The Impact of Redistricting Commission Design on Partisan Fairness Performance

Daniel Luce

April 3, 2024

Senior Honors Seminar 2023-2024

Faculty Advisor: Shigeo Hirano

Seminar Advisor: John Huber

Graduate Instructor: Martin Devaux

Thesis submitted to the Faculty of the Department of Political Science of Columbia University in partial fulfillment of the requirements for the undergraduate honors program

Abstract

This study evaluates the performance of independent redistricting commissions (IRCs) on various measures of partisan fairness using the Sequential Monte Carlo simulation method. It also evaluates the impact of institutional design features on partisan fairness by clustering IRCs into two categories based on these features. The findings suggest that IRCs as a whole outperform legislative redistricting on partisan-aware symmetric fairness metrics like the mean-median difference but are less consistent on partisan-aware proportionality metrics and partisan-blind metrics. However, IRC performance does appear to vary based on design criteria. “Type II” IRCs with balanced partisan composition, more unaffiliated commissioners, explicit considerations of partisanship, and non-majoritarian approval mechanisms consistently outperformed legislative redistricting and "Type I" IRCs that lacked such safeguards on partisan-aware mean-median fairness and potentially the partisan-aware efficiency gap. However, Type I IRCs outperformed legislative redistricting and Type II IRCs on the partisan-blind mean-median metric.

keywords: redistricting, partisan fairness, institutional design, redistricting commissions

Introduction

In 2011 and 2012, Colleen Mathis, the chair of the Arizona Independent Redistricting Commission, received numerous death threats and was nearly expelled from her position by the governor. As the commission's only member who was unaffiliated with the Democratic or Republican parties, Mathis had effectively been operating as the swing vote on the five-person, volunteer commission that had been entrusted with the typically hyperpartisan responsibility of redistricting. Each decision she made that aligned her with the commission's Democrats prompted "intense organized opposition from tea party-inspired residents and Republican lawmakers" due to perceptions that she was biased and secretly colluding to disempower Republicans (Vasilogambros 2019). In 2021, the commission's new chair was similarly accused of harboring secret partisan intentions, only those accusations came primarily from Democrats (Daley 2021). While many of the attacks received by Arizona commission chairs were unfounded and excessive, Best et al. (2021) have suggested that the state may have produced maps that are suspect from a partisan fairness perspective.

The Arizona experience is not universal, however. Other states with independent redistricting commissions (IRCs) like Michigan and Colorado have not experienced the same level of vitriol, although they have received plenty of criticism as well. These states have also been largely praised for their successful efforts to increase partisan fairness and increase competitiveness (Li 2021). Therefore, while Arizona, Michigan, and Colorado each utilize an independent commission for their redistricting process, these discrepancies in public response and redistricting outcomes spark critical concerns about the efficacy of these institutions and their ability to reduce or eliminate partisan gerrymandering. In the existing literature, these concerns about the efficacy of IRCs in achieving partisan fairness have not been thoroughly resolved, despite strong theoretical arguments in favor of IRCs. Studies by Nelson (2023) and Royden & Li (2017) tend to support the notion that IRCs produce more fair maps than legislative redistricting processes, particularly concerning metrics such as competitiveness and measures of partisan fairness. However, conflicting research, such as that of Henderson, Hamil, & Goldzimer (2018) and Miller & Groffman (2015), suggests that IRCs may not significantly differ from legislative redistricting in terms of encouraging competition. Meanwhile, as alluded to earlier, Best et al. (2021) reached mixed conclusions, finding that most IRCs outperform legislative redistricting on measures of partisan fairness, but not Arizona's.

In this paper, I posit that some of the discrepancies within the literature regarding the relative partisan fairness of IRCs compared to legislative redistricting may be partially attributable to two key factors: 1) differing conceptions of what constitutes partisan fairness, and 2) the tendency to treat all IRCs as a unified category despite meaningful variation in their institutional designs. Therefore, I first delineate between four distinct operational conceptions of partisan fairness that have been implicitly employed in the redistricting literature to date: partisan blind symmetry, partisan blind proportionality, partisan-aware symmetry, and partisan-aware proportionality. I then apply the work of scholars like Cain (2012) and Zhang (2021) to identify key institutional design differences between IRCs that could plausibly impact their relative ability to optimize for different partisan fairness goals, thus complicating the treatment of IRCs as a unified category of redistricting institution.

Subsequently, I compare the partisan fairness of enacted redistricting plans from both the 2010 and 2020 redistricting cycles (including congressional and state senate maps) against a series of partisan fairness baselines generated using the novel redistricting simulation algorithm introduced by McCartan & Imai (2023). The performance of redistricting plans from states utilizing IRCs is then systematically compared against the average performance of states still employing a legislative redistricting process along each conception of partisan fairness.

The findings suggest that when considered collectively, IRCs do tend to outperform legislative redistricting processes on metrics associated with partisan-aware symmetric conceptions of fairness, such as the mean-median difference. Moreover, analysis examining changes in states' fairness metrics before and after transitioning from a legislative redistricting process to an IRC also provisionally points towards general improvements in most partisan-aware fairness scores. However, the analysis also indicates that one particular cluster of IRCs exhibiting certain shared institutional design features – such as balanced partisan composition requirements for commissioner selection and non-majoritarian approval mechanisms – may correspond to even stronger adherence to both partisan-aware symmetric fairness baselines as well as some partisan-aware proportional fairness baselines like the efficiency gap, although the evidence is not conclusive. Meanwhile, the other cluster of IRCs appears to outperform legislative redistricting on partisan-blind symmetry metrics. Importantly, each of the fairness advantages of an IRC cluster also appears to potentially extend beyond just outperforming legislative redistricting plans, but also outperforming the other IRC cluster that lacks those particular institutional characteristics, although with limited statistical significance.

These findings may have significant implications given the growing momentum behind IRC reforms, as evidenced by upcoming ballot measures in Nevada and Ohio ("Redistricting Measures on the Ballot"). While IRCs generally appear better than partisan legislative processes, particularly for partisan-aware symmetric fairness, the particularities of their institutional design also appear to matter. Commissions without safeguards like balanced membership and consensus rules may struggle to consistently achieve partisan-aware fairness. Conversely, those with stringent partisanship considerations may not be able to achieve partisan blind fairness. As such, advocates and policymakers should carefully consider design factors, not just redistricting independence, when pursuing redistricting reforms aimed at improving partisan fairness.

1 Partisan Fairness

Normatively, redistricting aims to facilitate meaningful and effective voter participation by ensuring representation along several fundamental dimensions: party affiliation, geographic location, and potentially race or language group (Morrill 1987). The significance of partisanship in redistricting's effects on electoral outcomes is especially relevant, as voters' policy preferences are primarily reflected through their votes for candidates from specific political parties (Nagle 2019). This is especially true as parties become increasingly ideologically sorted and distinct, a tendency that has been well-established in the American political landscape over recent decades (Mason 2015; Levendusky 2009). Consequently, when redistricting plans skew in a partisan direction, the ideological midpoint of the legislature is also skewed, fundamentally altering how legislators represent their constituents in a way that threatens policy congruence between voters and their legislature (Stephanopoulos 2016; 2018; Kim & Urpelainen 2017). As such, partisan fairness has become a central focus of redistricting debates. However, there is no standard definition of partisan fairness and existing definitions tend to be constructed as what it is not: partisan gerrymandering nor a disproportionate advantage given to any political party via redistricting (Duchin & Schoenbach 2023).

1.1 Proportionality-Symmetry Fairness

Beyond these definitions by negation, though, many operational conceptions of partisan fairness explicitly or implicitly encompass two distinct dimensions: proportionality-symmetry, and partisan blind-partisan aware. The proportionality-symmetry dimension of fairness, which has been a focal point in existing redistricting literature, broadly concerns the balance between the distribution of seats in a legislative body and the distribution of votes among political parties. The proportional conception of partisan fairness relies on the assumption that proportionality in a legislature is desirable and therefore the proportion of votes for a particular political party should approximately correspond to the proportion of seats that party receives in the legislative body (Deford et al. 2021). However, "proportionality is not the neutral tendency of redistricting" and has not been held by the Supreme Court as an implication of the Equal Protection Clause (Deford et al. 2021; Cain et al. 2018). Therefore, the methodology for measuring more reasonable understandings of partisan fairness either related to or distinct from this goal of approximate proportionality remains a subject of contention among scholars (Deford et al. 2021). To this end, various alternative metrics have been proposed, although the efficiency gap, a transformation of the simple proportionality metric that gives a bonus to the winning party, is the most widely used and accepted (Stephanopoulos & McGhee 2015; Duchin & Schoenbach 2023).

At the other end of this spectrum is the symmetry-based conception of fairness, which assumes that a redistricting plan is fair between opposing parties when the number of seats won by one party at a specified vote share equals the number of seats the opposite party would win at that same vote share (Katz et al. 2020). These symmetry metrics rely on a counterfactual hypothetical election, often approximated using a uniform and linear swing of the realized election result, which may not reflect the complexities of real-world political dynamics (Deford et al. 2021). Metrics associated with symmetry include the widely-used partisan bias score, which is a measure of the discrepancy between the seat shares of two parties at 50% of the vote share (Katz et al. 2020).

To demonstrate the practical distinctions between these two conceptions of partisan fairness, consider an election in which Party A wins 51% of the vote share and all of the seats in the legislature. In this scenario, fairness as proportionality is strongly violated, Party A's share of seats far exceeds its share of the popular vote and likely indicates a significant disparity between electoral outcomes and voter preferences. However, if Party B also were to win all of the seats when they receive 51% of the vote share in a hypothetical election, fairness as symmetry would not be violated, as the parties receive an equal number of seats at that vote share.

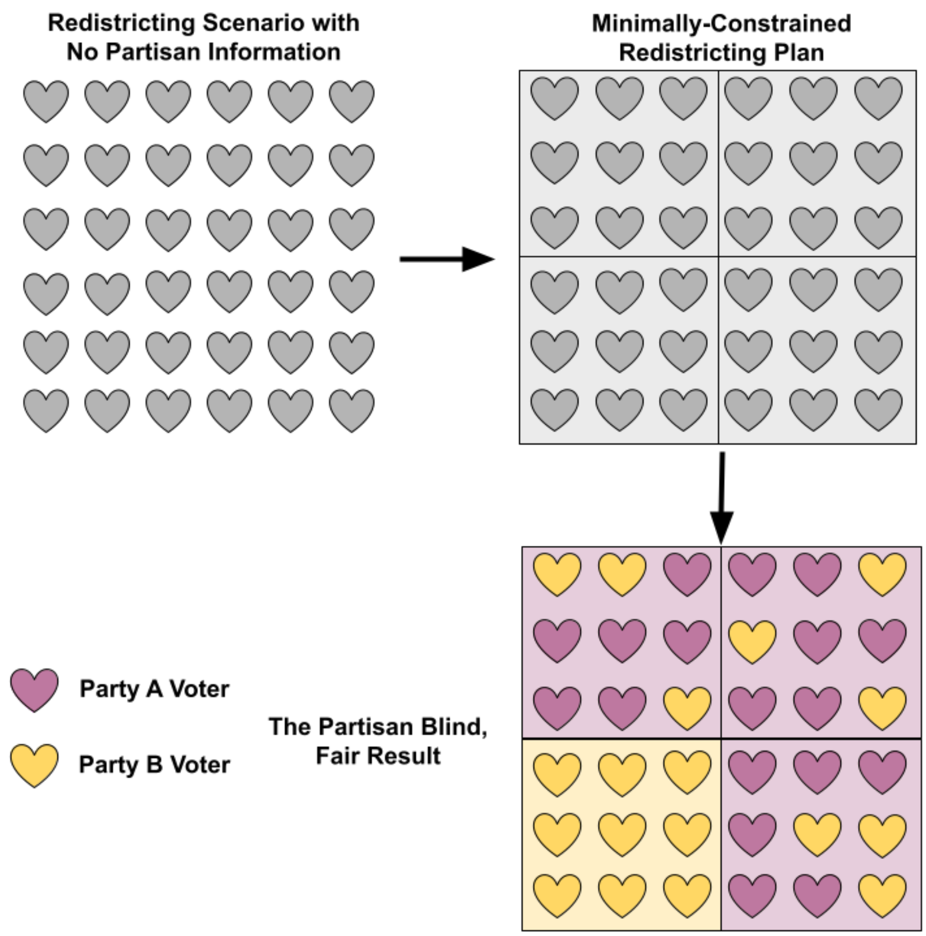
1.2 Partisan Blind-Partisan Aware Fairness

However, a much less discussed dimension of partisan fairness that is of significant importance is what might be referred to as the *partisan aware-partisan blind* dimension, using terminology borrowed from Deford, Duchin, & Solomon (2020). *Partisan blind* conceptions of fairness emphasize non-discrimination against any particular political party by restricting or prohibiting the consideration of partisanship and its associated political geography in redistricting, thereby mitigating claims of partisan gerrymandering, which require a demonstration of an intent to dilute the power of one political party (McDonald et al. 2018). This “process-oriented” conception of fairness considers any resulting partisan outcome and associated partisan fairness metrics to be incidental to the redistricting process (Pildes 2018). Partisan blind fairness is therefore achieved when measures of partisan symmetry or proportionality are within the range of what would be reasonably expected if redistricting were minimally constrained or adhered only to other redistricting criteria. However, what can be *reasonably expected* is itself a vague baseline for comparison. That said, both the political science literature and the courts have accepted that algorithmic ensemble methods that simulate thousands or tens of thousands of possible redistricting plans can be used to generate a distribution of what is reasonably expected for a given partisan fairness metric, although the strength of the constraints provided to these algorithms can alter this distribution (Cain et al. 2018).

Conversely, *partisan-aware* conceptions of fairness are "outcome-oriented" and prioritize the achievement of proportional or symmetric political representation by explicitly considering partisanship (Pildes 2018; Deford, Duchin, & Solomon 2020). This may or may not require the intentional violation of traditional redistricting criteria like compactness and the maintenance of political subdivisions to "correct" for political geographic realities that might confound the achievement of a given partisan fairness metric (Nagle 2019; Chen & Rodden 2013). However, even the most aggressive efforts to correct for a state's political geography may still preclude enforcement of perfect proportionality or symmetry. For example, in Massachusetts, where Republican candidates consistently receive a significant portion of the vote share but remain unrepresented in the House of Representatives, Duchin et al. (2019) have demonstrated that the distribution of votes across the state has structurally disadvantaged Republicans on partisan fairness metrics, regardless of redistricting. Additionally, this conception of fairness may not be robust to claims of partisan gerrymandering as an intent to dilute the power of a political party, as the optimization for a partisan fairness metric may require the structurally advantaged party to be penalized for the distribution of their supporters via cracking or packing.

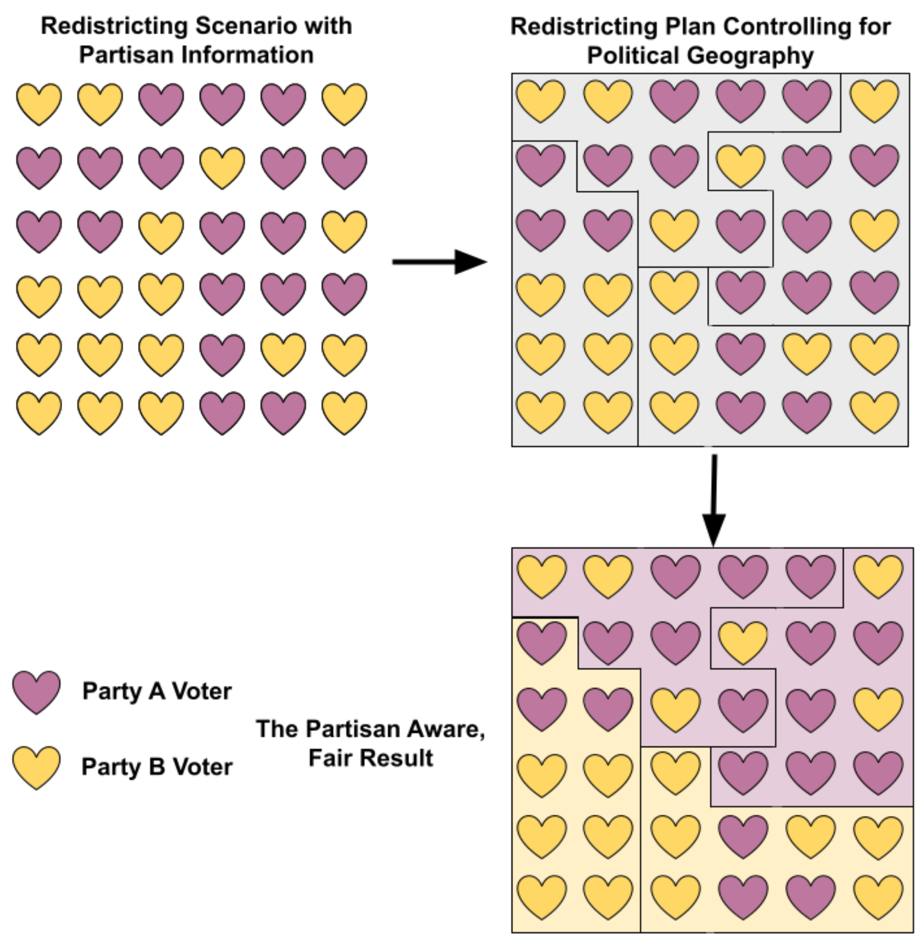
To illustrate the difference between partisan-aware and partisan blind approaches to fairness, consider a state in which Party A and Party B each receive approximately half of the vote, but Party B's supporters are particularly concentrated in one geographic area. A partisan blind approach to fairness would ignore this distribution of partisanship and instead focus on other redistricting factors. As a result, Party B might win an overwhelming majority in the seat where its voters are concentrated while Party A would win the rest of the seats, and more overall, due to its broader geographic spread. However, because this result is what is expected under partisan blind, minimally constrained redistricting, the redistricting map would be considered fair (see Figure 1 for a visual representation of this partisan blind approach).

