-box is relevant. Or it assumes the probabilities of the two error types to be equal: \(w = 1 - r\).\(^{10}\) In that case, the r-w-parameters of a single judge are confined to the diagonal that starts in the upper left corner of figure 1. With one of these implicit assumptions, the CHT would make the same claim as the CJT: if \(r > 0.5\) (and \(w = 0\)), then the chamber performs better than the single judge.

The generalized theorem reveals that, even if an individual judge is characterized by an error probability \(r < 0.5\), the chamber of three judges may be

\(^{10}\)This implicit assumption has been made in Gelfand/Solomon (1973, 272).
less likely to commit errors than each of its single members. This is the case in the small triangle labeled b) in the south of the “chamber” area (below the dotted line). Even though \( r < 0.5 \), the chamber of three judges performs better, as the higher probability of the first type error is outweighed by a small probability of a second-type error. Hence, it has been proven that \( r > 0.5 \) is not a necessary condition for the group decision to have higher quality.

Moreover, \( r > 0.5 \) is not a sufficient condition for the superiority of juries, either. Consider the case \( w > 0.5 \). Here, the likelihood of a second-type error is relatively great. In contradiction to the CJT, \( r > 0.5 \) is not a sufficient condition for a chamber to perform better. In fact, only in the \((r, w)\) combinations in the triangle labeled c) (to the right of the vertical dotted line) imply that the decision quality of the chamber is higher. To the right of this triangle, in area d), the single judge performs better than the chamber of three judges, despite of \( r > 0.5 \).

### 3.3 Juries of More than Three Judges

For juries which consist of \( k \) judges, with \( k = h - 1, h > 2 \) and \( h \in \mathbb{N} \), the computations become more awkward without generating qualitatively different insights. The assumption of homogenous judges is maintained.\textsuperscript{11} In the case of five members, e.g., condition (4) expands to

\[
r^2(10 - 15r + 6r^2) > w^2(10 - 15w + 6w^2)
\]

and \((1 - r)^2[6(1 - r)^2 + 15r - 5] < (1 - w)^2[6(1 - w)^2 + 15w - 5]\). With \( k = 7 \), the jury performs better than a single judge if, and only if

\[
r^3(35 - 84r + 70r^2 - 20r^3) > w^3(35 - 84w + 70w^2 - 20w^3)
\]

and

\[
(1 - r)^3[35 - 84(1 - r) + 70(1 - r)^2 - 20(1 - r)^3] < (1 - w)^3[35 - 84(1 - w) + 70(1 - w)^2 - 20(1 - w)^3]
\]

\textsuperscript{11} Karotkin/Paroush (2003) assume that individual decision skill may decrease if the size of the jury increases, which allows them to derive an optimal jury size. Mukhopadhaya (2003) introduces the idea of a free-rider problem in juries: an individual member may have less incentive to pay attention in court, the larger the jury. In this setting, larger juries may even perform worse than smaller ones.
hold. Finally, the condition for nine judges would be

\[ r^4(126 - 420r + 540r^2 - 315r^3 + 70r^4) \]

\[ > w^4(126 - 420w + 540w^2 - 315w^3 + 70w^4) \]  \hspace{1cm} (7)

and

\[ (1 - r)^4[126 - 420(1 - r) + 540(1 - r)^2 - 315(1 - r)^3 + 70(1 - r)^4] \]

\[ < (1 - w)^4[126 - 420(1 - w) + 540(1 - w)^2 - 315(1 - w)^3 + 70(1 - w)^4] \]

Both parts of condition (4) imply a figure which looks almost identical to figure 1 with just two differences: the lower one of the falling diagonals is now concave, the upper one convex. The main result, however, is still valid: two areas of \( r-w \)-combinations exist where the decision of the chamber has higher quality, even though \( r < 0.5 \) or \( w > 0.5 \), respectively.

Figure 2: The CHT with juries of 3, 5, 7, and 9 judges

Figure 2 displays the cases of \( k \in \{3; 5; 7; 9\} \). The downwards sloped lines (and the \( r = w \) line) limit the subset of \( r-w \)-combinations under which a jury decision is better than the decision of a single judge. The innermost pair of curves refers to the case \( k = 9 \), the next pair to \( k = 7 \), then \( k = 5 \), and the outermost (straight) lines are valid if \( k = 3 \). The lower (upper) lines start in
r = 0.5, w = 0 (r = 1, w = 0.5) and cross the main diagonal below (above) $r = w = 0.5$; r-w-combinations for which a group decision is better than an individual decision are found above this diagonal only.

Without general proof, the four examples demonstrate that the general claim made by the CHT is valid (with the necessary modifications) for juries with more than three members. In all these examples, r-w-combinations exist where the juries perform better even though the individual (and homogenous) members are characterized by one probability to make a correct decision which is smaller than 0.5 (either $r < 0.5$ or $w > 0.5$). While the subset of r-w-combinations with that property shrinks as the jury size grows, it remains non-empty for $n < \infty$.

4 Discussion

Judges and juries have to make binary decisions and, therefore, may commit two types of errors the probabilities of which are independent of each other. This fact is not taken into account in the original formulation of the Condorcet-Jury-Theorem. Acknowledging this makes a more complex definition of “decision quality” inevitable. A proposal for such a definition is the condition (4) above.

This condition allows for a generalization of the first statement of the Condorcet-Jury-Theorem: the Condorcet-Heiner-Theorem. According to the CHT, a jury of three homogenous members makes better decisions than each single member if they have positive detection skill and are neither too suspect-friendly nor too suspect-hostile (as defined above by r-w-combinations in the upper and lower left corner of the r-w-box). A corollary of the CHT is that combinations of the parameters $r$ and $w$ exist under which the probability of a correct decision is smaller than 0.5, nevertheless the group decision has a higher quality than an individual decision. In the areas of the r-w-box for which this result holds, the low probability of making a correct decision for one of the available options is outweighed by a high probability for choosing correctly the other option. Hence, the CHT demonstrates that a probability of a correct decision which is greater than 0.5 is neither sufficient, nor necessary for the jury decision to have higher quality than an individual
decision.

Just as the CJT, the CHT depends on strict assumptions. Some of them (homogenous judges, independent decision-making) have been dealt with in the literature that was mentioned in the introduction above. In particular, the optimistic results regarding juries is based on the assumption that the jury members are perfectly rational. Real-world actors, however, often violate the Bayes’ rule when integrating new information.\textsuperscript{12}

References


\textsuperscript{12}Bottom/Ladha/Miller (2002) have introduced boundedly rational jury members and shown that, under the majority rule, individual bias is not attenuated by group decisions.


