Abstract

The size of a supermajority required to change an existing contract varies widely in different settings. This paper analyzes the optimal supermajority requirement, determined by multilateral bargaining behind the veil of ignorance, where there are a continuum of possible policies. The optimum is determined by a tradeoff between reducing blocking power of small groups and reducing expropriation of minorities. We solve for the optimal supermajority requirement as a function of the distribution of voter types, the number of voters and the degree of importance of the decision. The findings are consistent with observed heterogeneity of supermajority requirements in different settings and jurisdictions.

Keywords: Supermajority, social contract.

JEL Classification Codes: D63, D72, D74.
1 Introduction

Almost all agreements contain provisions governing the process by which the terms of the agreement can be changed. Often these clauses require a supermajority (more than 50%) of the parties to agree in order to make a change. Constitutions of democratic countries are perhaps the most prominent example of this. Yet the phenomenon is far more widespread. Majority creditor clauses in corporate and sovereign debt contracts provide that a supermajority of the creditors can bind other creditors in renegotiations with the borrower. The size of the required supermajority varies. Filibuster rules mean that a 60% majority of the United States Senate is required to appoint a federal judge. Corporations laws in different countries specify different supermajorities, sometimes as high as 95%, required to compulsorily acquire (or “squeeze-out”) equity securities.\footnote{For instance, in Australia 90% of shareholders must accept a takeover offer for the bidder to be able to move to compulsory acquisition. In the UK the requirement is 75%, and it is a simple majority in the US. Since 2002 the requirement in Germany is 95%.} To change the International Monetary Fund Articles of Agreement requires an 85% vote of member nations. The five permanent members of the UN security council all have a veto–in effect 100% supermajority rule.

The variation in supermajority requirements\footnote{We shall use the terms “supermajority requirement” and “supermajority rule” interchangeably.} in different settings and jurisdictions raises the obvious question of why they differ. This paper presents a model of collective decision making in order to analyze supermajority rules. Individuals engage in multilateral over a contract behind the veil of ignorance (i.e. they know the distribution from which their preferences will be drawn, but not their actual preference). The contract has two elements: a default social choice (from the unit interval), and the proportion of voters who need to agree \textit{ex post} in order to change the social choice (i.e. a voting rule). As is common in voting models, we rule out the possibility of side-payments. The optimal supermajority requirement is determined by a trade-off between two factors. On the one hand a high supermajority is attractive because it reduces the problem of minorities being expropriated by the majority. On the other hand, a high supermajority is detrimental because it provides a small group of voters with blocking power in the sense that they can prevent the efficient action from being taken. It is the tradeoff between minority protection and blocking minimization which determines the optimal
supermajority requirement. We show how the rule depends on the distribution of voter preferences, the number of voters, and the risk-aversion of the voters.

Our most striking result is that the optimal supermajority rule declines in percentage terms as the number of voters increases. The logic of the result can be illustrated as follows. Suppose there are 3 voters and the voting rule is 2/3. Suppose (for simplicity) that voter preferences are uniformly distributed on $[0, 1]$. The status quo policy and supermajority rule are chosen before voters draw their preferences, so the optimal ex ante policy is clearly 1/2. To change this ex post (after preferences are realized) one needs at least 2 voters to want to change. This will (with probability 1) be the case because at least 2 voters will draw a preference greater or less than 1/2 (except in the zero probability event of drawing exactly 1/2). So there will always be amendment.

Now consider the same setting but with 6 voters. One now needs at least 4 voters to change the policy ex post. Now enumerate all the possible “events” of preference draws. One is that three voters draw preferences less than 1/2 and three draw preferences greater than 1/2. This happens with positive probability, and in this case there will be no change because any policy change makes at 3 voters worse off. Thus, the tradeoff between flexibility and minority protection is affected by the number of voters.

This logic comes from various properties of order statistics. One can view the preference draws as order statistics draw from a common parent distribution. The key is that the variance of a given order statistic is decreasing as the number of draws from the parent increases. For example, it is just very unlikely that the 2nd order statistic of 100 drawn on $[0, 1]$ is going to be less than, say, 1/2. on the other hand, it is “quite” possible that the 2nd of three is less than 1/2. It turns out that this logic extends to any order statistic from any continuous parent distribution.\(^3\)

Other results, such as the convergence properties of the optimal supermajority rule as the number of voters goes large, are more subtle. For instance, given the above logic one might conjecture that the optimal voting rule converges to 50 percent as the number of voters goes to infinity. This is, in fact, not true in general, but only for symmetric distributions of voter preferences.

\(^3\)Dixon and Holden (2012) provide evidence from US State constitutions to show, using a panel-data identification strategy, that states with large voting age populations have less stringent amendment provisions.
A contribution of the paper is to use techniques from the theory of order statistics to characterize the optimal supermajority rule in a fairly general setting.

This paper builds on a large literature in both economics and political science examining voting rules and on supermajority rules in particular. In economics, interest in supermajority voting rules can be traced to Black (1948). Attention to supermajority requirements has also been an important part of subsequent work in social choice theory. Arrow himself conjectured (Arrow (1951)) that a sufficient degree of social consensus could overcome his impossibility theorem. This conjecture was formalized by Caplin and Nalebuff (1988) and with greater generality by Caplin and Nalebuff (1991).

More recent work on incomplete contracts also develops a framework for analyzing optimal supermajority requirements in certain contexts. Aghion and Bolton (1992) show that some form of majority voting dominates a unanimity requirement in a world of incomplete social contracts. They highlight the fact that if a contract could be complete then the issue of supermajority requirements is moot if rules are chosen behind the veil of ignorance. Aghion et al. (2004) utilize a related framework, in the spirit of public good provision analyzed by Roemer and Rosenthal (1983). In a similar model, Erlenmaier and Gersbach (2001) consider “flexible” majority rules whereby the size of required supermajority depends on the proposal made by the agenda setter. Barbera and Jackson (2004) consider “self-stable” majority rules, in the sense that the required supermajority does not wish to change the supermajority rule itself ex post. A related paper is Maggi and Morelli (2006), which finds that unanimity, in certain settings, is usually optimal if there is imperfect enforcement. Finally, Harstad (2005) analyzes a model where members of a club invest in a joint project and then vote over its implementation—focusing on the interaction between voting rules and investment incentives.

