Article

Who controls U.S. politics? An analysis of major political endorsements in U.S. midterm elections

Ziang Huang[1], Grayson York[2]
[1]Shenzhen College of International Education, Shenzhen, China
[2]Duke University, Durham, North Carolina

SUMMARY

In this study, we investigated the underlying patterns, rationales, and effects of major political endorsements in the United States midterm elections, specifically focusing on the endorsement patterns and impacts of Donald Trump and Bernie Sanders in the 2018 and 2020 midterms. In order to uncover such patterns, we examined five demographic variables (Cook's Partisan Voter Index (PVI), bachelor's degree attainment, percentage white, percentage aged 65 and above, and median income) as well as two candidate-specific variables (incumbency and progressivity). Through employing density-based clustering methods and logistic regressions, we discovered a dichotomy of pragmatic and ideological endorsements in U.S. politics. We thus hypothesized that all political endorsements can be predicted by a combination of ideological and demographic factors, and fall somewhere along the dichotomy. Trump's endorsements were 97% predictable through logistic regressions on demographic variables alone, with PVI contributing the majority (55%) and education level contributing 28%; meanwhile, Sanders' endorsements were only 67% demographically predictable, but 80% predictable through the progressivity of individual candidates alone. We therefore concluded a trend of Trump basing his endorsements upon state demographics instead of individual candidates, and Sanders endorsing based on whether or not candidates align with a progressive ideology. Using a support vector machine model for binary classification, we found that Trump's and Sanders' endorsements contributed 11% and 6% to election victories of their respective endorsed candidates, significantly behind factors such as PVI (38%), bachelor's attainment (25%) and incumbency (14%), justifying trends such as Congressional stagnation, widening educational disparities in electorates, and political polarization.

INTRODUCTION

Former president Donald Trump endorsed over 200 candidates during the 2022 election cycle; in total, he has endorsed 551 candidates across both general and primary elections since he took office until today (1). Facilitated by the convenience of social media, other major political figures, including Bernie Sanders and Barack Obama, have endorsed similar numbers of candidates, mostly in the U.S. midterms (2). Given the extensive involvement and influence of major politicians in swaying the midterm electorate, it is crucial to determine what effects endorsements have, if any, on the political landscape of the United States today. The prolific endorsement patterns of Trump, Sanders, and many other politicians, alongside the emergence of social media as a platform for political influence, give ample opportunities to analyze both the rationale and the effects of endorsements by major U.S. politicians. Trump's endorsements, in particular, have come under significant public scrutiny. Touting a continued 73% approval rating and a 60% preference rate for the 2024 presidential primaries among Republicans even after the 2020 Capitol insurrection and several criminal investigations, Trump remains possibly the most influential figure in the Republican Party (3). He has reiterated this several times in public statements, especially concerning the 92% victory rate of his endorsed candidates in Republican primaries (4).

The true intent of Trump's endorsements alongside their actual impact on general elections, however, remains shrouded in mystery. From the inception of his political career onwards, Trump has remained a divisive figure, polarizing America's Democrats and Republicans towards either extreme hatred or unconditional support (3). With a record-low approval rating of 41% during his one-term presidency, Trump's support of a candidate is just as likely to galvanize opposition from Democratic voters as support from Republican voters, a phenomenon evidenced by research from Ballard *et al.* (3). Of particular interest are the driving factors behind Trump's proclaimed endorsement successes: of the 33 candidates within the 33-0 primary record he touted in March 2022, five were uncontested in their primaries and several others won with vote shares upwards of 80% in uncompetitive elections (5). Of the 235 races he has issued endorsements in overall, more than 200 saw his candidates facing either no opposition or opponents who were lacking in campaign funding (5). In general elections, Trump has generally shown severe reluctance in endorsing candidates in firmly Democratic states like Washington or Massachusetts, leading to claims that his endorsements serve little purpose besides padding his own record (4). Therefore, identifying the underlying patterns and demographic factors that drive Trump's endorsements as well as their actual impact on elections is key towards determining the extent to which Trump's influence lingers over American politics, even after the end of his presidency.

On the other side of the aisle, a rapidly surging wave of progressive social-democratic politics has appeared at the forefront of the Democratic Party. At the center of it lies Senator Bernie Sanders from Vermont, two-time runner-up for Democratic Nominee for President in 2016 and 2020 (though officially an independent) and the driving force behind a self-described "political revolution" (6). Since bringing left-wing policies such as Medicare for All and free higher education

to the forefront of the Democratic Party platform, Sanders has applied his considerable political leverage to advance the progressive wing of the Democratic Party, endorsing candidates such as Ocasio-Cortez and Omar in their election and reelection campaigns to Congress (2). Though Sanders is unlikely to contest the presidency in the future due to his advanced age, his platform as well as his personal influence remains pervasive over the Biden administration, with recent policy successes in student debt elimination and marijuana legalization echoing Sanders' rhetoric (7). Studying the effects and patterns in Sanders' endorsements serves as an effective avenue towards quantifying and analyzing the extent of Sanders' influence over American electoral politics, and by extension, the influence of the progressive left.

