SUPPLEMENT TO 'THE ECONOMICS OF PARTISAN GERRYMANDERING'

ANTON KOLOTILIN
School of Economics, UNSW Business School

ALEXANDER WOLITZKY
Department of Economics, MIT

APPENDIX A: PROOFS

A.1. *Proof of Lemma 2*

Lemma 2 follows from Theorem 4 in Kolotilin et al. (2025) for the translation-invariant subcase of the state-independent sender case. For completeness, we prove Lemmas 3 and 4, which immediately yield Lemma 2.

LEMMA 3: *For any optimal $\mathcal{H}$ and any $P, P' \in \operatorname{supp}(\mathcal{H})$ such that P contains types $s < s''$ and P' contains a type $s' \in (s, s'')$, we have $r^*(P) \geq r^*(P')$.*

PROOF OF LEMMA 3: Fix such a plan $\mathcal{H}$, districts P, P', voter types $s < s' < s''$, and district strengths $r = r^*(P)$ and $r' = r^*(P')$. By Lemma 1, we have

$$G(r) + \lambda(r)(Q(s-r) - Q(0)) \geq G(r') + \lambda(r')(Q(s-r') - Q(0)), \tag{7}$$

$$G(r') + \lambda(r')(Q(s'-r') - Q(0)) \geq G(r) + \lambda(r)(Q(s'-r) - Q(0)), \quad \text{and} \tag{8}$$

$$G(r) + \lambda(r)(Q(s''-r) - Q(0)) \geq G(r') + \lambda(r')(Q(s''-r') - Q(0)). \tag{9}$$

These inequalities imply that

$$\begin{aligned} 0 &\geq (Q(s''-r') - Q(s'-r'))(Q(s'-r) - Q(s-r)) \\ &\quad -(Q(s''-r) - Q(s'-r))(Q(s'-r') - Q(s-r')) \\ &= \int_{s'}^{s''} \int_{s}^{s'} q(\tilde{s}'-r')q(\tilde{s}-r)d\tilde{s}d\tilde{s}' - \int_{s'}^{s''} \int_{s}^{s'} q(\tilde{s}'-r)q(\tilde{s}-r')d\tilde{s}d\tilde{s}' \\ &= \int_{s'}^{s''} \int_{s}^{s'} \left(q(\tilde{s}'-r')q(\tilde{s}-r) - q(\tilde{s}'-r)q(\tilde{s}-r')\right) d\tilde{s}d\tilde{s}', \end{aligned}$$

where the first inequality holds by summing (7) multiplied by $Q(s''-r) - Q(s'-r)$, (8) multiplied by $Q(s''-r) - Q(s-r)$, and (9) multiplied by $Q(s'-r) - Q(s-r)$, and then dividing by $\lambda(r')$, which is strictly positive by Lemma 1. This inequality in turn implies that $r \geq r'$, as if $r < r'$ then the integrand is strictly positive by $\tilde{s} < \tilde{s}'$ and strict log-concavity of q. *Q.E.D.*

LEMMA 4: *For any optimal $\mathcal{H}$ and any $P, P' \in \operatorname{supp}(\mathcal{H})$ such that $r^*(P) = r^*(P')$, we have* $|\operatorname{supp}(P) \bigcup \operatorname{supp}(P')| \leq 2$.

Anton Kolotilin: akolotilin@gmail.com
Alexander Wolitzky: wolitzky@mit.com

PROOF OF LEMMA 4: It suffices to show that there cannot be a district $P \in \text{supp}(\mathcal{H})$ such that $|\text{supp}(P)| \geq 3$, because if $\mathcal{H}$ contains two districts P, P' such that $r^*(P) = r^*(P')$ and $|\text{supp}(P) \bigcup \text{supp}(P')| \geq 3$, they can be merged into a single district without affecting optimality. So, suppose by contradiction that some district $P \in \text{supp}(\mathcal{H})$ contains three types $s < s' < s''$ and $r^*(P) = r$. By Theorem 1 of Kolotilin et al. (2025), the function λ in Lemma 1 can be taken to be differentiable at $r^*(P)$ for all non-degenerate $P \in \text{supp}(\mathcal{H})$, where, for all $s \in \text{supp}(P)$, its derivative $\lambda'(r^*(P))$ satisfies

$$g(r^*(P)) - \lambda(r^*(P))q(s - r^*(P)) + \lambda'(r^*(P))\left(Q(s - r^*(P)) - Q(0)\right) = 0.$$[56]

We thus have

$$g(r) - \lambda(r)q(s-r) + \lambda'(r)\left(Q(s-r) - Q(0)\right) = 0, \tag{10}$$

$$g(r) - \lambda(r)q(s'-r) + \lambda'(r)\left(Q(s'-r) - Q(0)\right) = 0, \quad \text{and} \tag{11}$$

$$g(r) - \lambda(r)q(s''-r) + \lambda'(r)\left(Q(s''-r) - Q(0)\right) = 0. \tag{12}$$

This yields a contradiction because

$$\begin{aligned}
0 = \det &\begin{pmatrix} g(r) & q(s-r) & Q(s-r) - Q(0) \\ g(r) & q(s'-r) & Q(s'-r) - Q(0) \\ g(r) & q(s''-r) & Q(s''-r) - Q(0) \end{pmatrix} \\
&= g(r)(q(s'-r) - q(s-r))(Q(s''-r) - Q(s'-r)) \\
&\quad - g(r)(q(s''-r) - q(s'-r))(Q(s'-r) - Q(s-r)) \\
&= g(r)\left[\int_s^{s'} q'(\tilde{s}-r)d\tilde{s} \int_{s'}^{s''} q(\tilde{s}'-r)d\tilde{s}' - \int_{s'}^{s''} q'(\tilde{s}'-r)d\tilde{s}' \int_s^{s'} q(\tilde{s}-r)d\tilde{s}\right] \\
&> \frac{g(r)q'(s'-r)}{q(s'-r)}\left[\int_s^{s'} q(\tilde{s}-r)d\tilde{s} \int_{s'}^{s''} q(\tilde{s}'-r)d\tilde{s}' - \int_{s'}^{s''} q(\tilde{s}'-r)d\tilde{s}' \int_s^{s'} q(\tilde{s}-r)d\tilde{s}\right] = 0,
\end{aligned}$$

where the first equality is by (10)–(12), and the inequality is by strict log-concavity of q, which implies that the derivative of $\ln q$ is strictly decreasing, yielding

$$\frac{q'(\tilde{s}-r)}{q(\tilde{s}-r)} > \frac{q'(s'-r)}{q(s'-r)} > \frac{q'(\tilde{s}'-r)}{q(\tilde{s}'-r)}, \quad \text{for } \tilde{s} < s' < \tilde{s}'.$$ *Q.E.D.*

A.2. *Characterization of Segregate-Pair Districting*

LEMMA 5: *For any segregate-pair districting plan $\mathcal{H}$, there exists a bifurcation point $r^b \in (\underline{s}, \overline{s}]$, a decreasing function $s_1 : (r^b, \overline{s}) \to [\underline{s}, r^b)$, and an increasing function $s_2 : (r^b, \overline{s}) \to (r^b, \overline{s}]$ satisfying $s_1(r) < r < s_2(r)$, such that for each $P \in \text{supp}(\mathcal{H})$, we have $P = \delta_{r^*(P)}$ if $r^*(P) \leq r^b$ and $\text{supp}(P) = \{s_1(r^*(P)), s_2(r^*(P))\}$ if $r^*(P) > r^b$.*

PROOF OF LEMMA 5: Let $\mathcal{H}$ be a segregate-pair districting plan. Since $\mathcal{H}$ is strictly single-dipped, the support of each $P \in \text{supp}(\mathcal{H})$ has at most two elements and thus can be represented

[56] Intuitively, this is the first-order condition with respect to r of the designer's maximization problem in Lemma 1.

as $\{s_1(r^*(P)), s_2(r^*(P))\}$ with $s_1(r^*(P)) \le r^*(P) \le s_2(r^*(P))$. Moreover, for each $P, P' \in \text{supp}(\mathcal{H})$ with $r^*(P) < r^*(P')$, we have $s_2(r^*(P)) \le s_2(r^*(P'))$, as otherwise we would have $s_2(r^*(P')) \in (s_1(r^*(P)), s_2(r^*(P)))$, contradicting strict single-dippedness.

Assume that there exists $P \in \text{supp}(\mathcal{H})$ such that $s_1(r^*(P)) < s_2(r^*(P))$, as otherwise the lemma obviously holds with $r^b = \overline{s}$. Define $r^b = \inf\{r^*(\tilde{P}) : \tilde{P} \in \text{supp}(\mathcal{H}),\ s_1(r^*(\tilde{P})) < s_2(r^*(\tilde{P}))\}$, so that, for each $P \in \text{supp}(\mathcal{H})$ with $r^*(P) < r^b$, we have $\text{supp}(P) = \{r^*(P)\}$. Since $\text{supp}(\mathcal{H})$ is compact, there exists $P^b \in \text{supp}(\mathcal{H})$ with $r^*(P^b) = r^b$. It follows that $\text{supp}(P^b) = \{r^b\}$, as otherwise (i.e., if $s_1(r^*(P^b)) < r^b < s_2(r^*(P^b))$ voter types in $(r^b, s_2(r^*(P^b)))$ (which have strictly positive mass since f is strictly positive on $[\underline{s}, \overline{s}]$) cannot be segregated, as this would contradict strict single-dippedness, and also cannot be paired with other types, as this would contradict either strict single-dippedness or the definition of r^b.

Next, we show that, for each $P, P' \in \text{supp}(\mathcal{H})$ with $r^b < r^*(P) < r^*(P')$, we have $s_1(r^*(P)) \ge s_1(r^*(P'))$. Suppose by contradiction that $s_1(r^*(P)) < s_1(r^*(P'))$. Since $\mathcal{H}$ is a strictly single-dipped segregate-pair districting plan, by the definition of r^b, we have $s_1(r^*(P)) < r^*(P) < s_2(r^*(P)) \le s_1(r^*(P')) < r^*(P') < s_2(r^*(P'))$. Define $r^\dagger = \inf\{r^*(\tilde{P}) : \tilde{P} \in \text{supp}(\mathcal{H}),\ s_1(r^*(P')) \le s_1(r^*(\tilde{P})) < s_2(r^*(\tilde{P})) \le s_2(r^*(P'))\} \ge s_1(r^*(P'))$. By the same argument as in the previous paragraph, we have $\delta_{r^\dagger} \in \text{supp}(\mathcal{H})$, contradicting that $\mathcal{H}$ is segregate-pair.