Despite this literature there is no general model of optimal supermajority rules. In our model we consider a general formulation where the policy set is a continuum. This allows us to study the effect of risk and risk-aversion on the voting rule. As discussed in section 3, we consider a particularly strong form of incompleteness of the social con-

---

4Early works by economists using this notion include Vickrey (1945), Harsanyi (1953) and Harsanyi (1976). Mirrlees (1971) and, of course, Rawls (1999) analyze profound questions within this framework.

5“The solution of the social welfare problem may lie in some generalization of the unanimity condition...” (quoted in Caplin and Nalebuff (1988))
The social contract is not permitted to specify a state-contingent supermajority rule, nor are monetary transfers / side payments allowed. In the context of the model this means that the supermajority requirement cannot differ based on realized draws from the distribution of types.

The remainder of the paper is organized as follows. Section 2 is the heart of the paper: it states the problem and then analyzes two examples in detail which illustrate the approach to the analysis and foreshadow the general results which are obtained in the following subsection. Section 3 discusses some extensions and concludes.

2 The Model

2.1 Statement of the Problem

Let there be \( n \) voters, with \( n \) finite. The policy space is assumed to be the unit interval \([0, 1]\). Voters preferences over this policy space are drawn from the distribution function \( F(x) \).

Definition 1. A Social Decision is a scalar, \( \theta \in [0, 1] \).

Assumption A1. Each voter \( i \) has a utility function of the form

\[
U_i = -u(|\theta - x_i|),
\]

where \( u(\cdot) \) is an increasing function, and \( x_i \) is voter \( i \)'s preferred policy.

We are thus assuming that voters are \( \text{ex ante} \) identical, but not (generically) \( \text{ex post} \).

Definition 2. A Supermajority Rule is a scalar \( \alpha \in [\frac{1}{2}, 1] \) that determines the proportion of voters required to modify the social decision.

There are two time periods in the model. In period 1 voters know the distribution of preferences, \( F(x) \), but they do not know their draw from the distribution. In this period they determine, behind the veil of ignorance, a social choice and a supermajority rule. In period 2, after preferences are realized, the social decision can be changed if a coalition of at least \( \alpha n \) voters prefer a new social decision.
Note that, because of the timing of the realization of voter preferences, and because voters are \textit{ex ante} identical, we can treat this as a control problem, not a game.

As mentioned before, we restrict the (social) contracting space. State contingent supermajority rules are not permitted. An example of such a rule would be any kind of utilitarian calculus which would vary the supermajority requirement to change the status quo according to the aggregate utility to be gained \textit{ex post}. We also rule out monetary transfers / side payments. Let $\hat{\theta}$ be the \textit{ex ante} optimal social decision.

\textbf{Definition 3.} The \textit{ex post} optimal social decision is:

$$\theta^* = \arg\max_\theta \sum_{i=1}^n u(|\theta - x_i^*|).$$

With a finite number of voters the \textit{ex post} optimal decision may well differ from the \textit{ex ante} optimal decision because of the realized draws from $F(x)$. It is this wedge between \textit{ex ante} and \textit{ex post} optimality which creates complexity in the choice of the optimal supermajority rule.

We make the following technical assumption which enables us to avail ourselves of several useful results from the theory of order-statistics.

\textbf{Assumption A2.} The parent distribution of voter types $F(x)$ is absolutely continuous.

By using order-statistics we are able to fully characterize the aggregate expected utility of a given voter for an arbitrary distribution of the population, number of voters, degree of risk-aversion and supermajority rule. We are, therefore, able to determine which rule yields the highest expected utility, and is hence optimal.

There is an obvious issue of how the \textit{ex post} social decision is determined if a coalition has a sufficient number of members relative to the required supermajority who would be made better-off by a change to the \textit{ex ante} social decision. In principle, any \textit{ex post} social decision within the interval spanned by their preferences improves each of their payoffs. For simplicity we make the following assumption about how the bargaining power amongst members of such a coalition.

\textbf{Assumption A3.} If a coalition has the required supermajority \textit{ex post} then the social decision is
that preferred by the “final” member of the coalition. That is, the member of the coalition whose preference is closest to the ex ante social decision.

2.2 Examples

2.2.1 Example 1

Voters’ types are drawn from the uniform distribution on \([0, 1]\), \(u_i = -\exp\{\beta |\theta - x_i|\}\), and \(n = 5\).

First note that the ex ante optimal social decision is simply \(\theta^* = \frac{1}{2}\). First we focus on the outcome under majority rule, which is simply that the ex post social decision is the median of the voters’ draws. Consider voter \(i\) and let the other voters’ draws be:

\[
x_1^* \leq x_2^* \leq x_3^* \leq x_4^*
\]

where \(x_k^*\) is the \(k\)th order-statistic. Now note that the density of \((x_2^*, x_3^*)\) on \([0, 1] \times [0, 1]\) is:

\[
f(a_2, a_3) = 24a_2(1 - a_3)
\]

Note that in considering the median we need only be concerned with voter \(i\)’s position relative to \(x_2^*\) and \(x_3^*\). If they are between \(x_2^*\) and \(x_3^*\) then they are the median. If \(x_2^* \leq x_i^* \leq x_3^*\) then the expected loss is \(\int_{a_2}^{a_2} -\exp\{\beta |t - a_2|\} dt\) and if \(x_i^* \geq x_3^*\) it is \(\int_{a_3}^{1} -\exp\{\beta |t - a_3|\} dt\). If \(x_2^* \geq x_i^* \geq x_3^*\) then the expected loss is \(-\exp(0) = -1\). The expected utility of voter \(i\) is

\[\text{For an absolutely continuous population the joint density of two order statistics } i < j, \text{ from } n \text{ statistics, is given by:}\]

\[
\frac{n!}{(i-1)! (j-i-1)! (n-j)!} F(x_i)^{i-1} (F(x_j) - F(x_i))^{j-i-1} \left[(1 - F(x_j))^{n-j} f(x_i) f(x_j)\right]
\]