We viewed Sanders as nearly the perfect antithesis to Trump - both initially anti-establishment, populist politicians who campaigned upon championing the working-class cause, yet in radically different ways. Nowhere, perhaps, is this contrast more evident than Trump and Sanders' endorsement patterns. While many have accused Trump of making "safe" endorsements in securely Republican states to pad his own record, Sanders, in hundreds of endorsements throughout the past two election cycles (2018 and 2020), has indiscriminately endorsed any candidates who were in explicit support of the progressive political movement without regard for state demographics or political climate (2). Through focusing on their endorsed candidates and endorsement effects, we hypothesized that Trump and Sanders represent two juxtaposing patterns of political endorsements in the United States. While Trump embodies a predominantly pragmatic approach in which politicians endorse based on favorable state demographics, including PVI, education level, race distribution and income, in order to support their own credibility, Sanders pursues a predominantly ideological approach that determines support for candidates based on their individual ideologies and whether or not they correspond to a particular cause. Additionally, we hypothesized that endorsements by both Trump and Sanders would ultimately have relatively insignificant effects on the outcomes of both past (2018 and 2020) and future midterm general elections due to partisan polarization and divisiveness of American political figures galvanizing both support and opposition, especially in recent years. Indeed, in the past four decades, negative opinions from one party towards members of the other party increased by nearly 20 points on a 100-point scale, more than in any other liberal democracy examined, and an average of 30% of Americans believe their opposing party is an existential threat to America, compared to 16% in 1994 (8, 9).

Our results confirmed this using two sets of data on endorsed candidates' state and district demographics as well as their individual progressivity scores, with the demographic model serving as a 96% accurate predictor of Trump's endorsements in 2020 and 88% in 2022 and the ideological model as an 80% accurate predictor of Sanders' endorsements in 2018 and 2020 (13% higher than the demographic model for Sanders). We also found that Trump and Sanders' endorsements positively contributed 11% and 7% to electoral outcomes of endorsed candidates, partially disagreeing with our hypothesis; however, it was a relatively insignificant factor compared to PVI (~50%), education level (~25%) and incumbency (~15%). Thus, though we had sufficient evidence to uncover how Trump and Sanders embodied the

pragmatic-ideological dichotomy of endorsements and that their impacts on elections were fairly insignificant, whether or not this could be generalized to American politicians in general and elucidate their motives for endorsing candidates remains to be determined in future research.

RESULTS

To determine the degree to which Trump and Sanders' endorsements conform to fundamental or ideological variables of candidates and their states/districts, we identified several key fundamental variables relating each candidate's state or district (PVI, bachelor's degree attainment, percentage white, average income, and percentage aged 65 or older) for states in which candidates received endorsements from the two figures. Collectively, these variables formed a description of each state from socioeconomic (bachelor's attainment, average income), political (PVI, a measure of the average margin of victory of Democratic or Republicans in a state or district), and social (percentage 65 and over, and percentage white) perspectives, and thus we selected these as the key demographic factors for our analysis. We deemed progressivity in particular to be highly relevant to our hypothesis of Sanders' ideological endorsements, as progressivity provided a metric to confirm whether or not Sanders' endorsed candidates were "progressive" and thus in line with Sanders' wing of the Democratic party. Incumbency, meanwhile, contributes towards verifying our hypothesis of pragmatism, as incumbents generally maintain an above 90% chance of reelection in Congress and thus are safe, pragmatic choices for endorsements (10).