Figure 1: Partisan Blind Redistricting



In contrast, a partisan-aware approach would recognize the concentration of Party B's supporters and seek to distribute them across multiple districts to achieve either symmetry or proportionality across the entire state. This might involve intentionally dividing Party B's supporters or "packing" Party A's supporters into fewer districts to dilute their influence (see Figure 2 for a visual representation of this partisan-aware approach).

Figure 2: Partisan-Aware Redistricting



1.3 The Four Conceptions of Partisan Fairness

When the fairness dimensions of proportionality-symmetry and partisan blind-partisan aware are considered together, four distinct conceptions of partisan fairness emerge:

Table 1: Four Conceptions of Partisan Fairness

Conception of Partisan Fairness	Description	Example Studies
Partisan Blind Symmetry	The discrepancy in the number of seats won at a given vote share between parties should equal the discrepancy that would be expected if redistricting were minimally constrained or adhered only to other redistricting criteria	Best et al. (2021); McDonald & Best (2015)
Partisan Blind Proportionality	The difference between the proportion of *votes* won by a given political party and the proportion of *seats* won by that party should equal the difference that would be expected if redistricting were minimally constrained or adhered only to other redistricting criteria	Chen & Cottrell (2016)
Partisan Aware Symmetry	The discrepancy in the number of seats won at a given vote share between parties should be minimized as much as practicable through additional constraints that control for a state's political geography	Keena et al. (2019); Nagle (2019)
Partisan Aware Proportionality	The difference between the proportion of *votes* won by a given political party and the proportion of *seats* won by that party should be minimized as much as practicable through additional constraints that control for a state's political geography	Royden & Li (2017); Duchin & Schoenbach (2023)

These distinct conceptions of partisan fairness carry significant theoretical and practical implications for how redistricting and its effects on representation should be evaluated. Partisan blind measures are useful for determining if a redistricting plan exhibits partisan bias relative to what would be expected under minimal constraints. However, they provide little insight into whether the resulting redistricting plan achieves a normative ideal of fair partisan representation. Conversely, partisan-aware measures directly assess how well a redistricting plan achieves specific visions of fair representation but may conflict with other democratic principles like minority representation and adherence to traditional redistricting criteria.

Furthermore, the symmetry and proportionality conceptions reflect fundamentally different notions of what fair partisan representation should entail. An over-reliance on any one conception could lead to problematic conclusions about the fairness or unfairness of a redistricting plan. When engaging with the existing literature on redistricting and partisan fairness, it is, therefore, crucial to precisely identify which conception(s) of fairness are being operationalized, as combining or comparing findings that use different fairness conceptions could be misleading. For example, studies by Best et al. (2021) and Chen and Cottrell (2016) emphasize fairness in a partisan blind sense, while works like Royden and Li (2017) use a more partisan-aware approach focused on achieving specific representational ideals. Conflating partisan blind and partisan aware approaches, or symmetry and proportionality metrics, risks muddying substantive insights into how well various redistricting institutions and processes actually perform in facilitating fair political representation across their multiple possible definitions.

2 Design Features of Redistricting Institutions that Impact Partisan Fairness

The existing literature has identified several key dimensions along which the institutional design of redistricting processes can vary, with significant implications for their ability to achieve different conceptions of partisan fairness. Broadly, these design factors can be categorized into two interrelated dimensions: who draws the district lines, and what constrains them in doing so. Within the "who draws the lines" dimension, previous studies have highlighted the importance of the separation of the redistricting authority from the legislative branch, as well as the level of enactment authority granted to that redistricting body. However, for redistricting institutions structured as independent commissions, there are additional design features regarding the commission's composition that could affect partisan fairness, such as the appointment mechanism, requirements for partisan parity, and the involvement of politically unaffiliated members. Similarly, the "what constrains them" dimension has focused primarily on the constraints imposed upon legislative redistricting bodies by political realities like divided government and upon all redistricting institutions by traditional redistricting criteria, like compactness, the maintenance of political subdivisions, and the preservation of communities of interest. However, I also propose that independent commissions exhibit meaningful variation in the degree to which partisan fairness considerations are permitted to influence the redistricting process and the approval mechanisms required to enact final maps. Each of these design features carries important theoretical implications for how redistricting institutions may perform on the various conceptions of partisan fairness.

2.1 Who Draws the Lines

For most of American history, state legislatures were exclusively responsible for redistricting (Engstrom 2013). However, legislative control over redistricting poses challenges to achieving partisan fairness, as legislators may be incentivized to manipulate district boundaries for partisan advantage, undermining democratic principles of impartial representation (Barkow 2010; Kubin 1996). Therefore, to mitigate these conflicts of interest and promote partisan fairness, several states have shifted some or all redistricting authority away from the legislature by establishing redistricting commissions. However, it has been widely acknowledged that these redistricting commissions vary widely in their level of enactment authority and their degree of separation from the legislature. As such, the political science literature generally acknowledges four types of redistricting commissions: advisory, backup, bipartisan, and independent (Cain 2012; Levitt & Spencer 2024).

2.1.1 Enactment Authority and Separation from the Legislature

Independent redistricting commissions (IRCs) distinguish themselves from the other redistricting commissions in both their enactment authority and separation from the legislature. Unlike advisory and backup commissions, IRCs have primary enactment authority in redistricting and are therefore not subject to the same degree of potential legislative interference (Cain 2012). Meanwhile, bipartisan commissions do share the same primary enactment authority as IRCs, but the presence of sitting politicians on these commissions reintroduces the potential for partisan conflicts of interest during the redistricting process. Therefore, IRCs are the only commission structure that ostensibly removes legislators from the redistricting process entirely, resolving the conflict of interest concern (Cain 2012).

This institutional approach to redistricting is quite similar to other "politically neutral" institutions, such as independent bureaucratic agencies and appointed judiciaries, in that they are relatively autonomous from majoritarian democratic control and seek to promote unbiased decision-making (Iancu and Tănăsescu 2018; Iancu 2018; Cain 2012). Analyses of these comparable institutions support the theory that increased separation from electoral politics corresponds with greater institutional fairness. For example, multiple reviews of the evidence and individual studies regarding judicial impartiality largely find that appointed judges act less partisanly than their elected counterparts (Robertson 2018; Tabarrok & Helland 1999; Hanssen 1999). Similarly, bureaucratic agencies in the federal government have also consistently been demonstrated to act more impartially when they are more politically insulated, such as when their leadership consists of fewer political appointees and more civil service members (Lewis 2008; Gersen & Berry 2010). However, beyond the basic degree of separation from the legislature, additional differences in the "who draws the lines" dimension of the institutional design of IRCs may impact their relative ability to achieve partisan fairness. These relevant differences include the commissioner selection mechanism, whether or not the IRC has explicit partisan parity, and whether or not the commission has multiple unaffiliated commissioners.

2.1.2 Variation in the Selection Mechanism of IRCs

While the separation of redistricting authority from the legislature is a defining feature of IRCs, Cain (2012) highlights additional nuances in the degree of separation based on how commissioners are selected. Some IRCs utilize political appointment methods in which commissioners are directly chosen by state legislative leaders or other elected officials. Alternatively, some IRCs employ random selection methods in which commissioners are selected randomly from applicant pools, aiming to maximize independence from the legislature. This distinction in selection mechanisms has important implications for partisan fairness. Random selection is expected to reduce the influence of partisan biases by further insulating commissioners from legislative interests. In contrast, the political appointment of commissioners maintains closer ties to the legislature, which could undermine partisan fairness in the same vein as bipartisan commissions.

Montana, Washington, Idaho, Alaska, and New York each employ political appointment methods, with commissioners directly appointed by state legislative leaders or other elected officials. In all of these states except Alaska, the majority and minority members of each legislative chamber select a commissioner, or in the case of New York, two commissioners. In Montana, Washington, and New York, these appointed commissioners subsequently vote to select additional commissioners. In Idaho, the two major state party leaders also select commissioners. Alaska has the most distinct appointment process, with the majority leaders of each legislative chamber selecting commissioners, the governor selecting two commissioners, and the chief justice of the Alaska Supreme Court selecting the final commissioner.

Conversely, California, Colorado, and Michigan utilize random selection methods, where commissioners are chosen through a process that involves random drawing from applicant pools. In California and Michigan, state legislative leaders remain involved in the process insofar as they can strike a prescribed number of randomly selected applicants from the pool of final candidates. Meanwhile, in Colorado, the first half of the commissioners are chosen entirely by lot. The second half is narrowed from a pool of randomly drawn applicants by state legislative leaders, with the final selection conducted by a panel of retired judges. Finally, Arizona utilizes a hybrid regime in which the Commission on Appellate Court Appointments nominates a pool of candidates and from which each state legislative leader selects a member to serve on the IRC. For the state constitutional language defining each of these selection processes, see Appendix A, Table 1.

2.1.3 Variations in Partisan Parity and Composition among IRCs

Partisan composition within redistricting commissions is also likely to play a critical role in ensuring partisan fairness in the redistricting process. There are two existing designs for achieving the partisan composition of an IRC: allocation by office type and explicit party balance. The former design, found only in Alaska, allocates commission seats based on the political offices held by individuals, potentially resulting in an unbalanced commission. As Cain (2012) notes, this method may be a rough approximation of a state's existing party balance and can lead to one party being overrepresented, undermining the partisan fairness of the process. All other IRCs use a system with explicit party balance mandating equal representation from both major parties in their selection processes, ostensibly preventing partisan dominance and associated biases on these commissions (Cain 2012).

However, the IRCs with partisan parity vary in the presence of politically unaffiliated members, which can further influence the partisan fairness of the redistricting process. The IRCs in Idaho and Washington lack voting unaffiliated members. This absence may lead to deadlock within the commission, potentially undermining its intended insulation and fairness. Without unaffiliated members to mediate or break ties, the commission may struggle to reach a consensus, risking delays or external intervention. As Kubin (1996) notes, while perfect partisan parity aims for fairness, it may hinder negotiation and compromise, potentially compromising the effectiveness of these institutions. Having unaffiliated members, who can in some cases serve as tiebreakers, can prevent commissions from deadlocking and may encourage the partisan members to work in good faith, lest they push the tiebreaker to side with the opposing partisans (Kubin 1996). However, the success of this unaffiliated commissioner feature relies on the actual independence and neutrality of those commissioners (Kubin 1996).

Commissions with multiple unaffiliated members, such as those in California, New York, Michigan, and Colorado should be more likely to produce partisanly fair maps, as any potential individual partisan biases among unaffiliated members should approximate true neutrality as the number of those commissioners increases (Zhang 2021). In contrast, states with only one voting unaffiliated member, like Arizona and Montana, may face challenges in achieving partisan fairness, as the individual bias of the unaffiliated commissioner remains unchecked, potentially causing unintentional partisan imbalance within the commission. As the example of Colleen Mathis demonstrated perfectly, the Arizona IRC has regularly faced controversies regarding the neutrality of its single unaffiliated member, highlighting the vulnerability of citizen commissions to attacks on their political independence due to their composition (Zhang 2021). For the state constitutional language defining the partisan composition of each IRC, see Appendix A, Table 2.

2.2 What Constrains Them

All redistricting institutions, whether legislative or commission-based, are subject to external constraints and federal criteria that shape how they draw district lines. These criteria include contiguity, approximate population equality, and compliance with the Voting Rights Act. Legislative redistricting may face additional constraints such as split party control and gubernatorial veto power, which historically have served to mitigate partisan gerrymandering (Cain & Campagna, 1987; McDonald, 2004). However, with the decline in divided state governments, these political constraints may no longer be sufficient to ensure fairness in the redistricting process.

2.2.1 Traditional Redistricting Criteria

To curb the worst abuses of legislative redistricting, many states have established additional redistricting criteria in their constitutions, statutes, or legal precedents, aiming to enforce greater fairness. These traditional criteria typically include compactness, preservation of political subdivision boundaries, and safeguarding of communities of interest, aligning with normative representational goals and aiming to ensure representation based on geographic location, community, and identity rather than party affiliation (Nagle, 2019; Cain, 2012). These criteria ostensibly align with normative representational goals and align with partisan blind conceptions of fairness, aiming to ensure adequate representation based on geographic location, community, and identity, not party (Nagle 2019). Importantly, these traditional redistricting criteria have also been applied to all of the IRCs as well (Levitt & Spencer 2024).

2.2.2 Permitted Considerations of Partisanship and Partisan Fairness Metrics

In addition to traditional redistricting criteria, however, many IRCs are constrained by additional criteria that govern the extent to which they can consider, which varies significantly, influencing both the overall partisan fairness of the process and which conception of partisan fairness is emphasized. At one end of the IRC spectrum of permitted considerations of partisanship is Idaho, which strictly prohibits the use of any redistricting data beyond the population data provided by the US Census Bureau. This results in a completely partisan-blind process, where political affiliations are not factored into the redistricting decisions. Further along the spectrum, states like California and New York use negative regulatory language to prohibit discrimination against political parties without providing specific directives on how to consider partisanship. While these states do not forbid the consideration of partisan data, they do not provide clear guidelines on what defines “non-discrimination”, nor how its violation would be measured. In contrast, states like Washington and Arizona are more explicit in allowing the consideration of partisanship. They specify that partisanship can be utilized to craft competitive districts, acknowledging the role of political competition in enhancing partisan fairness while still avoiding overt discrimination in favor of a given political party. However, at the furthest end of the spectrum, Michigan explicitly requires the consideration of partisan fairness metrics in the redistricting process, likely corresponding to a strong, partisan-aware conception of fairness. For the state constitutional language defining the maximum allowable partisanship considerations of each IRC, see Appendix A, Table 3.

Notably, the state constitutions and enforceable governing statutes of some states, such as Alaska and Montana, remain silent on the consideration of partisanship in redistricting. This lack of specification grants these redistricting institutions a unique degree of autonomy, allowing them to determine the extent to which partisanship may or may not be considered. For example, in Montana, the commission-approved guidelines for the redistricting process specified that "No plan may be drawn to unduly favor a political party" and that "[t]he commission may consider competitiveness of districts when drawing plans" ("Criteria and Goals"). This suggests that even in the absence of explicit guidelines, partisanly balanced IRCs may self-impose a moderate standard of partisanship considerations.

This spectrum of permitted partisanship considerations not only influences the overall partisan fairness of the redistricting process among the IRCs but also potentially reflects differing operational conceptions of partisan fairness. States like Idaho, with their strict prohibition on considering partisanship, adhere to a partisan-blind conception of fairness, prioritizing geographic representation over political considerations, and would likely perform well on partisan-blind metrics. On the other hand, states like Michigan, with explicit mandates for the commission to consider partisan fairness metrics, embrace a partisan-aware approach..