(See Balakrishnan and Rao (1998)). For the uniform distribution this implies:

\[
f(x_i, x_j) = \frac{n!}{(i-1)! (j-i-1)! (n-j)!} u_i^{i-1} (u_j - u_i)^{j-i-1} (1 - u_j)^{n-j}
\]
therefore:

\[
E \left[ u_i^M \right] = \int_0^1 \int_0^{a_3} \left( \int_0^{a_2} - \exp \{ \beta(a_2 - t) \} \, dt ight) \left( \int_0^{a_2} (-1) \, dt ight) \left( \int_1^{a_3} - \exp \{ \beta(t - a_3) \} \, dt \right) 24a_2(1-a_3)da_2da_3 \\
= -\beta(\beta^4 - 10\beta^3 + 120\beta + 480) + 240e^\beta(\beta - 3) + 720 \quad \frac{5\beta^5}{\beta^5}.
\]

Now consider the expected utility of voter \( i \) if we require unanimity in order to change the social decision \( \text{ex post} \). Denote the \( \text{ex post} \) social decision as \( t \). Let \( B \) be the event where \( 0 \leq x_1^* \leq x_4^* < \frac{1}{2} \) and let \( B' \) be the event where \( \frac{1}{2} \geq x_1^* \geq x_4^* \geq 1 \). Let \( A = \Omega \setminus (B + B') \). It is clear that \( \Pr(A) = \frac{7}{8} \) and that \( \Pr(B) = \Pr(B') = \frac{1}{16} \). The expected utility of voter \( i \) conditional on event \( A \) is:

\[
E \left[ u_i^U | A \right] = 2 \int_0^{\frac{1}{2}} - \exp \left\{ \beta \left( \frac{1}{2} - t \right) \right\} \, dt \\
= 2 \left( 1 - e^{\beta/2} \right) / \beta.
\]

The remaining calculations for this case are contained in the appendix, where we show that the total expected utility under unanimity is:

\[
E \left[ u_i^U \right] = \frac{7}{8} E \left[ u_i^U | A \right] + \frac{1}{16} E \left[ u_i^U | B \right] + \frac{1}{16} E \left[ u_i^U | B' \right] \\
= -3840 + \beta^4(\beta - 160) + 10e^{\beta/2}(-384 + \beta(192 + \beta(\beta(15\beta + 8)48))) \quad \frac{80\beta^5}{\beta^5}.
\]

For majority rule to be preferable to unanimity therefore requires \( E \left[ u_i^M \right] > E \left[ u_i^U \right] \). Solving numerically shows that this is the case if and only if \( 0 \leq \beta \leq 3.9 \). Therefore when the decision is relatively unimportant majority rule dominates, but with a sufficiently high enough degree of importance unanimity is preferred.

Now consider the case where the social decision can be altered \( \text{ex post} \) if four voters agree. In this example with five voters this reflects the only supermajority which is greater than simple majority but less than unanimity.

Now define events \( B, B', C \) and \( C' \) as follows. \( B \) is the event where \( 0 \leq x_1^* \leq x_2^* \leq
$x_3^* \leq x_4^* \leq \frac{1}{2}$. $B'$ is the event where $\frac{1}{2} \leq x_1^* \leq x_2^* \leq x_3^* \leq x_4^* \leq 1$. $C$ is the event where $0 \leq x_1^* \leq \frac{1}{2} \leq x_2^* \leq x_3^* \leq x_4^* \leq 1$. $C'$ is the event where $0 \leq x_1^* \leq x_2^* \leq x_3^* \leq \frac{1}{2} \leq x_4^*$. Also, let $A = \Omega \setminus (B + B' + C + C')$.

Figure 1: Events for 80%-Supermajority Rule

Note that $\Pr(B) = \Pr(x_4^* \leq \frac{1}{2}) = \frac{1}{16} = \Pr(B')$. $\Pr(C') = \Pr(x_3^* \leq \frac{1}{2} \land x_4^* \geq \frac{1}{2}) = \frac{1}{4} = \Pr(C')$. Also note that $\Pr(A) = \frac{3}{8}$.

Again, we relegate the calculations to the appendix, where we show that the expected utility for 80% supermajority rule is:

$$E[u_i^S] = \left(\frac{-1920e^\beta + \beta^4(80 - 3\beta) + 1920(\beta + 3)}{-10e^{\beta/2}(284 + \beta(-192 + \beta(\beta + 4)(5\beta - 12)))}\right)\frac{40\beta^5}{40\beta^5}$$

For an 80% supermajority to be preferable to majority rule therefore requires $E[u_i^S] > E[u_i^M]$. Solving numerically shows that this is the case if and only if $\beta \gtrsim 2.69$. For una-
nimity to be superior to an 80% supermajority rule requires $E[u_i'] > E[u_i^S]$. Solving numerically reveals that this the case for $\beta \gtrapprox 9.02$. That is, the 80% supermajority rule dominates unanimity until the degree of importance becomes sufficiently large. For sufficiently large degrees of importance unanimity dominates because the fear of expropriation dominates and a veto provides them with insurance against this possibility. Therefore, in this example, for $0 \gtrapprox \beta \gtrapprox 2.69$ majority rule is optimal, for $2.69 \gtrapprox \beta \gtrapprox 9.02$ an 80% supermajority requirement is optimal, and for $\beta \gtrapprox 9.02$ a unanimity requirement is optimal. This is reflected in the following figure.

![Figure 1 – Voting rule comparison](image-url)
2.2.2 Example 2

Voters’ types are drawn from the uniform distribution on $[0, 1]$, $u_i = -\exp \{\beta |\theta - x_i|\}$, and $n = 3$.

This example illustrates that as the number of voters increases the optimal supermajority rule decreases. We again use the uniform distribution, but with 3 voters rather than 5.