We were generally unable to discover trends in Sanders-endorsed states and districts' fundamental variables, but found that Trump's endorsed states, on average, were relatively less educated, significantly Republican-leaning and predominantly white with a pronounced lower variance than national data (**Figures 1, 2, 3**). DBSCAN (Density-Based Spatial Clustering
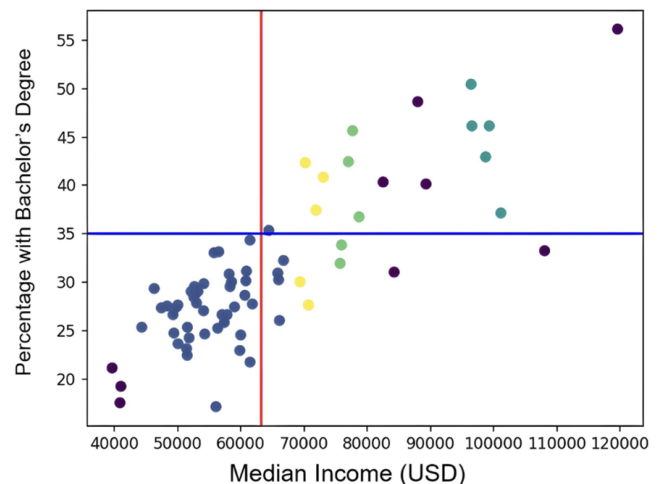
Figure 1: DBSCAN-clustered data for all Trump-endorsed states and districts' educational attainment and median incomes (14). Primary cluster shown in blue, with other clusters each represented in a different color (yellow, green, purple and cyan); colors do not have individual meaning and serve only to distinguish between data clusters. Blue and red solid lines represent the national average bachelor's attainment (35%) and median income (63,000 USD), respectively.
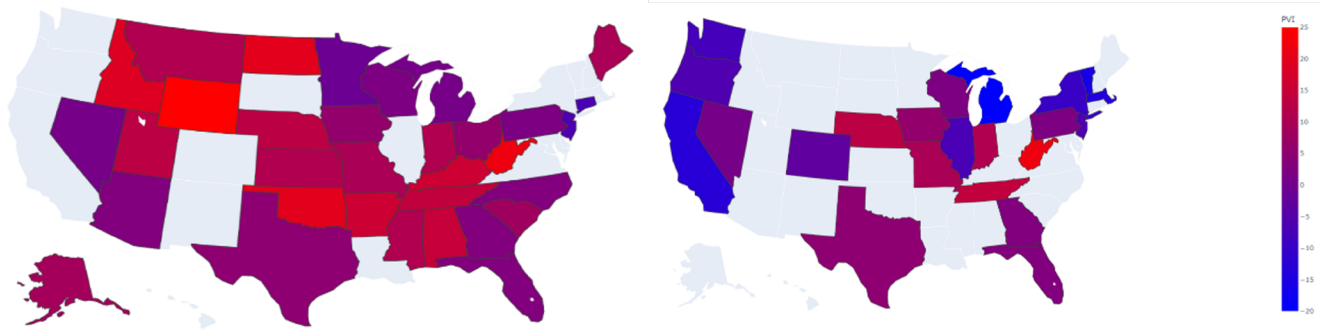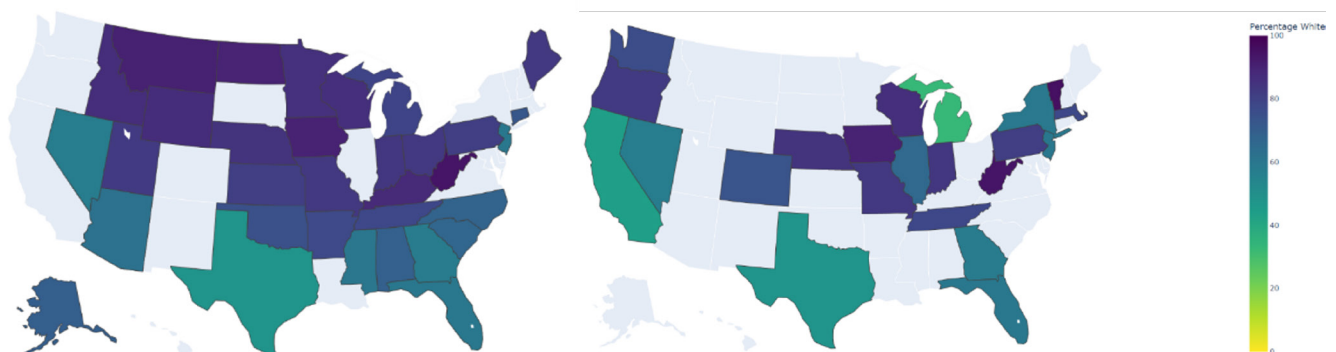
Figure 2: Statewide chloropleth map of Trump- and Sanders-endorsed states' political leaning, measured by PVI (14). Republican = red; Democrat = blue; Figure A = Trump; Figure B = Sanders. Numerical values of PVI represent average margin of victory for a certain party (positive = Republican; negative = Democrat; e.g. PVI = +2 means Republicans win by 2% on average)

of Applications with Noise) organizes data points into distinct clusters by maximizing the density of points in each cluster and minimizing the distance between data points to the center of their corresponding clusters (**Figure 1**). As shown, the primary data cluster (dark blue) lies below both the national averages for median income and bachelor's attainment. Figures 2 and 3 present state-by-state racial distributions and political leanings for Trump-endorsed states. Figure 2 reveals that nearly every single Trump-endorsed state had a positive (Republican-leaning) PVI; on a scale of -25 to 25 (blue to red, Democratic-leaning to Republican-leaning), the minimum PVI of a Trump-endorsed state was -6, with the majority of the states being visually represented as purple (neutral, slightly Republican-leaning) or red (firmly Republican-leaning). Notably, the states excluded from the map – states Trump refused to endorse candidates in – included Washington, Massachusetts, and other Democratic strongholds. Similarly, Figure 3 shows that the white population in Trump-endorsed states tended to fluctuate between 80 to 100%, but nearly never dropped below the national average (75%), suggesting an advantageous political climate for Trump support in these states. To confirm this, we conducted a logistic regression on Trump's and Sanders' midterm endorsements in 2018 and 2020 with endorsement status as the dependent variable. We found that PVI was the largest contributing factor towards Trump's endorsements (contributing >50% to whether Trump endorsed a candidate), with education level

contributing 28% (**Table 1**). In general, cross-validating the model by applying it to re-predict Trump's endorsements in 2018 and 2020 found that 97% of these endorsements were correctly predicted by fundamental state and district demographic factors alone. While the logistic regression performed on Sanders' endorsements yielded similar weights for each variable, it only offered a 67% accuracy rate.