2.2.3 Approval Mechanisms

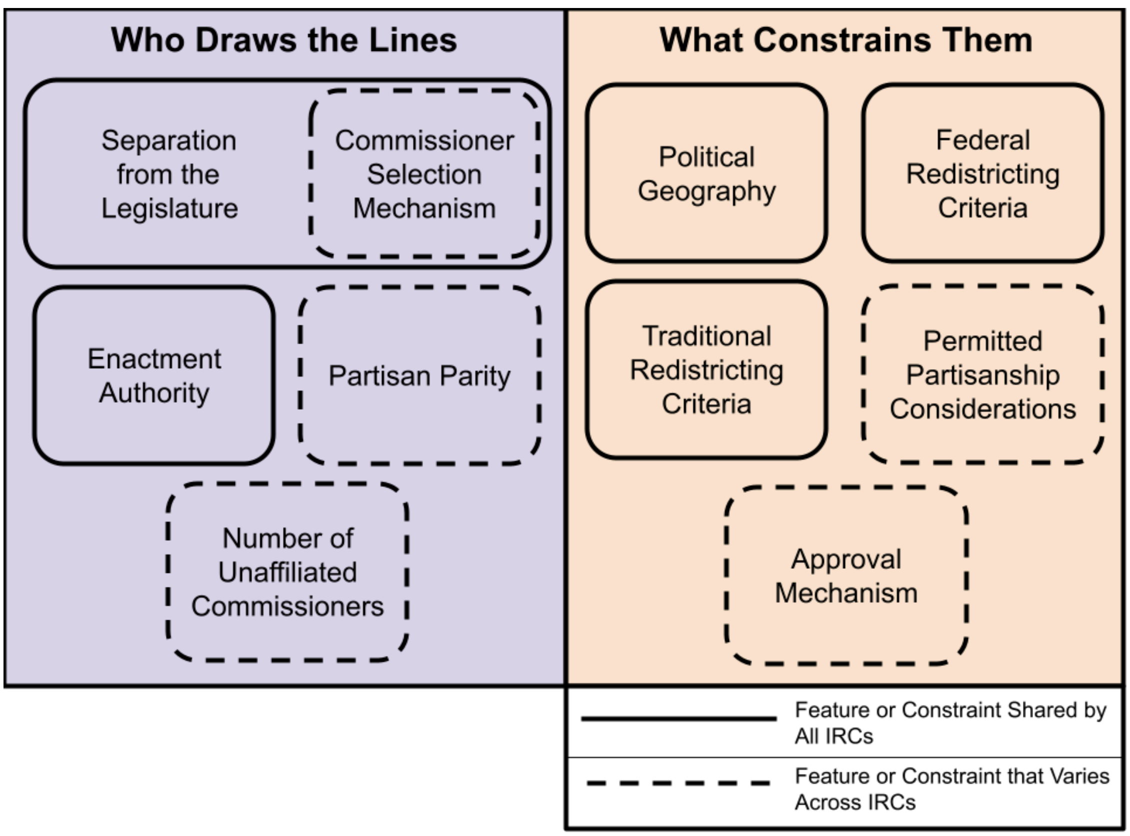
Finally, the variation in the approval mechanisms used by IRCs to enact a final redistricting plan may also influence the level of partisan fairness achieved in those plans. The approval mechanisms used by IRCs can be broadly categorized into majoritarian and non-majoritarian systems. Majoritarian systems, as seen in states like Arizona and Montana, allow a redistricting plan to pass with a simple majority vote, potentially enabling one party, in conjunction with the biases of unaffiliated commissioners, to dominate the process. However, it's worth noting that in some majoritarian systems, such as those in Washington and Idaho, an effective majority must include at least one member of the opposite party. This requirement for cross-party support introduces a degree of bipartisanship into the approval process, which may mitigate partisan bias and foster greater consensus.

Meanwhile, non-majoritarian requirements in states like Colorado, California, New York, and Michigan, mandate supermajorities and/or specific vote quotas from each partisan category for plan approval. By necessitating broader support across party lines or among commissioners of different affiliations, non-majoritarian systems are designed to compel maps to have partisan characteristics that broadly satisfy both major parties and unaffiliated commissioners (Miller & Grofman 2013). Consequently, states with non-majoritarian institutions are expected to exhibit the highest levels of partisan fairness, followed by states with majoritarian systems that require some cross-party support, like Washington and Idaho. Meanwhile, states with simple majoritarian approval mechanisms, such as Montana, Alaska, and Arizona, would demonstrate the least partisan fairness.

2.3 Summarizing Variations in IRC Design

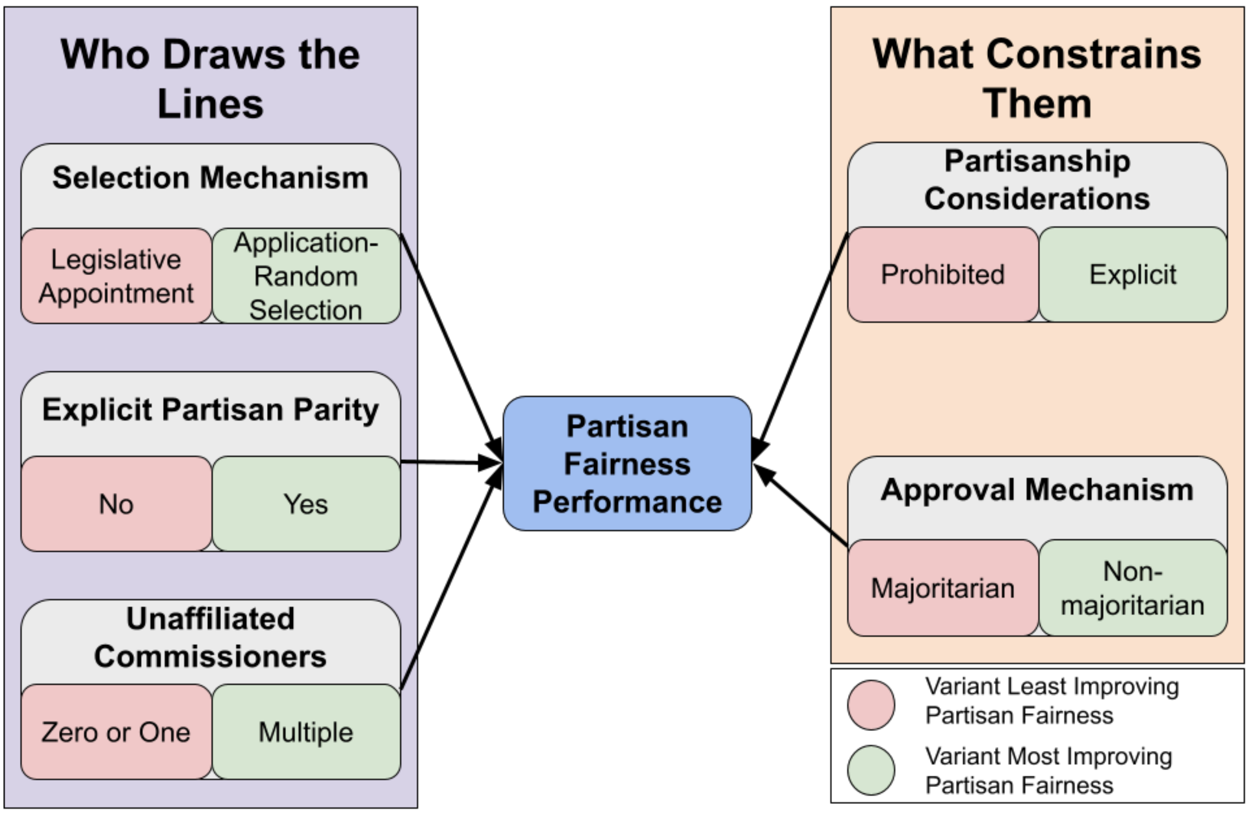
In brief, the institutional design of redistricting institutions, particularly IRCs, encompasses two critical dimensions that shape their ability to achieve partisan fairness - "who draws the lines" and "what constrains them" during the redistricting process. As illustrated below in Figure 3, the "who draws the lines" dimension covers features like separation from the legislature, partisan parity in composition, and the number of unaffiliated commissioners involved. Meanwhile, the "what constrains them" dimension includes factors such as the state's political geography, traditional redistricting criteria, permitted partisan considerations, and the approval mechanism. Importantly, while all IRCs share certain defining characteristics of separation from the legislature and primary enactment authority over redistricting, Figure 3 highlights the additional institutional design features within these two broad dimensions across which IRCs do vary.

Figure 3: Design Features and Constraints of Redistricting Institutions



Meanwhile, Figure 4 provides a visual representation of how these specific institutional design features that vary across IRCs are theoretically expected to influence performance on partisan fairness. The diagram illustrates the hypothesized impact of different variations in those design features, with variants in green representing features anticipated to enhance partisan fairness, and variants in red indicating features that may hinder the achievement of fairness. Note that for the purposes of this diagram, partisan fairness is assumed to be partisan-aware. If a partisan-blind conception of fairness were used, the expected impact of all design features would remain the same, barring allowed partisanship considerations, which would be reversed.

Figure 4: Design Feature Variation of IRCs and Theorized Impacts on Partisan Fairness



Finally, to fully demonstrate the diversity of institutional design features and their potential implications for political impartiality, a visual representation of the distribution of these features across the various IRCs is presented below in Table 2. Each feature variation is categorized based on its theoretical impact on partisan-aware fairness in a similar manner to Figure 4. In the table, green denotes features theorized to most improve partisan fairness, yellow signifies those theorized to modestly improve partisan fairness, and red indicates features theorized to least improve partisan fairness.

Table 2: Relevant Institutional Design Features by IRC

IRC State	Appointment Mechanism	Explicit Partisan Parity	Number of Unaffiliated Commissioners	Allowed Partisanship Considerations	Approval Mechanism
Alaska	Legislative Appointment	No	Varies	Unspecified	Majoritarian
Montana	Legislative Appointment	Yes	1	Unspecified	Majoritarian
Idaho	Legislative Appointment	Yes	0	None	Majoritarian (cross-party)
Washington	Legislative Appointment	Yes	0 (1 Non-Voting)	Competitiveness	Majoritarian (cross-party)
Arizona	Hybrid	Yes	1	Competitiveness	Majoritarian
New York	Legislative Appointment	Yes	2	Non-discrimination	Non-majoritarian
California	Application/ Random Selection	Yes	4	Non-discrimination	Non-majoritarian
Colorado	Application/ Random Selection	Yes	4	Competitiveness	Non-majoritarian
Michigan	Application/ Random Selection	Yes	5	Explicit Partisan Fairness	Non-majoritarian

Variant Least Improving Partisan Fairness

Intermediate Variant

Variant Most Improving Partisan Fairness

With the relevant institutional design features laid out in Table 2, it is clear that the features hypothesized to correlate with partisan impartiality are not randomly distributed among the IRCs. Instead, two distinct groups seem to emerge based on their institutional design features. One set of commissions ("Type I"), including Montana, Washington, Idaho, Alaska, and Arizona, exhibits more features thought to not improve partisan fairness as much. These features include legislative appointment, zero or one unaffiliated commissioners, less expansive partisanship considerations, and majoritarian approval mechanisms. Conversely, another set of commissions ("Type II"), comprising California, New York, Colorado, and Michigan, features institutional designs theorized to most improve partisan fairness. These commissions tend to utilize random selection methods, including more than one unaffiliated commissioner, consider partisanship more expansively, and employ non-majoritarian approval mechanisms.

However, there are some relative outliers within this framework. For example, while Arizona shares most of its characteristics with Type I commissions, it also exhibits some features more commonly associated with Type II commissions, such as heightened consideration of partisanship in the mapping process and a hybrid appointment regime. Meanwhile, New York aligns closely with Type II commissions but also has a majoritarian approval mechanism, which is more characteristic of Type I commissions. Without a random distribution of institutional design features among the IRC states, the influence of any individual feature cannot be accurately parsed in statistical analysis. Therefore, despite the lack of a purely dichotomous distribution, and to achieve the most rigorous examination and comparison possible, these commissions will be considered within these two semi-homogeneous categories of Type I and Type II. This convenient division of IRCs into two groups is not entirely novel and has been discussed previously (see Cain 2012). However, applying this categorization to all current redistricting commissions and considering multiple institutional design criteria in delineating these commission types represents a more comprehensive approach to the landscape of redistricting institutions than what has been explored previously.

3 Methodology

The empirical methodology employed in this study aims to understand if the identified clusters of design features of IRCs measurably influence their ability to improve performance on partisan fairness metrics when compared to legislative redistricting. This investigation utilizes a combination of fairness metrics and fairness baselines, including the partisan bias score, the mean-median difference, and the efficiency gap, to comprehensively evaluate redistricting plans along each of the four conceptions of partisan fairness. The Sequential Monte Carlo (SMC) redistricting simulation method is used to generate a distribution of possible redistricting plans and associated fairness metrics. ANOVA is employed to determine differences in performance between different redistricting institutions and their respective types, offering insights into their effectiveness. Additionally, a comparative analysis over the 2010 and 2020 redistricting cycles is conducted to assess changes over time and the impact of IRC adoption on partisan fairness performance.

3.1 Partisan Fairness Metrics

In this study, several commonly used partisan fairness metrics are utilized to comprehensively evaluate redistricting plans across the various conceptions of partisan fairness: partisan bias, the mean-median gap, and the efficiency gap. Partisan bias, as alluded to previously, is a symmetry-based metric that quantifies the deviation from perfect partisan symmetry at a given vote share, with the vote share of 0.5 given particular importance (Katz et al. 2020). This measure indicates whether one party is systematically advantaged at the inflection point of the popular vote (King 1990). The mean-median difference, another symmetry-based metric, is also being considered because it does not rely on a hypothetical partisan swing and captures a vote-denominated interpretation of partisan symmetry (McDonald & Best 2015; Katz et al. 2020). As such, the mean-median difference highlights the disparity between the average district vote proportion and the median district vote for a particular party, and fairness is achieved when this difference is zero, indicating how much more one party needs to earn in votes compared to the other to secure a seat (Katz et al. 2020).

Finally, the efficiency gap, as defined by Stephanopoulos & McGhee (2015) measures the difference between wasted votes of each party divided by the total number of votes cast. These wasted votes are those that do not contribute to the election of a candidate, including all votes for the party that did not win a particular legislative seat and all the additional votes for the party that won a particular legislative seat beyond what would have been necessary to win a bare majority. Unlike partisan bias, the efficiency gap does not require hypothetical simulations and is more closely tied to concepts of proportionality, albeit generally favoring the majority party (Stephanopoulos & McGhee 2015; Duchin & Schoenbach 2023). Therefore, when the adopted plans of a particular redistricting institution are analyzed for partisan fairness, each of these three metrics is calculated to achieve a more nuanced understanding of the symmetry-proportionality dimension of the redistricting process.

3.2 The SMC Simulation

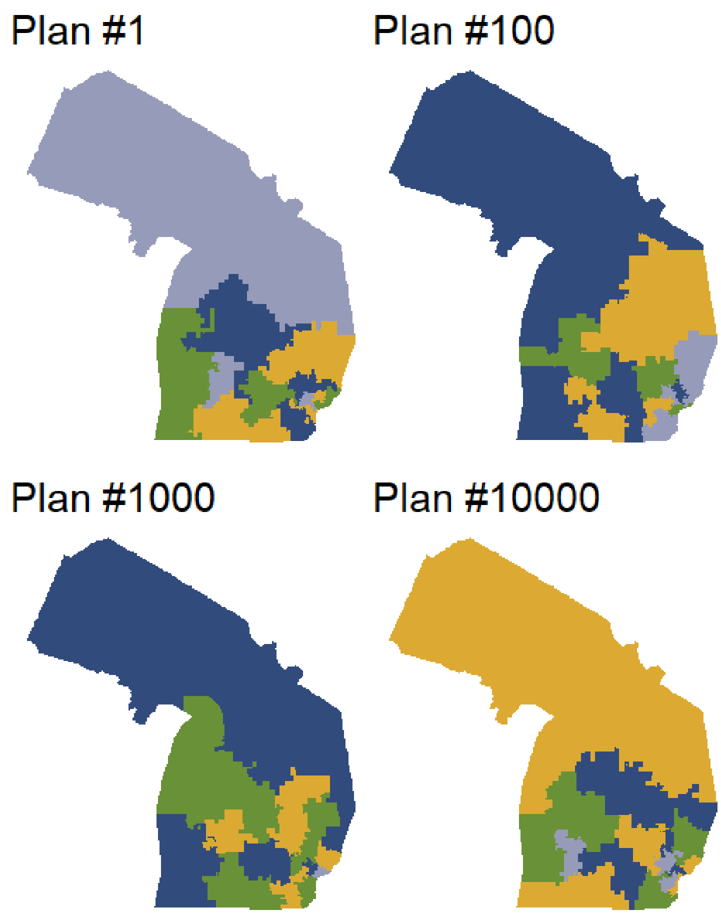
However, these static measures of partisan fairness are not particularly useful on their own, as they cannot control for the confounders of political geography that may altogether prevent a perfect partisan fairness score (Duchin & Schoenbach 2023). Additionally, they reveal no information about the partisan blind-partisan aware dimension of fairness, which requires an understanding of what partisan fairness metrics can be “reasonably expected”. As such, a simulation approach offers a powerful tool to determine the distribution of what partisan fairness scores are both possible and/or could be “reasonably expected”. These simulation methods aim to generate a representative sample of redistricting plans that adhere to requirements such as contiguity, population equality, and compactness, allowing researchers to identify gerrymandered plans by measuring their deviation from this sample according to partisan fairness measures (Fifield et al. 2020). This is particularly relevant given the infeasibility of enumerating all possible redistricting plans in a state, which would result in an astronomical number of unique plans (Fifield et al. 2020). They provide a practical alternative to this full enumeration by approximating the sample space of possible redistricting plans, allowing for the comparison of implemented plans against this distribution.