Finally, we have (i) $r^b > \underline{s}$, as otherwise all voter types above $\underline{s}$ are paired with $\underline{s}$, contradicting that F has no atom at $\underline{s}$, and (ii) $\sup\{r^*(\tilde{P}) : \tilde{P} \in \text{supp}(\mathcal{H}),\ s_1(r^*(\tilde{P})) < s_2(r^*(\tilde{P}))\} < \overline{s}$ when $r^b < \overline{s}$, as otherwise $\delta_{\overline{s}} \in \text{supp}(\mathcal{H})$ by the same argument as in the second paragraph, contradicting that $\mathcal{H}$ is segregate-pair. Then we can extend the functions s_1 and s_2 from the set $\tilde{R} = \{r^*(\tilde{P}) : \tilde{P} \in \text{supp}(\mathcal{H}),\ s_1(r^*(\tilde{P})) < s_2(r^*(\tilde{P}))\} \subset (r^b, \underline{s})$ to the interval $(r^b, \underline{s})$ by setting $s_1(r) = \inf\{s_1(\tilde{r}) : \tilde{r} \in \tilde{R},\ \tilde{r} < r\}$ and $s_2(r) = \sup\{s_2(\tilde{r}) : \tilde{r} \in \tilde{R},\ \tilde{r} < r\}$ for all $r \in (r^b, \overline{s}) \setminus \tilde{R}$. By construction, the extended functions s_1 and s_2 are as required. *Q.E.D.*

A.3. *Auxiliary Lemmas*

Lemmas 6–9 are used to prove Theorems 1 and 2.

LEMMA 6: *If for all $s < r < s'$ such that*

$$G(r) + \lambda(r)\left(Q(s-r) - Q(0)\right) \ge G(s), \tag{13}$$

where

$$\lambda(r) = \frac{g(r)(Q(s'-r)) - Q(s-r))}{(Q(s'-r) - Q(0))q(s-r) - (Q(s-r) - Q(0))q(s'-r)}, \tag{14}$$

we have, for all $s'' \ge s'$,

$$G(r) + \lambda(r)(Q(s-r) - Q(0)) < G(s'') + \frac{g(s'')}{q(0)}(Q(s-s'') - Q(0)), \tag{15}$$

then there is a unique optimal districting plan, which is segregate-pair.

PROOF OF LEMMA 6: Suppose by contradiction that there exists an optimal non-segregate-pair plan $\mathcal{H}$. By Lemma 2, $\mathcal{H}$ is strictly single-dipped. Consequently, since $\mathcal{H}$ is not segregate-pair, there exist $s < r < s' \le s''$ and $P, P' \in \text{supp}(\mathcal{H})$ such that $r^*(P) = r$, $\text{supp}(P) = \{s, s'\}$, and $\text{supp}(P') = \{s''\}$. By Lemma 1, condition (13) holds and condition (15) fails, yielding a contradiction.[57] Finally, for uniqueness, by Theorem 7 in Kolotilin et al. (2025), it suffices to show that $\mathcal{H}$ is *regular*, in that for each $P \in \text{supp}(\mathcal{H})$, there exists $\varepsilon > 0$ such that either (i) $|\text{supp}(\tilde{P})| = 1$ for all $\tilde{P} \in \text{supp}(\mathcal{H})$ satisfying $r^*(\tilde{P}) \in (r^*(P) - \varepsilon, r^*(P))$, or (ii) $|\text{supp}(\tilde{P})| = 2$ for all $\tilde{P} \in \text{supp}(\mathcal{H})$ satisfying $r^*(\tilde{P}) \in (r^*(P) - \varepsilon, r^*(P))$. But each segregate-pair plan $\mathcal{H}$ is clearly regular, with any $\varepsilon > 0$ for $r^*(P) \leq r^b$ and with any $\varepsilon \in (0, r^*(P) - r^b)$ for $r^*(P) > r^b$. *Q.E.D.*

LEMMA 7: *If $s < r < s'$, then $\lambda(r)$ given by (14) satisfies $\lambda(r) > g(r)/q(0)$.*

PROOF OF LEMMA 7: Follows from (14) and q being uniquely maximized at 0. *Q.E.D.*

LEMMA 8: *If $\eta \geq 1$ and $s < r < s'$ satisfy (13), with $\lambda(r)$ given by (14), then $r > 0$.*

PROOF OF LEMMA 8: If $r \leq 0$, then (13) fails, because

$$G(r) - G(s) = \int_s^r g(x)dx \leq \frac{g(r)}{g(0)} \int_s^r g(x-r)dx = \frac{g(r)}{g(0)}(G(0) - G(s-r))$$

$$= \frac{g(r)}{\eta q(0)}(Q(\eta 0) - Q(\eta(s-r)) \leq \frac{g(r)}{q(0)}(Q(0) - Q(s-r)) < \lambda(r)\left(Q(0) - Q(s-r)\right),$$

where the first inequality is by $s < r \leq 0$ and strict log-concavity of g on $[s, 0]$, the second inequality is by $\eta \geq 1$ and strict convexity of Q on $[s-r, 0]$, and the last inequality is by Lemma 7. *Q.E.D.*

LEMMA 9: *If $\mathcal{H}$ is optimal and $\delta_s, \delta_{s'} \in \text{supp}(\mathcal{H})$ with $s < s'$, then $s < 0$.*

PROOF OF LEMMA 9: Follows from (4). *Q.E.D.*

A.4. *Proof of Theorem 1*

Given Lemmas 7 and 8, the argument given in the text shows that if $s < r < s'$ satisfy (13), with $\lambda(r)$ given by (14), then (15) holds for all $s'' \geq s'$. The theorem then follows from Lemma 6.

A.5. *Proof of Theorem 2*

Part 1. Let $\mathcal{H}$ be an optimal strictly single-dipped plan. By Lemma 9, there do not exist $s < s'$ in $[\underline{s}, \overline{s}]$ such that $\delta_s, \delta_{s'} \in \text{supp}(\mathcal{H})$. Then, by Theorem 6 in Kolotilin et al. (2025), $\mathcal{H}$ is NAD.

Part 2. Suppose by contradiction that there exist an optimal strictly single-dipped plan $\mathcal{H}$ and $P \in \text{supp}(\mathcal{H})$ such that $r^*(P) = r$ and $\text{supp}(P) = \{s, s'\}$ with $s < r < s'$. By Lemma 1, (13) holds with $\lambda(r)$ given by (14). So, by Lemma 8, $r^*(P) > 0$, contradicting that $r < s' \leq \overline{s} \leq 0$.

Part 3. Since f is strictly positive on $[\underline{s}, \overline{s}]$ and $\underline{s} < \overline{s}$, we have $\underline{s} < r^*(F) = 0 < \overline{s}$, so segregation is suboptimal by Lemma 9.

---

[57] Intuitively, (13) says that the designer prefers not to move a few type-s voters from district P to district δ_s, and (15) says that the designer strictly prefers to move a few type-s voters from district P to district $\delta_{s''}$.

Suppose by contradiction that there exists an optimal NAD plan $\mathcal{H}$. By Lemma 5, for each $P \in \operatorname{supp}(\mathcal{H})$ except for δ_{r^b}, we have $s_1(r^*(P)) < r^*(P) < s_2(r^*(P))$, where s_1 is decreasing and s_2 is increasing. Note that $r^b < r^*(F) = 0$, because

$$\int Q(s - r^*(F))dF(s) = Q(0) = \iint Q(s - r^*(P))dP(s)d\mathcal{H}(P)$$
$$< \iint Q(s - r^b)dP(s)d\mathcal{H}(P) = \int Q(s - r^b)dF(s),$$

where the first two equalities hold by the definition of $r^*(F)$ and $r^*(P)$, the inequality holds by $r^*(P) > r^b$ for all $P \in \operatorname{supp}(\mathcal{H})$ except for $P = \delta_{r^b}$, and the last equality holds by $\int P d\mathcal{H}(P) = F$. Since f is strictly positive on $[\underline{s}, \overline{s}]$, we have $\lim_{r \downarrow r^b} s_1(r) = \lim_{r \downarrow r^b} s_2(r) = r^b$, as otherwise voter types in $(\lim_{r \downarrow r^b} s_1(r), \lim_{r \downarrow r^b} s_2(r))$ are not assigned to any district. Thus, for any $\varepsilon > 0$, there exists $P \in \operatorname{supp}(\mathcal{H})$ such that, for $r = r^*(P)$, $s = s_1(r)$, and $s' = s_2(r)$, we have $r^b - \varepsilon \le s < r < s' \le r^b + \varepsilon$, and

$$G(r) + \frac{g(r)}{q(0)}(Q(s-r) - Q(0)) > G(r) + \lambda(r)(Q(s-r) - Q(0)) \ge G(s),$$

where the first inequality is by Lemma 7, and the second inequality is by Lemma 1. But, for sufficiently small $\varepsilon \in (0, -r^b)$, this contradicts inequality (5) in the text.

A.6. *Proof of Theorem 3*

Theorem 3 follows from Lemmas 10–14.

For each r, let $\mathcal{R}(r)$ be the set of optimal plans $\mathcal{H}$ when the aggregate shock is sure to be r (no aggregate uncertainty). Lemma 10 characterizes $\mathcal{R}(r)$. If $r^*(F) \ge r$, then $\mathcal{H} \in \mathcal{R}(r)$ assigns all voters to districts that the designer wins. If $r^*(F) < r$, then $\mathcal{H} \in \mathcal{R}(r)$ assigns all voter types above $s^*(r)$ to cracked districts that the designer wins with exactly 50% of the vote and packs the remaining voters arbitrarily.

