

Online Appendix for Designing Checks and Balances

Tiberiu Dragu^{*}, Xiaochen Fan,^{*} and James Kuklinski^{**}

^{*}New York University

^{**}University of Illinois at Urbana-Champaign

The Appendix contains two sections. In Appendix A, we present the proofs of our main results. In Appendix B, we extend our framework to a situation in which there are $N > 2$ decision-makers who must agree on changing an existing policy.

Appendix A

Proof of Lemma 1. First, we show that a strategy-proof mechanism $x(c, g)$ is weakly increasing in c for all g . Let us assume that for some g , $x(c, g)$ is not weakly increasing; that is, without loss of generality, let $c_2 > c_1$ and $x(c_1, g) > x(c_2, g)$. Given this assumption, we consider the following four cases, which exhaust all possibilities:

i) $c_1 \geq x(c_1, g) > x(c_2, g)$. Then $c_2 > x(c_1, g) > x(c_2, g)$; the player C with type c_2 prefers $x(c_1, g)$ to $x(c_2, g)$ and will have incentives to misrepresent its type. Thus $x(c, g)$ is not strategy-proof.

ii) $x(c_1, g) > x(c_2, g) \geq c_1$. Then by the single-peakedness of $U_C(\cdot)$, the player C with type c_1 prefers $x(c_2, g)$ to $x(c_1, g)$ and will have incentives to misrepresent its type. Thus $x(c, g)$ is not strategy-proof.

iii) $c_2 \geq x(c_1, g) > c_1 > x(c_2, g)$. Then by the single-peakedness of $U_C(\cdot)$, the player C with type c_2 prefers $x(c_1, g)$ to $x(c_2, g)$ and will have incentives to misrepresent its type. Thus $x(c, g)$ is not strategy-proof.

iv) $x(c_1, g) > c_2 > c_1 > x(c_2, g)$. Incentive compatibility at c_1 implies that $U_C(x(c_1, g), c_1) \geq U_C(x(c_2, g), c_1)$. Single-crossing property further implies that $U_C(x(c_1, g), c_2) > U_C(x(c_2, g), c_2)$. As a result, the player C with type c_2 will prefer $x(c_1, g)$ to $x(c_2, g)$; thus $x(c, g)$ is not strategy-proof.

Second, we show that for arbitrary g , if $x(c, g)$ strictly increasing in c on an open interval (c_1, c_2) , then $x(c, g) = c$ on (c_1, c_2) . Suppose not. Without loss of generality consider the case in which $x(c^*, g) < c^*$ for some $c^* \in (c_1, c_2)$. If $x(c, g)$ is continuous at c^* , there exists

an $\epsilon > 0$ such that $x(c^*, g) < x(c^* + \epsilon, g) < c^*$ and thus $U_C(x(c^* + \epsilon, g), c^*) > U_C(x(c^*, g), c^*)$ so $x(\cdot, g)$ is not strategy-proof.

If $x(c, g)$ is not continuous at c^* , since $x(c, g)$ is strictly increasing in c on (c_1, c_2) , then $x(c, g)$ can only have jump discontinuities and there are at most countable discontinuity points. If $\lim_{c \rightarrow c^*+} x(c, g) < c^*$, we can use the preceding argument to show a contradiction (i.e, $x(\cdot, g)$ is not strategy-proof).

If $\lim_{c \rightarrow c^*+} x(c, g) = c^*$, since U_C is continuous there exists $\epsilon > 0$ such that $U_C(x(c^* + \epsilon, g), c^*) > U_C(x(c^*, g), c^*)$ and thus $x(\cdot, g)$ is not strategy-proof.

If $\lim_{c \rightarrow c^*+} x(c, g) > c^*$, let us denote $\lim_{c \rightarrow c^*+} x(c, g) \equiv c^* + \Delta$ where $\Delta > 0$. Then for any $c \in (c^*, \min\{c^* + \Delta, c_2\})$, we have $x(c, g) > \lim_{c \rightarrow c^*+} x(c, g) = c^* + \Delta > c$. Since $x(c, g)$ is strictly increasing on $(c^*, \min\{c^* + \Delta, c_2\})$, there exists $c^{**} \in (c^*, \min\{c^* + \Delta, c_2\})$ such that $x(c, g)$ is continuous at c^{**} , and from the previous line we also have $x(c^{**}, g) > c^{**}$. Then there exists $\epsilon > 0$ such that $x(c^{**}, g) > x(c^{**} - \epsilon, g) > c^{**}$, and therefore $U_C(x(c^{**} - \epsilon, g), c^{**}) > U_C(x(c^{**}, g), c^{**})$. As a result, $x(\cdot, g)$ is not strategy-proof.

A similar argument holds if $x(c^*, g) > c^*$ for some $c^* \in (c_1, c_2)$.

As a result, for any g , any mechanism $x(c, g)$ that is strategy-proof for player C is weakly increasing in c , and if $x(c, g)$ is strictly increasing in c on an open interval (c_1, c_2) , $x(c, g) = c$ on (c_1, c_2) .

A similar argument can be used to characterize mechanisms that are strategy-proof for G , given an arbitrary type of C . □

Before proceeding to prove our main result, we first prove the following lemmas.

Lemma 2. *Let $x(c, g)$ be a strategy-proof mechanism. Then for any g , if $x(c, g) = c$ on (c_1, c_2) , then $x(c, g)$ is continuous at both c_1 and c_2 . Similarly, for any c , if $x(c, g) = g$ on (g_1, g_2) , then $x(c, g)$ is continuous at both g_1 and g_2 .*

Proof. We prove the first half of the lemma, and the proof for the second half is completely analogous. Without loss of generality we prove that $x(c, g)$ is continuous at c_1 . The case

of c_2 is analogous. Suppose to the contrary that $x(c, g)$ is discontinuous at c_1 . Since $x(c, g)$ is (weakly) increasing in c , $\lim_{c \rightarrow c_1^-} x(c, g) < \lim_{c \rightarrow c_1^+} x(c, g) = c_1$. Since U_C is continuous there exists small enough $\epsilon > 0$ such that $x(c_1 - \epsilon, g) \leq \lim_{c \rightarrow c_1^-} x(c, g) < c_1 - \epsilon$ and $U_C(x(c_1 + \epsilon, g), c_1 - \epsilon) > U_C(x(c_1 - \epsilon, g), c_1 - \epsilon)$. That is, the player C with type $c_1 - \epsilon$ can do better by reporting $c_1 + \epsilon$ and thus a contradiction to $x(c, g)$ being strategy-proof. \square