Simulation approaches also offer valuable insights into the redistricting process by permitting the calculation of two distinct baselines, one partisan-aware and one partisan blind, for assessing partisan fairness metrics. The partisan-aware baseline represents the fairness score for the simulation-produced redistricting plan produced by the simulation that is closest to the “perfect” value of that particular fairness metric, thereby representing the “best possible” map that can be reasonably obtained given the constraints of redistricting (Katz et al. 2020; Deford, Duchin, & Solomon 2020). In contrast, the partisan blind baseline reflects the average partisan fairness metric over the set of simulated redistricting plans (Deford, Duchin, & Solomon 2020; McDonald et al. 2018). These baselines can diverge significantly, highlighting the trade-offs between achieving partisan-aware fairness and adhering to a partisan-blind approach (Deford, Duchin, & Solomon 2020).

Various simulation methods, including edge-based techniques and Markov-Chain Monte Carlo (MCMC) algorithms, have been employed to simulate redistricting plans while adhering to specific criteria (Fifield et al. 2020; McCartan & Imai 2023). However, these methods often struggle to approximate the full sampling distribution of redistricting plans accurately and suffer from computational inefficiency and large standard errors (McCartan & Imai 2023). In contrast, the Sequential Monte Carlo (SMC) simulation method, as outlined by McCartan and Imai (2023), offers several advantages. Notably, it provides a more reliable approximation of the sampling distribution, enhances computational efficiency, and enables better incorporation of additional constraints. Moreover, the SMC method benefits from an open-source package in R, facilitating convenient implementation. The algorithm's authors also offer extensive guidance on best practices, ensuring maximum accuracy within the constraints of a limited sample size. Thus, SMC emerges as the most robust choice among simulation approaches for redistricting analysis and Figure 5 shows four example redistricting plans sampled using this method.

Figure 5: Sample SMC Redistricting Plans

(Michigan 2020 Congressional)



In this investigation, the simulation approach encompasses two redistricting cycles: 2020 and 2010, chosen for their rich data availability, including electoral data, block-level population data accessibility, and the widespread availability of geometric shapefiles for blocks and precincts. For the 2020 redistricting cycle, simulations have been conducted for most of the IRC states at both the congressional and state senate levels. Congressional simulations were only performed on states with more than two congressional districts, as accurate partisan fairness scores cannot be calculated with only two districts. Additionally, effective state senate simulations could not be produced for the IRC states of New York and Montana due to computational constraints and the difficulty of approximating the sampling distributions for these particular states. For the 2010 redistricting cycle, simulations were again conducted for most current IRC states, regardless of whether they had an IRC in 2010. State senate simulations were again not possible in New York, Montana, and Alaska due to computational, algorithmic, or data availability constraints. As such, the IRC state of Montana is excluded entirely from the analysis, as it has only had one or two congressional districts across the redistricting cycles of interest and state senate simulations could not be performed. Simulations have also been run for many legislative redistricting states for the 2020 redistricting cycle, including most congressional and state senate plans, to serve as a baseline for comparison with IRC states. Some state senate districts were not simulated for certain legislative redistricting states due to the larger number of districts at these legislative levels and the difficulty in accurately approximating the sampling space.

Population data, block-level shape files, and implemented districting plans are drawn from the US Census Bureau. Precinct shape files and electoral data are primarily sourced from Dave's Redistricting and the Harvard ALARM Project. Additional electoral data sources include the Harvard Election Data Project, the MIT Election Data and Science Lab, and the King County Elections Department.

District-level partisanship was calculated by determining the average Democratic vote share across the past two presidential election cycles for each geographic unit. This approach serves as a proxy for latent district partisanship and has been shown to possess relatively robust predictive capabilities for lower-level elections, particularly in an increasingly nationalized political environment (Levendusky, Pope, & Jackman 2008). The choice of using past elections as the basis for partisanship calculation rather than realized election results is grounded in the assumption that redistricting institutions cannot accurately forecast future electoral trends and instead rely on historical data readily available to them (Yoshinaka & Murphy 2009). Meanwhile, the decision to focus on past presidential elections rather than other contests such as House of Representatives, state legislative, gubernatorial, or senatorial races stems from several considerations. Many of these seats often witness uncontested races, complicating the estimation of partisanship. Common approaches like assigning a fixed vote share or imputation based on contested races have been found inadequate (Plener Cover 2018; Wilson & Grofman 2022). Moreover, contested races for governor and sometimes for the US Senate do not nationalize evenly across the country, potentially skewing perceptions of support in congressional and state legislative elections, which have generally become more nationalized, albeit with some exceptions (Sievert & McKee 2018; Jacobson 2015; Zingher & Richman 2018).

The decision to use precincts as the foundational units for constructing districts rather than other potential geographic units like blocks, counties, or municipalities is primarily driven by the availability of authoritatively accurate election outcomes at the precinct level. Precincts provide the smallest level of granularity at which such election results are reliably accessible and while electoral results can be disaggregated to smaller units, this approach may compromise accuracy to a significant extent (Deford, Duchin, & Solomon 2020). However, precinct boundaries often undergo frequent changes, which can pose challenges for maintaining consistency in analysis over time. To address this issue, a disaggregation-reaggregation approach is employed in the datasets sourced from Dave's Redistricting, VEST, and the Harvard ALARM project. This method involves disaggregating electoral data from previous elections, which may have been based on different precinct boundaries, down to the census block level using voting-age population weights. Subsequently, the data is reaggregated to align with the most recent precinct boundaries. In some states, however, precinct-level shapes and electoral data are not readily available. In such cases, electoral data and shapes are provided using block groups instead of precincts. For the 2020 redistricting cycle, states utilizing block group data include California, Hawaii, Oregon, and West Virginia. Similarly, for the 2010 cycle, states relying on block group data are California and Montana. For the 2004 election, in which precinct-level data was largely unavailable, a similar disaggregation-aggregation approach was used to approximate precinct-level electoral results using available county-level or municipality-level data, an approach that has been validated by other researchers (Best et al. 2021).

The construction of the geometric dataset for precincts involved several technical steps to ensure accuracy and contiguity. Geometric data for precincts in each state was obtained from either DRA or ALARM using the *geomander* package in R, a tool extensively utilized in the data preparation process. Given that not all precincts are contiguous, additional steps were necessary to generate an accurate adjacency matrix. First, electoral data was disaggregated down to the block level using voting-age population as the weighting variable, with shapes and data sourced from the US Census Bureau using the *tidycensus* package in R. Subsequently, precinct shapes were separated into their constituent components, and the disaggregated electoral and population data were reaggregated back into contiguous precincts. In states where the precinct adjacency matrix was not entirely contiguous due to islands or disconnected precincts, each disconnected precinct was assigned a neighbor, which was determined as the single closest precinct. This process was repeated for non-contiguous components until the entire adjacency matrix was contiguous. Utilizing the redist package, map objects were then produced for each state and relevant legislative level using the now-contiguous precincts and adjacency matrix. The number of legislative districts for each map object was specified accordingly, along with population deviation tolerance levels based on federal standards. For congressional districts, a tolerance of 1% was set, following the precedent set by *Tennant v. Jefferson County Commission* (2012), while a tolerance of 10% was specified for state legislative districts, as established in *Brown v. Thomson* (1983). These technical procedures were essential to ensure the accuracy and validity of the geometric dataset for precincts, laying the groundwork for subsequent redistricting analysis.

Setting the hyperparameters for Sequential Monte Carlo (SMC) simulations involves careful consideration of various factors to ensure an accurate representation of the sampling space of possible redistricting plans. The compactness parameter was set to the default value of 1, as suggested by McCartan and Imai (2023), to loosely constrain the compactness of the resulting districts. To enhance diversity in the sampling space, adjustments were made to the population tempering (pop_temper) and sequence alpha (seq_alpha) hyperparameters (McCartan & Imai 2023). Higher values of pop_temper and seq_alpha can promote greater diversity among sampled plans, preventing the algorithm from getting stuck on final splits and down-weighting unlikely plans, respectively. However, altering these parameters must be balanced to avoid sacrificing accuracy. The sequence alpha was typically maintained near 0.5, with minor adjustments made to increase diversity or prevent bottlenecks in the algorithm. Population tempering was also adjusted accordingly to prevent the algorithm from converging prematurely.

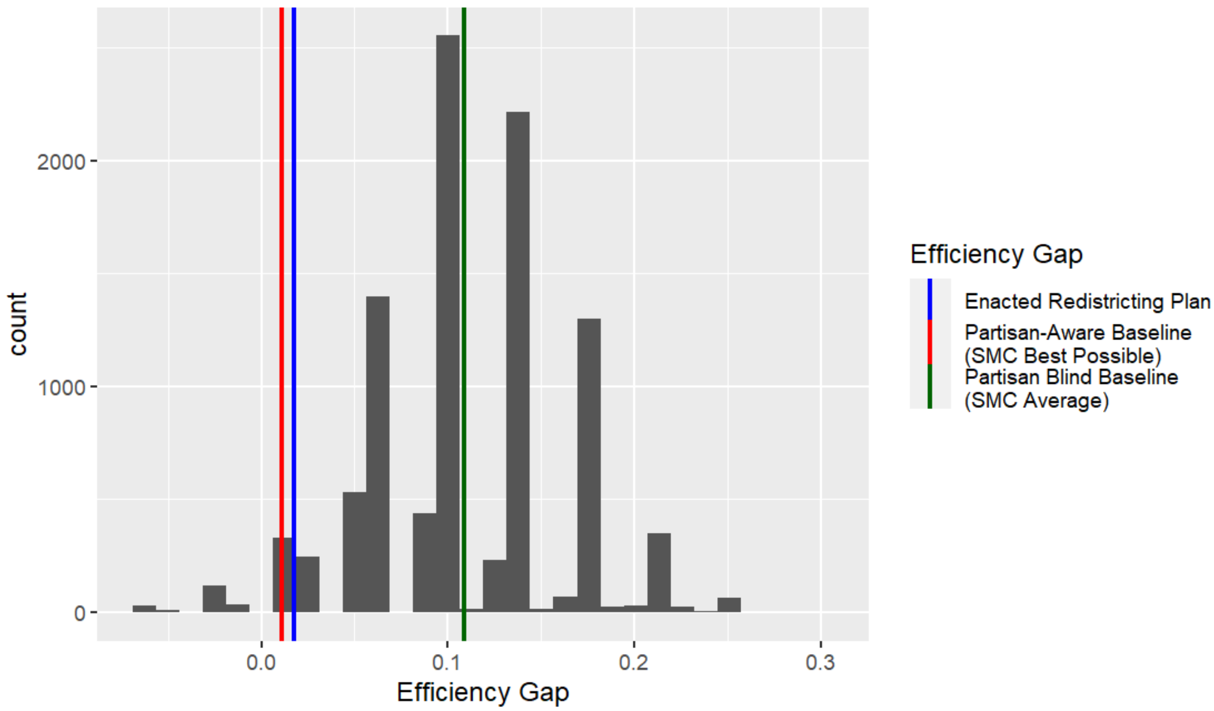
For each simulation, I produced a minimum of 10,000 sample plans, ensuring a robust representation of the sampling space. In cases where minor hyperparameter tuning was insufficient to achieve adequate diversity or convergence in the target metric, hyperparameter adjustments were combined with an increase in the number of sampled plans to maintain accuracy without sacrificing diversity. Detailed information on individual simulations, including hyperparameter settings and diagnostics, is provided in Appendix B for reference.

3.3 Comparative Methodology

In comparing enacted redistricting plans to the SMC distributions and associated baselines, each sampled plan underwent an assessment of partisan fairness metrics using the redist package. The partisan bias, mean-median difference, and efficiency gap metrics were computed based on the previous two presidential election results and then averaged between the two elections to ensure equal weighting. This approach was preferred over simply totaling votes by district across the two election cycles because it prevents one election cycle from disproportionately influencing fairness metrics like the efficiency gap. Subsequently, sampling distributions of each fairness metric were generated, allowing for the calculation of two distinct baselines. The partisan-blind baseline was determined by the average of the partisan fairness metric over the entire distribution (SMC average), while the partisan-aware baseline (SMC best possible) was identified as the closest partisan fairness metric to its ideal value within the distribution. Enacted redistricting plans, sourced from the US Census Bureau, underwent similar fairness metric calculations, although population and electoral data were considered at the block level rather than the precinct level due to precinct-level splits in many states' plans. Figure 6 illustrates an example sampling distribution of efficiency gap scores, along with the enacted redistricting plan's efficiency gap score, the partisan-blind baseline, and the partisan-aware baseline.

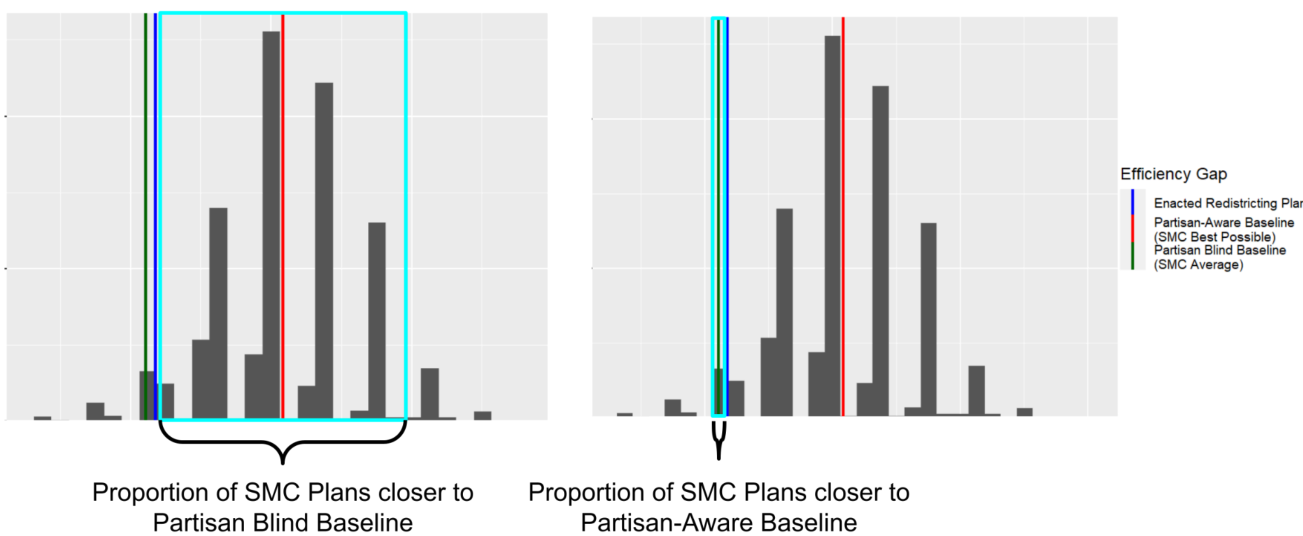
Figure 6: Example Sampling Distribution of the Efficiency Gap

(Michigan 2020 Congressional Simulation)



Comparisons between enacted plans and the calculated partisan fairness baselines were based on the percentage of simulated plans closer to each baseline than the enacted plan. This metric was chosen for its interpretability and resilience to outliers, as partisan-aware baselines frequently occupy the tail ends of distributions. When partisan-aware baselines represent outlier values, distance measures provide limited insight. Moreover, these outlier partisan-aware baselines may indicate strong violations of other redistricting criteria to an undesirable degree. In such cases, while the raw distance to the baseline could be substantial, the proportion of plans outperforming the enacted plan could remain relatively small. By evaluating the percentage of simulated plans closer to each fairness baseline, this approach robustly assesses enacted plan performance within the overall distribution. It offers a valuable perspective on both partisan-blind reasonability and the degree of difficulty in satisfying partisan-aware outlier fairness targets. Thus, for each partisan fairness metric, each enacted redistricting plan yielded two scores: the percentage of simulated plans outperforming it on the partisan-aware baseline of that metric and the percentage outperforming it on the partisan-blind baseline of that metric. Figure 7 illustrates how the percentage of simulated plans closer to each fairness baseline than the enacted plan is calculated, using the efficiency gap distribution as an example. Additionally, the full table of the proportions of simulated plans closer to each partisan fairness metric baseline can be found in Appendix C.