The expected utility under majority rule (here 2 out of three voters) is\(^7\):

$$E[U_M] = \int_0^1 \int_0^{a_2} \left[ \int_0^{a_1} - \exp \{\beta (a_1 - t)\} dt + \int_0^{a_2} - \exp \{\beta (t - a_2)\} dt \right] 2 da_1 da_2$$

$$= \frac{12 - 12 e^\beta (12 - \beta (\beta - 6))}{3 \beta^3}.$$

The expected utility under a unanimity requirement is:

$$E[U_U] = \int_0^{1/2} - \exp \{\beta (1/2 - t)\} dt + (\int_0^{1/2} \left( \int_0^{a_2} - \exp \{\beta (a_2 - t)\} dt - 1 \int_0^{1/2} dt \right) 8 a_2 da_2)$$

$$= \frac{-48 + \beta^2 (\beta - 12) + 6e^{\beta/2}(-8 + \beta (\beta + 4))}{6 \beta^3}.$$

Now consider $\beta = 5$. In this case, where $n = 3$, a unanimity requirement is optimal and yields expected utility of approximately $-3.44$. Where $n = 5$ (ie. example 1) and $\beta = 5$ an 80% supermajority is optimal and the expected utility is approximately $-4.12$. For majority rule under $n = 3$ expected utility is $-4.50$. This illustrates the general point made in Theorem 1, that unless adjusted downward, any given supermajority rule will over-protect against the danger of oppression, as compared to blocking, as the voting population increases: or be subject to a form of constitutional inflation, which increases the effective hurdle to achieving constitutional change.

\(^7\)Note that the joint density of $(x_1, x_2)$ where there are just two order statistics is simply 2.
2.3 General Results

**Theorem 1.** Assume A1-A3. Then the optimal supermajority rule is (weakly) decreasing in the number of voters, \( n \).

**Proof.** Let event 1 be the event that \( x_1^* \leq ... \leq x_n^* \leq \hat{\theta} \), where \( x_i^* \) is the \( i \)th order statistic. Let event 2 be the event that \( x_1^* \leq ... \leq x_{n-1}^* \leq \hat{\theta} \leq x_n^* \), and so on up to event \( n+1 \). Note then that the probability of event \( j \) occurring is given by

\[
\pi_j = \Pr \left( \text{Event } j \right) = F^{(j-1)} \left( \hat{\theta} \right) \left[ 1 - F \left( \hat{\theta} \right) \right]^{(n-j+1)}.
\]

When the number of voters required to change the social decision ex post is \( \lambda = \psi \left( \alpha n \right) \), where \( \psi \) is the ceiling function which rounds its argument up to the nearest integer, utility conditional on draws \( x_1^*, ..., x_n^* \) is

\[
\bar{V} = \sum_{i=1}^{n} \sum_{j=1}^{n+1} \pi_j u \left( |\theta^* - x_i^*| \right).
\]

By A3, the ex post social choice under supermajority rule \( \lambda \) is \( x_\lambda^* \) for \( j \leq \left( (n+1)/2 \right) \) and \( x_{n+1-\lambda}^* \) for \( j > \left( (n+1)/2 \right) \). We can thus write \( \bar{V} \) as

\[
\bar{V} = -\sum_{i=1}^{n} \left( \sum_{j=1}^{(n+1)/2} \pi_j u \left( |x_\psi^\prime \left( \alpha n \right) - x_i^*| \right) + \sum_{j=\left( (n+1)/2 \right)+1}^{n+1} \pi_j u \left( |x_{n+1-\psi \left( \alpha n \right)}^* - x_i^*| \right) \right).
\]

Expected utility involves integrating over all possible realizations of the order statistics—that is, over their joint pdf. Thus, expected utility is

\[
E[\bar{V}] = -\int \cdots \int \sum_{i=1}^{n} \left( \sum_{j=1}^{(n+1)/2} \pi_j u \left( |x_\psi^\prime \left( \alpha n \right) - x_i^*| \right) + \sum_{j=\left( (n+1)/2 \right)+1}^{n+1} \pi_j u \left( |x_{n+1-\psi \left( \alpha n \right)}^* - x_i^*| \right) \right) f \left( a_1, ..., a_n \right) da_1...da_n.
\]

This can be simplified by noting that the joint pdf of all \( n \) order statistics is \( n! \), since the unordered sample has density equal to 1 and there are \( n! \) different permutations of the
sample corresponding to the same sequence of order statistics. Thus we have

\[ E[V] = - \int \cdots \int \sum_{i=1}^{n} \left( \frac{\pi_j}{\sum_{j=1}^{n+1}} \right) \left( \sum_{j=1}^{n+1/2} \pi_j \left( \left| x_{\psi(n)}^{*} - x_i^{*} \right| \right) + \sum_{j=(n+1)/2+1}^{n+1} \pi_j \left( \left| x_{n+1-\psi(n)}^{*} - x_i^{*} \right| \right) \right) n! \, da_1 \cdots da_n. \]  

(1)

Denote the optimal supermajority rule as \( \alpha^{*} = \arg \max_{\alpha} \{ E[V] \} \). By the Monotonicity Theorem of Milgrom and Shannon (1994), a necessary and sufficient condition for \( \alpha^{*} \) to be nonincreasing in \( n \) is that \( E[V] \) have decreasing differences in \( (\alpha, n) \). This requires that for all \( n' \geq n \), \( E[V](n', \alpha) - E[V](n, \alpha) \) is nonincreasing in \( \alpha \). Assuming for simplicity that \( n \) and \( n' \) are odd (the generalization to even integers is simply a matter of notation) this entails

\[ \int \cdots \int \sum_{i=1}^{n} \left( \frac{\pi_j}{\sum_{j=1}^{n+1}} \right) \left( \sum_{j=1}^{n+1/2} \pi_j \left( \left| x_{\psi(n)}^{*} - x_i^{*} \right| \right) + \sum_{j=(n+1)/2+1}^{n+1} \pi_j \left( \left| x_{n+1-\psi(n)}^{*} - x_i^{*} \right| \right) \right) n! \, da_1 \cdots da_n, \]  