In other words, the only discontinuity points of a strategy-proof mechanism $x(c, g)$ are the points that connect two flat segments. The next lemma further characterizes the discontinuity points of $x(c, g)$:

Lemma 3. *Let $x(c, g)$ be a strategy-proof mechanism. Then for any g , if \hat{c} is a discontinuity point of $x(c, g)$, then $\hat{c} - \lim_{c \rightarrow \hat{c}^-} x(c, g) > 0$ and $\lim_{c \rightarrow \hat{c}^+} x(c, g) - \hat{c} > 0$. Similarly, for any c , if \hat{g} is a discontinuity point of $x(c, g)$, then $\hat{g} - \lim_{g \rightarrow \hat{g}^-} x(c, g) > 0$ and $\lim_{g \rightarrow \hat{g}^+} x(c, g) - \hat{g} > 0$.*

Proof. Again we prove the first half of the lemma. Let us show that both terms are positive. There are only three other possibilities: (i) $\hat{c} - \lim_{c \rightarrow \hat{c}^-} x(c, g) \leq 0$ and $\lim_{c \rightarrow \hat{c}^+} x(c, g) - \hat{c} \leq 0$, but this is impossible because $x(c, g)$ is increasing and discontinuous at \hat{c} ; (ii) $\hat{c} - \lim_{c \rightarrow \hat{c}^-} x(c, g) \leq 0$ and $\lim_{c \rightarrow \hat{c}^+} x(c, g) - \hat{c} > 0$; and (iii) $\hat{c} - \lim_{c \rightarrow \hat{c}^-} x(c, g) > 0$ and $\lim_{c \rightarrow \hat{c}^+} x(c, g) - \hat{c} \leq 0$. We show a contradiction for case (ii), and the argument for case (iii) is analogous.

Suppose $\hat{c} - \lim_{c \rightarrow \hat{c}^-} x(c, g) \leq 0$ and $\lim_{c \rightarrow \hat{c}^+} x(c, g) - \hat{c} > 0$. Since $x(c, g)$ is increasing and is discontinuous at \hat{c} , we have $\lim_{c \rightarrow \hat{c}^+} x(c, g) > \lim_{c \rightarrow \hat{c}^-} x(c, g) \geq \hat{c}$. Since U_C is continuous, there exists $\epsilon > 0$ such that $U_C(x(\hat{c} - \epsilon, g), \hat{c} + \epsilon) > U_C(x(\hat{c} + \epsilon, g), \hat{c} + \epsilon)$, so the player C with type $\hat{c} + \epsilon$ will have an incentive to misreport his type as $\hat{c} - \epsilon$ and thus $x(c, g)$ is not strategy-proof. \square

Lemma 4. *Let $x(c, g)$ be a strategy-proof mechanism. Then for any g , if for some \hat{c} , $x(\hat{c}, g) = a \neq \hat{c}$, then $x(c, g) = a$ for all $c \in (\min\{a, \hat{c}\}, \max\{a, \hat{c}\})$. Similarly, for any c , if for some \hat{g} , $x(c, \hat{g}) = a \neq \hat{g}$, then $x(c, g) = a$ for all $g \in (\min\{a, \hat{g}\}, \max\{a, \hat{g}\})$.*

Proof. We prove the first half of the lemma for the case $a < \hat{c}$. The proof for $a > \hat{c}$ and the proof for the second half are analogous. Suppose there exists $\tilde{c} \in (a, \hat{c})$ such that $x(\tilde{c}, g) \neq a$, then since $x(c, g)$ is (weakly) increasing, $x(\tilde{c}, g) < a$. Thus $x(\tilde{c}, g) < x(\hat{c}, g) = a < \tilde{c}$, and since $U_C(\cdot, \tilde{c})$ is single-peaked with peak \tilde{c} , we have $U_C(x(\hat{c}, g), \tilde{c}) > U_C(x(\tilde{c}, g), \tilde{c})$, a contradiction to $x(c, g)$ being strategy-proof. \square

Proof of Proposition 1. By the definition of $m(c, g)$, if $c, g \geq 0$, then $m(c, g) = \mathbf{min}\{c, g\}$.

First, it is easy to check that the mechanism $x(c, g) = \mathbf{min}\{c, g\}$ satisfies the checks condition and strategy-proofness condition.

Next note that ex-post efficiency is equivalent to stating that for any (c, g) , $x(c, g) \in [\mathbf{min}\{c, g\}, \mathbf{max}\{c, g\}]$. Therefore the mechanism $x(c, g) = \mathbf{min}\{c, g\}$ is ex-post efficient and if $c = g$, the unique ex-post efficient outcome is $x(c, g) = c = g = \mathbf{min}\{c, g\}$.

Now suppose that for $c, g \geq 0$, there exists another mechanism $\tilde{x}(c, g)$ that satisfies all three conditions, such that for some $\tilde{c} \neq \tilde{g}$, $\tilde{c}, \tilde{g} \geq 0$, $\tilde{x}(\tilde{c}, \tilde{g}) \neq \mathbf{min}\{\tilde{c}, \tilde{g}\}$, and without loss of generality let $\tilde{c} < \tilde{g}$. Ex-post efficiency implies that $\tilde{x}(\tilde{c}, \tilde{g}) > \tilde{c}$.