We thus introduced a second ideological model for predicting endorsements through collecting progressivity scores (i.e., proportion of Senate/House votes in alignment with "progressive" positions) for each incumbent candidate, with non-incumbent candidates being assigned scores equal to their corresponding party's average; a similar logistic regression using only progressivity scores for Sanders yielded an 80% accuracy rating for 2018 and 2020, while a combination of ideological and fundamental factors yielded 78% accuracy. Furthermore, we also quantified the precise impact endorsements by Trump and Sanders had on the electoral outcomes of their respective endorsed candidates. We achieved this through applying a support-vector machine binary classification model, which functions by separating the data points into two categories with a line (or a higher-dimensional space for more variables; see Methods and Materials section) that maximizes the total distance from all data points. This model took both endorsement status (endorsed/not endorsed by either Trump or Sanders) and the aforementioned fundamental variables as independent

Table 1: Statewide chloropleth map of Trump- and Sanders-endorsed states' race distribution, measured by percentage white (15). 100% = dark purple; 0% = yellow; Figure A = Trump; Figure B = Sanders. Numerical values (from 20 to 40 percent) represent percentage of the state's adult population with a bachelor's degree.

Variable	Coefficient/Weight	Percentage Weight (Absolute)
PVI	0.925	55.0%
Percentage White	-0.00942	0.560%
Incumbent?	0.0221	1.31%
Percentage w/ Bachelor's Degree	-0.474	28.2%
Percentage Aged 65 or Older	-0.248	14.7%
Median Income	0.000398	0.0237%

Table 1: Weights and percentage weights of factors contributing towards Trump's endorsements. Positive coefficients represent a positive contribution to securing Trump's endorsement for a candidate and vice versa, with the magnitude of the coefficient determining the extent of the contribution; note that percentages do not add up precisely to 100% due to rounding.

variables, and election result as the dependent variable, allowing us to quickly classify any data point as either an election victory or an election defeat. We found that Trump's and Sanders' endorsements contributed 11% and 6% to electoral outcomes, respectively, which was a relatively insignificant contribution compared to PVI (~40%), education level (25%), and incumbency (14%) (**Table 2**, **Table 3**).

DISCUSSION

Our primary goal in this study was to shed greater light on what best characterized both the rationales and the impacts of Trump and Sanders' endorsements. From DBSCAN clustering and chloropleth map visualization, we were able to discover evident patterns suggestive of an underlying rationale in demographic variables of Trump-endorsed states. The vast majority of Trump's endorsed districts lie below the national average in bachelor's degree attainment and average income **(Figure 1)**, corroborated by the racial distributions and political leanings of Trump-endorsed states showing a predominantly Republican-leaning and mostly white subset of the nation **(Figure 2, Figure 3)**. The implications of such trends in all four variables combined are clear: Trump has endorsed in states with political atmospheres that overwhelmingly support the Republican Party and himself. Every single demographic variable considered here – the Republican-leaning PVI, the lower levels of education and median incomes – seem to be correlated with higher levels of Republican support (11). PVI naturally contributes towards Republican support, but Trump's electorate in 2016 and 2020 were both less educated and somewhat poorer than the Democratic electorate (11). This conclusion is further justified by Trump's aversion to endorsing candidates in Democratic strongholds like Massachusetts and Washington **(Figure 2)**. We thus observe that Trump endorses in a very pragmatic manner, always in states where Republican candidates are favored to win elections and never in Democratic stronghold states. Furthermore, he seems to do so in a way that prioritizes his own influence over the Republican Party rather than the tangible benefits they offer to candidates, as demonstrated by the 200 virtually unopposed primary candidates Trump endorsed (5).The results from our logistic regression confirm this **(Table 1)**.

The dominance of PVI in Trump's endorsement-making process (contributing 55%) heavily reinforces the narrative of a pragmatic, demographic-driven approach from Trump. The PVI of a state is the variable that most clearly reflects its political leaning and thus whether Republican candidates are likely to win, demonstrating how Trump's endorsements

stem primarily from political leanings of states rather than individual candidates. The majority of the remaining contribution originates from educational attainment (28%), corroborating the initial findings of our DBSCAN clustering of a strong correlation between Trump's endorsed states and below-average bachelor attainment rates. This also lends evidence towards the hypothesis that Trump's endorsements are strongly decided by advantageous state demographics, as lower levels of educational attainment were strongly correlated with Trump support in both the 2016 and 2020 elections and have traditionally played a large part in the Republican electorate. This was further validated by the 97% accuracy rate our model achieved when re-predicting 2018 endorsements, showing that demographic variables alone served as good predictors of Trump's endorsements (**Table 2**).

