Factors Affecting Support for Autocratic Governance Among U.S. Adults

##### Sam Minard1 & Efrain Ramirez2

*1Department of Social Sciences & Global Studies, California State University, Monterey Bay, Seaside, CA* *2Department of Psychology, California State University, Monterey Bay, Seaside, CA*

# Abstract

This study employs multiple logistic regression to analyze data from Pew Research Center’s American Trends Panel, exploring how perceived macroeconomic conditions and social media news use preferences relate to support for autocratic leadership in the United States. While prior research suggests that economic anxiety and social media reliance fuel autocratic sentiment, results indicate that perceiving the economy as “good” is associated with higher odds of autocratic support, contrary to expectations. Social media preference, frequency of use, and its interaction with political ideology were not statistically significant. However, several demographic factors, including income, race, gender, and party affiliation, emerged as significant predictors, though not always in the expected direction.

# Introduction

As of 2021, 70% of the world’s population, or 5.4 billion people, live under an autocratic regime (Boese et al., 2022). Autocracies are a form of government “by a single person or small group that has unlimited power or authority” (Cambridge University Press & Assessment, n.d.), and they often arise from populist movements. Populism is an exclusionary, anti-elitist, anti-pluralist political ideology based on the idea that the “true people” are under threat from some kind of elite enemy, be it an economic, political, racial, or foreign one (Berlin et al., 1968). The ideology can fuel support for autocratic leaders who persuade the public that only they can protect them from perceived enemies, thus justifying the consolidation of executive power without accountability (Al Waroi et al., 2024; Baturo et al., 2024). Donald Trump’s populist messaging has galvanized many Americans against a variety of perceived enemies, including racial and gender minorities, immigrants, and other vaguely defined “elites.” Since regaining the presidency in 2025, Trump has taken steps to consolidate his power and remove checks and balances, thus bringing the United States closer to an authoritarian autocracy– and many Americans wholeheartedly support this transformation. Therefore, this paper seeks to understand the factors associated with support for autocratic leadership among Americans.

Populist movements and the autocratic leaders they support are often fueled by anxiety and subjective perceptions of economic insecurity, irrespective of whether objective measures of hardship are present (Gidron & Hall, 2017). In other words, actual economic insecurity is not necessary, only the perception of insecurity. Moreover, it is the perception of economic insecurity at the macro level that significantly predicts these attitudes, not at the micro level (Watson et al., 2021). Severe real-world threats can be especially impactful in this regard. For example, through their analysis of the Covid-19 pandemic and the subsequent economic downturn in Italy, Rocatto et al. (2020) found that times of uncertainty and hardship can drive people toward anti-democratic, authoritarian, and autocratic leaders because they offer a restored sense of control over their social world. Given that a recent Pew Research Center poll found Americans’ current economic ratings to be predominantly negative (Cerda, 2025), it is important to test the relationship between subjective economic insecurity and anti-democratic sentiment in the US. Therefore, our first research question is as follows: *How does perceived macro level economic insecurity relate to support for autocratic leadership?*

As the information channels American society relies on continue to evolve, it is important to consider the relationship between media type preferences and anti-democratic sentiments that encourage support for autocratic leadership. Notably, mainstream media trust among conservatives has declined, leading them to seek alternatives like social media, where they may encounter misinformation and populist rhetoric (Hutchens et al., 2025; Mourão, 2017). While social media has been touted as a democratic panacea because of its potential for disseminating anti-regime information and mobilizing collective action, many autocratic regimes engage in social media censorship (Gunitsky, 2015). However, Gunitsky (2015) argues that many regimes have since moved from suppression to co-optation of social media as a tool to maintain autocratic stability– a far more insidious tactic. Co-opting social media allows autocratic regimes to shape the discourse on social media, promote counter-mobilization, and bolster regime legitimacy. Given autocratic co-optation alongside the transition to social media as the public’s primary source of news and information, it is crucial to investigate the relationship between social media use and support for autocratic governance among American adults. Furthermore, the disparity in media preferences between conservatives and liberals means that political ideology must be considered as a mediating factor in this relationship. As a result, we propose the following research question: *Is a preference for getting news and information primarily from social media associated with support for autocratic leadership? How does this relationship vary depending on political ideology?*

By investigating these research questions, this paper will contribute to our understanding of the factors associated with support for autocratic leadership among Americans. First, we hypothesize that perceiving the economy as “bad” has predictive power of support for autocratic leadership, increasing the odds Second, we hypothesize that there is a stronger relationship between the preference for getting news and information primarily from social media and support for autocratic leadership among conservatives compared to liberals.

# Methods

## Data Collection

This study conducts secondary analysis of the American Trends Panel Wave 124 data set from Pew Research Center, also known as ATP Wave 124. The survey was administered to a nationally representative, stratified random sample of U.S. adults from March 20 to March 26, 2023. Of the 4,058 panelists sampled, 3,576 responded, resulting in a response rate of 88%. Hispanic men, non-Hispanic Black men, and non-Hispanic Asian were oversampled then weighted to provide more precise estimates.

## Variable Creation

The primary dependent variable in this study is support for autocratic leadership, which was measured through the following question: “Would each of the following political systems be a good or bad way of governing this country? A system in which a strong leader can make decisions without interference from Congress or the courts.” Response options included a range from “Very good” to “Very bad,” and “Refused.” This was re-coded into “Supports” and “Does not support” to ensure a binary response variable.