(2)

\[ - \int \cdots \int \sum_{i=1}^{n'} \left( \frac{\pi_j}{\sum_{j=1}^{n'+1}} \right) \left( \sum_{j=1}^{n'+1/2} \pi_j \left( \left| x_{\psi(n')}^{*} - x_i^{*} \right| \right) + \sum_{j=(n'+1)/2+1}^{n'+1} \pi_j \left( \left| x_{n'+1-\psi(n')}^{*} - x_i^{*} \right| \right) \right) (n')! \, da_1 \cdots da_{n'} \]

to be nonincreasing in \( \alpha \) for all \( n' \geq n \) and all \( \alpha \). An increase in \( \alpha \) makes the term \( \sum_{j=1}^{n+1/2} \pi_j \left( \left| x_{\psi(n)}^{*} - x_i^{*} \right| \right) \) larger, since \( \left| x_{\psi(n)}^{*} - x_i^{*} \right| \) increases and the probabilities \( \pi_j = F(j-1) (\theta) \left[ 1 - F(\theta) \right]^{(n-j+1)} \) are unchanged. Similarly

\[ \sum_{j=(n+1)/2+1}^{n+1} \pi_j \left( \left| x_{n+1-\psi(n)}^{*} - x_i^{*} \right| \right) \]

is larger and so the first line of (2) is overall larger. Note, however, that the second line of (2) increases by more than the first line for a given change in \( \alpha \). The first term inside the parentheses in the second line increases by more than its corresponding term in the first line since each term \( u(\cdot) \) is weakly larger in the second line by construction of the ordering of the order statistics, probabilities sum to 1, and \( n'! > n! \). This argument is true for all \( \alpha \) and \( n' > n \), which we establish directly in the following

13
Claim 1.

\[ u \left( \left| x^*_{\psi(\alpha + \delta)n'} - x_i^* \right| \right) - u \left( \left| x^*_{\psi(\alpha)n'} - x_i^* \right| \right) \geq u \left( \left| x^*_{\psi(\alpha + \delta)n} - x_i^* \right| \right) - u \left( \left| x^*_{\psi(\alpha)n} - x_i^* \right| \right), \]

for \( \delta \geq 0 \), and

\[ u \left( \left| x^*_{\psi(n' + 1 - (\alpha + \delta)n')} - x_i^* \right| \right) - u \left( \left| x^*_{\psi(n' + 1 - \alpha n')} - x_i^* \right| \right) \geq u \left( \left| x^*_{\psi(n + 1 - (\alpha + \delta)n)} - x_i^* \right| \right) - u \left( \left| x^*_{\psi(n + 1 - \alpha n)} - x_i^* \right| \right), \]

for \( \delta \geq 0 \).

Proof. Part 1: If \( n' \geq n \) then \( \alpha n' \geq \alpha n \) and \( (\alpha + \delta)n' \geq (\alpha + \delta)n \) for \( \alpha \in [0.5, 1] \) and \( \delta \geq 0 \). Therefore \( \| x^*_{\psi(n')} - x_i^* \| \geq \| x^*_{\psi(n)} - x_i^* \| \). Now, given Assumption A1, \( u(\cdot) \) is increasing and therefore

\[ u(\| x^*_{\psi(n')} - x_i^* \|) \geq u(\| x^*_{\psi(n)} - x_i^* \|). \]  

(3)

Similarly,

\[ u \left( \left| x^*_{\psi((\alpha + \delta)n)} - x_i^* \right| \right) \geq u \left( \left| x^*_{\psi((\alpha + \delta)n)} - x_i^* \right| \right) \]  

(4)

Therefore, by subtracting (4) from (3) and Assumption A1, we have

\[ u \left( \left| x^*_{\psi((\alpha + \delta)n') - x_i^*} \right| \right) - u \left( \left| x^*_{\psi(n') - x_i^*} \right| \right) \geq u \left( \left| x^*_{\psi(n + 1 - \alpha n) - x_i^*} \right| \right) - u \left( \left| x^*_{\psi(n + 1 - \alpha n) - x_i^*} \right| \right), \]

for \( \delta \geq 0 \), as required.

Part 2: If \( n' \geq n \) then \( \alpha n' \geq \alpha n \) and \( (\alpha + \delta)n' \geq (\alpha + \delta)n \) for \( \alpha \in [0.5, 1] \) and \( \delta \geq 0 \). Now, since \( n' + 1 \geq n + 1 \) it follows that \( n' + 1 - \alpha n' \geq n + 1 - \alpha n \) and \( n' + 1 - (\alpha + \delta)n' \geq n + 1 - (\alpha + \delta)n \). Therefore \( \| x^*_{\psi(n' + 1 - \alpha n')} - x_i^* \| \geq \| x^*_{\psi(n + 1 - \alpha n)} - x_i^* \| \).

Now, given Assumption A1, \( u(\cdot) \) is increasing and therefore

\[ u(\| x^*_{\psi(n' + 1 - \alpha n')} - x_i^* \|) \geq u(\| x^*_{\psi(n + 1 - \alpha n)} - x_i^* \|). \]  

(5)

Similarly,

\[ u(\| x^*_{\psi(n' + 1 - (\alpha + \delta)n')} - x_i^* \|) \geq u(\| x^*_{\psi(n + 1 - (\alpha + \delta)n)} - x_i^* \|). \]  

(6)
Therefore, by subtracting (6) from (5) and Assumption A1, we have
\[
\begin{align*}
&u \left( \| x^*_{\psi(n'+1-(\alpha+\delta)n')} - x^*_i \| \right) \\
&- u \left( \| x^*_{\psi(n'+1-\alpha n')} - x^*_i \| \right) \geq u \left( \| x^*_{\psi(n+1-(\alpha+\delta)n)} - x^*_i \| \right) - u \left( \| x^*_{\psi(n+1-\alpha n)} - x^*_i \| \right),
\end{align*}
\]
for \( \delta \geq 0 \), as required.

Hence the proof is complete.

As the number of voters increases, the probability of being part of an expropriated minority decreases. The benefit gained from avoiding blocking, however, is unchanged and the probability of this increases. The risk-averse agents therefore require less insurance and hence the optimal supermajority rule decreases.

**Theorem 2.** Assume A1-A3. Then the optimal supermajority rule is (weakly) increasing in the coefficient of importance/risk-aversion, \( \beta \equiv -u''(\cdot) / u'(\cdot) \).