Checks condition for player C with ideal point 0 implies that $\tilde{x}(0, \tilde{g}) = 0$. Lemma 1 then implies that $\tilde{x}(c, \tilde{g})$ has at least one discontinuous point on $[0, \tilde{c}]$. Now if $\tilde{x}(c, \tilde{g})$ is discontinuous at $c = 0$, i.e. if $\mathbf{lim}_{c \rightarrow 0^+} \tilde{x}(c, \tilde{g}) > 0$, then there exists $\epsilon > 0$ such that $U_C(0, \epsilon) > U_C(\tilde{x}(\epsilon, \tilde{g}), \epsilon)$. But then player C with ideal point ϵ has an incentive to deviate and report type 0, a contradiction to $\tilde{x}(c, g)$ being strategy-proof. Therefore, $\tilde{x}(c, \tilde{g})$ is continuous at $c = 0$ and therefore $\tilde{x}(c, \tilde{g})$ has at least one discontinuous point on $(0, \tilde{c}]$.

Let $\hat{c} \in (0, \tilde{c}]$ be a discontinuous point of $\tilde{x}(c, \tilde{g})$. From Lemma 3, we have $\hat{c} - \mathbf{lim}_{c \rightarrow \hat{c}^-} \tilde{x}(c, \tilde{g}) > 0$.

Therefore, there exists $c_1 \in (0, \hat{c}]$ such that $\tilde{x}(c_1, \tilde{g}) < c_1$. But since $c_1 \leq \hat{c} \leq \tilde{c} < \tilde{g}$, $\tilde{x}(c_1, \tilde{g})$ is not ex-post efficient. We have a contradiction.

Therefore the unique mechanism that satisfies checks, strategy-proofness and efficiency is $x(c, g) = \mathbf{min}\{c, g\} = m(c, g)$ for $c, g \geq 0$. \square

Proof of Proposition 2. By the definition of $m(c, g)$, if $c, g \geq 0$, then $m(c, g) = \mathbf{min}\{c, g\}$.

First, it is easy to check that the mechanism $x(c, g) = \mathbf{min}\{c, g\}$ satisfies checks, strategy-proofness and responsiveness.

Next, we show that any mechanism that is strategy-proof and responsive must be that $x(c, g) \in \{c, g\}$ for all (c, g) . Suppose not, then there exist (\hat{c}, \hat{g}) and $a \notin \{\hat{c}, \hat{g}\}$ such that $x(\hat{c}, \hat{g}) = a$. Then $\mathbf{min}\{a, \hat{c}\} \neq \mathbf{max}\{a, \hat{c}\}$ and $\mathbf{min}\{a, \hat{g}\} \neq \mathbf{max}\{a, \hat{g}\}$. Lemma 4 implies that $x(c, \hat{g}) = a$ for all $c \in [\mathbf{min}\{a, \hat{c}\}, \mathbf{max}\{a, \hat{c}\}]$, and $x(\hat{c}, g) = a$ for all $g \in [\mathbf{min}\{a, \hat{g}\}, \mathbf{max}\{a, \hat{g}\}]$. This further implies that $x(c, g) = a$ for all $c \in [\mathbf{min}\{a, \hat{c}\}, \mathbf{max}\{a, \hat{c}\}]$ and $g \in [\mathbf{min}\{a, \hat{g}\}, \mathbf{max}\{a, \hat{g}\}]$. Therefore $x(\mathbf{min}\{a, \hat{c}\}, \mathbf{min}\{a, \hat{g}\}) = x(\mathbf{max}\{a, \hat{c}\}, \mathbf{max}\{a, \hat{g}\}) = a$ but $\mathbf{max}\{a, \hat{c}\} > \mathbf{min}\{a, \hat{c}\}$ and $\mathbf{max}\{a, \hat{g}\} > \mathbf{min}\{a, \hat{g}\}$, a contradiction to $x(c, g)$ being responsive.

Therefore, if $x(c, g)$ is strategy-proof and responsive, $x(c, g) \in \{c, g\}$ for all (c, g) . We next prove that for $c, g \geq 0$, the unique mechanism that satisfies checks, strategy-proofness and responsiveness is $x(c, g) = \mathbf{min}\{c, g\}$. Suppose there exists another strategy-proof and responsive mechanism $\tilde{x}(c, g)$. From what we show above, since $\tilde{x}(c, g)$ is strategy-proof and responsive, $\tilde{x}(c, g) \in \{c, g\}$. For any $c = g$, $\tilde{x}(c, g) \in \{c, g\}$ implies $\tilde{x}(c, g) = \mathbf{min}\{c, g\}$, so there exists $\tilde{c} \neq \tilde{g}$ such that $\tilde{x}(\tilde{c}, \tilde{g}) \neq \mathbf{min}\{\tilde{c}, \tilde{g}\}$. Without loss of generality let $\tilde{c} < \tilde{g}$, then $\tilde{x}(\tilde{c}, \tilde{g}) = \tilde{g}$. Since $\tilde{x}(c, g)$ satisfies the checks condition, $\tilde{x}(0, \tilde{g}) = 0$. By Lemma 1 and 3, $\tilde{x}(0, \tilde{g}) = 0$ and $\tilde{x}(\tilde{c}, \tilde{g}) = \tilde{g} > \tilde{c}$ imply that there exists $c_1 \in (0, \tilde{c}]$ such that $\tilde{x}(c, \tilde{g})$ is discontinuous at c_1 , and there exists $c_2 < c_1$ such that $\tilde{x}(c_2, \tilde{g}) < c_2$. Since $c_2 < c_1 \leq \tilde{c} < \tilde{g}$, $\tilde{x}(c_2, \tilde{g}) \notin \{c_2, \tilde{g}\}$, a contradiction.

Therefore the unique mechanism that satisfies checks, strategy-proofness and responsiveness is $x(c, g) = \mathbf{min}\{c, g\} = m(c, g)$ for $c, g \geq 0$.