Analyzing the factors driving Sanders' endorsements revealed an interesting and noteworthy contrast with Trump's pragmatism. In contrast with Trump's chloropleth map in Figure 2, no discernible pattern was revealed through visually representing Sanders-endorsed states' PVIs and bachelor's attainment levels. Instead, the PVI map showed a large degree of variance in the political leaning of states, with several states that were firmly Republican, such as West Virginia and Tennessee, seeing representation. Similarly, education levels were widely distributed with some Sanders-endorsed states such as Massachusetts and Colorado seeing above average (e.g., 40%) bachelor's attainment rates typical of Democratic-leaning states, but an even larger number seeing far-below-average levels, such as Indiana with 24.7% and West Virginia with 19.2% (**Figure 1**). This indicates that there may be other underlying factors regarding individual candidates that influence Sanders' endorsements more strongly than demographic factors do, which contrasts with Trump's endorsement patterns.

After introducing a progressivity variable for each candidate and performing a logistic regression two more times, first with progressivity and demographic patterns combined, then solely with progressivity, we discovered an interesting result: the accuracy the model achieved in predicting 2020 endorsements with only progressivity was 80%, 13% higher than the demographic model and 5% higher than both combined. This would suggest that our predictions for Sanders' endorsements are made

Variable	Weight	Percentage Contribution
Trump Endorsed?	0.419	11.0%
Incumbent?	0.518	13.6%
PVI	1.44	37.8%
Percentage White	0.00772	0.203%
Percentage w/ Bachelor's Degree	-0.956	25.1%
Percentage Aged 65 or Older	-0.192	5.04%
Median Income	0.278	7.30%

Table 2: Weights and percentage weights of factors contributing to Republican midterm election victories, including Trump's endorsements. Positive coefficients represent a positive contribution to election victory and vice versa, with the magnitude of the coefficient determining the extent of the contribution; note that percentages do not add up precisely to 100% due to rounding.

Variable	Weight	Percentage Contribution
Sanders Endorsed?	0.261	6.17%
Incumbent?	0.863	20.4%
PVI	-1.15	27.2%
Percentage White	-0.333	7.87%
Percentage w/ Bachelor's Degree	0.562	13.3%
Percentage Aged 65 or Older	0.389	9.20%
Median Income	-0.672	15.9%

Table 3: Weights and percentage weights of factors contributing to Democratic midterm election victories, including Sanders' endorsements. Positive coefficients represent a positive contribution to election victory and vice versa, with the magnitude of the coefficient determining the extent of the contribution; note that percentages do not add up precisely to 100% due to rounding.

less accurate with the inclusion of demographic factors, likely because they offer nearly no underlying pattern; progressivity, meanwhile, boosts the accuracy of the model upon its inclusion and predicts endorsements better on its own than when combined with demographics. We therefore conclude that the individual political leaning of candidates, as quantified by progressivity, serves as a more effective predictor of Sanders' endorsements than demographic variables. This fits in line with Sanders' unique ideological position in American politics and the progressive movement, as we would expect Sanders' endorsements to go towards candidates who embrace progressivism regardless of their state's political leaning. Another reason behind the contrast between Sanders' and Trump's endorsements is that Trump, as an incumbent president and the de facto leader of the Republican Party, may be expected to apply as much of his considerable power and leverage as possible to influence outcomes in swing states and states where Republicans are actually able to win, leading to pragmatic demographic conditions. Sanders, in comparison, has no such obligations.

The results we found for endorsement impacts on electoral outcomes also proved to be of interest. Our support-vector machine (SVM) model found that, on average, Trump and Sanders' endorsements contributed 11% and 6% respectively to election victories, behind factors such as PVI, incumbency, and education level (**Table 2**). PVI (~30%) is expected to be the most significant predictor of election victories because it simply describes the average margin of victory of the Republican Party within a given state's elections; therefore, a higher value would contribute importantly to the likelihood of a Republican victory, and vice versa. Notably, bachelor's attainment (~20%) was the second most important variable; the importance of bachelor's attainment is primarily corroborated by Trump's electorate and the Democratic electorate, including Hillary Clinton's and Biden's, having an unusually large disparity in education levels: in the 2018 midterms, college graduates leaned Democratic to Republican 54 to 39 percent, a complete inversion of the proportion from 25 years ago (11). The Republican electorate had shifted towards increasingly uneducated white voters by 2018 as Trump's presidency continued, leading to a more significant contribution by education level than any other factor. The role of incumbency (~15%) was also notable, which serves to justify the ongoing

trend of Congressional stagnation in which incumbents are overwhelmingly reelected; in recent years, the incumbent reelection rate had been up to 90% in the House, but it was lower in the senate (10). Simultaneously, many Congressional elections occur with both candidates as non-incumbents due to retirements; the "turn-over rate" following the 2020 midterms, or the percentage of Congresspeople who had served 12 years or less, was 72% in the House and 65% in the Senate (10). This explains why incumbency contributes towards election victories to a large extent for elections with incumbents, but is not the largest deciding factor in general due to not every election having an incumbent.