LEMMA 10: *The following hold.*
1. *Let $r^*(F) \ge r$. Then $\mathcal{H} \in \mathcal{R}(r)$ iff, for each $P \in \operatorname{supp}(\mathcal{H})$, we have $r^*(P) \ge r$.*
2. *Let $r^*(F) < r$. Then $\mathcal{H} \in \mathcal{R}(r)$ iff, for each $P \in \operatorname{supp}(\mathcal{H})$, we have either* $\operatorname{supp}(P) \subset [\underline{s}, s^*(r)]$ *or* $\operatorname{supp}(P) \subset [s^*(r), \overline{s}]$ *and $r^*(P) = r$.*

PROOF OF LEMMA 10: *Part 1.* Since $r^*(F) \ge r$, $s^*(r) = \underline{s}$ and $\delta_F \in \mathcal{R}(r)$ is optimal. Hence, $\mathcal{H} \in \mathcal{R}(r)$ iff $\int \mathbf{1}\{r \le r^*(P)\} d\mathcal{H}(P) = 1$, which is equivalent to $r^*(P) \ge r$ for all $P \in \operatorname{supp}(\mathcal{H})$, because the set $\{P \in \Delta[\underline{s}, \overline{s}] : r^*(P) \ge r\}$ is closed by the continuity of r^*, which follows from the continuity and strict monotonicity of Q.

Part 2. Assume that $r^*(F) < r < \overline{s}$, as if $r \geq \overline{s}$ then $s^*(r) = \overline{s}$, so part 2 holds trivially. For each plan $\mathcal{H}$, we have

$$\begin{aligned}\int \mathbf{1}\{r \leq r^*(P)\} d\mathcal{H}(P) &= \int \mathbf{1}\left\{\mathbb{E}_P[Q(s-r) - Q(0)] \geq 0\right\} d\mathcal{H}(P) \\ &\leq \int \max\left\{0, \frac{\mathbb{E}_P[Q(s-r)] - Q(s^*(r)-r)}{Q(0) - Q(s^*(r)-r)}\right\} d\mathcal{H}(P) \\ &\leq \iint \max\left\{0, \frac{Q(s-r) - Q(s^*(r)-r)}{Q(0) - Q(s^*(r)-r)}\right\} dP(s) d\mathcal{H}(P) \qquad (16) \\ &= \int \max\left\{0, \frac{Q(s-r) - Q(s^*(r)-r)}{Q(0) - Q(s^*(r)-r)}\right\} dF(s) \\ &= \int_{s^*(r)}^{\overline{s}} \frac{Q(s-r) - Q(s^*(r)-r)}{Q(0) - Q(s^*(r)-r)} dF(s) = 1 - F(s^*(r)),\end{aligned}$$

where the first equality is by the definition of $r^*(P)$, the first inequality is by pointwise dominace of the integrands, the second inequality is by Jensen's inequality, the second equality is by $\int P d\mathcal{H}(P) = F$, the third equality is by strict monotonicity of Q, and the last equality is by the definition of $s^*(r)$. Hence, $\mathcal{H} \in \mathcal{R}(r)$ iff, for a measure-1 set of districts P under $\mathcal{H}$, we have (a) $\mathbb{E}_P[Q(s-r)] \leq Q(s^*(r)-r)$ or $\mathbb{E}_P[Q(s-r)] = Q(0)$ (as otherwise the first inequality in (16) is strict) and (b) $\text{supp}(P) \subset [\underline{s}, s^*(r)]$ or $\text{supp}(P) \subset [s^*(r), \overline{s}]$ (as otherwise the second inequality in (16) is strict), or equivalently, either (i) $\text{supp}(P) \subset [\underline{s}, s^*(r)]$ (which implies that $\mathbb{E}_P[Q(s-r)] \leq Q(s^*(r)-r)$) or (ii) $\text{supp}(P) \subset [\underline{s}, s^*(r)]$ and $r^*(P) = r$ (which is equivalent to $\mathbb{E}_P[Q(s-r)] = Q(0)$). Finally, as in the proof of part 1, continuity implies that properties (i) or (ii) hold for all $P \in \text{supp}(\mathcal{H})$, rather than just for a measure-1 set. *Q.E.D.*

Lemma 11 shows that pack-and-crack districting is approximately optimal. An upper bound on the designer's optimal expected seat share V_η can be obtained by allowing the designer to choose $\mathcal{H}_r \in \mathcal{R}(r)$ after observing each realization r,

$$\overline{V}_\eta = \int (1 - F(s^*(r))) dG_\eta(r).$$

A lower bound on V_η can be obtained by restricting attention to $\mathcal{H}_{\tilde{r}} \in \mathcal{R}(\tilde{r})$ for some $\tilde{r}$,

$$\underline{V}_\eta(\tilde{r}) = \int (1 - F(s^*(\tilde{r}))) \mathbf{1}\{r \leq \tilde{r}\} dG_\eta(r).$$

LEMMA 11: *For all η and all $\tilde{r}$, we have $\underline{V}_\eta(\tilde{r}) \leq V_\eta \leq \overline{V}_\eta$. Moreover, if $\eta \to \infty$, then $\overline{V}_\eta \to 1 - F(s^*(0))$, $\underline{V}_\eta(\tilde{r}) \to 1 - F(s^*(\tilde{r}))$ for all $\tilde{r} > 0$, and $V_\eta \to 1 - F(s^*(0))$.*

PROOF OF LEMMA 11: Let $\mathcal{H}_\eta$ be the optimal plan and let $\mathcal{H}_r$ be any districting plan in $\mathcal{R}(r)$. On the one hand, we have

$$\begin{aligned}V_\eta &= \iint \mathbf{1}\{r \leq r^*(P)\} d\mathcal{H}_\eta(P) dG_\eta(r) \\ &< \iint \mathbf{1}\{r \leq r^*(P)\} d\mathcal{H}_r(P) dG_\eta(r) = \int (1 - F(s^*(r))) dG_\eta(r) = \overline{V}_\eta,\end{aligned}$$

where the inequality holds because $\int \mathbf{1}\{r \leq r^*(P)\} d\mathcal{H}_\eta(P) \leq \int \mathbf{1}\{r \leq r^*(P)\} d\mathcal{H}_r(P)$ for all r by the definition of $\mathcal{H}_r$.

On the other hand, for any $\tilde{r}$, we have

$$V_\eta = \iint \mathbf{1}\{r \leq r^*(P)\} d\mathcal{H}_\eta(P) dG_\eta(r)$$

$$\geq \iint \mathbf{1}\{r \leq r^*(P)\} d\mathcal{H}_{\tilde{r}}(P) dG_\eta(r) \geq \int (1 - F(s^*(\tilde{r}))) \mathbf{1}\{r \leq \tilde{r}\} dG_\eta(r) = \underline{V}_\eta(\tilde{r}),$$

where the first inequality holds because $\mathcal{H}_\eta$ is optimal and $\mathcal{H}_{\tilde{r}}$ is feasible, and the second inequality holds by Lemma 10.

Suppose now that $\eta \to \infty$, which implies that $G_\eta \to \delta_0$. By the implicit function theorem, $F(s^*(r))$ is continuous in r, so $\overline{V}_\eta \to 1 - F(s^*(0))$. For $\tilde{r} > 0$, $\underline{V}_\eta(\tilde{r}) \to 1 - F(s^*(\tilde{r}))$, which converges to $1 - F(s^*(0))$ as $\tilde{r} \downarrow 0$, implying that $V_\eta \to 1 - F(s^*(0))$. *Q.E.D.*

Lemma 12 shows that limit points of optimal plans $H_n = H_{\eta_n}$, for $\eta_n \to \infty$, belong to $\mathcal{P}(0)$.

LEMMA 12: *Let $\mathcal{H}_n \to \mathcal{H}$ as $\eta_n \to \infty$. Then $\mathcal{H} \in \mathcal{R}(0)$.*

PROOF OF LEMMA 12: Suppose by contradiction that there exists a sequence $\eta_n \to \infty$ such that an optimal plan $\mathcal{H}_n$ converges weakly to $\mathcal{H} \notin \mathcal{R}(0)$. Then we have

$$1 - F(s^*(0)) = \lim_{n \to \infty} \int Q(\eta_n r^*(P)) d\mathcal{H}_n(P) \leq \int \mathbf{1}\{r^*(P) \geq 0\} d\mathcal{H}(P) < 1 - F(s^*(0)),$$

where the equality is by Lemma 11, the first inequality is by the Portmanteau theorem, and the second inequality is by $\mathcal{H} \notin \mathcal{R}(0)$ and Lemma 10. *Q.E.D.*

Lemma 13 shows that, in the limit, all districts are equally strong when $r^*(F) \geq 0$.

LEMMA 13: *Let $r^*(F) \geq 0$ and $\mathcal{H}_n \to \mathcal{H}$ as $\eta_n \to \infty$. Then, for each $P \in \operatorname{supp}(\mathcal{H})$, we have $r^*(P) = r^*(F)$.*

PROOF OF LEMMA 13: If $r^*(F) = 0$, then, for each $P \in \operatorname{supp}(\mathcal{H})$, we have $r^*(P) \geq 0$ by Lemma 10, so $r^*(P) = 0$ by $\int P d\mathcal{H}(P) = F$. So suppose that $r^*(F) > 0$. Moreover, suppose by contradiction that there exists $\varepsilon \in (0, r^*(F))$, $\delta \in (0,1)$, and a sequence $\eta_n \to \infty$ such that $\int \mathbf{1}\{r^*(P) \leq r^*(F) - \varepsilon\} d\mathcal{H}_n(P) \geq \delta$ for all n. We obtain a contradiction for sufficiently large n, because

$$\int Q(\eta_n r^*(P)) d\mathcal{H}_n(P) \leq \delta Q(\eta_n (r^*(F) - \varepsilon)) + (1 - \delta) < Q(\eta_n r^*(F)),$$

where the first inequality is by the supposition and the second inequality is by

$$\frac{1 - Q(\eta r^*(F))}{1 - Q(\eta (r^*(F) - \varepsilon))} \to 0, \quad \text{as } \eta \to \infty, \tag{17}$$

which we now establish.

Denote $c = q'(r^*(F)-\varepsilon)/q(r^*(F)-\varepsilon)$. Since $q'(0)=0$ and q is strictly log-concave, for all $\eta > 1$, we have

$$0 = \frac{q'(0)}{q(0)} > c = \frac{q'(r^*(F)-\varepsilon)}{q(r^*(F)-\varepsilon)} > \frac{q'(\eta(r^*(F)-\varepsilon))}{q(\eta(r^*(F)-\varepsilon))} > \frac{q'(x)}{q(x)}, \quad \text{for all } x > \eta(r^*(F)-\varepsilon).$$

Hence Gronwall's inequality gives $\lim_{\eta\to\infty} q(\eta r^*(F))/q(\eta(r^*(F)-\varepsilon)) \le \lim_{\eta\to\infty} e^{c\varepsilon\eta} = 0$, so, by L'Hopital's rule, we have

$$\lim_{\eta\to\infty} \frac{1-Q(\eta r^*(F))}{1-Q(\eta(r^*(F)-\varepsilon))} = \lim_{\eta\to\infty} \frac{q(\eta r^*(F))r^*(F)}{q(\eta(r^*(F)-\varepsilon))(r^*(F)-\varepsilon)} = 0,$$

establishing (17). *Q.E.D.*

Lemma 14 shows that, in the limit, types below $s^*(0)$ are segregated and types above $s^*(0)$ are paired in a negatively assortative manner.