□

Proof of Proposition 1'. First, it's easy to check that the mechanism $x(c, g) = m(c, g)$ if $c, g \geq 0$ or $c, g \leq 0$, and $x(c, g) = 0$ if $c < 0 < g$ or $g < 0 < c$, satisfies both the checks and (ex-post) Pareto efficiency conditions. To show that it is also strategy-proof on the general domain $\mathbf{R} \times \mathbf{R}$, it suffices to show the following three types of inequalities: (1) for $c, g \geq 0$,

$U_C(x(c, g), c) \geq U_C(x(\tilde{c}, g), c)$ for any $\tilde{c} < 0$, and $U_G(x(c, g), g) \geq U_G(x(c, \tilde{g}), g)$ for any $\tilde{g} < 0$. The former is true because U_C is single-peaked and $x(c, g) = \min\{c, g\} \in [0, c]$ since $c, g \geq 0$ while $x(\tilde{c}, g) = 0$ since $\tilde{c} < 0 \leq g$. The latter inequality can be shown analogously; (2) for $c, g \leq 0$, $U_C(x(c, g), c) \geq U_C(x(\tilde{c}, g), c)$ for any $\tilde{c} > 0$, and $U_G(x(c, g), g) \geq U_G(x(c, \tilde{g}), g)$ for any $\tilde{g} > 0$. The proof is the same as in (1) above; and (3) for $c < 0 < g$ or $g < 0 < c$, $U_C(x(c, g), c) \geq U_C(x(\tilde{c}, g), c)$ for any \tilde{c} , and $U_G(x(c, g), g) \geq U_G(x(c, \tilde{g}), g)$ for any \tilde{g} . To show this, we show that when $c < 0 < g$, $U_C(x(c, g), c) \geq U_C(x(\tilde{c}, g), c)$ for any \tilde{c} , since all other cases are analogous. This is true because for any $\tilde{c} \leq 0$, $x(\tilde{c}, g) = x(c, g) = 0$, so $U_C(x(c, g), c) \geq U_C(x(\tilde{c}, g), c)$ trivially. Also, for any $\tilde{c} > 0$, $x(\tilde{c}, g) = \min\{\tilde{c}, g\} > 0$, so $U_C(x(c, g), c) \geq U_C(x(\tilde{c}, g), c)$ since U_C is single-peaked, $c < 0$, and $x(c, g) = 0$.

We next show that the mechanism $x(c, g) = m(c, g)$ if $c, g \geq 0$ or $c, g \leq 0$, and $x(c, g) = 0$ if $c < 0 < g$ or $g < 0 < c$ is the unique mechanism that satisfies checks, strategy-proofness and efficiency on $\mathbf{R} \times \mathbf{R}$. Since $x(c, g) = m(c, g)$ is the unique mechanism that satisfies the three conditions on $\mathbf{R}_+ \times \mathbf{R}_+$, any mechanism that satisfies the three conditions on $\mathbf{R} \times \mathbf{R}$ must equal to $m(c, g)$ on $\mathbf{R}_+ \times \mathbf{R}_+$. Similarly, any mechanism that satisfies the three conditions on $\mathbf{R} \times \mathbf{R}$ must equal to $m(c, g)$ on $\mathbf{R}_- \times \mathbf{R}_-$, and equal to 0 on $(\mathbf{R}_- \times \mathbf{R}_+) \cup (\mathbf{R}_+ \times \mathbf{R}_-)$. \square

Proof of Proposition 5. First, it is easy to check that the above issue-by-issue moderate rule is indeed strategy-proof, minimally efficient and satisfies checks condition.

For uniqueness, first note that according to Border and Jordan (1983), a mechanism $\mathbf{x}(p^C, p^G) : \mathbf{R}^{2n} \rightarrow \mathbf{R}^n$ is strategy-proof and minimally efficient¹ if and only if there are mechanisms $x_1, x_2, \dots, x_n : \mathbf{R}^2 \rightarrow \mathbf{R}$ which are strategy-proof and minimally efficient such that

$$\mathbf{x}(p^C, p^G) = (x_1(p_1^C, p_1^G), x_2(p_2^C, p_2^G), \dots, x_n(p_n^C, p_n^G)),$$

where p_j^i is player i 's ideal point on issue dimension j for $i \in \{C, G\}$ and $j = 1, 2, \dots, n$.

For any j , $x_j(p_j^C, p_j^G)$ is strategy-proof and minimally efficient, these imply that $x_j(p_j^C, p_j^G)$

¹In Border and Jordan (1983), minimal efficiency is labeled as unanimity.

is Pareto efficient, that is, $\min\{p_j^C, p_j^G\} \leq x_j(p_j^C, p_j^G) \leq \max\{p_j^C, p_j^G\}$.² To show this, suppose first to the contrary that there exists (p_j^C, p_j^G) such that $x_j(p_j^C, p_j^G) > \max\{p_j^C, p_j^G\}$, without loss of generality suppose $p_j^C \leq p_j^G$ so that $\max\{p_j^C, p_j^G\} = p_j^G$. But then $p_j^C \leq x_j(p_j^C, p_j^G) = p_j^G < x_j(p_j^C, p_j^G)$, where the equality is due to minimal efficiency. This is a violation of strategy-proofness. Therefore $x_j(p_j^C, p_j^G) \leq \max\{p_j^C, p_j^G\}$. Similarly we can show that $x_j(p_j^C, p_j^G) \geq \min\{p_j^C, p_j^G\}$, therefore $x_j(p_j^C, p_j^G)$ is Pareto efficient.

Therefore, $\mathbf{x}(p^C, p^G)$ is strategy-proof and minimally efficient if and only if there exist $x_1, x_2, \dots, x_n : \mathbf{R}^2 \rightarrow \mathbf{R}$ which are strategy-proof and Pareto efficient such that $\mathbf{x}(p^C, p^G) = (x_1(p_1^C, p_1^G), x_2(p_2^C, p_2^G), \dots, x_n(p_n^C, p_n^G))$.

Next we show that if $\mathbf{x}(p^C, p^G)$ satisfies the checks condition in addition to strategy-proofness and minimal efficiency, then the above $x_j(p_j^C, p_j^G)$ must also satisfy checks at 0 for all j , that is, $x_j(p_j^C, p_j^G) = 0$ if $p_j^i = 0$ for some $i \in \{C, G\}$. Since $\mathbf{x}(p^C, p^G)$ satisfies the checks condition, $\mathbf{x}(p^C, p^G) = \mathbf{0}$ if $p^i = \mathbf{0}$ for some $i \in \{C, G\}$. Let $p^C = \mathbf{0}$, then for any p^G , $\mathbf{x}(\mathbf{0}, p^G) = (x_1(0, p_1^G), x_2(0, p_2^G), \dots, x_n(0, p_n^G)) = \mathbf{0}$, therefore $x_j(0, p_j^G) = 0$ for any j and any p_j^G . Similarly, $x_j(p_j^C, 0) = 0$ for any j and any p_j^C .