Figure 7: Calculation of the Proportion of Simulated Plans closer to each Fairness Baseline (Efficiency Gap of Michigan 2020 Congressional Plan)



To compare the proportions of plans closer to each baseline across different redistricting institutions, two phases of analysis were conducted using ANOVA. First a comparison between legislative redistricting and IRCs as a unified category was performed. This phase aimed to assess whether IRCs, as a collective group, demonstrate superior performance compared to traditional legislative redistricting. Secondly, a comparison was made between legislative redistricting and different types of IRCs, categorized as Type I and Type II. This phase aimed to ascertain whether the performance of IRCs relative to legislative redistricting varies depending on the type of IRC in place and whether there are significant differences in performance between IRC types. The decision to utilize ANOVA in these analyses was made to statistically determine significant differences in performance between the different redistricting institutions and their respective types, providing robust insights into the performance of each institution.

Table 3 provides a concise overview of the entire cross-sectional methodological process, from data collection and generation of redistricting plans to the final comparative analysis and ANOVA tests.

Table 3: Abbreviated Methodological Process

Step	Description
1. Data Collection	- Collect population data, block-level shapefiles, and enacted districting plans from the US Census Bureau. - Obtain precinct shapefiles and electoral data from sources like Dave's Redistricting, Harvard ALARM Project, Harvard Election Data Project, MIT Election Data and Science Lab, and King County Elections Department.
2. Precinct-Level Partisanship Calculation	- Calculate district-level partisanship as the average Democratic vote share across the past two presidential election cycles for each geographic unit (precinct or block group). - Use disaggregation-reaggregation to align electoral data with precinct boundaries.
3. Geometric Dataset Construction	- Disaggregate electoral data to the block level. - Separate precinct shapes into constituent components. - Reaggregate disaggregated electoral data into contiguous precincts. - Assign neighbors to non-contiguous precincts. - Produce map objects for each state and legislative level using the contiguous precincts.
4. SMC Simulation Setup	- Set hyperparameters for Sequential Monte Carlo (SMC) simulations, including compactness (ρ=1), population tempering (pop_temper), and sequence alpha (seq_alpha). - Generate a minimum of 10,000 sample redistricting plans for each state-election-year combination, adjusting hyperparameters as needed to ensure diversity and convergence.
5. Fairness Metric Calculation	- Calculate partisan bias, mean-median difference, and efficiency gap for each simulated redistricting plan and the enacted plan, averaged across the two presidential election cycles.
6. Baseline Determination	- Calculate the partisan-blind baseline as the average of each fairness metric over the entire distribution of simulated plans (SMC average). - Identify the partisan-aware baseline as the simulated plan closest to the ideal value of each fairness metric (SMC best possible).
7. Comparative Analysis	- Compare the enacted plan's fairness metrics to the calculated baselines. - Calculate the percentage of simulated plans closer to each baseline than the enacted plan.
8. ANOVA Analysis	- Conduct ANOVA to compare the proportions of plans closer to each baseline between legislative redistricting and IRCs as a unified category. - Conduct ANOVA to compare the proportions of plans closer to each baseline between legislative redistricting, Type I IRCs, and Type II IRCs.

3.4 Comparison Across Time

Finally, to assess changes over time and plausible claims of causality regarding the adoption of IRCs and partisan fairness performance, a comparative analysis over the 2010 and 2020 redistricting cycles was conducted, focusing on select states. Specifically, attention was given to states that transitioned from legislative redistricting to an IRC between 2010 and 2020, compared to those that maintained legislative redistricting over the same period. In this context, New York served as an important case study, offering the possibility of examining the impact of IRC adoption in comparison with realized, counterfactual redistricting plans. In 2010, New York underwent legislative redistricting under divided control, whereas in 2020, three distinct redistricting plans were produced: one approved by the unified legislature, another approved via judicial intervention, and the most recent plan approved by an independent commission. This approach allowed for a comparative assessment of the impact of different redistricting methods on partisan fairness over time. The analysis also extended to other states, including those that maintained the IRC approach in both redistricting cycles (such as Montana, Washington, Idaho, Arizona, and California) and additional states transitioning from legislative redistricting to IRC Type II (such as Michigan and Colorado). By considering each of these scenarios, the study aims to discern potential causal relationships between redistricting mechanisms and their performance relative to partisan fairness score baselines, thus providing valuable insights into the effectiveness of different redistricting approaches beyond a cross-sectional approach.

4 Results

Table 4 below presents the results of a cross-sectional t-test comparing the performance of IRCs as a unified category against legislative redistricting institutions (LR) on partisan fairness metrics. This particular test evaluates if IRCs as a unified category outperform legislative redistricting institutions on each of the different conceptions of partisan fairness. The comparison is based on the proportion of simulated redistricting plans that are closer to the partisan-blind and partisan-aware baselines than the enacted plans. Negative estimates indicate that IRCs perform better than LR (i.e., have a lower proportion of simulated plans closer to the given baseline), while positive estimates suggest the opposite.

Table 4: Proportion of Simulated Redistricting Plans Closer to Partisan Fairness Baselines: IRC compared to Legislative Redistricting

Metric	Baseline[a]	Estimate[b]	Standard Error	t-value	p-value
Partisan Bias	Partisan Blind	0.01918	0.09029	0.212	0.832
Partisan Bias	Partisan-Aware	-0.06334	0.09354	-0.677	0.5
Mean-Median	Partisan Blind	-0.09668	0.08037	-1.203	0.232
Mean-Median	Partisan-Aware	-0.23536	0.08591	-2.74	0.007516**
Efficiency Gap	Partisan Blind	0.01350	0.08948	0.151	0.88
Efficiency Gap	Partisan-Aware	-0.06281	0.09412	-0.667	0.506

Notes:

a. The fairness baseline against which the enacted plan is compared.
b. Estimate of the difference in the percentage of plans closer to the baseline than the enacted redistricting plan. Negative estimates indicate better performance by IRCs compared to LR.

* Indicates statistical significance at the 0.1 level.

** Indicates statistical significance at the 0.05 level.

These results demonstrate that IRCs, when considered as a uniform category, substantively and statistically significantly outperform legislative redistricting methods on the mean-median fairness metrics when compared to the partisan-aware baseline. However, IRCs are not demonstrated to outperform legislative redistricting on this metric when considering the partisan-blind baseline. Additionally, on partisan bias and the efficiency gap, IRCs are not demonstrated to outperform legislative redistricting when compared to either baseline with statistical significance. This suggests that while IRCs may excel in optimizing redistricting plans for mean-median fairness when considering partisan awareness, they do not exhibit a consistent advantage over legislative redistricting methods across all fairness metrics.

Table 5, meanwhile, presents the results of the cross-sectional ANOVA that compares the performance of legislative redistricting to Type I IRCs and Type II IRCs on the partisan fairness metrics. The comparison aims to determine if the performance of IRCs relative to legislative redistricting varies depending on the type of IRC in place, and whether there are significant differences in performance between the IRC types as theorized. The analysis compares the proportion of simulated redistricting plans that are closer to the partisan-blind and partisan-aware baselines than the enacted plans. The table reports the estimated differences in proportions between each pair of redistricting institutions, with negative values indicating that the comparison institution outperforms the reference institution. The associated p-values are included in parentheses.

Table 5: Proportion of Simulated Redistricting Plans Closer to Partisan Fairness Baselines: IRC Type I v IRC Type II v Legislative Redistricting

	Partisan Bias		Mean-Median Difference		Efficiency Gap	
	Partisan Blind	Partisan-Aware	Partisan Blind	Partisan-Aware	Partisan Blind	Partisan-Aware
LR[a] vs IRC I[b]	-0.04901[c] (0.9050)[d]	0.0326 (0.9594)	-0.2273 (0.0672)*	-0.1497 (0.3608)	-0.0211 (0.9815)	0.0806 (0.7744)
LR vs IRC II	0.1026 (0.6944)	-0.1806 (0.3494)	0.0629 (0.8346)	-0.3400 (0.0150)**	0.0558 (0.8964)	-0.2381 (0.1608)
IRC I vs IRC II	0.1516 (0.6083)	-0.2132 (0.3995)	0.2902 (0.0979)*	-0.1903 (0.4205)	0.0769 (0.8783)	-0.3187 (0.1300)
F-Statistic	0.477 (0.622)	1.075 (0.346)	2.939 (0.0584)*	4.575 (0.013)**	0.129 (0.879)	2.141 (0.124)

Notes:

a. The first institution listed is the reference category
b. The second institution listed is the comparison category
c. The difference in the proportion of simulated plans closer to the given baseline for the particular fairness metric. Negative values indicate that the comparison category outperforms the reference category
d. The p-value of the estimated difference in proportions

* Indicates statistical significance at the 0.1 level.

** Indicates statistical significance at the 0.05 level.

The F-statistics reveal that overall differences among the redistricting institutions are strongly statistically significant for the mean-median difference metric on the partisan blind baseline (F=2.939, p=0.0584) and the mean-median difference on the partisan-aware baseline (F=4.575, p=0.013). Examining the comparisons within each of these columns reveals that the institution likely responsible for the difference is not the same for the two conceptions of fairness. For the partisan-blind mean-median difference, Type I IRCs seem to be causing the improvement, with those commissions seeing a statistically significant 23% improvement in that metric compared to legislative redistricting. There is also some statistical evidence, albeit relatively weak (p=0.098), suggesting that Type I IRCs also perform better on this particular metric when compared to Type II IRCs. Meanwhile, for the partisan aware mean-median difference, Type II IRCs seem to be causing the improvement, with those commissions seeing a statistically convincing (p=0.015) improvement over legislative redistricting of 34%.

While the F-statistic is not quite significant for the efficiency gap metric across both baselines, the effect sizes observed, particularly on the partisan-aware baseline, are noteworthy. The coefficient of -0.2381 suggests that IRC Type II outperforms legislative redistricting, with a lower proportion of plans closer to the partisan-aware efficiency gap baseline, although this difference does not reach statistical significance at conventional levels. Interestingly, when analysis is restricted to the 2020 redistricting cycle, the efficiency gap emerges as the most statistically significant metric, with IRC Type II strongly outperforming both legislative redistricting and IRC Type I (see Appendix D for a table of these exploratory results)[1].

Finally, the partisan fairness outcomes for select states over time are presented below in Table 6. The table presents changes in partisan fairness metrics across redistricting cycles for current IRC states that either maintained their redistricting institution or transitioned from legislative redistricting to an IRC between 2010 and 2020. For each analyzed state, in both the 2020 and 2010 redistricting cycles, the partisan fairness metrics of its congressional performance are averaged with its state senate performance, if both exist. Then, the 2020 values are simply subtracted from the 2010 values to determine the changes presented in the table.

[1] There are many potential reasons for this discrepancy. For one, California is the only state that utilized a Type II IRC in 2010, leading to their overrepresentation in the calculator of the group mean. Additionally, electoral data is not as precise for the 2010 redistricting cycle, potentially leading to less accurate approximations of district-level partisanship. Finally, it could very well be that there is significant variation among Type II IRCs and within individual IRC states across redistricting cycles that complicate statistical significance.

Table 6: Change in Partisan Fairness Metrics Across Redistricting Institutions: 2010-2020 Comparisons for Current IRC States

	Redistricting Institution		Change in Partisan Bias[a]		Change in Mean-Median Difference		Change in Efficiency Gap	
State	2010	2020	Partisan Blind	Partisan-Aware	Partisan Blind	Partisan-Aware	Partisan Blind	Partisan-Aware
NY[b]	LR	LR	-0.32678	0.24885	0.374075	-0.14495	-0.00352	0.995925
NY[c]	LR	IRC II	0.032325	-0.24435	0.374075	-0.14495	-0.62293	0.280625
MI	LR	IRC II	-0.32227	-0.76826	-0.0296	-0.91023	0.451908	-0.32714
CO	LR	IRC II	0.550975	0.56355	-0.48975	-0.27393	-0.9876	-0.75868
WA	IRC I	IRC I	0.23709	-0.33953	0.553927	-0.07724	0.536853	0.3637
ID	IRC I	IRC I	-0.42825	-0.70435	-0.21535	0.029125	-0.13618	0.0697
AZ	IRC I	IRC I	-0.23913	-0.19483	-0.22624	-0.13738	-0.52525	-0.30422
CA	IRC II	IRC II	0.2569	0.63114	0.40852	0.18062	0.0852	-0.8494

Notes:

a. Change is measured as the difference in the average proportion of plans closer to the given baseline for the given partisan fairness metric from 2010 to 2020. Negative values indicate an improvement on that particular metric.
b. This row corresponds to the counterfactual legislatively-implemented redistricting plan that was adopted by the New York legislature in 2022.
c. This row corresponds to the IRC-proposed redistricting plan that was produced by the commission in 2024.

For states that changed redistricting institutions, the adoption of an IRC seems to be associated with improvements in certain partisan fairness metrics compared to their previous legislative redistricting plans. New York's transition from LR to a Type II IRC resulted in improvements on the partisan-aware baselines for partisan bias (-0.24435) and the mean-median difference (-0.14495), suggesting the IRC adoption enhanced these conceptions of fairness. Notably, comparing New York's IRC II plan to its LR counterfactual highlights potential causal effects, with marked improvements in partisan-aware partisan bias (improvement of 24.4% vs a decline in performance of 24.9%) and a relative tempering of performance decline on the efficiency gap (0.280625 vs 0.995925). Michigan and Colorado, which also adopted Type II IRCs in 2020, also saw a common improvement in their partisan-aware efficiency gaps(-0.32714 and -0.75868, respectively) and partisan-aware mean-median differences (-0.91023 and -0.27393, respectively) compared to their 2010 LR plans. Among states that maintained their redistricting institution, like Washington, Idaho, Arizona, and California, there is no clear pattern of improvement or deterioration on the partisan fairness metrics between the redistricting cycles.

5 Discussion

The results of this investigation yield several important insights regarding the relationship between independent redistricting commissions, their institutional design features, and different conceptions of partisan fairness in redistricting. Broadly, the findings indicate that IRCs, when treated as a unified category, do not consistently outperform legislative redistricting processes across all metrics and baselines associated with partisan fairness. However, examining IRC performance through the lens of the various conceptions of partisan fairness articulated in this study reveals more nuanced dynamics.

When measuring partisan fairness from a partisan-aware symmetry perspective, as operationalized by the mean-median fairness metric compared to its partisan-aware baseline, IRCs as a unified category do exhibit a statistically significant advantage over legislative redistricting. This aligns with theoretical expectations, as IRCs are designed to mitigate conflicts of interest and intentional gerrymandering that could undermine fair treatment between parties. However, IRCs do not demonstrate clear superiority on metrics associated with partisan-aware proportionality, like the efficiency gap score, nor partisan-blind fairness metrics in general. These results suggest that while IRCs may enhance the ability of redistricting processes to achieve specific representational ideals tied to symmetry when partisan considerations are embraced, their performance is less consistent when optimizing for other fairness conceptions.

Meanwhile, the results demonstrate that not all IRCs are equally effective at facilitating the various partisan fairness goals. When dividing IRCs into two broad categories based on their composition and approval mechanisms, important distinctions emerge. This is most clearly demonstrated via the partisan blind mean-median difference. When IRCs are considered as a unified category, they are not demonstrated to significantly outperform legislative redistricting on this particular metric. However, when the IRCs are not considered as a monolith, statistically significant differences do emerge, as noted by the F-test for this metric in Table 5, with only Type I IRCs outperforming legislative redistricting. Meanwhile, the opposite is true for the partisan-aware mean-median difference where Type II IRCs outperform both legislative redistricting and Type I IRCs. However, it is important to note that the difference in performance between the two types of IRCs on this metric is not statistically significant, likely due to the low sample sizes for each of these groups. That said, taken together, these results suggest that Type I IRCs perform better relative to partisan-blind conceptions of fairness while Type II IRCs perform better relative to partisan-aware conceptions.

Additionally, while the full cross-sectional analysis did not find statistically significant advantages for either IRC type over legislative redistricting on the efficiency gap, exploratory analyses restricted to just the 2020 redistricting cycle revealed stronger effects. Specifically, Type II IRCs demonstrated a statistically significant outperformance compared to both legislative redistricting and Type I IRCs when considering the partisan-aware efficiency gap baseline in 2020 (see Appendix D). One possibility is that the refined techniques, data resources, and heightened scrutiny around partisan fairness metrics in recent years may have enabled Type II commissions to better optimize for the efficiency gap's focus on wasted votes and partisan-proportional representation than in 2010. Alternatively, the 2020 results could simply reflect a cohort effect, wherein the Type II IRC states in 2020 produced particularly fair maps for the efficiency gap that are being down-weighted in the full analysis because of California, which was the only Type II state during the 2010 redistricting cycle. Therefore, while these exploratory efficiency gap findings raise intriguing possibilities regarding the impacts of Type II IRC design and evolving prioritization of fairness metrics, they should be interpreted cautiously given the multiple plausible explanations and lack of statistical significance in the full ANOVA results. However, this efficiency gap finding does underscore the importance of not treating IRCs as a monolithic entity and potentially the limitations of the two-category division of IRCs that is employed for statistical purposes in the study. Instead, it may be more important to closely examine variation across individual states and redistricting cycles.