LEMMA 14: *Let $\mathcal{H}_n \to \mathcal{H}$ as $\eta_n \to \infty$.*
1. *For any $P \in \operatorname{supp}(\mathcal{H})$ with $r^*(P) \le s^*(0)$, we have $|\operatorname{supp}(P)| = 1$.*
2. *For any $P, P' \in \operatorname{supp}(\mathcal{H})$ with $r^*(P) = r^*(P') \ge 0$, we have $\operatorname{supp}(P) = \{s_1(P), s_2(P)\}$ and $\operatorname{supp}(P') = \{s_1(P'), s_2(P')\}$ with $s_1(P) \le s_2(P)$, $s_1(P') \le s_2(P')$, and $(s_2(P') - s_2(P))(s_1(P) - s_1(P')) \ge 0$.*

PROOF OF LEMMA 14: Denote $\Lambda_n = \operatorname{supp}(\mathcal{H}_n)$. Since the set of compact subsets of a compact set is compact (in the Hausdorff topology), taking a subsequence if necessary, Λ_n converges to some compact set Λ. By Box 1.13 in Santambrogio (2015), we have $\operatorname{supp}(\mathcal{H}) \subset \Lambda$. Since $\mathcal{H}_n$ is strictly single-dipped by Lemma 2, we have $|\operatorname{supp}(P_n)| \le 2$ for all $P_n \in \Lambda_n$, and thus $|\operatorname{supp}(P)| \le 2$ for all $P \in \Lambda$.

Suppose part 2 fails. Then, by Lemmas 10, 12, and 13, there exist $P, P' \in \operatorname{supp}(\mathcal{H})$ such that $s_1(P') < s_1(P) < r_+^*(F) < s_2(P') < s_2(P)$. But then since $\Lambda_n \to \Lambda$, there exist n and $P_n, P_n' \in \Lambda_n$ such that $\operatorname{supp}(P_n) = \{s_1(P_n), s_2(P_n)\}$, $\operatorname{supp}(P_n') = \{s_1(P_n'), s_2(P_n')\}$, and $s_1(P_n') < s_1(P_n) < r_+^*(F) < s_2(P_n') < s_2(P_n)$, contradicting that $\mathcal{H}_n$ is strictly single-dipped.

Suppose part 1 fails. Then, by Lemmas 10 and 12, there exists $P \in \operatorname{supp}(\mathcal{H})$ such that $\operatorname{supp}(P) = \{s, s'\}$ with $\underline{s} \le s < s' \le s^*(0)$. Moreover, by Lemmas 10 and 12 and part 2, there exists $P' \in \operatorname{supp}(\mathcal{H})$ such that $\operatorname{supp}(P') = \{s^*(0), \overline{s}\}$. But then since $\Lambda_n \to \Lambda$, there exist n and $P_n, P_n' \in \Lambda_n$ with $(s_1(P_n), s_2(P_n), s_1(P_n'), s_2(P_n'))$ close to $(s, s', s^*(0), \overline{s})$. Then, by Lemma 5, $\mathcal{H}_n$ cannot be segregate-pair, contradicting Theorem 1. *Q.E.D.*

To complete the proof of Theorem 3, note that Lemmas 10 (for $r = 0$), 12, 13, and 14 show that if a sequence of optimal plans $\mathcal{H}_\eta$ converges to $\mathcal{H}$, then $\mathcal{H}$ must segregate types below $s^*(0)$ and pair types above $s^*(0)$ in a negatively assortative manner in equally strong districts. The unique such plan is $\mathcal{H} = \mathcal{H}^*$. Finally, since every convergent sequence $\mathcal{H}_n$ converges to $\mathcal{H}^*$, compactness of $\Delta\Delta[\underline{s}, \overline{s}]$ implies that $\mathcal{H}_\eta$ also converges to $\mathcal{H}^*$.

A.7. *Proof of Theorem 4*

Theorem 4 follows from Lemmas 15–18.

Let $\mathcal{T}$ be the set of optimal plans $\mathcal{H}$ when each voter's idiosyncratic shock is sure to be 0 (no idiosyncratic uncertainty). Lemma 15 characterizes $\mathcal{T}$: it shows that $\mathcal{H} \in \mathcal{T}$ iff each district

$P \in \text{supp}(H)$ contains 50% voters with some type $s^P \geq s^m$ and 50% voters with types $s \leq s^m$ (so the designer wins district P iff $r \leq s^P$).

For $P \in \Delta[\underline{s}, \overline{s}]$, define $\overline{P}(r) = \int \mathbf{1}\{s \geq r\} dP(s) = 1 - P(r_-)$ for all r. The designer wins district P iff the aggregate shock r satisfies $r \leq r_0^*(P) = \{\max \tilde{r} : \overline{P}(\tilde{r}) \geq 1/2\}$. For $\mathcal{H} \in \Delta\Delta[\underline{s}, \overline{s}]$, define $\overline{H}(r) = \int \mathbf{1}\{r_0^*(P) \geq r\} d\mathcal{H}(P)$ for all r.

LEMMA 15: *$\mathcal{H} \in \mathcal{T}$ iff, for each $P \in \text{supp}(\mathcal{H})$, there exists $s^P \geq s^m$ such that $P(s) = 1$ for all $s \geq s^P$, $P(s) = 1/2$ for all $s \in [s^m, s^P)$, and $P(s) \leq 1/2$ for all $s < s^m$.*

PROOF OF LEMMA 15: For each $r \geq s^m$, we have

$$\begin{aligned}\overline{F}(r) &= \int \overline{P}(r) d\mathcal{H}(P) = \int \mathbf{1}\{\overline{P}(r) \geq \tfrac{1}{2}\} \overline{P}(r) d\mathcal{H}(P) + \int \mathbf{1}\{\overline{P}(r) < \tfrac{1}{2}\} \overline{P}(r) d\mathcal{H}(P) \\ &\geq \int \mathbf{1}\{\overline{P}(r) \geq \tfrac{1}{2}\} \tfrac{1}{2} d\mathcal{H}(P) = \int \mathbf{1}\{r_0^*(P) \geq r\} \tfrac{1}{2} d\mathcal{H}(P) = \tfrac{1}{2}\overline{H}(r).\end{aligned} \tag{18}$$

So, any feasible $\mathcal{H}$ satisfies $\overline{H}(r) \leq \overline{H}^*(r)$ for all r, where

$$\overline{H}^*(r) = \begin{cases} 1, & \text{if } r \leq s^m, \\ 2\overline{F}(r), & \text{if } r > s^m. \end{cases}$$

Thus, the designer's expected seat share for any feasible plan is $\int \overline{H}(r) dG(r) \leq \int \overline{H}^*(r) dG(r)$, with strict inequality if $\overline{H}(r) < \overline{H}^*(r)$ for some r (and thus on some interval (r', r) with $r' < r$, by continuity of $\overline{H}^*$ and monotonicity and left-continuity of $\overline{H}$), because $G(r)$ is strictly increasing in r. Hence, a districting plan $\mathcal{H}$ is optimal iff it induces $\overline{H} = \overline{H}^*$. In turn, $\overline{H} = \overline{H}^*$ iff, for each $r \geq s^m$, the inequality in (18) holds with equality, or equivalently, $\int \mathbf{1}\{\overline{P}(r) = 1/2\} d\mathcal{H}(P) = 2\overline{F}(r)$ and $\int \mathbf{1}\{\overline{P}(r) = 0\} d\mathcal{H}(P) = 1 - 2\overline{F}(r)$. Finally, this holds for all $r \geq s^m$ iff, for each $P \in \text{supp}(\mathcal{H})$, there exists $s^P \geq s^m$ such that $\overline{P}(s) = 0$ for all $s > s^P$, $\overline{P}(s) = 1/2$ for all $s \in (s^m, s^P]$, and $\overline{P}(s) \geq 1/2$ for all $s \leq s^m$. *Q.E.D.*

Lemma 16 characterizes the optimal seat share as idiosyncratic uncertainty vanishes.

LEMMA 16: *If $\eta \to 0$, then $V_\eta \to 2\int_{s^m}^{\overline{s}} G(r) dF(r)$.*

PROOF OF LEMMA 16: Let $\mathcal{H}_\eta$ be the optimal plan and let $\mathcal{H}_r$ be any districting plan in $\mathcal{R}_\eta(r)$. We have

$$\begin{aligned} V_\eta &= \iint \mathbf{1}\{r \leq r_\eta^*(P)\} d\mathcal{H}_\eta(P) dG(r) \\ &\leq \iint \mathbf{1}\{r \leq r_\eta^*(P)\} d\mathcal{H}_r(P) dG(r) = \int (1 - F(s_\eta^*(r))) dG(r) = \overline{V}_\eta, \end{aligned}$$

where the inequality holds because $\int \mathbf{1}\{r \leq r_\eta^*(P)\} d\mathcal{H}_\eta(P) \leq \int \mathbf{1}\{r \leq r_\eta^*(P)\} dH_r(P)$ for all r by the definition of $\mathcal{H}_r$.

Let $\mathcal{H}_q^*$, with $q \in (0, 1/2)$, be NAD with a q-$1-q$ split in each district. Formally, $\mathcal{H}_q^*$ is the unique plan $\mathcal{H}$ such that, for any $P \in \text{supp}(\mathcal{H})$, we have either (a) $\text{supp}(P) = \{s^q\}$ with

$s^q = F^{-1}(q)$ or (b) $\text{supp}(P) = \{s_1(P), s_2(P)\}$ such that $\underline{s} \leq s_1(P) < s^q < s_2(P) \leq \overline{s}$, and $(1-q)F(s_1(P)) = q(1-F(s_2(P)))$. We have

$$V_\eta = \int G(r^*_\eta(P))d\mathcal{H}_\eta(P) \geq \int G(r^*_\eta(P))d\mathcal{H}^*_q(P) = \underline{V}_\eta(q),$$

where the inequality holds because $\mathcal{H}_\eta$ is optimal and $\mathcal{H}^*_q$ is feasible.