Therefore, $\mathbf{x}(p^C, p^G)$ is strategy-proof, minimally efficient and satisfies the checks condition if and only if there exist $x_1, x_2, \dots, x_n : \mathbf{R}^2 \rightarrow \mathbf{R}$ which are strategy-proof, Pareto efficient and satisfy checks at 0, such that $\mathbf{x}(p^C, p^G) = (x_1(p_1^C, p_1^G), x_2(p_2^C, p_2^G), \dots, x_n(p_n^C, p_n^G))$.

Finally, we show that if $\mathbf{x}(p^C, p^G)$ is strategy-proof, minimally efficient and satisfies the checks condition, $\mathbf{x}(p^C, p^G) = (\min\{p_1^C, p_1^G\}, \min\{p_2^C, p_2^G\}, \dots, \min\{p_n^C, p_n^G\})$. Since we have shown that $\mathbf{x}(p^C, p^G)$ is strategy-proof, minimally efficient and satisfies the checks condition if and only if there exist $x_1, x_2, \dots, x_n : \mathbf{R}^2 \rightarrow \mathbf{R}$ which are strategy-proof, Pareto efficient and satisfy checks at 0, such that $\mathbf{x}(p^C, p^G) = (x_1(p_1^C, p_1^G), x_2(p_2^C, p_2^G), \dots, x_n(p_n^C, p_n^G))$, it suffices to show that $x_j(p_j^C, p_j^G) = \min\{p_j^C, p_j^G\}$ for all $j = 1, \dots, n$.

Consider an arbitrary $j \in \{1, 2, \dots, n\}$. If $p_j^C = p_j^G$, minimal efficiency implies $x_j(p_j^C, p_j^G) = \min\{p_j^C, p_j^G\}$ trivially.

²Note that $x_j(p_j^C, p_j^G)$ being Pareto efficient for all j is a necessary but not sufficient condition for $\mathbf{x}(p^C, p^G)$ to be Pareto efficient.

Suppose that there exists a mechanism $\tilde{x}_j(p_j^C, p_j^G) \neq \mathbf{min}\{p_j^C, p_j^G\}$ that will make $\mathbf{x}(p^C, p^G)$ satisfy the checks, strategy-proofness and minimal efficiency conditions. That is, $\tilde{x}_j(p_j^C, p_j^G)$ is strategy-proof, Pareto efficient, and satisfies checks at 0, and for some $(\tilde{p}_j^C, \tilde{p}_j^G)$ such that $\tilde{p}_j^i \geq 0$ for $i \in \{C, G\}$, and $\tilde{p}_j^C \neq \tilde{p}_j^G$, $\tilde{x}_j(\tilde{p}_j^C, \tilde{p}_j^G) \neq \mathbf{min}\{\tilde{p}_j^C, \tilde{p}_j^G\}$. Pareto efficiency then implies that $\tilde{x}_j(\tilde{p}_j^C, \tilde{p}_j^G) > \mathbf{min}\{\tilde{p}_j^C, \tilde{p}_j^G\}$. Without loss of generality let $\tilde{p}_j^C \leq \tilde{p}_j^G$, so $\tilde{p}_j^C = \mathbf{min}\{\tilde{p}_j^C, \tilde{p}_j^G\}$.

Checks at 0 implies that $\tilde{x}_j(0, \tilde{p}_j^G) = 0$.

Since $\tilde{x}_j(p_j^C, p_j^G)$ is strategy-proof, Lemma 1 through Lemma 4 in our one-dimension analysis applies. Lemma 1 then implies that $\tilde{x}_j(p_j^C, \tilde{p}_j^G)$ has at least one discontinuous point on $p_j^C \in [0, \tilde{p}_j^C]$. Now if $\tilde{x}_j(p_j^C, \tilde{p}_j^G)$ is discontinuous at $p_j^C = 0$, i.e. if $\lim_{p_j^C \rightarrow 0^+} \tilde{x}_j(p_j^C, \tilde{p}_j^G) > 0$, then there exists $\epsilon > 0$ such that $(0 - \epsilon)^2 < (\tilde{x}_j(\epsilon, \tilde{p}_j^G) - \epsilon)^2$. This means the player C with ideal point ϵ has an incentive to misreport to type 0, a contradiction to $\tilde{x}_j(p_j^C, p_j^G)$ being strategy-proof. Therefore, $\tilde{x}_j(p_j^C, \tilde{p}_j^G)$ is continuous at $p_j^C = 0$ and therefore $\tilde{x}_j(p_j^C, \tilde{p}_j^G)$ has at least one discontinuous point on $(0, \tilde{p}_j^C]$.

Let $\hat{p}_j^C \in (0, \tilde{p}_j^C]$ be a discontinuous point of $\tilde{x}_j(p_j^C, \tilde{p}_j^G)$. From Lemma 3, we have $\hat{p}_j^C - \lim_{p_j^C \rightarrow \hat{p}_j^C} \tilde{x}_j(p_j^C, \tilde{p}_j^G) > 0$.

Therefore, there exists $q \in (0, \hat{p}_j^C]$ such that $\tilde{x}_j(q, \tilde{p}_j^G) < q$. But since $q \leq \hat{p}_j^C \leq \tilde{p}_j^C \leq \tilde{p}_j^G$, $\tilde{x}_j(q, \tilde{p}_j^G) < q \leq \mathbf{min}\{q, \tilde{p}_j^G\}$ and therefore is not Pareto efficient. We have a contradiction.

Therefore, if $\mathbf{x}(p^C, p^G)$ is strategy-proof, minimally efficient and satisfies the checks condition, then

$$\mathbf{x}(p^C, p^G) = \mathbf{m}(p^C, p^G) = (\mathbf{min}\{p_1^C, p_1^G\}, \mathbf{min}\{p_2^C, p_2^G\}, \dots, \mathbf{min}\{p_n^C, p_n^G\}).$$

□

Appendix B

In this section, we extend our main results to a situation in which there are $N > 2$ decision-makers who must agree on changing an existing policy. Similar to the main analysis, each of these N players has veto power in the sense that each has a reservation payoff given by the status quo policy, $q = 0$.