These results potentially help to reconcile contradictory findings within the existing literature regarding IRC performance. Studies that have identified partisan fairness improvements under IRCs may have been primarily capturing the partisan-aware advantages of Type II IRC institutions or the partisan-blind advantages of Type I IRCs. In contrast, research suggesting limited differences between IRCs and legislative redistricting could reflect a failure to distinguish between IRC types, obscuring the meaningful performance disparities this study identifies. Nevertheless, it is crucial to underscore the inherent limitations of this study's correlational design. While suggestive patterns arise regarding IRC adoption and subsequent fairness improvements when examining states over time, particularly for partisan-aware proportionality metrics, no definitive causal claims can be made. Numerous confounding variables, including broader political shifts, could plausibly contribute to observed fairness changes. Establishing causality would require more robust methodological approaches, such as synthetic control methods or regression discontinuity designs leveraging the staggered adoption of IRCs as a treatment.

Furthermore, it is important to acknowledge that this investigation's findings do not resolve all debates surrounding redistricting institutions and partisan fairness. Even Type II IRCs, despite their relative performance advantages, do not consistently optimize for all partisan fairness metrics and only appear to improve both on partisan-aware baselines. This reality underscores the inherent trade-offs and limitations in achieving a unified conception of perfectly fair representation through redistricting alone. Ultimately, these results reinforce the importance of carefully considering both the specific conceptions of partisan fairness that are being prioritized in redistricting reforms and their connection to the particulars of IRC institutional design. Solely pursuing redistricting independence without robust institutional safeguards like balanced membership, non-majoritarian approval mechanisms, and explicit partisanship considerations may be insufficient to consistently enhance partisan-aware fairness metrics. On the other hand, these design features also appear to reduce the ability of an IRC to consistently enhance partisan blind fairness metrics. As such, advocates and policymakers should carefully evaluate which fairness goals are paramount and implement IRC designs tailored to those priorities.

Conclusion

This study evaluated the performance of independent redistricting commissions (IRCs) in achieving partisan fairness across various conceptions, while also examining the potential impact of specific institutional design features. The findings suggest that IRCs, when considered collectively, outperform legislative redistricting processes on metrics associated with partisan-aware symmetric conceptions of fairness, such as the mean-median difference. However, their performance is less consistent on metrics tied to partisan-aware proportionality, like the efficiency gap, and partisan-blind fairness measures. Importantly, the analysis revealed that IRC performance can vary significantly based on specific institutional design characteristics. IRCs categorized as "Type II," which tended to exhibit features like balanced partisan composition, a heightened presence of unaffiliated commissioners, permitted considerations of partisanship, and non-majoritarian approval mechanisms, consistently outperformed both legislative redistricting and "Type I" IRCs on key partisan fairness metrics. This advantage was most pronounced on partisan-aware mean-median fairness and, potentially, the partisan-aware efficiency gap. However, it remains unclear whether the distinctions between Type I and Type II IRCs are inherent to their institutional designs or artifacts of the particular states and redistricting cycles examined. The findings do suggest, however, that IRC design elements, not just redistricting independence, are important when pursuing reforms aimed at improving partisan fairness. Solely pursuing an independent redistricting institution without robust institutional safeguards may be insufficient to consistently enhance partisan fairness, particularly partisan-aware conceptions.

Future research could explore several avenues to build upon these findings. First, more robust causal inference methods leveraging techniques like regression discontinuity designs could provide definitive evidence on the causal impact of IRC adoption and specific design features on partisan fairness outcomes. Additionally, as partisan fairness metrics continue to evolve and gain traction, examining how IRC performance varies across additional fairness metrics and measures of competitiveness would be valuable. Finally, investigating the potential trade-offs between optimizing for a given partisan fairness conception and other critical redistricting goals, such as minority representation or adherence to traditional criteria like compactness and respect for political subdivisions, could offer additional insights.

References

Barkow, Rachel E. 2010. “Insulating Agencies: Avoiding Capture through Institutional Design.” *Texas Law Review* 89 (1): 15–80.

Best, Robin E., Steve B. Lem, Daniel B. Magleby, and Michael D. McDonald. 2021. “Do Redistricting Commissions Avoid Partisan Gerrymanders?” *American Politics Research* 50 (3): 379–95. https://doi.org/10.1177/1532673X211053216.

“Brown v. Thomson, 462 U.S. 835 (1983).” Justia Law. Accessed February 28, 2024. https://supreme.justia.com/cases/federal/us/462/835/.

Cain, Bruce. 2012. “Redistricting Commissions: A Better Political Buffer?” *Yale Law Journal*, January. https://openyls.law.yale.edu/handle/20.500.13051/10007.

Cain, Bruce E., and Janet C. Campagna. 1987. “Predicting Partisan Redistricting Disputes.” *Legislative Studies Quarterly* 12 (2): 265–74. https://doi.org/10.2307/439923.

Cain, Bruce E, Wendy K Tam Cho, Yan Y Liu, and Emily R Zhang. 2018. “A Reasonable Bias Approach to Gerrymandering: Using Automated Plan Generation to Evaluate Redistricting Proposals.” *William and Mary Law Review* 59, no. 5 (2018): 1521–57. https://scholarship.law.wm.edu/wmlr/vol59/iss5/2.

Chen, Jowei, and David Cottrell. 2016. “Evaluating Partisan Gains from Congressional Gerrymandering: Using Computer Simulations to Estimate the Effect of Gerrymandering in the U.S. House.” *Electoral Studies* 44 (December): 329–40. https://doi.org/10.1016/j.electstud.2016.06.014.

Chen, Jowei, and Jonathan Rodden. 2013. “Unintentional Gerrymandering: Political Geography and Electoral Bias in Legislatures.” *Quarterly Journal of Political Science* 8 (3): 239–69. https://doi.org/10.1561/100.00012033.

“Criteria and Goals for Congressional Districts.” 2021. *Montana Districting and Apportionment Commission.*

https://leg.mt.gov/content/Districting/2020/Topics/Criteria/adopted-criteria-congressional-dac-july-2021.pdf.

Lewis, David E. *The Politics of Presidential Appointments: Political Control and Bureaucratic Performance*. Princeton University Press, 2008. https://doi.org/10.2307/j.ctt7rnqz.

Daley, David. "Voters Had Their Say. Partisans Ignored Them." *The New York Times*, September 29, 2021. https://www.nytimes.com/2021/09/29/opinion/redistricting-commissions-democracy.html.

DeFord, Daryl, Natasha Dhamankar, Moon Duchin, Varun Gupta, Mackenzie McPike, Gabe Schoenbach, and Ki Wan Sim. 2021. "Implementing Partisan Symmetry: Problems and Paradoxes." *Political Analysis* 31 (3): 305–24. https://doi.org/10.1017/pan.2021.49.

DeFord, Daryl, Moon Duchin, and Justin Solomon. 2020. "A Computational Approach to Measuring Vote Elasticity and Competitiveness." *Statistics and Public Policy* 7 (1): 69–86. https://doi.org/10.1080/2330443X.2020.1777915.

Duchin, Moon, Taissa Gladkova, Eugene Henninger-Voss, Ben Klingensmith, Heather Newman, and Hannah Wheelen. 2019. "Locating the Representational Baseline: Republicans in Massachusetts." *Election Law Journal: Rules, Politics, and Policy* 18 (4): 388–401. https://doi.org/10.1089/elj.2018.0537.

Duchin, Moon, and Gabe Schoenbach. 2023. "Redistricting for Proportionality." *The Forum* 20 (3–4): 371–93. https://doi.org/10.1515/for-2022-2064.

Engstrom, Erik J. 2013. *Partisan Gerrymandering and the Construction of American Democracy*. Legislative Politics and Policy Making. Ann Arbor: University of Michigan Press. https://catalog.hathitrust.org/Record/100163709.

Fifield, Benjamin, Kosuke Imai, Jun Kawahara, and Christopher T. Kenny. 2020. "The Essential Role of Empirical Validation in Legislative Redistricting Simulation." *Statistics and Public Policy* 7 (1): 52–68. https://doi.org/10.1080/2330443X.2020.1791773.

Gersen, Jacob, and Christopher Berry. 2010. “Agency Design and Distributive Politics.” *Public Law & Legal Theory*, October. https://chicagounbound.uchicago.edu/public_law_and_legal_theory/23.

Hanssen, F. Andrew. 1999. “The Effect of Judicial Institutions on Uncertainty and the Rate of Litigation: The Election Versus Appointment of State Judges.” *The Journal of Legal Studies* 28 (1): 205–32. https://doi.org/10.1086/468050.

Henderson, John A., Brian T. Hamel, and Aaron M. Goldzimer. 2018. “Gerrymandering Incumbency: Does Nonpartisan Redistricting Increase Electoral Competition?” *The Journal of Politics* 80 (3): 1011–16. https://doi.org/10.1086/697120.

Iancu, Bogdan. 2018. “Government and Governance: The Constitutional Politics of Institutional Neutrality.” In *Governance and Constitutionalism*. Routledge.

Iancu, Bogdan, and Elena Simina Tănăsescu. 2018. “Introduction.” In *Governance and Constitutionalism*. Routledge.

Jacobson, Gary C. 2015. “It’s Nothing Personal: The Decline of the Incumbency Advantage in US House Elections.” *The Journal of Politics* 77 (3): 861–73. https://doi.org/10.1086/681670.

Katz, Jonathan N., Gary King, and Elizabeth Rosenblatt. 2020. “Theoretical Foundations and Empirical Evaluations of Partisan Fairness in District-Based Democracies.” *American Political Science Review* 114 (1): 164–78. https://doi.org/10.1017/S000305541900056X.

Keena, Alex, Michael Latner, Charles Anthony Smith, and Anthony McGann. 2019. “Here’s How to Fix Partisan Gerrymandering, Now That the Supreme Court Kicked It Back to the States.” *The Washington Post*, July 2. https://www.washingtonpost.com/politics/2019/07/02/heres-how-fix-partisan-gerrymandering-now-that-supreme-court-kicked-it-back-states/.

Kenny, Christopher T., and Cory McCartan. 2021. “ALARM Project: 2020 Redistricting Data Files,” August. https://github.com/alarm-redist/census-2020/.

Kim, Sung Eun, and Johannes Urpelainen. 2017. "The Polarization of American Environmental Policy: A Regression Discontinuity Analysis of Senate and House Votes, 1971–2013." *Review of Policy Research* 34 (4): 456–84. https://doi.org/10.1111/ropr.12238.

King, Gary. 1990. "Electoral Responsiveness and Partisan Bias in Multiparty Democracies." *Legislative Studies Quarterly* 15 (2): 159–81. https://doi.org/10.2307/440124.

Kubin, Jeffrey C. 1996. "Case for Redistricting Commissions." *Texas Law Review* 75 (4): 837–72.

Levendusky, Matthew. 2009. *The Partisan Sort: How Liberals Became Democrats and Conservatives Became Republicans*. Chicago Studies in American Politics. Chicago, IL: University of Chicago Press. https://press.uchicago.edu/ucp/books/book/chicago/P/bo8212972.html.

Levendusky, Matthew S., Jeremy C. Pope, and Simon D. Jackman. 2008. "Measuring District-Level Partisanship with Implications for the Analysis of U.S. Elections." *The Journal of Politics* 70 (3): 736–53. https://doi.org/10.1017/S0022381608080729.

Levitt, Justin, and Doug Spencer. 2024. All About Redistricting. https://redistricting.lls.edu.

Lewis, David E. 2008. *The Politics of Presidential Appointments: Political Control and Bureaucratic Performance*. Princeton University Press. https://doi.org/10.2307/j.ctt7rnqz.

Li, Michael. 2021. "Early Lessons from the Current Redistricting Round | Brennan Center for Justice." Accessed February 28, 2024. https://www.brennancenter.org/our-work/analysis-opinion/early-lessons-current-redistricting-round.

Mason, Lilliana. 2015. "'I Disrespectfully Agree': The Differential Effects of Partisan Sorting on Social and Issue Polarization." *American Journal of Political Science* 59 (1): 128–45. https://doi.org/10.1111/ajps.12089.

McCartan, Cory, and Kosuke Imai. 2023. "Sequential Monte Carlo for Sampling Balanced and Compact Redistricting Plans." *The Annals of Applied Statistics* 17 (4). https://doi.org/10.1214/23-AOAS1763.

McDonald, Michael D., and Robin E. Best. 2015. "Unfair Partisan Gerrymanders in Politics and Law: A Diagnostic Applied to Six Cases Symposium: Common Cause Redistricting Competition Winners." *Election Law Journal* 14 (4): 312–30.

McDonald, Michael D., Daniel B. Magleby, Jonathan Krasno, Shawn J. Donahue, and Robin Best. 2018. "Making a Case for Two Paths Forward in Light of Gill v. Whitford." *Election Law Journal: Rules, Politics, and Policy* 17 (4): 315–27. https://doi.org/10.1089/elj.2018.0527.

McDonald, Michael P. 2004. "A Comparative Analysis of Redistricting Institutions in the United States, 2001–02." *State Politics & Policy Quarterly* 4 (4): 371–95. https://doi.org/10.1177/153244000400400402.

Miller, Peter, and Bernard Grofman. 2013. "Redistricting Commissions in the Western United States Symposium Issue: Foxes, Henhouses, and Commissions: Assessing the Nonpartisan Model in Election Administration, Redistricting, and Campaign Finance." *UC Irvine Law Review* 3 (3): 637–68.

Morrill, Richard L. 1987. "Redistricting, Region and Representation." *Political Geography Quarterly* 6 (3): 241–60. https://doi.org/10.1016/S0260-9827(87)80003-4.

Nagle, John F. 2019. "What Criteria Should Be Used for Redistricting Reform?" *Election Law Journal: Rules, Politics, and Policy* 18 (1): 63–77. https://doi.org/10.1089/elj.2018.0514.

Nelson, Matthew. 2023. "Independent Redistricting Commissions Are Associated with More Competitive Elections." *PS: Political Science & Politics* 56 (2): 207–12. https://doi.org/10.1017/S104909652200124X.

Pildes, Richard H. "Redistricting Reform and the 2018 Elections." 2018. *Harvard Law Review* (blog essay), October 26. https://harvardlawreview.org/blog/2018/10/redistricting-reform-and-the-2018-elections/.

Plener Cover, Benjamin. 2018. "Quantifying Partisan Gerrymandering: An Evaluation of the Efficiency Gap Proposal." *Stanford Law Review* 70, April. https://digitalcommons.law.uidaho.edu/faculty_scholarship/123.

"Redistricting Measures on the Ballot." n.d. Ballotpedia. Accessed February 28, 2024. https://ballotpedia.org/Redistricting_measures_on_the_ballot.

Robertson, Cassandra Burke. 2018. "Judicial Impartiality in a Partisan Era." *Florida Law Review* 70 (4): 739–76.

Sievert, Joel, and Seth C. McKee. 2019. "Nationalization in U.S. Senate and Gubernatorial Elections." *American Politics Research* 47 (5): 1055–80. https://doi.org/10.1177/1532673X18792694.

Stephanopoulos, Nicholas. 2016. "Race, Place, and Power." *Public Law & Legal Theory*, July. https://chicagounbound.uchicago.edu/public_law_and_legal_theory/624.

Stephanopoulos, Nicholas, and Eric McGhee. 2015. "Partisan Gerrymandering and the Efficiency Gap." *University of Chicago Law Review* 82 (March): 831–900.

Stephanopoulos, Nicholas O. 2018. "The Causes and Consequences of Gerrymandering." *William & Mary Law Review* 59 (5): 2115–58. https://scholarship.law.wm.edu/wmlr/vol59/iss5/14/.

Tabarrok, Alexander, and Eric Helland. 1999. "Court Politics: The Political Economy of Tort Awards." *The Journal of Law & Economics* 42 (1): 157–88. https://doi.org/10.1086/467421.

"Tennant v. Jefferson County, 567 U.S. 758 (2012)." Justia Law. Accessed February 28, 2024. https://supreme.justia.com/cases/federal/us/567/758/.

Vasilogambros, Matt. 2019. "The Tumultuous Life of an Independent Redistricting Commissioner." *Stateline* (blog), November 26. https://stateline.org/2019/11/26/the-tumultuous-life-of-an-independent-redistricting-commissioner/.

Wilson, Mark C., and Bernard N. Grofman. 2022. "Models of Inter-Election Change in Partisan Vote Share." *Journal of Theoretical Politics* 34 (4): 481–98. https://doi.org/10.1177/09516298221123263.

Yoshinaka, Antoine, and Chad Murphy. 2009. "Partisan Gerrymandering and Population Instability: Completing the Redistricting Puzzle." *Political Geography* 28 (8): 451–62. https://doi.org/10.1016/j.polgeo.2009.10.011.