Suppose now that $\eta \to 0$, which implies that $Q_\eta \to \delta_0$. For each r, $1 - F(s^*_\eta(r)) \to \overline{H}^*(r)$, so, by the dominated convergence theorem and integration by parts, $\overline{V}_\eta \to \int \overline{H}^*(r)dG(r) = 2\int_{s^m}^{\overline{s}} G(r)dF(r)$. For $q < 1/2$ and $s < s'$, $r^*_\eta(q\delta_s + (1-q)\delta_{s'}) \to s'$, so, by the dominated convergence theorem, $\underline{V}_\eta(q) \to \int_{s^q}^{\overline{s}} G(r)dF(r)/(1-q)$, which converges to $2\int_{s^m}^{\overline{s}} G(r)dF(r)$ as $q \uparrow 1/2$. *Q.E.D.*

Lemma 17 shows that limit points of optimal plans $H_n = H_{\eta_n}$, for $\eta_n \to 0$, belong to $\mathcal{T}$.

LEMMA 17: *Let $\mathcal{H}_n \to \mathcal{H}$ as $\eta_n \to \infty$. Then $\mathcal{H} \in \mathcal{T}$.*

PROOF OF LEMMA 17: Suppose by contradiction that there exists a sequence $\eta_n \to 0$ such that $\mathcal{H}_n$ converges weakly to $\mathcal{H} \notin \mathcal{T}$. Then we have

$$2\int_{s^m}^{\overline{s}} G(r)dF(r) = \lim_{n\to\infty} \int G(r^*_{\eta_n}(P))d\mathcal{H}_n(P) \leq \int \overline{H}(r)dG(r) < 2\int_{s^m}^{\overline{s}} G(r)dF(r),$$

where the equality is by Lemma 16, the first inequality is by the Portmanteau theorem and integration by parts, and the second inequality is by $\mathcal{H} \notin \mathcal{T}$ and Lemma 15. *Q.E.D.*

Lemma 18 shows that, in the limit, all types are paired in a negatively assortative manner.

LEMMA 18: *Let $\mathcal{H}_n \to \mathcal{H}$ as $\eta_n \to 0$. For any $P, P' \in \text{supp}(\mathcal{H})$, we have $\text{supp}(P) = \{s_1(P), s_2(P)\}$ and $\text{supp}(P') = \{s_1(P'), s_2(P')\}$ with $s_1(P) \leq s_2(P)$, $s_1(P') \leq s_2(P')$, and $(s_2(P') - s_2(P))(s_1(P) - s_1(P')) \geq 0$.*

PROOF OF LEMMA 18: Denoting $\Lambda_n = \text{supp}(\mathcal{H}_n)$, the same argument as in the proof of Lemma 14 implies that there exists Λ such that, up to a subsequence, $\Lambda_n \to \Lambda$, $\text{supp}(\mathcal{H}) \subset \Lambda$, and $|\text{supp}(P)| \leq 2$ for all $P \in \Lambda$.

By Lemmas 15 and 17, if the conclusion of the lemma fails, there must exist $P, P' \in \text{supp}(\mathcal{H})$ such that $\text{supp}(P) = \{s_1(P), s_2(P)\}$ and $\text{supp}(P') = \{s_1(P'), s_2(P')\}$ with $s_1(P') < s_1(P) < s^m < s_2(P') < s_2(P)$. Then, since $\Lambda_n \to \Lambda$, there exist n and $P_n, P'_n \in \Lambda_n$ such that $\text{supp}(P_n) = \{s_1(P_n), s_2(P_n)\}$, $\text{supp}(P'_n) = \{s_1(P'_n), s_2(P'_n)\}$, and $s_1(P'_n) < s_1(P_n) < s^m < s_2(P'_n) < s_2(P_n)$, contradicting that $\mathcal{H}_n$ is strictly single-dipped. *Q.E.D.*

To complete the proof of Theorem 4, note that Lemmas 15, 17, and 18 show that if a sequence of optimal plans $\mathcal{H}_\eta$ converges to $\mathcal{H}$, then $\mathcal{H}$ must pair all types in a negatively assortative manner, with 50% mass on the higher type. Clearly, the unique such plan is $\mathcal{H} = \mathcal{H}^{**}$. Since every convergent sequence $\mathcal{H}_n$ converges to $\mathcal{H}^{**}$, compactness of $\Delta\Delta[\underline{s}, \overline{s}]$ implies that $\mathcal{H}_\eta$ also converges to $\mathcal{H}^{**}$.

A.8. *Proof of Theorem 5*

The proof derives three necessary conditions for optimal Y-districting to involve a bifurcation point at r and shows that these conditions imply that r must equal 0 and γ must lie in the specified range. The first condition (equation (19)) says that it is optimal to pair voter types just below and just above r. The second condition (equation (20)) says that it is optimal to segregate types just below r. The third condition (equation (21)) says that the proportions of favorable and unfavorable voters in each district P with $r^*(P) = r'$ just above r actually induce the desired cutoff r'.

Formally, by Theorem 1 of Kolotilin et al. (2025), the function λ in Lemma 1 can be taken to have a derivative $\lambda'(r)$ at each $r \in (r^b, r^b + \varepsilon]$ satisfying

$$g(r) - \lambda(r)q(s_2(r) - r) + \lambda'(r)\left(Q(s_2(r) - r) - Q(0)\right) = 0,$$
$$g(r) - \lambda(r)q(s_1(r) - r) + \lambda'(r)\left(Q(s_1(r) - r) - Q(0)\right) = 0.$$

Solving for $\lambda(r)$ and $\lambda'(r)$ yields, for all $r \in (r^b, r^b + \varepsilon]$,

$$\lambda(r) = \frac{g(r)[Q(s_2(r) - r) - Q(s_1(r) - r)]}{\left(Q(s_2(r) - r) - Q(0)\right)q(s_1(r) - r) - \left(Q(s_1(r) - r) - Q(0)\right)q(s_2(r) - r)},$$
$$\lambda'(r) = \frac{g(r)[q(s_2(r) - r) - q(s_1(r) - r)]}{\left(Q(s_2(r) - r) - Q(0)\right)q(s_1(r) - r) - \left(Q(s_1(r) - r) - Q(0)\right)q(s_2(r) - r)}.$$

Since λ' is the derivative of λ, we have $d\lambda(r)/dr = \lambda'(r)$ for all $r \in (r^b, r^b + \varepsilon]$. Since s_1 and s_2 are twice differentiable and satisfy $\lim_{r\downarrow r^b} s_1(r) = \lim_{r\downarrow r^b} s_2(r) = r^b$, we can apply L'Hopital's rule to evaluate $d\lambda(r)/dr = \lambda'(r)$ in the limit $r \downarrow r^b$ to obtain

$$\frac{g'(r^b)q(0)}{(q(0))^2} = \frac{g(r^b)q'(0)}{(q(0))^2},$$

which implies that $r^b = 0$, because $G(r) = Q(\eta r)$ for all r and $q'(r) = 0$ iff $r = 0$. Denote $\lim_{r\downarrow r^b} s_1'(r) = 1 - \beta_1$ and $\lim_{r\downarrow r^b} s_2'(r) = 1 + \beta_2$, where $\beta_1 \geq 1$ (because s_1 is decreasing) and $\beta_2 \geq 0$ (because $s_2(r) > r$). Differentiating $d\lambda(r)/dr = \lambda'(r)$ with respect to r and taking the limit $r \downarrow 0$, we get

$$\frac{\eta q''(0)(\eta^2 - \beta_2\beta_1)}{q(0)} = \frac{\eta q''(0)(\beta_2 - \beta_1)}{2q(0)},$$

and hence

$$2\eta^2 = 2\beta_2\beta_1 + \beta_2 - \beta_1. \tag{19}$$

Since, for small enough $r > 0$, type $s_1(r)$ is assigned to both district $\delta_{s_1(r)}$ and district P with $r^*(P) = r$ and $\text{supp}(P) = \{s_1(r), s_2(r)\}$, we must have, by Lemma 1,

$$Q(\eta s_1(r)) = Q(\eta r) + \lambda(r)\left(Q(s_1(r) - r) - Q(0)\right).$$

In the limit $r \downarrow 0$, the values and the derivatives up to order 2 of both sides always coincide, while the third derivatives coincide iff

$$q''(0)\eta^3(-\beta_1 + 1)^3 = q''(0)\eta^3 - 3q''(0)\eta^3\beta_1 + 3q''(0)\eta\beta_2\beta_1^2 - q''(0)\eta\beta_1^3,$$

which simplifies to

$$-\eta^2\beta_1 + 3\eta^2 = 3\beta_2 - \beta_1. \tag{20}$$

Since, for small enough $r > 0$, type $s_1(r)$ is assigned to both district $\delta_{s_1(r)}$ and district P with $r^*(P) = r$, while type $s_2(r)$ is assigned only to district P, we have

$$f(s_1(r))s_1'(r)\left(Q(s_1(r)-r) - Q(0)\right) \geq f(s_2(r))s_2'(r)\left(Q(s_2(r)-r) - Q(0)\right).$$

In the limit $r \downarrow 0$, both sides are equal, and hence their derivatives must satisfy

$$-f(0)q(0)\beta_1(1-\beta_1) \geq f(0)q(0)\beta_2(\beta_2+1),$$

which, given that $\beta_1 + \beta_2 > 0$, simplifies to

$$\beta_1 \geq \beta_2 + 1. \tag{21}$$

Recalling that $\gamma = \eta^2/(1+\eta^2)$, equations (19) and (20) have two solutions

$$(\beta_1, \beta_2) = \left(\tfrac{3\eta^2}{(2(\eta^2-1))}, \tfrac{\eta^2}{2}\right) = \left(\tfrac{3\gamma}{2(2\gamma-1)}, \tfrac{\gamma}{2(1-\gamma)}\right) \quad \text{and} \quad (\beta_1', \beta_2') = \left(1, \tfrac{(2\eta^2+1)}{3}\right) = \left(1, \tfrac{\gamma+1}{3(1-\gamma)}\right),$$

unless $\gamma = 1/2$, in which case (19) and (20) have only one solution $(\beta_1, \beta_2) = (1, 1)$. The solution (β_1', β_2') never satisfies (21) and thus is discarded. Moreover, for the solution (β_1, β_2), condition $\beta_1 \geq 1$ yields $\gamma > 1/2$, and condition (21) yields $\gamma \leq \sqrt{3} - 1$. Thus, for Y-districting to be optimal, we must have $\gamma \in (1/2, \sqrt{3} - 1]$.