Formally, for $i = 1, 2, \dots, N$, we denote the ideal policy of player i by p_i . A mechanism $x(p_1, p_2, \dots, p_N)$ specifies an outcome $x \in X$ as a function of all members' reported types. A mechanism $x(p_1, p_2, \dots, p_N)$ is a checks and balances mechanism if $U_i(x(p_1, p_2, \dots, p_N), p_i) \geq U_i(q, p_i)$ for all i . We restate the following properties for the case of $N > 2$ players:

Checks: A mechanism $x(c, g)$ satisfies the checks condition if and only if $U_G(x(p_i, p_{-i}), p_i) \geq U_G(0, p_i)$ for all i , p_i , and p_{-i} .

Strategy-proofness: A mechanism $x(p_1, p_2, \dots, p_N)$ is strategy-proof if and only if $U_i(x(p_i, p_{-i}), p_i) \geq U_i(x(\tilde{p}_i, p_{-i}), p_i)$ for all i , p_i , \tilde{p}_i and p_{-i} .

Pareto Efficiency: A mechanism $x(p_1, p_2, \dots, p_N)$ is (ex-post) Pareto efficient if and only if its outcome is Pareto efficient. Formally, for any (p_1, p_2, \dots, p_N) , there is not another outcome x' such that $U_i(x', p_i) \geq U_i(x(p_1, p_2, \dots, p_N), p_i)$ for all i , and $U_j(x', p_j) > U_j(x(p_1, p_2, \dots, p_N), p_j)$ for some j .

Responsiveness: A mechanism $x(p_1, p_2, \dots, p_N)$ is (strictly) responsive to players' preferences if and only if for any $(p'_1, p'_2, \dots, p'_N)$ and (p_1, p_2, \dots, p_N) such that $p'_i > p_i \geq 0$ for all i , $x(p'_1, p'_2, \dots, p'_N) > x(p_1, p_2, \dots, p_N)$, and for any $(p'_1, p'_2, \dots, p'_N)$ and (p_1, p_2, \dots, p_N) such that $p'_i < p_i \leq 0$ for all i , $x(p'_1, p'_2, \dots, p'_N) < x(p_1, p_2, \dots, p_N)$.

We also generalize the definition for the moderate player's most preferred outcome for the case of $N > 2$ players, $m(p_1, p_2, \dots, p_N)$:

Definition. For any (p_1, p_2, \dots, p_N) where all p_i 's have the same signs, the outcome $m(p_1, p_2, \dots, p_N)$ is defined as the most preferred policy of the player whose ideal policy is the closest to the sta-

tus quo $q = 0$; that is, $m(p_1, p_2, \dots, p_N) = \min_{i=1}^N p_i$ if $p_i \geq 0$ for all i , and $m(p_1, p_2, \dots, p_N) = \max_{i=1}^N p_i$ if $p_i \leq 0$ for all i .

Before proving the results with multiple decision-makers, we generalize Lemma 1 through Lemma 4 to the model with N players.

Lemma 5. *For any i and any p_{-i} , any mechanism $x(p_i, p_{-i})$ that is strategy-proof for player i is weakly increasing in p_i , and if $x(p_i, p_{-i})$ strictly increasing in p_i on an open interval (p_i^1, p_i^2) , $x(p_i, p_{-i}) = p_i$ on (p_i^1, p_i^2) .*

Lemma 6. *Let $x(p_1, p_2, \dots, p_N)$ be a strategy-proof mechanism. Then for any i and any p_{-i} , if $x(p_i, p_{-i}) = p_i$ on (p_i^1, p_i^2) , then $x(p_i, p_{-i})$ is continuous at both p_i^1 and p_i^2 .*

Lemma 7. *Let $x(p_1, p_2, \dots, p_N)$ be a strategy-proof mechanism. Then for any i and any p_{-i} , if \hat{p}_i is a discontinuity point of $x(p_i, p_{-i})$, then $\hat{p}_i - \lim_{p_i \rightarrow \hat{p}_i^-} x(p_i, p_{-i}) > 0$ and $\lim_{p_i \rightarrow \hat{p}_i^+} x(p_i, p_{-i}) - \hat{p}_i > 0$.*

Lemma 8. *Let $x(p_1, p_2, \dots, p_N)$ be a strategy-proof mechanism. Then for any i and any p_{-i} , if for some \hat{p}_i $x(\hat{p}_i, p_{-i}) = a \neq \hat{p}_i$, then $x(p_i, p_{-i}) = a$ for all $p_i \in (\min\{a, \hat{p}_i\}, \max\{a, \hat{p}_i\})$.*

The proof of the above lemmas are exactly the same as the proof of Lemma 1, 2, 3, and 4.

Just as in the main analysis, we state the propositions for the case where $p_i \in \mathbf{R}_+$ for all i . We have the following results:

Proposition 7. *The unique mechanism that satisfies the checks, strategy-proofness and Pareto efficiency conditions on \mathbf{R}_+^N is $x(p_1, p_2, \dots, p_N) = m(p_1, p_2, \dots, p_N)$.*

Proof. When $p_i \geq 0$ for all i , $m(p_1, p_2, \dots, p_N) = \min_{i=1}^N p_i$.

In our model with single-peaked preferences, an outcome x is efficient if and only if $\min_{i=1}^N p_i \leq x \leq \max_{i=1}^N p_i$.

First, it is easy to check that the mechanism $x(p_1, p_2, \dots, p_N) = \min_{i=1}^N p_i$ satisfies checks, strategy-proofness and efficiency.

Now we show that for (p_1, p_2, \dots, p_N) where $p_i \geq 0$ for all i , $x(p_1, p_2, \dots, p_N) = \min_{i=1}^N p_i$ is the unique mechanism that satisfies all three conditions.

Note that if $p_i = p_j$ for all $i \neq j$, the unique efficient outcome is $x(p_1, p_2, \dots, p_N) = \min_{i=1}^N p_i$.