Zhang, Emily Rong. 2021. "Bolstering Faith with Facts: Supporting Independent Redistricting Commissions with Redistricting Algorithms Democracy Reform Symposium." *California Law Review* 109 (3): 987–1018.

Zingher, Joshua N., and Jesse Richman. 2019. "Polarization and the Nationalization of State Legislative Elections." *American Politics Research* 47 (5): 1036–54. https://doi.org/10.1177/1532673X18788050.

Appendix A State Constitutional Language Governing IRC Design

Table 1: IRC Commission Appointment Mechanism

State	Commission Appointment Mechanism
Montana	**Political appointment**: “The majority and minority leaders of each house shall each designate one commissioner [...] the four commissioners shall select the fifth member [...] If the four members fail to select the fifth member within the time prescribed, a majority of the supreme court shall select him” (Mont. Const. art. V, § 14)
Washington	**Political appointment:** “The legislative leader of the two largest political parties in each house of the legislature shall appoint one voting member to the commission [...] the four appointed members, by an affirmative vote of at least three, shall appoint the remaining member [...] who shall be nonvoting [...] If any appointing authority fails to make the required appointment by the date established by this subsection, within five days after that date the supreme court shall make the required appointment (Wash. Const. art. II, § 43(2))
Idaho	**Political appointment:** “The leaders of the two largest political parties of each house of the legislature shall each designate one member and the state chairmen of the two largest political parties, determined by the vote cast for governor in the last gubernatorial election, shall each designate one member. In the event any appointing authority does not select the members [...], such members shall be appointed by the Supreme Court. No member of the commission may be an elected or appointed official in the state of Idaho at the time of designation or selection” (Idaho Const. art. III, § 2(2))
Alaska	**Political appointment:** “The governor shall appoint two members of the board. The presiding officer of the senate, the presiding officer of the house of representatives, and the chief justice of the supreme court shall each appoint one member of the board (Alaska Const. art. VI, § 8)”
Arizona	**Hybrid:** “The commission on appellate court appointments shall nominate candidates for appointment to the independent redistricting commission [...] The pool of candidates shall consist of twenty-five nominees, with ten nominees from each of the two largest political parties in Arizona based on party registration, and five who are not registered with either of the two largest political parties [...] the highest ranking officer elected by the Arizona house of representatives shall make one appointment to the independent redistricting commission from the pool of nominees, followed by one appointment from the pool made in turn by each of the following: the minority party leader of the Arizona house of representatives, the highest ranking officer elected by the Arizona senate, and the minority party leader of the Arizona senate. [...] the four independent redistricting commission members shall select by majority vote from the nomination pool a fifth member who shall not be registered with any party already represented on the independent redistricting commission and who shall serve as chair” (Ariz. Const. art. IV, pt. 2, § 1(4)-(8))
California	**Random selection:** “The State Auditor shall initiate an application process [...] From the applicant pool, the Applicant Review Panel shall select 60 of the most qualified applicants, including 20 who are registered with the largest political party in California based on registration, 20 who are registered with the second largest political party in California based on registration, and 20 who are not registered with either of the two largest political parties

State	Commission Appointment Mechanism
	in California based on registration [...] the President pro Tempore of the Senate, the Minority Floor Leader of the Senate, the Speaker of the Assembly, and the Minority Floor Leader of the Assembly may each strike up to two applicants from each subpool of 20 for a total of eight possible strikes per subpool [...] the State Auditor shall randomly draw eight names from the remaining pool of applicants [...] the eight commissioners shall review the remaining names in the subpools of applicants and appoint six applicants to the commission (Cal. Gov't Code § 8252(b)-(g))"
New York	**Political appointment:** "two members shall be appointed by the temporary president of the Senate [...] two members shall be appointed by the speaker of the assembly [...] two members shall be appointed by the minority leader of the senate [...] two members shall be appointed by the minority leader of the assembly [...] two members shall be appointed by the [above] eight members (N.Y. Const. art. III, §§ 5-b(a)-(c))"
Colorado	**Random selection:** "any person who seeks to serve on the commission must submit a completed application to nonpartisan staff [...] the chief justice of the Colorado supreme court shall designate a panel to review the applications [...] the panel, in a public meeting, shall randomly select by lot from all of the applicants who were found to meet the qualifications specified [...] , the panel shall identify fifty applicants who are affiliated with the state's largest political party, fifty applicants who are affiliated with the state's second largest political party, and fifty applicants who are unaffiliated with any political party [...] from the applicants identified in subsection (8)(a) of this section, the panel shall choose by lot six applicants to serve on the commission as follows: Two commissioners who are not affiliated with any political party; Two commissioners who are affiliated with the state's largest political party; and Two commissioners who are affiliated with the state's second largest political party (Colo. Const. art. V, §§ 44.1)"
Michigan	**Random selection:** "applications for commissioner [are made] available to the general public [and] [...] mail applications for commissioner to ten thousand Michigan registered voters, selected at random [and then] [...] randomly select 60 applicants from each pool of affiliating applicants and 80 applicants from the pool of non-affiliating applicants [...] the majority leader of the senate, the minority leader of the senate, the speaker of the house of representatives, and the minority leader of the house of representatives may each strike five applicants from any pool or pools, up to a maximum of 20 total strikes by the four legislative leaders [...] the secretary of state shall randomly draw the names of four commissioners from each of the two pools of remaining applicants affiliating with a major party, and five commissioners from the pool of remaining non-affiliating applicant (Mich. Const. art. IV, § 6(2))"

Table 2: IRC Partisan Composition

State	Effective Partisan Composition
Montana	**5 Members (2 Democrats, 2 Republicans, 1 Unaffiliated):** “A commission of five citizens [...], shall be selected to prepare a plan for redistricting [...]. The majority and minority leaders of each house shall each designate one commissioner. [...], the four commissioners shall select the fifth member, who shall serve as chairman of the commission” (Mont. Const. art. V, § 14)
Washington	**5 Members (2 Democrats, 2 Republicans, 1 Unaffiliated (non-voting)):** “The legislative leader of the two largest political parties in each house of the legislature shall appoint one voting member to the commission [...] the four appointed members, by an affirmative vote of at least three, shall appoint the remaining member [...] who shall be nonvoting” (Wash. Const. art. II, § 43(2))
Idaho	**6 Members (3 Democrats, 3 Republicans):** “The leaders of the two largest political parties of each house of the legislature shall each designate one member and the state chairmen of the two largest political parties, determined by the vote cast for governor in the last gubernatorial election, shall each designate one member” (Idaho Const. art. III, § 2(2))
Alaska	**5 Members (In 2020: 3 Republicans, 2 unaffiliated; In 2010: 4 Republicans, 1 Democrat):** “The governor shall appoint two members of the board. The presiding officer of the senate, the presiding officer of the house of representatives, and the chief justice of the supreme court shall each appoint one member of the board (Alaska Const. art. VI, § 8)”
Arizona	**5 Members (2 Democrats, 2 Republicans, 1 Unaffiliated):** “The independent redistricting commission shall consist of five members. No more than two members of the independent redistricting commission shall be members of the same political party. Of the first four members appointed, no more than two shall reside in the same county. Each member shall be a registered Arizona voter who has been continuously registered with the same political party or registered as unaffiliated with a political party for three or more years immediately preceding appointment” (Ariz. Const. art. IV, pt. 2, § 1(3))
California	**14 Members (5 Democrats, 5 Republicans, 4 Unaffiliated):** “the State Auditor shall randomly draw eight names from the remaining pool of applicants as follows: three from the remaining subpool of applicants registered with the largest political party in California based on registration, three from the remaining subpool of applicants registered with the second largest political party in California based on registration, and two from the remaining subpool of applicants who are not registered with either of the two largest political parties in California based on registration [...] the eight commissioners shall review the remaining names in the subpools of applicants and appoint six applicants to the commission as follows: two from the remaining subpool of applicants registered with the largest political party in California based on registration, two from the remaining subpool of applicants registered with the second largest political party in California based on registration, and two from the remaining subpool of applicants who are not registered with either of the two largest political parties in California based on registration” (Cal. Gov’t Code § 8252(f)(g))
New York	**10 Members (4 Democrats, 4 Republicans, 2 Unaffiliated):** “two members shall be appointed by the temporary president of the Senate [...] two members shall be appointed by the speaker of the assembly [...] two members shall be appointed by the minority leader of the senate [...] two members shall be

State	Effective Partisan Composition
	appointed by the minority leader of the assembly [...] two members shall be appointed by the [above] eight members (N.Y. Const. art. III, §§ 5-b(a)-(c))"
Colorado	**12 Members (4 Democrats, 4 Republicans, 4 Unaffiliated)** "The commission consists of twelve members [...] The panel of judges must ensure that the commission includes four commissioners who are not affiliated with any political party, four commissioners who are affiliated with the state's largest political party, and four commissioners who are affiliated with the state's second largest political party" (Colo. Const. art. V, §§ 44.1)
Michigan	**13 Members (4 Democrats, 4 Republicans, 5 Unaffiliated):** "the secretary of state shall randomly draw the names of four commissioners from each of the two pools of remaining applicants affiliating with a major party, and five commissioners from the pool of remaining non-affiliating applicants (Mich. Const. art. IV, § 6(2))"

Table 3: Maximum Partisanship Considerations

State	Maximum Partisanship Considerations
Montana	**Unspecified**
Washington	**Competitiveness:** “The commission's plan shall not be drawn purposely to favor or discriminate against any political party or group”; “The commission shall exercise its powers to provide fair and effective representation and to encourage electoral competition. The commission's plan shall not be drawn purposely to favor or discriminate against any political party or group (Rev. Code Wash. § 44.05.090)”
Idaho	**Prohibited:** “The total state population as reported by the U.S. census bureau, and the population of subunits determined therefrom, shall be exclusive permissible data (Idaho Code § 72-1506)”
Alaska	**Unspecified**
Arizona	**Competitiveness:** “To the extent practicable, competitive districts should be favored where to do so would create no significant detriment to the other goals” [...] Party registration and voting history data shall be excluded from the initial phase of the mapping process but may be used to test maps for compliance with the above goals ([Ariz. Const. art. IV, pt. 2, § 1)”
California	**Non-discrimination:** “Districts shall not be drawn for the purpose of favoring or discriminating against an incumbent, political candidate, or political party (Cal. Const. art. XXI, § 2(e))”
New York	**Non-discrimination:** “Districts shall not be drawn to discourage competition or for the purpose of favoring or disfavoring incumbents or other particular candidates or political parties (N.Y. Const. art. III, § 4(c)(5))”
Colorado	**Competitiveness:** “Thereafter, the commission shall, to the extent possible, maximize the number of politically competitive districts [...] "competitive" means having a reasonable potential for the party affiliation of the district's representative to change at least once between federal decennial censuses. Competitiveness may be measured by factors such as a proposed district's past election results, a proposed district's political party registration data, and evidence based analyses of proposed districts (Colo. Const. art. V, §§ 44(3))”
Michigan	**Partisan Fairness:** “Districts shall not provide a disproportionate advantage to any political party. A disproportionate advantage to a political party shall be determined using accepted measures of partisan fairness” (Mich. Const. art. IV, § 6(13)(d)).

Table 4: Redistricting Plan Final Approval Requirements

State	Approval Requirements
Montana	**Majoritarian:** “A motion may be adopted only on the affirmative vote of a majority of commission members” (“Operating Procedures”)
Washington	**Majoritarian (w/cross-party support):** “A motion may be adopted only on the affirmative vote of a majority of commission members” (Wash. Const. art. II, § 43(6))
Idaho	**Majoritarian (w/cross-party support):** “Any final action of the commission shall be by a vote of two-thirds (2/3) [4/6 members: a true majority] of the full membership of the commission” (Idaho Code § 72-1505(5)
Alaska	**Majoritarian:** “Adoption of a final redistricting plan shall require the affirmative votes of three members of the Redistricting Board (Alaska Const. art. VI § 10(b))”
Arizona	**Majoritarian:** “Three or more affirmative votes are required for any official action” (Ariz. Const. art. IV, pt. 2, § 1(12))
California	**Non-majoritarian:** 9/14 votes (3 Democrats, 3 Republicans, 3 Unaffiliated) “The four final redistricting maps must be approved by at least nine affirmative votes which must include at least three votes of members registered from each of the two largest political parties in California based on registration and three votes from members who are not registered with either of these two political parties (Cal. Const. art. XXI, § 2(c)(5))”
New York	**Non-majoritarian:** 7/10 votes (Minimum 1 Democrat and 1 Republican*) “In the event that the speaker of the assembly and the temporary president of the senate are members of the same political party, approval of a redistricting plan and implementing legislation by the commission for submission to the legislature shall require the vote in support of its approval by at least seven members including at least one member appointed by each of the legislative leaders. (2) In the event that the speaker of the assembly and the temporary president of the senate are members of two different political parties, approval of a redistricting plan by the commission for submission to the legislature shall require the vote in support of its approval by at least seven members including at least one member appointed by the speaker of the assembly and one member appointed by the temporary president of the senate” (N.Y. Const. art. III, § 5-b(f))
Colorado	**Non-majoritarian:** 8/12 votes (incl. 2 Unaffiliated) “Adoption of the final plan for submission to the supreme court and the adoption of a revised plan after a plan is returned to the commission from the supreme court requires the affirmative vote of at least eight commissioners, including the affirmative vote of at least two commissioners who are unaffiliated with any political party” (Colo. Const. art. V, §§ 44.2(2))
Michigan	**Non-majoritarian:** 7/13 votes (2 Democrats, 2 Republicans, 2 Unaffiliated) “A final decision of the commission to adopt a redistricting plan requires a majority vote of the commission, including at least two commissioners who affiliate with each major party, and at least two commissioners who do not affiliate with either major party” (Mich. Const. art. IV, § 6(14)(c))

Appendix B Hyperparameter Settings and Diagnostics for SMC Simulations

State	Legislative Level	Cycle	Districts n [a]	Simulation n [b]	Population Tempering [c]	Seq. a [d]	Plan Diversity Lower [e]	Plan Diversity Upper [f]	Bias $\hat{R}$ [g]	Mean-Median $\hat{R}$ [h]	Efficiency Gap $\hat{R}$ [i]
AL	Congress	2020	7	10000	0	0.5	0.718	0.890	1.003	1.004	1.001
AL	State Senate	2020	35	15000	0.01	0.5	0.955	1.060	1.011	1.049	1.047
AK	State Senate	2020	20	15000	0.02	0.5	0.730	0.884	1.007	1.015	1.041
AR	Congress	2020	4	10000	0	0.5	0.562	0.851	1.000	1.000	1.000
AR	State Senate	2020	35	20000	0	0.5	0.937	1.038	1.028	1.006	1.024
AZ	Congress	2020	9	10000	0	0.5	0.916	1.107	1.006	1.000	1.002
AZ	Congress	2010	9	15000	0	0.5	0.907	1.080	1.003	1.002	1.006
AZ	State Senate	2020	30	15000	0.02	0.5	1.000	1.118	1.001	1.027	1.012
AZ	State Senate	2010	30	20000	0	0.5	1.002	1.111	1.018	1.017	1.006
CA	State Senate	2020	40	50000	0	0.5	0.957	1.058	1.010	1.034	1.028
CA	State Senate	2010	40	50000	0	0.5	0.977	1.065	1.003	1.007	1.052
CA	Congress	2020	52	50000	0.01	0.5	0.945	1.041	1.040	1.064	1.093
CA	Congress	2010	53	50000	0	0.5	0.924	1.047	1.202	1.004	1.146
CO	Congress	2010	7	10000	0	0.5	0.854	1.059	1.009	1.001	1.001
CO	Congress	2020	8	10000	0	0.5	0.876	1.073	1.006	1.001	1.002
CO	State Senate	2020	35	20000	0.02	0.5	1.005	1.121	1.018	1.049	1.010
CT	Congress	2020	5	10000	0	0.5	0.676	0.921	1.000	1.001	1.001
CT	State Senate	2020	36	20000	0.01	0.5	0.927	1.057	1.046	1.003	1.025
DE	State Senate	2020	21	20000	0.01	0.5	0.798	0.944	1.034	1.032	1.008
FL	Congress	2020	28	20000	0	0.5	0.846	0.974	1.007	1.037	1.003
FL	State Senate	2020	40	20000	0.015	0.5	0.933	1.013	1.002	1.047	1.015
GA	Congress	2020	14	10000	0	0.5	0.961	1.097	1.001	1.004	1.014