**Proof.** By a similar argument to the proof of the above theorem we require \( E[\mathcal{V}] \) to have increasing differences in \( (\beta, \alpha) \). This requires that for all \( \beta' \geq \beta \), \( E[\mathcal{V}(\beta', \alpha)] - E[\mathcal{V}(\beta, \alpha)] \) is nondecreasing in \( \alpha \). That is

\[
\int \cdots \int \sum_{i=1}^{n} \left( \begin{array}{c}
\sum_{j=1}^{(n+1)/2} \pi_j (n) u_{\beta} \left( \| x^*_{\psi(\alpha n)} - x^*_i \| \right) + \\
\sum_{j=(n+1)/2+1}^{n+1} \pi_j (n) u_{\beta} \left( \| x^*_{\psi(\alpha n) - x^*_i} \| \right)
\end{array} \right) n! \, da_1 \cdots da_n,
\]

is nondecreasing in \( \alpha \) for all \( \beta' > \beta \). The first line of the above is identical to the second except for the differences in the function \( u \). As in the proof of Theorem 1 an increasing in \( \alpha \) makes the term \( \sum_{j=1}^{(n+1)/2} \pi_j (n) u \left( \| x^*_{\psi(\alpha n)} - x^*_i \| \right) \) larger, since \( \| x^*_{\psi(\alpha n)} - x^*_i \| \) increases and the probabilities \( \pi_j = F^{j-1}(\hat{\theta}) \left[ 1 - F(\hat{\theta}) \right]^{(n-j+1)} \) are unchanged and similarly for the term

\[
\sum_{j=(n+1)/2+1}^{n+1} \pi_j (n) u \left( \| x^*_{\psi(\alpha n) - x^*_i} \| \right).
\]
The magnitude of this change is larger for \( u'_{\beta} \) than \( u_\beta \) by Jensen’s inequality, and thus the result follows. \( \square \)

As the coefficient of importance/risk-aversion increases voters are progressively more concerned with being expropriated. They essentially purchase insurance against this by requiring that the size of the majority required to expropriate them be large, thereby reducing the probability of that event occurring. In fact, when the coefficient of importance is sufficiently high a unanimity requirement is always optimal. If there is the prospect of a sufficiently bad payoff then voters require a veto in order to insure themselves against this outcome.\(^8\)

Before stating the next result, the following definition is necessary.

**Definition 4.** A distribution \( \hat{F}(\cdot) \) is Rothschild-Stiglitz Riskier than another distribution \( F(\cdot) \) if either (i) \( F(\cdot) \) Second Order Stochastically Dominates \( \hat{F}(\cdot) \), (ii) \( \hat{F}(\cdot) \) is a Mean Preserving Spread of \( F(\cdot) \), or (iii) \( \hat{F}(\cdot) \) is an Elementary Increase in Risk from \( F(\cdot) \).

As is well known, Rothschild and Stiglitz (1970) showed that these three statements are equivalent.

**Theorem 3.** Assume A1-A3. Then the optimal supermajority rule is (weakly) larger for a distribution of voter types, \( \hat{F}(x) \) than for the distribution \( F(x) \) if \( \hat{F}(x) \) is Rothschild-Stiglitz Riskier than \( F(x) \).

**Proof.** Trivial, since a Rothschild-Stiglitz increase in risk has the same effect as an increase in the coefficient of importance. \( \square \)

This result obtains for reasons closely related to those of the two previous theorems. As the spread of voter types increases more insurance is desired, which is effected by requiring the supermajority rule to be higher. This is, however, only the case if the voters’ utility is more than proportionally decreasing as the social decision moves away from their ideal point (i.e. \( \beta > 0 \)).

It is also desirable to know something about the limiting properties of supermajority rules as the number of voters becomes large.

\(^8\)And as \( \beta \to \infty \) expected utility \( \to -\infty \).
Theorem 4. As $n \to \infty$ the optimal supermajority rule converges to

$$\bar{\alpha} = \max \{F(E[F(x)]), 1 - F(E[F(x)])\}.$$ 

Proof. Recall from (1) that the expected payoff under supermajority rule $\alpha$ is

$$E[V] = -\int \ldots \int \sum_{i=1}^{n} \left( \frac{\sum_{j=1}^{(n+1)/2} \pi_j u \left( x_{\psi(\alpha n)}^* - x_i^* \right)}{\sum_{j=(n+1)/2+1}^{n+1} \pi_j u \left( x_{n+1-\psi(\alpha n)}^* - x_i^* \right)} \right) n! \, da_1 \ldots da_n. \quad (8)$$

Observe that as $n \to \infty$ the ex post distribution of voter preferences converges almost surely to $F(x)$. Now, let us normalize the measure of events to 1 and write the limit of (8) for $\bar{\alpha}$ as

$$\lim_{n \to \infty} (E[V]|\alpha = \bar{\alpha}) = -\int \ldots \int \sum_{i=1}^{n} \left( \sum_{j=1}^{(n+1)/2} \pi_j u \left( |x_{\bar{\alpha n}}^* - x_i^*| \right) \right) n! \, da_1 \ldots da_n$$

$$+ \int \ldots \int \sum_{i=1}^{n} \left( \sum_{j=(n+1)/2+1}^{n+1} \pi_j u \left( |x_{1-\bar{\alpha}}^* - x_i^*| \right) \right) n! \, da_1 \ldots da_n,$n

where we keep the sum over events finite for the moment.

By an appropriate law of large numbers, a sufficient condition for the theorem to hold is

$$\int_0^x u \left( |x_{\bar{\alpha}}^* - x| \right) f(x) \, dx + \int_x^1 u \left( |x_{1-\bar{\alpha}}^* - x| \right) f(x) \, dx = \int_0^1 u \left( |E[F(x)] - x| \right) f(x) \, dx,$$n

which holds by the definition of $\bar{\alpha}$, and hence $\alpha = \bar{\alpha}$ maximizes expected utility.

Corollary 1. If $F(x)$ is symmetric then as $n \to \infty$ the optimal supermajority rule $\alpha^*$ converges to 1/2.

Remark 1. For $n$ very large, one does not want the social decision altered ex post since (by an appropriate law of large numbers) the ex post distribution of voter preferences converges to $F(x)$. This seems to imply that a unanimity rule would (also) be optimal for $n$ very large. Yet, for any finite $n$ the convergence of the ex post distribution to $F(x)$ will not be complete. It is clear from (8) that there exist events where voters are made very badly off by the change being blocked. Thus strict unanimity cannot be optimal for $n$ finite.
We now show that if voters are sufficiently risk-averse, then, for a finite number of
voters, unanimity is the optimal supermajority rule. Intuitively, with a finite number of
voters there is the possibility that the ex post distribution of voter preferences will differ
from the ex ante distribution. Moreover, the ex post distribution can be such that, for any
rule short of unanimity, a coalition of voters may want to change the social choice to
move it further away from the bliss point of certain other voters. When those voters are
sufficiently risk averse the loss from this is large, and the veto power that stems from
unaninimity rule is preferable to the possibility of adaptation. This is the same intuition
present in example 1, above. The following theorem shows that the intuition generalizes
beyond the specific functional forms of example 1.