The relatively low impacts of Trump and Sanders' endorsements can be interpreted in several ways. First, we can interpret it as an indication that endorsements – Trump's endorsements in particular – are a good predictor of, but not a good contributor towards, electoral victories. Trump's endorsements appear to contribute fairly little towards election results, with any favorability in Trump-endorsed candidates' election results being nearly entirely explained (89%) by state demographic factors; this corresponds with how Trump's endorsements select Republican-leaning states where candidates have better chances to win. Second, we can explain the insignificant impact of endorsements in American politics through the growing trend of political polarization. Consider Trump as an example; as of May 2023, though he has maintained a net -10% approval rating throughout his presidency, he continues to retain more than 90% support from Republicans and single-digit support from Democrats (12). As such, the weight carried by Trump's endorsement is equally as unfavorable among Democrats as it is favorable among Republicans, galvanizing opposition and support for the candidate equally. Therefore, we expect that while Trump's endorsements are impactful in Republican primaries where the majority of the electorate are Republicans, their net effect is less significant in general elections.

Though this study illustrates a dichotomy of ideological and pragmatic endorsements through Sanders and Trump and estimates their effects on midterm elections, our framework continues to contain several limitations and areas for future research. First, general elections are a poor descriptor of the true impacts of endorsements. Though endorsements may not contribute greatly to general elections - a trend supported by the phenomenon that voters from the opposing party may also be galvanized to vote against the endorsed candidate - they can play a major role in primary elections, in which Trump's supported candidates have won 92% of the time (3, 5). This is due to the majority of primary voters being affiliated with the party of the candidates, making Trump or Sanders or the influence of any political figure far greater. Future research could attempt to apply the same binary classification model alongside fundamental factors in U.S. midterm primary election.

Furthermore, the fundamental variables selected in our model - age, race, and socioeconomic factors - may not necessarily be the most optimal combination of factors. In fact, our results have yielded some evidence that may suggest how some of these variables will negatively affect the accuracy of the predictions made. When all fundamental variables were excluded in favor of progressivity in predicting Sanders' endorsements, the model's predictive accuracy jumped from 67% to 80%; when age was excluded and all other fundamental

factors were included, accuracy also increased to 70%. Age thus appears to be less relevant than other factors to the predictions and may have decreased the overall accuracy of the model. As the model combined data from Congressional districts and states, several Congressional districts with small populations also serve as major outliers with high incomes, educational attainment, or extreme PVIs (e.g., California's 17th district has a median income of $142,408, just shy of triple the national average, and a bachelor's attainment rate of 60%) (13). Future research could benefit from using a weighting system considering the population of each district to prevent large outliers from affecting accuracy.

In terms of accuracy, the model's described trends and predictions did achieve fairly satisfying results (>80%, with Trump's endorsement model close to 90% in both 2018-2020 and 2022), but crucially lacked reliable predictions in swing states such as Nevada or Georgia. Indeed, all inaccuracies of the model in 2018 and 2020 stemmed from states identified as swing states such as Nevada and Wisconsin. Our 2022 prediction model also did not consider Sanders' endorsements due to a lack of endorsed candidates. Additionally, the probabilities of victory generated by the model were occasionally too moderate, especially when compared to expert predictions or forecasts generated by FiveThirtyEight, potentially due to information loss from specific polls being unavailable. For instance, Republican candidate for U.S. Senate in Illinois, Kathy Salvi, is unlikely to have a probability of victory as large as 15.1% and was nearly certain to lose the election at the time of the study due to Illinois having a +7 Democratic PVI and being considered one of the most Democratic states in the nation, alongside California and New York (14). Future models can improve upon the accuracy of our results by considering a range of individualized variables for each candidate (e.g., recent scandals, events, or donations), as fundamental variables are static over time and do not reflect important political trends. Furthermore, only Trump and Sanders were considered in this paper as two case studies of diametrically opposed endorsement patterns; in future research, the same model may be applied to study and examine patterns in other political figures' endorsements, such as Obama or Biden, and classify them as pragmatic, ideological, or potentially outside this simple dichotomy.

Our results and conclusions are generally validated by the present literature but provide several novel implications that could prove to be of further interest. In Ballard *et al.*, both the rationale and effects of Trump's midterm endorsements were investigated, concluding that Trump was more likely to endorse candidates poised to win their elections without his endorsement and that Trump's endorsements provide financial boosts to both the candidates themselves and their opponents, culminating in a lower likelihood of winning (3). The strategic nature of Trump's endorsements as well as his propensity towards endorsing winning candidates remains consistent. Though the finding that Trump's endorsements had a net negative effect could not be replicated in our paper due to information loss from a lack of election-wise polling data, the underlying political polarization behind Trump's endorsements potentially galvanizing opposition voters and thus having little positive effect was referenced in both this study and (3). In Gadarian *et al.*, which investigated the effect of Trump's endorsements of partisan COVID-19 policies such as mask bans on the general population, the