HI	State Senate	2020	25	20000	0.02	0.5	0.689	0.836	1.034	1.080	1.014
IA	Congress	2020	4	10000	0	0.5	0.565	0.863	1.002	1.000	1.001
ID	State Senate	2020	35	40000	0.01	0.5	0.883	0.999	1.015	1.286	1.054
ID	State Senate	2010	35	40000	0	0.5	0.875	0.992	1.001	1.085	1.007
IL	Congress	2020	17	10000	0	0.5	1.000	1.118	1.008	1.011	1.004
IN	Congress	2020	9	10000	0	0.5	0.770	0.956	1.017	1.010	1.001
IN	State Senate	2020	50	40000	0.01	0.5	0.472	1.100	1.013	1.205	1.003
KS	Congress	2020	4	10000	0	0.5	0.498	0.825	1.000	1.000	1.001
KS	State Senate	2020	40	10000	0	0.5	0.962	1.055	1.001	1.054	1.039
KY	Congress	2020	6	10000	0	0.5	0.614	0.842	1.002	1.000	1.006
KY	State Senate	2020	38	15000	0.01	0.5	0.964	1.052	1.004	1.028	1.041
LA	Congress	2020	6	10000	0	0.5	0.611	0.877	1.001	1.001	1.000
LA	State Senate	2020	39	40000	0.01	0.5	1.010	1.106	1.036	1.013	1.082
MA	Congress	2020	9	10000	0	0.5	0.771	0.947	1.022	1.004	1.024
MA	State Senate	2020	40	40000	0.01	0.5	0.998	1.101	1.003	1.048	1.123
MD	Congress	2020	8	10000	0	0.5	0.816	0.989	1.000	1.000	1.000
ME	Congress	2020	35	15000	0.01	0.5	0.282	0.899	1.088	1.131	1.053
ME	State Senate	2020	35	10000	0	0.5	0.713	0.890	1.036	1.056	1.065
MI	Congress	2020	13	10000	0	0.5	0.936	1.080	1.006	1.001	1.030
MI	Congress	2010	14	15000	0	0.5	0.968	1.105	1.002	1.006	1.024
MI	State Senate	2020	38	20000	0.01	0.5	1.054	1.132	1.035	1.026	1.010
MI	State Senate	2010	38	30000	0.01	0.5	1.040	1.129	1.048	1.059	1.012
MN	Congress	2020	8	10000	0	0.5	0.886	1.058	1.001	1.004	1.000
MO	Congress	2020	8	10000	0	0.5	0.706	0.907	1.002	1.001	1.002
MO	State Senate	2020	34	10000	0	0.5	0.965	1.057	1.023	1.002	1.058
MS	Congress	2020	4	10000	0	0.5	0.480	0.748	1.001	1.000	1.000
NC	Congress	2020	14	10000	0	0.5	0.872	1.004	1.003	1.009	1.021
NC	State Senate	2020	50	10000	0.01	0.35	0.510	1.087	1.001	1.036	1.045
NE	Congress	2020	3	10000	0	0.5	0.397	0.726	1.001	1.001	1.001

NE	State Senate	2020	49	10000	0	0.5	0.896	1.011	1.086	1.089	1.008
NH	State Senate	2020	24	10000	0	0.5	0.696	0.888	1.048	1.074	1.023
NJ	Congress	2020	12	10000	0	0.5	0.879	1.023	1.006	1.005	1.007
NM	Congress	2020	3	10000	0	0.5	0.429	0.821	1.000	1.002	1.001
NM	State Senate	2020	42	10000	0	0.5	0.869	0.987	1.008	1.019	1.022
NV	Congress	2020	4	10000	0	0.5	0.552	0.741	1.003	1.004	1.000
NV	State Senate	2020	21	10000	0	0.5	0.972	1.109	1.007	1.025	1.031
NY	Congress	2020	26	10000	0	0.5	0.888	1.007	1.028	1.003	1.045
NY	Congress	2010	27	40000	0.02	0.5	0.920	1.035	1.004	1.008	1.011
OH	Congress	2020	15	10000	0	0.5	0.948	1.085	1.009	1.018	1.026
OH	State Senate	2020	33	10000	0	0.5	1.037	1.125	1.005	1.014	1.007
OK	Congress	2020	5	10000	0	0.5	0.816	1.019	1.000	1.000	1.000
OK	State Senate	2020	48	10000	0	0.5	0.238	1.093	1.004	1.018	1.012
OR	Congress	2020	6	10000	0	0.5	0.774	0.996	1.003	1.014	1.004
OR	State Senate	2020	30	10000	0	0.5	0.937	1.064	1.003	1.012	1.008
PA	Congress	2020	17	10000	0	0.5	0.916	1.043	1.008	1.014	1.013
PA	State Senate	2020	50	30000	0.01	0.5	1.024	1.113	1.057	1.043	1.064
RI	State Senate	2020	38	10000	0	0.5	0.867	0.988	1.023	1.005	1.022
SC	Congress	2020	7	15000	0	0.5	0.697	0.916	1.000	1.000	1.001
SC	State Senate	2020	46	30000	0.01	0.5	0.985	1.081	1.017	1.068	1.118
SD	State Senate	2020	35	30000	0.01	0.5	0.768	0.892	1.010	1.038	1.062
TN	Congress	2020	9	10000	0	0.5	0.709	0.884	1.002	1.002	1.002
TN	State Senate	2020	33	30000	0.01	0.5	0.962	1.058	1.003	1.005	1.032
TX	State Senate	2020	31	10000	0	0.5	0.990	1.089	1.041	1.023	1.047
UT	Congress	2020	4	10000	0	0.5	0.656	0.968	1.004	1.003	1.004
UT	State Senate	2020	29	30000	0.01	0.5	0.983	1.095	1.033	1.038	1.041
VA	Congress	2020	11	10000	0	0.5	0.801	0.960	1.016	1.029	1.004
VT	State Senate	2020	13	30000	0.01	0.5	0.709	0.903	1.003	1.009	1.005

WA	Congress	2020	10	10000	0	0.5	0.811	0.993	1.003	1.005	1.001
WA	Congress	2010	10	15000	0	0.5	0.820	1.000	1.006	1.005	1.003
WA	State Senate	2020	49	50000	0	0.5	1.047	1.124	1.013	1.001	1.042
WI	Congress	2020	8	10000	0	0.5	0.808	0.977	1.009	1.006	1.002
WI	State Senate	2020	33	10000	0	0.5	1.006	1.106	1.009	1.019	1.018
WV	State Senate	2020	17	10000	0	0.5	0.757	0.887	1.011	1.019	1.003

Notes:

a. The number of legislative districts that are drawn in the particular model.
b. The number of simulations produced to approximate the sampling space of possible redistricting plans.
c. The chosen value of the population tempering hyperparameter. Adjusted to improve the accuracy and diversity of the simulations.
d. The chosen value of the sequence alpha hyperparameter. Adjusted to improve the accuracy and diversity of the simulations. See McCartan & Imai (2023) for more details.
e. The lower end of the 90% confidence interval for the plan diversity diagnostic test. Values close to or above one indicate that the redistricting plans simulated are sufficiently distinct from one another for analysis. See McCartan & Imai (2023) for more details.
f. The upper end of the 90% confidence interval for the plan diversity diagnostic test. See McCartan & Imai (2023) for more details.
g. The Gelman–Rubin $\hat{R}$ statistic for the partisan bias metric. This measures the convergence of the metric across the redistricting simulations that were run in parallel. Values below 1.05 are preferred, although due to computational constraints, values near or below 1.1 are also considered to demonstrate sufficient convergence. See McCartan & Imai (2023) for more details.
h. The Gelman–Rubin R statistic for the mean-median difference metric. See McCartan & Imai (2023) for more details.
i. The Gelman–Rubin $\hat{R}$ statistic for the efficiency gap metric. See McCartan & Imai (2023) for more details.

Appendix C Proportion of SMC Plans closer to Each Fairness Metric and Baseline by State, Year, and Level

			Partisan Bias		Mean-Median Difference		Efficiency Gap	
State	Cycle	Legislative Level	Partisan Blind	Partisan-Aware	Partisan Blind	Partisan-Aware	Partisan Blind	Partisan-Aware
AL	2020	Congress	0.999	0.999	0.932	0.958	0.835	0.155
AZ	2020	Congress	0.766	0.787	0.424	0.614	0.307	0.624
AR	2020	Congress	0.000	0.202	0.000	0.357	0.000	0.041
CA	2020	Congress	1.000	0.652	1.000	0.460	0.982	0.010
CO	2020	Congress	0.622	0.622	0.539	0.664	0.021	0.021
CT	2020	Congress	0.000	0.000	0.201	0.023	0.251	0.497
FL	2020	Congress	0.996	0.999	0.788	0.882	0.999	0.999
GA	2020	Congress	0.187	0.551	0.999	1.000	0.402	0.670
IL	2020	Congress	0.000	0.253	0.263	0.663	1.000	0.983
IN	2020	Congress	0.907	0.934	0.862	0.922	0.554	0.786
IA	2020	Congress	0.544	0.555	0.584	0.794	0.947	0.947
KS	2020	Congress	0.887	0.887	0.692	0.705	0.842	0.842
KY	2020	Congress	0.141	0.521	0.744	0.880	0.355	0.447
LA	2020	Congress	0.911	0.969	1.000	1.000	0.587	0.144
MD	2020	Congress	0.885	0.476	0.909	0.613	0.677	0.339
MA	2020	Congress	0.702	0.835	0.821	0.922	0.000	0.000
MI	2020	Congress	0.000	0.212	0.751	0.108	0.906	0.023
MN	2020	Congress	0.131	0.509	0.450	0.714	0.107	0.636
MS	2020	Congress	0.863	0.863	0.999	0.999	0.484	0.214
MO	2020	Congress	0.048	0.315	0.928	0.975	0.592	0.843
NE	2020	Congress	0.000	0.035	0.186	0.341	0.170	0.358
NV	2020	Congress	0.832	0.600	0.959	0.941	0.134	0.485
NJ	2020	Congress	0.140	0.000	0.578	0.293	0.549	0.808
NM	2020	Congress	0.841	0.000	0.857	0.005	0.929	0.929
NC	2020	Congress	0.439	0.000	0.047	0.405	0.513	0.026
OH	2020	Congress	0.574	0.726	0.195	0.466	0.717	0.812
OK	2020	Congress	0.557	0.092	0.417	0.247	0.000	0.029
OR	2020	Congress	0.506	0.519	0.330	0.249	0.978	0.980
PA	2020	Congress	0.840	0.012	0.967	0.017	0.997	0.008
SC	2020	Congress	0.973	0.973	0.833	0.911	0.803	0.938

TN	2020	Congress	0.000	0.147	0.000	0.459	1.000	1.000
UT	2020	Congress	0.939	0.544	0.471	0.084	0.854	0.856
VA	2020	Congress	0.533	0.190	0.748	0.733	0.590	0.216
WA	2020	Congress	0.067	0.435	0.511	0.146	0.294	0.481
WI	2020	Congress	0.555	0.729	0.658	0.828	0.734	0.900
AL	2020	State Senate	0.992	1.000	0.993	1.000	0.699	0.017
AK	2020	State Senate	0.589	0.753	0.793	0.876	0.309	0.651
AZ	2020	State Senate	0.377	0.610	0.237	0.644	0.354	0.643
AR	2020	State Senate	0.956	0.974	0.593	0.803	0.325	0.289
CA	2020	State Senate	1.000	0.682	0.977	0.309	0.915	0.018
CO	2020	State Senate	0.888	0.914	0.302	0.693	0.000	0.458
CT	2020	State Senate	0.960	0.972	0.802	0.873	0.286	0.411
DE	2020	State Senate	0.919	0.967	0.583	0.806	0.334	0.629
FL	2020	State Senate	0.778	0.879	1.000	1.000	0.819	0.934
HI	2020	State Senate	0.737	0.798	0.880	0.947	1.000	0.000
ID	2020	State Senate	0.469	0.215	0.182	0.601	0.728	0.095
IN	2020	State Senate	1.000	1.000	0.394	0.716	0.932	0.961
KS	2020	State Senate	0.887	0.952	0.041	0.559	0.935	0.946
KY	2020	State Senate	0.964	0.002	0.979	0.019	0.522	0.762
LA	2020	State Senate	0.997	0.997	1.000	1.000	0.919	0.022
ME	2020	State Senate	0.312	0.571	0.448	0.229	0.968	0.947
MA	2020	State Senate	0.318	0.241	0.415	0.684	0.960	0.009
MI	2020	State Senate	0.847	0.038	0.998	0.002	0.999	0.001
MO	2020	State Senate	0.982	0.002	0.800	0.910	0.259	0.624

NE	2020	State Senate	0.729	0.000	0.172	0.270	0.210	0.567
NV	2020	State Senate	0.955	0.976	0.746	0.855	0.129	0.554
NH	2020	State Senate	0.998	0.998	1.000	1.000	1.000	1.000
NM	2020	State Senate	0.999	0.999	1.000	1.000	0.496	0.217
NY (LR)	2020	Congress	0.134	0.602	0.853	0.061	0.996	0.996
NY (IRC)	2020	Congress	0.493	0.109	0.853	0.061	0.377	0.281
NC	2020	State Senate	0.000	0.584	0.266	0.611	0.031	0.551
OH	2020	State Senate	0.000	0.430	0.585	0.756	0.704	0.169
OK	2020	State Senate	0.640	0.074	0.986	0.007	0.034	0.524
OR	2020	State Senate	0.932	0.000	0.969	0.015	0.698	0.151
PA	2020	State Senate	0.455	0.094	0.682	0.136	0.591	0.201
RI	2020	State Senate	0.000	0.402	0.766	0.856	0.741	0.124
SC	2020	State Senate	1.000	1.000	1.000	1.000	0.992	0.004
SD	2020	State Senate	0.342	0.573	0.430	0.699	0.725	0.125
TN	2020	State Senate	0.193	0.550	0.619	0.156	0.313	0.650
TX	2020	State Senate	1.000	1.000	1.000	1.000	0.995	0.998
UT	2020	State Senate	0.437	0.109	0.938	0.953	0.833	0.920
VT	2020	State Senate	0.499	0.000	0.251	0.224	0.000	0.590
WA	2020	State Senate	0.994	0.209	0.995	0.174	0.813	0.891
WV	2020	State Senate	0.770	0.805	0.303	0.650	0.000	0.051
WI	2020	State Senate	0.724	0.792	0.993	0.999	0.741	0.815
AZ	2010	Congress	0.757	0.881	0.369	0.643	0.852	0.964
CA	2010	Congress	0.557	0.072	0.176	0.405	0.754	0.857
CO	2010	Congress	0.204	0.204	0.910	0.952	0.998	0.998

ID	2010	Congress	0.000	0.000	0.000	0.000	1.000	1.000
MI	2010	Congress	0.499	0.794	0.992	0.996	1.000	0.554
NY	2010	Congress	0.461	0.353	0.479	0.205	1.000	0.000
WA	2010	Congress	0.293	0.662	0.199	0.237	0.016	0.322
AZ	2010	State Senate	0.865	0.906	0.745	0.890	0.860	0.911
CA	2010	State Senate	0.929	0.000	0.984	0.002	0.973	0.869
ID	2010	State Senate	0.898	0.920	0.397	0.572	0.864	0.025
MI	2010	State Senate	0.993	0.993	0.816	0.934	0.001	0.125

Notes: The values in this table reflect the proportion of simulated redistricting plans that fall closer to the given partisan fairness baseline for the given metric. Low values indicate that the enacted districting plan performs well on the given particular fairness metric and baseline, while high values indicate that it performs poorly.

Appendix D Analysis of IRC Performance Restricted to the 2020 Redistricting Cycle

	Partisan Bias		Mean-Median Difference		Efficiency Gap	
	Partisan Blind	Partisan-Aware	Partisan Blind	Partisan-Aware	Partisan Blind	Partisan-Aware
LR[a] vs IRC I[b]	-0.0616[c] (0.9158)[d]	-0.0431 (0.9592)	-0.1362 (0.5613)	-0.122 (0.6740)	-0.1257 (0.6704)	0.0376 (0.9650)
LR vs IRC II	0.0874 (0.8160)	-0.0836 (0.8352)	0.1145 (0.6247)	-0.3039 (0.0691)*	0.0068 (0.9986)	-0.4106 (0.0107)**
IRC I vs IRC II	0.1490 (0.7387)	-0.0405 (0.9784)	0.2508 (0.3184)	-0.1812 (0.6024)	0.1324 (0.7689)	-0.4482 (0.0577)*
F-Statistic	0.29 (0.749)	0.188 (0.829)	1.059 (0.352)	2.744 (0.071)*	0.374 (0.689)	4.591 (0.0133)**

Notes:

a. The first institution listed is the reference category
b. The second institution listed is the comparison category
c. The difference in the proportion of simulated plans closer to the given baseline for the particular fairness metric. Negative values indicate that the comparison category outperforms the reference category
d. The p-value of the estimated difference in proportions

* Indicates statistical significance at the 0.1 level.

** Indicates statistical significance at the 0.05 level.