**Theorem 5.** For any finite $n$ and $F(x)$ there exists a $\beta$ such that for all $\beta > \beta$ the optimal
supermajority rule is unanimity.

**Proof.** Since $n$ is finite, the ex post distribution of voter preference does not converge
almost surely to $F(x)$. Thus, there exist events $j$ under which the social decision is
changed from $E[F(x)]$, to (say) $\theta'$ after preferences are realized. Notice from (8) that for
$\beta$ sufficiently large, there exist $x_i$'s such that $E[V]$ can be made arbitrarily negative. But
under $\theta = E[F(x)]$, $E[V]$ is bounded. The fact that in such cases unanimity leads to
$\theta = E[F(x)]$, together with the monotonicity of $u(\cdot)$ in $\beta$, establishes the result.

3 Discussion and Conclusion

The veil of ignorance setup is a useful device, but there is a real question as to its ac-
curacy in describing decision making processes. In reality, decision makers have some
understanding of their preferences, even if they are imperfect, when at the ex ante stage.
In this sense they are not ex ante identical. One could capture this by assuming that
each voter receives an imperfect signal of their subsequent draw from the distribution.
This would be a non-trivial change to the mode of analysis employed in this paper, since
once one departs from voters being ex ante identical one can no longer make use of the
notion of a representative individual.

Enriching the contracting space to consider state-contingent supermajority rules would
be another direction for future research. Although such rules are rarely observed in their
pure form, many democratic institutions are a proxy for such rules. A non-trivial portion of democratic decision making is based on notions of interpersonal utility comparisons and aggregate welfare gain, despite the difficulties inherent in such comparisons.

In this paper we have abstracted from the possibility of monetary transfers between voters in order to affect the ex post social decision. There are circumstances where this implies significant inefficiencies in the sense that aggregate welfare could be improved by a different ex post decision. While a simple compensation principle can lead to obvious Pareto improvements in certain circumstances, this relies strongly on the verifiability of voters preferences. If each voter’s draw from the distribution remains private information, which seems realistic, then compensation and the associated Pareto improvements are not so straightforward. An incentive compatible mechanism would be required so that voters reveal their type in equilibrium. Understanding the existence and properties of such a mechanism may prove useful in understanding the practicalities of designing supermajority rules in the presence of private information. We feel that the issues of monetary transfers and private information regarding voter’s preferences are inextricably linked if one seeks a deeper understanding of actual voting systems.

Finally, we have imposed an exogenous allocation of bargaining power between parties in coalitions. An interesting question is how the analysis might change if the formation of coalitions and the distribution of the surplus within them was determined simultaneously. Ray and Vohra (1999) analyze this problem generally and consider stationary perfect equilibria of games where the negotiation process is conducted via alternating offers with costly delay, à la Rubinstein. A number of issues complicate the application of their approach to this setting, however. Firstly, they require binding agreements within coalitions, though there is non-cooperative play between coalitions. Though complete contracts between coalitions after the realization of the state of nature is not inconsistent with an ex ante incomplete social contract, the familiar questions regarding the completeness of the within coalition contract arise. More problematic, however, is that their approach requires transferable utility - that is, a linear frontier of coalition payoffs. This is not consistent with our specification. Finally, whilst with symmetric players Ray and Vohra derive an algorithm which generates a unique coalition structure, with asymmetric players this uniqueness necessarily breaks down. Each of these difficulties would need to be overcome in order to consider endogenously gener-
4 Appendix

4.1 Example 1 calculations

4.1.1 Unanimity

The density\(^9\) of \(x_4^*\) is \(f(a_4) = 4(a_4)^3\). We now need the density of \(x_4^*\) on \([0, \frac{1}{2}]\), which is found by applying the Change of Variables Theorem, yielding \(g(a_4) = 2 \times 4(2a_4)^3 = 64(a_4)^3\). Therefore:

\[
E[u_i^U|B] = \int_{\frac{1}{2}}^{1} - \exp \left\{ \beta \left( t - \frac{1}{2} \right) \right\} dt
+ \int_{0}^{\frac{1}{2}} \left( \int_{0}^{a_4} - \exp \left\{ \beta (a_4 - t) \right\} dt - \int_{a_4}^{\frac{1}{2}} 1dt \right) 64 (a_4)^3 da_4
= \frac{2}{\beta} \left( 1 - e^{\beta/2} \right) + \frac{1}{\beta} - \frac{8(48 + e^{\beta/2}(\beta^3 - 6\beta^2 + 24\beta - 48))}{\beta^5} - \frac{1}{10},
\]

where \( \int_{\frac{1}{2}}^{1} - \exp \left\{ \beta \left( t - \frac{1}{2} \right) \right\} dt \) is the term associated with \( x_i \geq \frac{1}{2} \) and the term associated with \( x_i \leq \frac{1}{2} \) is \( \int_{0}^{\frac{1}{2}} \left( \int_{0}^{a_4} - \exp \left\{ \beta (a_4 - t) \right\} dt - \int_{a_4}^{\frac{1}{2}} 1 dt \right) 64 (a_4)^3 da_4 \).