Now suppose that there exists another mechanism $\tilde{x}(p_1, p_2, \dots, p_N)$ that satisfies all three conditions. That is, for some $(\tilde{p}_1, \tilde{p}_2, \dots, \tilde{p}_N)$ such that $\tilde{p}_i \geq 0$ for all i , and $\tilde{p}_i \neq \tilde{p}_j$ for some $i \neq j$, $\tilde{x}(\tilde{p}_1, \tilde{p}_2, \dots, \tilde{p}_N) \neq \min_{i=1}^N \tilde{p}_i$. Efficiency implies that $\tilde{x}(\tilde{p}_1, \tilde{p}_2, \dots, \tilde{p}_N) > \min_{i=1}^N \tilde{p}_i$. Let $i^* \in \{1, 2, \dots, N\}$ be such that $\tilde{p}_{i^*} = \min_{i=1}^N \tilde{p}_i$.

Checks condition for player i^* with ideal point 0 implies that $\tilde{x}(0, \tilde{p}_{-i^*}) = 0$. Lemma 5 then implies that $\tilde{x}(p_{i^*}, \tilde{p}_{-i^*})$ has at least one discontinuous point on $[0, \tilde{p}_{i^*}]$. Now if $\tilde{x}(p_{i^*}, \tilde{p}_{-i^*})$ is discontinuous at $p_{i^*} = 0$, i.e. if $\lim_{p_{i^*} \rightarrow 0^+} \tilde{x}(p_{i^*}, \tilde{p}_{-i^*}) > 0$, then there exists $\epsilon > 0$ such that $U_{i^*}(0, \epsilon) > U_{i^*}(\tilde{x}(\epsilon, \tilde{p}_{-i^*}), \epsilon)$. But then the player i^* with ideal point ϵ has an incentive to deviate and report type 0, a contradiction to $\tilde{x}(p_1, p_2, \dots, p_N)$ being strategy-proof. Therefore, $\tilde{x}(p_{i^*}, \tilde{p}_{-i^*})$ is continuous at $p_{i^*} = 0$ and therefore $\tilde{x}(p_{i^*}, \tilde{p}_{-i^*})$ has at least one discontinuous point on $(0, \tilde{p}_{i^*}]$.

Let $\hat{p}_{i^*} \in (0, \tilde{p}_{i^*}]$ be a discontinuous point of $\tilde{x}(p_{i^*}, \tilde{p}_{-i^*})$. From Lemma 7, we have $\hat{p}_{i^*} - \lim_{p_{i^*} \rightarrow \hat{p}_{i^*}^-} \tilde{x}(p_{i^*}, \tilde{p}_{-i^*}) > 0$.

Therefore, there exists $p_{i^*}^1 \in (0, \hat{p}_{i^*}]$ such that $\tilde{x}(p_{i^*}^1, \tilde{p}_{-i^*}) < p_{i^*}^1$. But since $p_{i^*}^1 \leq \hat{p}_{i^*} \leq \tilde{p}_{i^*} \leq \tilde{p}_j$ for all $j \neq i^*$, $\tilde{x}(p_{i^*}^1, \tilde{p}_{-i^*}) < p_{i^*}^1 \leq \min\{p_{i^*}^1, \tilde{p}_{-i^*}\}$ and therefore is not efficient. We have a contradiction.

Therefore the unique mechanism that satisfies checks, strategy-proofness and Pareto efficiency is $x(p_1, p_2, \dots, p_N) = \min_{i=1}^N p_i = m(p_1, p_2, \dots, p_N)$. \square

Proposition 8. *The unique mechanism that satisfies the checks, strategy-proofness and responsiveness conditions on \mathbf{R}_+^N is $x(p_1, p_2, \dots, p_N) = m(p_1, p_2, \dots, p_N)$.*

Proof. When $p_i \geq 0$ for all i , $m(p_1, p_2, \dots, p_N) = \min_{i=1}^N p_i$.

First, it is easy to check that the mechanism $x(p_1, p_2, \dots, p_N) = \min_{i=1}^N p_i$ satisfies checks, strategy-proofness and responsiveness.

Next, we show that any mechanism that is strategy-proof and responsive must be that $x(p_1, p_2, \dots, p_N) \in \{p_1, p_2, \dots, p_N\}$ for all (p_1, p_2, \dots, p_N) . Suppose not, then there exist $(\hat{p}_1, \hat{p}_2, \dots, \hat{p}_N)$ and $a \notin \{\hat{p}_1, \hat{p}_2, \dots, \hat{p}_N\}$ such that $x(\hat{p}_1, \hat{p}_2, \dots, \hat{p}_N) = a$. Then $\min\{a, \hat{p}_i\} \neq \max\{a, \hat{p}_i\}$ for all i . Lemma 8 implies that $x(p_i, \hat{p}_{-i}) = a$ for all i and $p_i \in [\min\{a, \hat{p}_i\}, \max\{a, \hat{p}_i\}]$. This further implies that $x(p_1, p_2, \dots, p_N) = a$ for all $p_i \in [\min\{a, \hat{p}_i\}, \max\{a, \hat{p}_i\}]$. In particular, $x(\min\{a, \hat{p}_1\}, \min\{a, \hat{p}_2\}, \dots, \min\{a, \hat{p}_N\}) = x(\max\{a, \hat{p}_1\}, \max\{a, \hat{p}_2\}, \dots, \max\{a, \hat{p}_N\}) = a$ but $\max\{a, \hat{p}_i\} > \min\{a, \hat{p}_i\}$ for all i , a contradiction to $x(p_1, p_2, \dots, p_N)$ being responsive.