effect of Trump's endorsements on attracting public support was also deemed to be statistically insignificant (neither positive nor negative) due to political polarization (15). Finally, our conclusions diverge from the existing literature in two ways. First, while Ballard's study primarily involves only PVI and political leaning as variables on which Trump's endorsements depend, our investigation found evidence for a correlation between Trump's endorsements and lower bachelor's attainment rates, as well as other variables that are characteristic of Republican-leaning states. Second, our paper introduces the characterization of endorsements as ideological through the case of Sanders and provides a metric through which ideological endorsements can be tested for (progressivity). Given the evidence we provide for Trump and Sanders' endorsement patterns, corroborated by existing research, we believe it is a distinct possibility that many U.S. politicians ascribe to either a pragmatic or ideological approach, making this framework potentially broadly useful.

MATERIALS AND METHODS
Variables and Data Collection
We first identified four fundamental variables we believed were most relevant to electoral outcomes: bachelor's degree attainment by state, race distribution of state (measured by percentage white), median income, and percentage aged 65 or older, obtained from the US Census Bureau's statistics in 2020 (13). In order to quantify the political leaning of each state and district, we also included the PVI, as obtained from the 2017 Cook Political Report (16). To construct our ideological/individualized model, we included each individual candidate's incumbency and progressivity scores, obtained from ProgressivePunch.com (17). We quantified progressivity as the proportion of an incumbent candidate's Congressional votes that aligned with progressive positions (e.g. universal healthcare, raising the minimum wage to 15$ and cancelling student debt). For non-incumbent candidates, we assumed their progressivity was roughly the same as their party's average in the U.S. House or Senate.

Data Processing and Analysis
Before analyzing and classifying the data, we standardized the data set with the StandardScaler() function in Python *scikit learn*:

$$\text{standardized data } z = \frac{x-\mu}{\sigma},$$

$$\text{mean } \mu = \frac{\sum_{i=1}^{N} x_i}{N},$$

$$\text{and standard deviation } \sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}}.$$

We separated the six variables used into three groups based on their calculated correlation with one another, with median income and bachelor's attainment (Pearson correlation $r = 0.823$) both representing socioeconomic status, percentage white and percentage aged 65 or older ($r = 0.525$) both representing miscellaneous demographic distributions, and PVI representing political leaning. We then clustered our data set based on these groups using DBSCAN clustering (18), calibrating the parameter ε (minimum distance between points required for a cluster) using k-Nearest Neighbor algorithms in Python's *scikit learn* package (19). DBSCAN clustering relies on defining clusters of points reachable from one another with a parameter ε. If a point P is within a

distance ε of point Q, then Q is considered *directly reachable* from P. If Q can be reached from P through the sequence $(P, P_1), (P_1, P_2), \cdot, P_n, Q)$ where each pair of points (P_i, P_{i+1}), are directly reachable from each other, then Q is *reachable* from P. A cluster with core point P is then defined as a set of points, at minimum *minPt* points, that are reachable from a point P. Finally, we plotted each variable on a statewide basis on a chloropleth map of the United States, with the patterns we discovered motivating our binary classification models.

Binary Classification

To quantify the contribution of each of the six variables to endorsement decisions and electoral outcomes, we employed two binary classification models: logistic regression models and support-vector machine models. Logistic regression models a binary variable (0 or 1) as a sigmoid S-shaped function: (returns a value between 0 and 1); if a certain set of values of independent variables output a value less than 0.5 in the

$$\frac{e^{\sum_{i=1}^{n} a_i x_i}}{1 + e^{\sum_{i=1}^{n} a_i x_i}}$$

logistic regression, it is classified as 0, and classified as 1 otherwise. We performed logistic regression with endorsement status (0 = not endorsed, 1 = endorsed) as the dependent variable, and interpreted the coefficient of each independent variable in the logistic regression as its weight. This was performed twice: once with only fundamental demographic variables and once with demographic variables combined with individual ideological variables. We calculated the proportion of the total weight each variable contributes to endorsement status to reach a conclusion regarding the deciding factors and patterns behind endorsements. We finally assessed the performance of both the fundamental demographic model and the ideological model through inputting our 2018 and 2020 data separately into the model and checking for accuracy. The SVM, taking election result as the dependent variable (0 = lost election, 1 = won election), generated a plane in hyperdimensional space that separated data points of candidates winning their elections and candidates losing their elections as optimally as possible, maximizing the distance between the closest points to the hyperplane: This distance is also known as the decision function; its sign determines whether a candidate is predicted to win or lose.

$$d(x) = \frac{|w^T \cdot x + b|}{|w|^2} \quad (20).$$

We assessed our SVM model through incorporating Platt scaling, a logistic regression with the decision function as the dependent variable, generating a probability of victory for each candidate. We then predicted electoral results for the 2022 Senate elections using our SVM, generated corresponding probabilities, and compared our results to FiveThirtyEight's predictions, among other sources (14).