Under event \( B' \) the expected utility is given by:

\[
E[u_i^U|B'] = \int_{0}^{\frac{1}{2}} - \exp \left\{ \beta \left( \frac{1}{2} - t \right) \right\} dt
+ \int_{\frac{1}{2}}^{1} \left( \int_{a_1}^{1} - \exp \left\{ \beta (t - a_1) \right\} dt - \int_{\frac{1}{2}}^{a_1} 1 dt \right) 64 (1 - a_1)^3 da_1
= \frac{2}{\beta} \left( 1 - e^{\beta/2} \right) + \frac{1}{\beta} - \frac{8(48 + e^{\beta/2}(\beta^3 - 6\beta^2 + 24\beta - 48))}{\beta^5} - \frac{1}{10}.
\]

\(^9\)For the uniform distribution the density of the \( i \)th order statistic is:

\[
f_i(u) = \frac{n!}{(i-1)! (n-i)!} u^{i-1} (1-u)^{n-i}
\]
Therefore the total expected utility under unanimity is:

\[
E \left[ u_i^U \right] = \frac{7}{8} E \left[ u_i^U | A \right] + \frac{1}{16} E \left[ u_i^U | B \right] + \frac{1}{16} E \left[ u_i^U | B' \right] \\
= -3840 + \beta^4(\beta - 160) + 10e^{\beta/2}(-384 + \beta(192 + \beta(15\beta + 8)48)) \frac{1}{80\beta^5}
\]

### 4.1.2 80% supermajority rule

As before, if event \( A \) occurs then there is no change to the \textit{ex ante} social decision and hence the expected utility of voter \( i \) is:

\[
E \left[ u_i^S | A \right] = 2 \int_0^{\frac{1}{2}} - \exp \left\{ \beta \left( \frac{1}{2} - t \right) \right\} dt \\
= \frac{2(1 - e^{\beta/2})}{\beta}.
\]

Note\(^{10}\) that the density of \( x_2^* \) conditional on event \( C \) is simply the density of the first order-statistic of three on \([\frac{1}{2}, 1]\). In fact, order statistics from a continuous parent form a Markov Chain. It follows that the density of the first-order statistic of three on \( U[0, 1] \) is \( 3(1 - a_2)^2 \). By a change of variables, the density on \([\frac{1}{2}, 1]\) is therefore \( 24(1 - a_2)^2 \). Hence the expected utility conditional on event \( C \) is:

\[
E \left[ u_i^S | C \right] = \int_\frac{1}{2}^1 \left( -1 - \int_{\frac{1}{2}}^{a_2} dt + \int_0^{\frac{1}{2}} - \exp \left\{ \beta \left( \frac{1}{2} - t \right) \right\} dt \right) 24(1 - a_2)^2 da_2 \\
+ \int_0^{\frac{1}{2}} - \exp \left\{ \beta \left( \frac{1}{2} - t \right) \right\} dt \\
= -\frac{1}{8} + \frac{1}{\beta} + \frac{1 - e^{\beta/2}}{\beta} - \frac{6(-8 + e^{\beta/2}(8 - 4 + \beta^2))}{\beta^4}.
\]

The density of \( x_3^* \) conditional on event \( C' \) is the third of three uniformly distributed

\(^{10}\)This fact is quite general. The conditional pdf of an order-statistic is given by:

\[
f_{X_r | X_{r+1}}(x) = \frac{(s-1)!}{(r-1)!(s-r-1)!} f(x) F(x)^{r-1}(F(v) - F(x))^{s-r-1} \frac{F(v)^{s-1}}{F(v)^{s-1}}
\]
order-statistics on \([0, \frac{1}{2}]\), which is \(g(a_2|C') = \frac{1}{24} (a_3)^2\). Hence the expected utility conditional on event \(C'\) is:

\[
E \left[ u^S_i | C' \right] = \int_0^{\frac{1}{2}} \left( -\int_{a_3}^{x_3} dt + \int_0^{a_3} -\exp \{\beta (a_3 - t)\} dt \right) \frac{1}{24} (a_3)^2 da_3 \\
+ \int_{1/2}^{1} -\exp \left\{ \beta \left( t - \frac{1}{2} \right) \right\} dt \\
= -\frac{1}{8} + \frac{1}{\beta} + \frac{1 - e^{\beta/2}}{\beta} - \frac{6(-8 + e^{\beta/2}(8 - 4 + \beta^2))}{\beta^4}.
\]

Now note that the joint density of \((x_3^*, x_4^*)\) on \([0, 1]\) is \(f(x_3, x_4) = 12 (a_3)^2\) and so on \([0, \frac{1}{2}]\) it is \(192 (a_3)^2\). The expected utility conditional on event \(B\) is therefore:

\[
E \left[ u^S_i | B \right] = \int_0^{\frac{1}{2}} \int_0^{a_3} \left( -\int_{a_3}^{x_3} dt + \int_0^{a_3} -\exp \{\beta (a_3 - t)\} dt \right) \frac{1}{192} (a_3)^2 da_3 da_4 \\
= -3840 e^{\beta/2} - \beta^4 (\beta - 20) + 1920 (\beta + 6) + 80 e^{\beta/2} (\beta^4 + 72\beta - 96) \\
10^5.
\]

The joint density of \((x_1^*, x_2^*)\) on \([0, 1]\) is \(f(x_1, x_2) = 12 (1 - a_2)^2\) and so on \([\frac{1}{2}, 1]\) it is \(192 (1 - a_2)^2\). The expected utility conditional on event \(B'\) is:

\[
E \left[ u^S_i | B' \right] = \int_0^{1} \int_0^{a_1} \left( -\int_{a_1}^{x_1} dt + \int_0^{a_1} -\exp \{\beta (a_1 - t)\} dt \right) (-192 (1 - a_2)^2) da_2 da_1 \\
= -3840 e^{\beta/2} - \beta^4 (\beta - 20) + 1920 (\beta + 6) + 80 e^{\beta/2} (\beta^4 + 72\beta - 96) \\
10^5.
\]

Therefore the total expected utility under a supermajority of four voters (ie. 80% supermajority) is:

22
\[
E \left[ u_i^S \right] = \frac{3}{8} E \left[ u_i^S | A \right] + \frac{1}{16} E \left[ u_i^S | B \right] + \frac{1}{16} E \left[ u_i^S | B' \right] \\
+ \frac{1}{4} E \left[ u_i^S | C \right] + \frac{1}{4} E \left[ u_i^S | C' \right].
\]

Which, upon simplification, is:

\[
E \left[ u_i^S \right] = \frac{\left( -1920e^\beta + \beta^4(80 - 3\beta) + 1920(\beta + 3) \\
-10\beta^2(284 + \beta(-192 + \beta(\beta + 4)(5\beta - 12))) \right)}{40\beta^5}
\]
References


_, *Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility* Essays on Ethics, Social Behavior, and Scientific Explanation, Springer,