APPENDIX B: ESTIMATORS

In this section, we formally define our estimators and show that they satisfy standard statistical properties. Fix a US state. We assume that there is a large number of voters, so that the vote share in a precinct n with type s_n in district d and election y with aggregate shock r_{dy} is given by $v_{ny} = \Phi((s_n - r_{dy})/\sqrt{\gamma})$. Let μ_s and σ_s^2 be the mean and variance of the distribution of precinct types, defined by $\mu_s = \mathbb{E}_F[s]$ and $\sigma_s^2 = Var_F[s]$. For convenience, we repeat some definitions from the text. Let $w_{ny} = \Phi^{-1}(v_{ny})$, T the number of elections, D the number of districts, and $\mathcal{N}_{dy}$ the set of precincts in district d and election y. Define

$$w_{dy} = \frac{\sum_{n\in\mathcal{N}_{dy}} k_{ny}w_{ny}}{\sum_{n\in\mathcal{N}_{dy}} k_{ny}}, \quad w_{d\bullet} = \frac{\sum_y w_{dy}}{T}, \quad w_{\bullet y} = \frac{\sum_d w_{dy}}{D}, \quad w_{\bullet\bullet} = \frac{\sum_{d,y} w_{dy}}{DT},$$

$$e_n^2 = \frac{1}{DT}\sum_{d,y}\frac{\sum_{n\in\mathcal{N}_{dy}} k_{ny}(w_{ny} - w_{\bullet y})^2}{\sum_{n\in\mathcal{N}_{dy}} k_{ny}},$$

$$e_d^2 = \frac{\sum_{d,y}(w_{dy} - w_{d\bullet})^2}{D(T-1)}, \quad e^2 = \frac{\sum_y (w_{\bullet y} - w_{\bullet\bullet})^2}{T-1},$$

$$cov = \frac{\sum_{y,d,d'>d}(w_{dy} - w_{d\bullet})(w_{d'y} - w_{d'\bullet})}{\frac{D(D-1)}{2}(T-1)} = \frac{De^2 - e_d^2}{D-1},$$

where the last equality follows from

$$e^2 = \frac{\sum_y \left(\sum_d \frac{1}{D}(w_{dy} - w_{d\bullet})\right)^2}{(T-1)}$$

$$= \frac{1}{D}\frac{\sum_{d,y}(w_{dy} - w_{d\bullet})^2}{D(T-1)} + \frac{D-1}{D}\frac{\sum_{y,d,d'>d}(w_{dy} - w_{d\bullet})(w_{d'y} - w_{d'\bullet})}{\frac{D(D-1)}{2}(T-1)}$$

$$= \frac{1}{D}e_d^2 + \frac{D-1}{D}cov.$$

To construct our estimators, we use the following proposition.

PROPOSITION 1: *In our empirical model,*

$$\mathbb{E}e_d^2 = \frac{1-\gamma}{\gamma}, \quad \mathbb{E}cov = \rho\frac{1-\gamma}{\gamma}, \quad \mathbb{E}w_{\bullet\bullet} = \frac{\mu_s}{\sqrt{\gamma}}, \quad \textit{and} \quad \mathbb{E}e_n^2 = \frac{\sigma_s^2}{\gamma} + (1-\rho)\frac{D-1}{D}\frac{1-\gamma}{\gamma},$$

and

$$e_d^2 \overset{d}{=} \frac{1-\gamma}{D(T-1)\gamma}\left[(1-\rho)\chi^2_{(D-1)(T-1)} + (1+(D-1)\rho)\chi^2_{T-1}\right],$$

where $\overset{d}{=}$ *denotes equality in distribution, and* $\chi^2_{(D-1)(T-1)}$ *and* χ^2_{T-1} *denote independent* χ^2 *random variables with* $(D-1)(T-1)$ *and* $T-1$ *degrees of freedom, respectively.*

Consider the following point estimators of γ, ρ, μ_s, and σ_s:

$$\widehat{\gamma} = \frac{1}{1+e_d^2}, \quad \widehat{\rho} = \frac{cov}{e_d^2}, \quad \widehat{\mu}_s = \frac{w_{\bullet\bullet}}{\sqrt{1+e_d^2}}, \quad \text{and} \quad \widehat{\sigma}_s = \sqrt{\frac{e_n^2 - \frac{D-1}{D}(e_d^2 - cov)}{1+e_d^2}}.$$

By Proposition 1, $1/\widehat{\gamma}$, $\widehat{\rho}/\widehat{\gamma} - \widehat{\rho}$, $\widehat{\mu}_s/\sqrt{\widehat{\gamma}}$, and $\widehat{\sigma}_s^2/\widehat{\gamma}$ are unbiased estimators of $1/\gamma$, $\rho/\gamma - \rho$, $\mu_s/\sqrt{\gamma}$, and σ_s^2/γ. Moreover, by the law of large numbers for $D(T-1) \to \infty$, we have that $\widehat{\gamma}$, $\widehat{\rho}$, $\widehat{\mu}_s$, and $\widehat{\sigma}_s$ are consistent estimators of γ, ρ, μ_s, and σ_s.

Proposition 1 also gives a confidence interval for γ. Specifically, for any $\alpha \in (0,1)$, let q_α be the α-quantile for $(1-\widehat{\rho})\chi^2_{(D-1)(T-1)} + (1+(D-1)\widehat{\rho})\chi^2_{T-1}$. Then, a one-sided $1-\alpha$ confidence interval for γ is $(\widehat{\gamma}_\alpha, 1)$ where

$$\widehat{\gamma}_\alpha = \frac{1}{1+\frac{D(T-1)}{q(\alpha)}e_d^2}.$$

PROOF OF PROPOSITION 1: Denote

$$r_{d\bullet} = \frac{\sum_y r_{dy}}{T}, \quad r_{\bullet y} = \frac{\sum_d r_{dy}}{D}, \quad s_{dy} = \frac{\sum_{n\in\mathcal{N}_{dy}} k_{ny}s_n}{\sum_{n\in\mathcal{N}_{dy}} k_{ny}}, \quad s_{\bullet y} = \frac{\sum_d s_{dy}}{D}.$$

First, we have

$$\mathbb{E}w_{\bullet\bullet} = \mathbb{E}\frac{1}{DT}\sum_{d,y}\frac{\sum_{n\in N_{dy}} k_{ny}(s_n - r_{dy})}{\sqrt{\gamma}\sum_{n\in N_{dy}} k_{ny}} = \mathbb{E}\frac{\sum_{n\in N_{dy}} k_{ny}s_n}{\sqrt{\gamma}\sum_{n\in N_{dy}} k_{ny}} = \frac{\mu_s}{\sqrt{\gamma}},$$

where the first equality is by $v_{ny} = \Phi((s_n - r_{dy})/\sqrt{\gamma})$ and the definition of v_{ny} and $w_{\bullet\bullet}$, the second is by $\mathbb{E}[r_{dy}] = 0$ and district equipopulation, and the fourth is by the definition of μ_s. Second, we have

$$\mathbb{E}e_d^2 = \mathbb{E}\frac{\sum_{d,y}\left(\frac{T-1}{T}r_{dy} - \frac{1}{T}\sum_{y'\neq y} r_{dy'}\right)^2}{D(T-1)\gamma} = \frac{DT\left[\left(\frac{T-1}{T}\right)^2 + \frac{T-1}{T^2}\right](1-\gamma)}{D(T-1)\gamma} = \frac{1-\gamma}{\gamma},$$

where the first equality is by $v_{ny} = \Phi((s_n - r_{dy})/\sqrt{\gamma})$, the definition of w_{dy} and $w_{d\bullet}$, and rearrangement, the second is by $Var[r_{dy}] = 1-\gamma$ and $Cov[r_{dy}, r_{dy'}] = 0$ for $y \neq y'$, and the third is by rearrangement. Third, we have

$$\mathbb{E}cov = \mathbb{E}\frac{\sum_{y,d,d'>d}(r_{dy} - r_{d\bullet})(r_{d'y} - r_{d'\bullet})}{\frac{D(D-1)}{2}(T-1)\gamma} = \rho\frac{1-\gamma}{\gamma},$$

where the first equality is again by $v_{ny} = \Phi((s_n - r_{dy})/\sqrt{\gamma})$ and the definition of w_{dy} and $w_{d\bullet}$, and the second is by $Cov[r_{dy}, r_{d'y}] = \rho(1-\gamma)$ for $d \neq d'$, $Cov[r_{dy}, r_{d'y'}] = 0$ for $y \neq y'$, and rearrangement. Fourth, we have

$$\begin{aligned}
\mathbb{E}e_n^2 &= \mathbb{E}\frac{1}{DT}\sum_{d,y}\frac{\sum_{n\in\mathcal{N}_{dy}} k_{ny}(s_n - s_{\bullet y} + r_{dy} - r_{\bullet y})^2}{\gamma\sum_{n\in\mathcal{N}_{dy}} k_{ny}} = \mathbb{E}\frac{\sum_{n\in\mathcal{N}_{dy}} k_{ny}(s_n - s_{\bullet y})^2}{\gamma\sum_{n\in\mathcal{N}_{dy}} k_{ny}} \\
&\quad + \mathbb{E}\frac{\sum_d (r_{dy} - r_{\bullet y})^2}{\gamma D} = \frac{\sigma_s^2}{\gamma} + \mathbb{E}\frac{\sum_d\left(\frac{D-1}{D}r_{dy} - \frac{1}{D}\sum_{d'\neq d} r_{d'y}\right)^2}{\gamma D} \\
&= \frac{\sigma_s^2}{\gamma} + \mathbb{E}\frac{\sum_d\left[\left(\frac{D-1}{D}\right)^2 r_{dy}^2 + \frac{1}{D^2}\sum_{d'\neq d} r_{d'y}^2 - \frac{2(D-1)}{D^2}r_{dy}r_{d'y} + \frac{2}{D^2}\sum_{d'\neq d, d''>d'} r_{d'y}r_{d''y}\right]}{\gamma D} \\
&= \frac{\sigma_s^2}{\gamma} + \left[\left(\frac{D-1}{D}\right)^2 + \frac{D-1}{D^2} - \rho\frac{2(D-1)}{D^2} + \rho\frac{(D-1)(D-2)}{D^2}\right]\frac{1-\gamma}{\gamma} \\
&= \sigma_s^2 + (1-\rho)\frac{D-1}{D}\frac{1-\gamma}{\gamma},
\end{aligned}$$

where the first equality is by $v_{ny} = \Phi((s_n - r_{dy})/\sqrt{\gamma})$, the definition of w_{ny} and $w_{\bullet y}$, and rearrangement, the second is by independence across elections and district equipopulation, the third is by the large number of voters and rearrangement of the second term, the fourth is by quadratic expansion, the fifth is by $\mathbb{E}[r_{dy}^2] = 1-\gamma$ and $\mathbb{E}[r_{dy}r_{d'y}] = \rho(1-\gamma)$ for $d' \neq d$, and the sixth is by rearrangement.