REFERENCES

1. Ballotpedia Editor. "Endorsements by Donald Trump." *Ballotpedia*, 28 Sept. 2022, ballotpedia.org/Endorsements_by_Donald_Trump. Accessed 30 Jan. 2023.
2. Ballotpedia Editor. "Endorsements by Bernie Sanders." *Ballotpedia*, 2 Sept. 2022, ballotpedia.org/Endorsements_by_Bernie_Sanders. Accessed 30 Jan. 2023.
3. Ballard, Andrew O., et al. "Be Careful What You Wish for: The Impacts of President Trump's Midterm Endorsements." *Legislative Studies Quarterly*, vol. 46, no. 2, 19 May 2020, pp. 459–491., doi:10.1111/lsq.12284.
4. Cillizza, Chris. "Why Donald Trump's '33-0' Record Is a Lot Less Impressive than It Looks | CNN Politics." *CNN*, Cable News Network, 4 Mar. 2022, www.cnn.com/2022/03/04/politics/trump-texas-republican-primary-endorsement/index.html. Accessed 30 Jan. 2023.
5. Kalich, Sydney, and Mitch Carr. "What Does a Trump Endorsement Mean during the Primaries?" *NewsNation*, NewsNation, 17 Aug. 2022, www.newsnationnow.com/politics/trump-endorsement-primaries-election/. Accessed 30 Jan. 2023.
6. Sanders, Bernie. "Our Revolution". *Bernie Sanders Official Website,* Friends of Bernie Sanders, 2020, berniesanders.com/get-involved/our-revolution/.
7. Semler, Ashley. "'Bernie Sanders Has Real Influence': Vermont's Longtime Outsider Has Become a Trusted Voice in the Biden White House | CNN Politics." *CNN*, Cable News Network, 9 June 2021, edition.cnn.com/2021/06/08/politics/bernie-sanders-gloria-borger-interview-cnntv/index.html. Accessed 30 Jan. 2023.
8. Kimball, Jill. "U.S. Is Polarizing Faster than Other Democracies, Study Finds." *Brown University*, 21 Jan. 2020, www.brown.edu/news/2020-01-21/polarization. Accessed 30 Jan. 2023.
9. Geiger, Abigail. "Political Polarization in the American Public." *Pew Research Center - U.S. Politics & Policy*, Pew Research Center, 9 Apr. 2021, www.pewresearch.org/politics/2014/06/12/political-polarization-in-the-american-public/. Accessed 30 Jan. 2023.
10. Murse, Tom. "Do Members of Congress Ever Lose Re-Election?" ThoughtCo, *ThoughtCo*, 10 Dec. 2020, www.thoughtco.com/do-congressmen-ever-lose-re-election-3367511. Accessed 10 Feb. 2023.
11. Willingham, AJ. "25 Years Ago, Republicans Were More Educated than Democrats. Now It's the Opposite, Study Says | CNN Politics." *CNN*, Cable News Network, 22 Mar. 2018, www.cnn.com/2018/03/22/politics/democrat-republican-education-poll-women-pew-research-trnd/index.html. Accessed 30 Jan. 2023.
12. Draca, M; Schwarz, C. (2021). How Polarized are Citizens? Measuring Ideology from the Ground-Up. *SSRN Electronic Journal*
13. US Census Bureau. "Voting Statistics: State Electorate Profiles." *Census.gov*, 8 Oct. 2021, www.census.gov/library/visualizations/2016/comm/electorate-profiles.html. Accessed 25 Apr. 2023.
14. Rakich, Nathaniel, et al. FiveThirtyEight, 2 May 2023, https://fivethirtyeight.com/.

15. Gadarian, Shana *et al.* "Partisan endorsement experiments do not affect mass opinion on COVID-19". *Journal of Elections*, Public Opinion and Parties, vol. 31, 15 Jun 2021, pp. 122-131, doi:10.1080/17457289.2021.1924727

16. Cook Political Report. "Introducing the 2017 Cook Political Report Partisan Voter Index." *Cook Political Report*, 2017, www.cookpolitical.com/introducing-2017-cook-political-report-partisan-voter-index. Accessed 25 Apr. 2023.

17. Grossman, Joshua, and Ben Barcellos. "Progressive Punch". ProgressivePunch, Apr. 2023, www.progressivepunch.org. Accessed 25 Apr. 2023.

18. Ester, Martin, et al. "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise." *Association for the Advancement of Artificial Intelligence*, 1996, pp. 226–231., doi.org/10.5120/739-1038.

19. Sharma, Ankush, and Amit Sharma. "KNN-DBSCAN: Using K-Nearest Neighbor Information for Parameter-Free Density Based Clustering." *2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)*, 2017, doi.org/10.1109/icicict1.2017.8342664.

20. Cortes, Corinna, and Vladimir Vapnik. "Support-Vector Networks." *Machine Learning*, vol. 20, no. 3, 1995, pp. 273–297., doi.org/10.1007/bf00994018.