Therefore, if $x(p_1, p_2, \dots, p_N)$ is strategy-proof and responsive, $x(p_1, p_2, \dots, p_N) \in \{p_1, p_2, \dots, p_N\}$ for all (p_1, p_2, \dots, p_N) . We next prove that for (p_1, p_2, \dots, p_N) where $p_i \geq 0$ for all i , the unique mechanism that satisfies checks, strategy-proofness and responsiveness is $x(p_1, p_2, \dots, p_N) = \min_{i=1}^N p_i$. Suppose there exists another mechanism $\tilde{x}(p_1, p_2, \dots, p_N)$ that satisfies all three conditions. From what we show above, since $\tilde{x}(p_1, p_2, \dots, p_N)$ is strategy-proof and responsive, $\tilde{x}(p_1, p_2, \dots, p_N) \in \{p_1, p_2, \dots, p_N\}$ for all (p_1, p_2, \dots, p_N) . For any (p_1, p_2, \dots, p_N) such that $p_i = p_j$ for all $i \neq j$, $\tilde{x}(p_1, p_2, \dots, p_N) \in \{p_1, p_2, \dots, p_N\}$ implies $\tilde{x}(p_1, p_2, \dots, p_N) = \min_{i=1}^N p_i$. Therefore there exists $(\tilde{p}_1, \tilde{p}_2, \dots, \tilde{p}_N)$ with $\tilde{p}_i \neq \tilde{p}_j$ for some $i \neq j$, such that $\tilde{x}(\tilde{p}_1, \tilde{p}_2, \dots, \tilde{p}_N) \neq \min_{i=1}^N \tilde{p}_i$. Let i^* be such that $\tilde{p}_{i^*} = \min_{i=1}^N \tilde{p}_i$, then $\tilde{x}(\tilde{p}_{i^*}, \tilde{p}_{-i^*}) > \tilde{p}_{i^*}$. Since $\tilde{x}(p_1, p_2, \dots, p_N)$ satisfies the checks condition for player i^* , $\tilde{x}(0, \tilde{p}_{-i^*}) = 0$. By Lemma 5 and 7, $\tilde{x}(0, \tilde{p}_{-i^*}) = 0$ and $\tilde{x}(\tilde{p}_{i^*}, \tilde{p}_{-i^*}) > \tilde{p}_{i^*}$ imply that there exists $p_{i^*}^1 \in (0, \tilde{p}_{i^*}]$ such that $\tilde{x}(p_{i^*}^1, \tilde{p}_{-i^*})$ is discontinuous at $p_{i^*}^1$, and there exists $p_{i^*}^2 < p_{i^*}^1$ such that $\tilde{x}(p_{i^*}^2, \tilde{p}_{-i^*}) < p_{i^*}^2$. Since $p_{i^*}^2 < p_{i^*}^1 \leq \tilde{p}_{i^*} \leq \tilde{p}_j$ for all $j \neq i^*$, $\tilde{x}(p_{i^*}^2, \tilde{p}_{-i^*}) \notin \{p_{i^*}^2, \tilde{p}_{-i^*}\}$, a contradiction.

Therefore the unique mechanism that satisfies checks, strategy-proofness and responsiveness is $x(p_1, p_2, \dots, p_N) = \min_{i=1}^N p_i = m(p_1, p_2, \dots, p_N)$. \square

Similar to what we found in the application to constitutional review section, the mechanism that implements the preferred policy of the more moderate player can be obtained

as the unique equilibrium outcome by means of a simple institutional arrangement. In this institution, $N - 1$ players choose a legal limit $\ell_i \in \mathbf{R}_+$ in a fixed sequential order,³ and the remaining player chooses a policy $x \in \mathbf{R}_+$. A policy x is legal if and only if x does not exceed the lowest of the $N - 1$ legal limits. In other words, a policy is legal if and only if it is within all the legal bounds set by the previous $N - 1$ players. The outcome is x if x is legal and the status quo, $q = 0$, otherwise. We have the following proposition:

Proposition 9. *The unique equilibrium outcome in the institution in which $N - 1$ players define the legal limits $\ell_i \in \mathbf{R}_+$ (in a fixed sequential order), and then the remaining player chooses a policy $x \in \mathbf{R}_+$, is $m(p_1, p_2, \dots, p_N)$.*

Proof. Without loss of generality re-order the players such that the player with ideal points p_i is the i -th one to make a choice. That is, player 1 through $N - 1$ each chooses a legal limit in the order $1, 2, \dots, N - 1$, and then player N (observing all the legal limits chosen by the previous $N - 1$ players) chooses a policy. We claim that the strategy profile $\ell_i(\ell_1, \dots, \ell_{i-1}, p_i) = p_i$ for $i = 1, 2, \dots, N - 1$ and $x_N(\ell_1, \dots, \ell_{N-1}, p_N) = \min\{\min_{i=1}^{N-1} \ell_i, p_N\}$ is a Bayesian Nash equilibrium strategy and it gives as the unique Bayesian Nash equilibrium outcome the ideal policy of the most moderate player. We prove this claim by backward induction.

In the last stage, the effective legal limit is $\min_{i=1}^{N-1} \ell_i$, i.e. if player N chooses a policy not exceeding $\min_{i=1}^{N-1} \ell_i$ the outcome is this chosen policy, otherwise the outcome is the status quo. If the effective legal limit chosen by the previous $N - 1$ players is $\min_{i=1}^{N-1} \ell_i$, for any strategy of the previous $N - 1$ players and any beliefs of player N , $x_N(\ell_1, \dots, \ell_{N-1}, p_N) = \min\{\min_{i=1}^{N-1} \ell_i, p_N\}$ is the unique optimal strategy for player N .

In the second-to-last stage, if the legal limits chosen by the previous $N - 2$ players are $\ell_1, \dots, \ell_{N-2}$, for any strategy of the previous $N - 2$ players and any beliefs of player $N - 1$, truth-telling (i.e. $\ell_{N-1}(\ell_1, \dots, \ell_{N-2}, p_{N-1}) = p_{N-1}$) is optimal for player $N - 1$. This is the case since if player $N - 1$ deviates to $\ell' < p_{N-1}$, she either does not change the outcome

³The identities of these $N - 1$ players can be arbitrary.

and thus receives the same payoff, or changes the outcome from somewhere in between ℓ' and p_{N-1} to ℓ' . Since players have single-peaked preferences, this change will make player $N - 1$ worse off. A similar argument shows that the player has no incentive to deviate to $\ell' > p_{N-1}$ either. Iterating this argument back to the first stage proves our claim.

□