Finally, let $r = (r_{11}, \ldots, r_{1T}, \ldots, r_{D1}, \ldots, r_{DT})'$. Then we can write

$$\sum_{d,y} (r_{dy} - r_{d\bullet})^2 = r' A r$$

where

$$A = \begin{pmatrix} \frac{T-1}{T} & \ldots & -\frac{1}{T} & \ldots & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ -\frac{1}{T} & \ldots & \frac{T-1}{T} & \ldots & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & \ldots & \frac{T-1}{T} & \ldots & -\frac{1}{T} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & \ldots & -\frac{1}{T} & \ldots & \frac{T-1}{T} \end{pmatrix}.$$

Note that

$$\frac{\mathbb{E}[rr']}{1-\gamma} = \Sigma = \begin{pmatrix} 1 \ldots 0 \ldots \rho \ldots 0 \\ \vdots \ddots \vdots \ddots \vdots \ddots \vdots \\ 0 \ldots 1 \ldots 0 \ldots \rho \\ \vdots \ddots \vdots \ddots \vdots \ddots \vdots \\ \rho \ldots 0 \ldots 1 \ldots 0 \\ \vdots \ddots \vdots \ddots \vdots \ddots \vdots \\ 0 \ldots \rho \ldots 0 \ldots 1 \end{pmatrix}.$$

By the spectral theorem, there is an orthogonal matrix P (so that $PP' = P'P = I$) and a diagonal matrix Λ with positive diagonal elements $\lambda_1, \ldots, \lambda_{DT}$ such that $\Sigma^{1/2} A \Sigma^{1/2} = P' \Lambda P$. Define $u = P\Sigma^{-1/2} r / \sqrt{1-\gamma}$ (so that $r = \Sigma^{1/2} P' u \sqrt{1-\gamma}$). Then

$$\frac{r'Ar}{1-\gamma} = u' P \Sigma^{1/2} A \Sigma^{1/2} P' u = u' P P' \Lambda P P' u = u' \Lambda u = \sum_{i=1}^{DT} \lambda_i u_i^2$$

where $u \sim N(0, I)$, and $\lambda_1, \ldots \lambda_{DT}$ are the roots of the characteristic equation

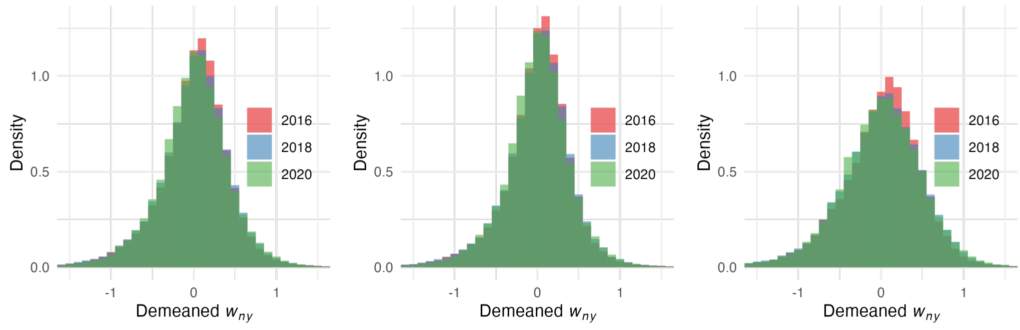
$$|\Sigma^{1/2} A \Sigma^{1/2} - \lambda I| = 0 \iff |A\Sigma - \lambda I| = 0.$$

Note that

$$A\Sigma = \begin{pmatrix} \frac{T-1}{T} & \ldots & -\frac{1}{T} & \ldots & \rho\frac{T-1}{T} & \ldots & -\rho\frac{1}{T} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ -\frac{1}{T} & \ldots & \frac{T-1}{T} & \ldots & -\rho\frac{1}{T} & \ldots & \rho\frac{T-1}{T} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \rho\frac{T-1}{T} & \ldots & -\rho\frac{1}{T} & \ldots & \frac{T-1}{T} & \ldots & -\frac{1}{T} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ -\rho\frac{1}{T} & \ldots & \rho\frac{T-1}{T} & \ldots & -\frac{1}{T} & \ldots & \frac{T-1}{T} \end{pmatrix}.$$

After some algebra, we obtain

$$|A\Sigma - \lambda I| = (-1)^{DT} \lambda^D (\lambda - 1 + \rho)^{(D-1)(T-1)} (\lambda - 1 - (D-1)\rho)^{T-1},$$



(a) Normal (b) Logistic (c) Uniform

FIGURE C.1.—Histograms of demeaned w_{ny}

showing that $r'Ar/(1-\gamma) \overset{d}{=} (1-\rho)\chi^2_{(D-1)(T-1)} + (1+(D-1)\rho)\chi^2_{T-1}$, and hence

$$e_d^2 = \frac{r'Ar}{D(T-1)\gamma} \overset{d}{=} \frac{1-\gamma}{D(T-1)\gamma}\left[(1-\rho)\chi^2_{(D-1)(T-1)} + (1+(D-1)\rho)\chi^2_{T-1}\right]. \qquad \textit{Q.E.D.}$$

APPENDIX C: ROBUSTNESS AND GOODNESS OF FIT

We analyze the robustness and goodness of fit of our empirical model. We focus on our key conclusion that idiosyncratic uncertainty is much larger than aggregate uncertainty.

Within the context of our maintained model with one-dimensional precinct types and one-dimensional aggregate shocks, our estimate of γ and the model fit do not appear sensitive to distributional assumptions. Table C.I reproduces Table I assuming that G and Q are logistic or uniform rather than normal. The estimates are similar, indicating that our estimate of γ is not sensitive to distributional assumptions. As noted in footnote 43, this is because γ is close to 1, so Q is approximately uniform over the relevant range, regardless of the assumed overall shape of Q.

Table C.II and Figures C.1 and C.2 show that the model's goodness of fit is similar for the normal, logistic, or uniform distribution. To see this, note that the model can fit any distribution of (normalized) vote shares w_{ny} in a single election y by varying the precinct type distribution F, so the model's fit can be measured by the extent to which the distribution of demeaned w_{ny} is constant across elections. Table C.II and Figures C.1 and C.2 show that the variation in this distribution across elections is similar for the normal, logistic, or uniform distribution.[58] Overall, the model appears to fit the data relatively well.

Our conclusion that idiosyncratic uncertainty is much larger than aggregate uncertainty can also be seen in a model-free manner. This is illustrated in Figure C.3. The first two panels show that idiosyncratic uncertainty is much larger than aggregate uncertainty in a model-free sense. Figure C.3(a) shows the probability density of voters in the United States who live in a precinct with Republican vote share v, with bin breaks $\{0, 0.05, \dots, 0.95, 1\}$, averaging over elections $y \in \{2016, 2018, 2020\}$ and over all districts included in the analysis. The figure shows that the distribution of v_{ny} is unimodal, with a large majority (72%) of the mass on $v \in [0.25, 0.75]$.

[58]The rows in Table C.II show the standard deviation, 25% quantile, median, 75% quantile, interquartile range, skewness, kurtosis, and the Kolmogorov-Smirnov distance for the other two elections (e.g., the 2016 column shows the KS distance between the 2018 and 2020 distributions).

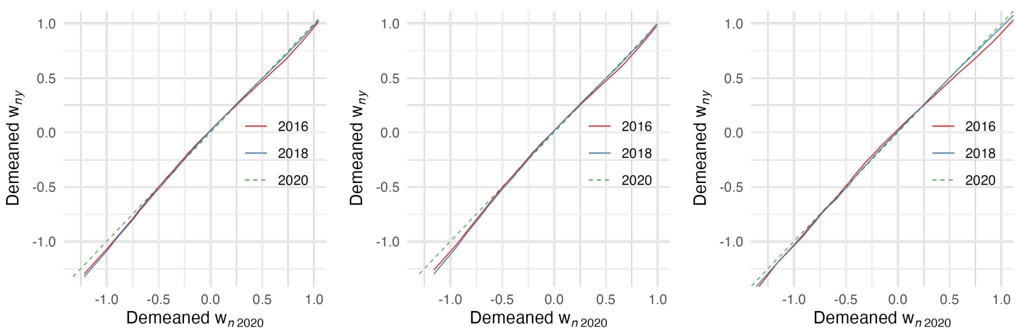
	Logistic					Uniform				
US	γ	σ_s	V	$\underline{V}$	$\mathcal{H}$	γ	σ_s	V	$\underline{V}$	$\mathcal{H}$
HI	0.976	0.170	0.004	0.003	Seg	0.966	0.202	0.000	0.000	Seg
NY	0.960	0.822	0.421	0.420	SOP	0.976	0.819	0.421	0.421	SOP
MD	0.990	0.704	0.460	0.460	SOP	0.991	0.773	0.481	0.481	SOP
RI	0.992	0.283	0.221	0.216	SOP	0.984	0.346	0.146	0.146	SOP
CT	0.996	0.358	0.357	0.357	SOP	0.992	0.417	0.240	0.240	SOP
ME	0.994	0.284	0.262	0.259	SOP	0.988	0.381	0.205	0.205	SOP
MA	0.998	0.215	0.164	0.161	SOP	0.996	0.278	0.090	0.090	SOP
DE	0.992	0.465	0.480	0.480	SOP	0.984	0.545	0.399	0.399	SOP
IL	0.986	0.726	0.576	0.576	SOP	0.977	0.750	0.532	0.533	SOP
NJ	0.983	0.573	0.512	0.512	SOP	0.973	0.629	0.439	0.439	SOP
CA	0.993	0.443	0.517	0.517	SOP	0.987	0.588	0.508	0.508	SOP
NM	0.982	0.500	0.559	0.558	SOP	0.970	0.652	0.541	0.541	SOP
NH	0.998	0.235	0.584	0.584	SOP	0.995	0.342	0.580	0.580	SOP
NV	0.998	0.408	0.741	0.741	SOP	0.996	0.560	0.771	0.771	SOP
MN	0.989	0.398	0.639	0.639	SOP	0.978	0.534	0.613	0.613	SOP
CO	0.991	0.482	0.698	0.698	SOP	0.983	0.643	0.707	0.707	SOP
OR	0.989	0.456	0.670	0.670	SOP	0.981	0.608	0.678	0.678	SOP
VA	0.987	0.503	0.735	0.734	SOP	0.976	0.662	0.737	0.737	SOP
WA	0.990	0.337	0.702	0.702	SOP	0.978	0.482	0.680	0.680	SOP
MI	0.992	0.577	0.814	0.814	SOP	0.984	0.647	0.810	0.810	SOP
GA	0.988	0.683	0.831	0.829	SOP	0.977	0.793	0.831	0.831	SOP
TX	0.991	0.599	0.845	0.844	SOP	0.984	0.759	0.871	0.871	SOP
IA	0.989	0.336	0.775	0.775	SOP	0.977	0.472	0.763	0.764	SOP
NC	0.967	0.521	0.756	0.755	SOP	0.945	0.653	0.719	0.719	SOP
AZ	0.992	0.363	0.877	0.877	SOP	0.983	0.510	0.891	0.891	SOP
FL	0.995	0.407	0.950	0.950	NAD	0.991	0.545	0.990	0.990	NAD
OH	0.987	0.595	0.915	0.914	NAD	0.977	0.729	0.930	0.934	SOP
AK	0.997	0.356	0.983	0.983	NAD	0.992	0.508	1.000	1.000	NAD
MT	0.994	0.448	0.971	0.971	NAD	0.990	0.603	1.000	1.000	NAD
SC	0.995	0.584	0.987	0.987	NAD	0.991	0.715	1.000	1.000	NAD
AR	0.987	0.583	0.953	0.952	NAD	0.979	0.741	0.980	0.980	SOP
NE	0.992	0.407	0.977	0.977	NAD	0.985	0.549	1.000	1.000	NAD
KY	0.992	0.514	0.991	0.991	NAD	0.986	0.634	1.000	1.000	NAD
MO	0.995	0.666	0.999	0.999	NAD	0.993	0.782	1.000	1.000	NAD
KS	0.982	0.425	0.967	0.967	NAD	0.967	0.560	1.000	1.000	NAD
IN	0.986	0.487	0.987	0.987	NAD	0.975	0.614	1.000	1.000	NAD
WI	0.991	0.271	0.995	0.995	NAD	0.983	0.384	1.000	1.000	NAD
AL	0.973	0.636	0.982	0.982	NAD	0.970	0.756	1.000	1.000	NAD
WV	0.975	0.310	0.983	0.983	NAD	0.958	0.418	1.000	1.000	NAD
UT	0.990	0.547	0.999	0.999	NAD	0.987	0.678	1.000	1.000	NAD
TN	0.992	0.657	1.000	1.000	NAD	0.989	0.758	1.000	1.000	NAD
MS	0.993	0.632	1.000	1.000	NAD	0.993	0.752	1.000	1.000	NAD
ID	0.989	0.432	1.000	1.000	NAD	0.984	0.534	1.000	1.000	NAD
OK	0.986	0.421	0.999	0.999	NAD	0.977	0.534	1.000	1.000	NAD
ND	0.974	0.400	0.997	0.997	NAD	0.958	0.481	1.000	1.000	NAD
WY	0.991	0.449	1.000	1.000	NAD	0.989	0.542	1.000	1.000	NAD
LA	0.974	0.570	0.999	0.999	NAD	0.977	0.640	1.000	1.000	NAD
WS	0.987	0.528	0.760	0.759	SOP	0.981	0.641	0.774	0.776	SOP
US	0.987	0.608	0.781	0.780	SOP	0.981	0.728	0.801	0.801	SOP

TABLE C.I: Estimates

This indicates that idiosyncratic uncertainty is much larger than aggregate uncertainty, because if $\gamma = 1$ the figure would show a unimodal density with mode around $v = 1/2$, while if $\gamma \to 0$

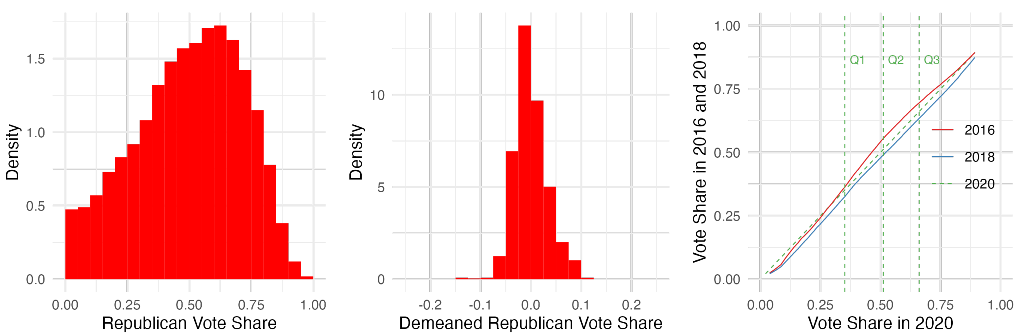
Statistic	Normal			Logistic			Uniform		
	2016	2018	2020	2016	2018	2020	2016	2018	2020
SD	0.433	0.446	0.432	0.411	0.424	0.404	0.490	0.505	0.502
Q1	-0.218	-0.233	-0.239	-0.197	-0.208	-0.217	-0.266	-0.295	-0.302
Median	0.038	0.033	0.019	0.034	0.033	0.018	0.045	0.030	0.022
Q3	0.259	0.270	0.259	0.238	0.248	0.236	0.312	0.328	0.322
IQR	0.477	0.503	0.499	0.435	0.456	0.453	0.578	0.623	0.624
Skewness	-0.567	-0.491	-0.278	-0.608	-0.566	-0.309	-0.543	-0.385	-0.249
Kurtosis	5.053	4.805	4.522	5.708	5.518	4.947	4.304	3.891	3.880
KS	0.019	0.026	0.013	0.022	0.025	0.012	0.012	0.030	0.019

TABLE C.II: Statistics



(a) Normal (b) Logistic (c) Uniform

FIGURE C.2.—Q-Q plots of demeaned w_{ny}



(a) Precinct Vote Shares (b) District Vote Swings (c) Q-Q Plot of Vote Shares

FIGURE C.3.—Distributions of Precinct Vote Shares and District Vote Swings

it would show a bimodal distribution with all mass at 0 and 1, and the former case is a much better approximation.[59]

[59]We have also reproduced Figure C.3(a) at the state level. The distribution appears to be well-approximated by a unimodal distribution for all states except Illinois and New York.

Similarly, Figure C.3(b) shows the probability density of *(district, election)* pairs where the district-wide Republican vote share deviated from its mean over the three elections we consider by x, with bin breaks $\{-0.25, -0.225, \ldots, 0.225, 0.25\}$.[60] This histogram gives another way of showing that aggregate shocks are small: the distribution is centrally unimodal, and most of the mass (59%) is on $x \in [-0.025, 0.025]$. In contrast, if aggregate shocks were very large, we would again have a bimodal distribution with all mass far from 0.

Finally, Figure C.3(c) shows the empirical distribution of vote shares v_{ny} across precincts n (weighted by the number of votes in each precinct and averaged over all districts), for each election y, normalized by the empirical vote-share distribution in 2020. Thus, the green curve is the $45°$ line; the red curve is the 2016 Republican vote share for a precinct with a given 2020 Republican vote share; and the blue curve is the analogous curve for 2018.[61] The ordering of the curves (except for the lowest-vote-share precincts, discussed below) reflects the fact that, among the 2016, 2018, and 2020 elections, 2018 was the best year for Democrats, 2016 was the best year for Republicans, and 2020 was in the middle. We can use these curves to assess the realism of our assumption that q is log-concave (so moderate voters are swingier than extremists). Under log-concavity of q, the red curve should be concave and everywhere above the green curve, and the blue curve should be convex and everywhere below the green curve. Figure C.3(c) shows that this is not exactly true in our data, because the red and blue curves are "too low" for the left-most districts (a small minority of districts, lying well into the lowest quartile of the vote-share distribution, as indicated in the figure). This small deviation from log-concavity likely reflects an unusually strong performance by Republicans in urban districts in 2020, due to a well-documented shift in the minority vote toward Republicans (e.g., Igielnik et al. 2021, Kolko and Monkovic 2021). Such demographic-specific shocks are outside our model, but could be explored in future work. Overall, we believe Figure C.3(c) is well-explained by a combination of our assumptions and an unexpected shift toward Republicans in urban areas in 2020.

REFERENCES

KOLOTILIN, ANTON, ROBERTO CORRAO, AND ALEXANDER WOLITZKY (2025): Persuasion and Matching: Optimal Productive Transport, *Journal of Political Economy*, 133 (4), 1334–1381. [4, 7, 8, 11, 15, 1, 2]

SANTAMBROGIO, FILIPPO (2015): *Optimal Transport for Applied Mathematicians*, vol. 55, Springer. [8]

IGIELNIK, RUTH, SCOTT KEETER, AND HANNAH HARTIG (2021): Behind Biden's 2020 Victory, Pew Research Center. [19]

KOLKO, JED AND TONI MONKOVIC (2021): The Places that had the Biggest Swings Toward and Against Trump, New York Times, https://www.nytimes.com/2020/12/07/upshot/trump-election-vote-shift. [19]

[60]This histogram is compiled at the district level because precincts are not matched across elections.

[61]More precisely, since we cannot match precincts across elections, the red curve is the 2016 Republican vote share for a precinct *at the same quantile of the vote share distribution* as a precinct with a given 2020 Republic vote share, and similarly for the green curve.