Independent variables we considered included demographics such as age, race, gender, and family income. In addition, we used economic perception (“How would you describe the current economic situation in the U.S.?”) to measure perceived economic insecurity at the macro level. Categories were collapsed into “Good” and “Bad.” Finally, multiple variables measured respondents’ news source and social media preferences, including how often respondents get news, facts, and opinions about issues or events from social media sites, and whether respondents prefer getting various kinds of information from news outlets, social media, both, or neither. The five variables concerning whether respondents preferred news outlets or social media were re-coded to contain three levels (2: “Prefers social media”; 1: “Prefers both”; and 0: “Prefers news outlets or neither”). Then, the values for each were summed within each row to assign an index score for social media preference to each respondent. With a range of 0-10, higher scores indicated a stronger preference for obtaining information from social media. For all variables, “Refused” responses were treated as “NA” and dropped. The final sample size included in the model was 2065 observations.

## Analytic Methods

Frequency tables were used to summarize the counts and proportions for control variables including income, race, gender, and political party categories. They were also used to summarize the response and some explanatory variables. The social media preference index score distribution for each autocratic support category was illustrated with ridgelines and boxplots. Multiple logistic regression was used to predict support for autocratic leadership. We included the aforementioned theory-driven variables as well as income, race (“White” or “Non-White”), gender, and political leaning in the final model. Control variables were refined through stepwise selection, which removed only age.

# Results

## Exploratory Data Analysis

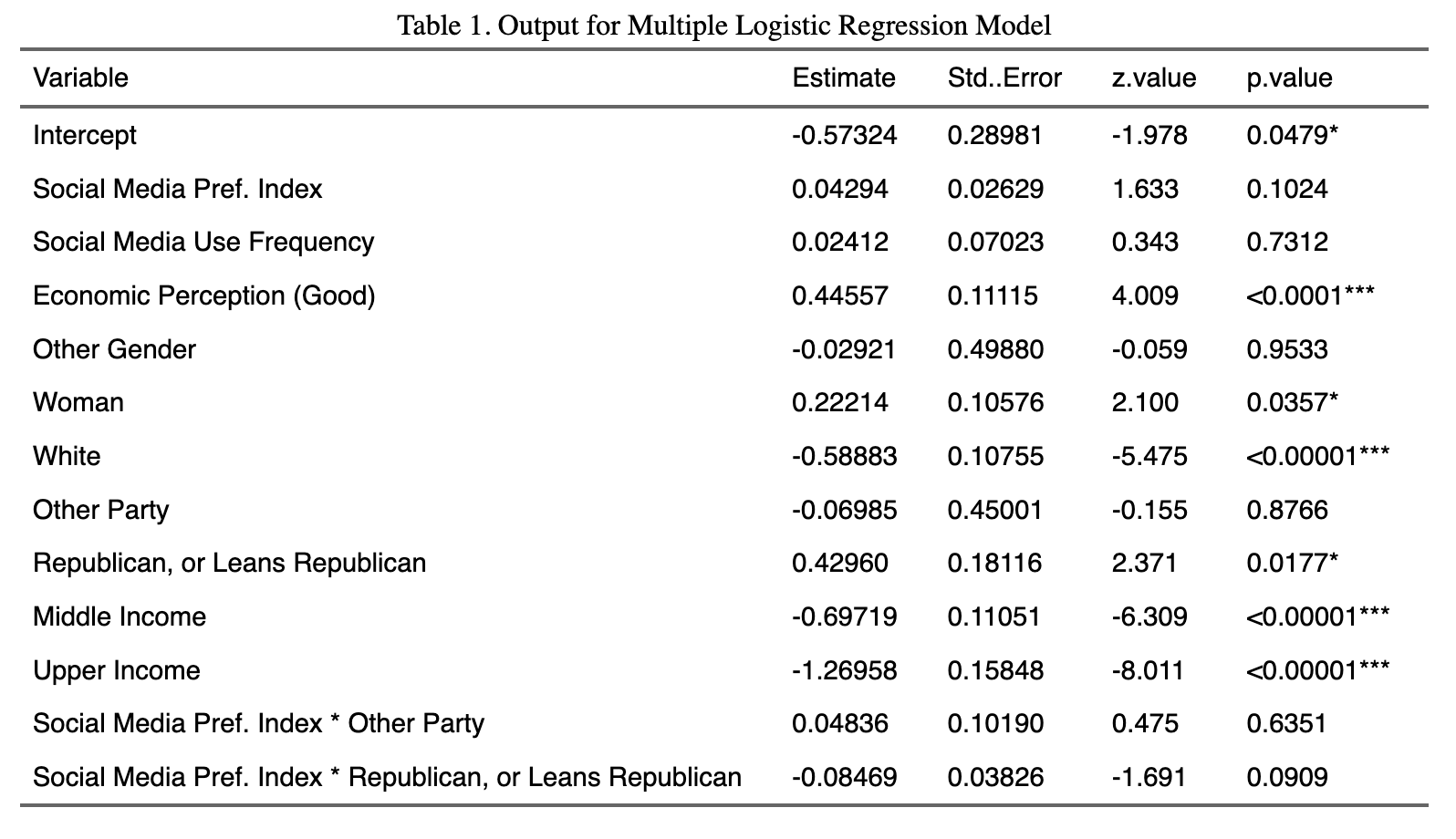
Figure 1 shows that social media preference index score distributions are similar between those who do and do not support autocratic leadership. However, the supportive group’s mean score (3.672) was marginally higher than the non-supportive group’s (3.510).

###### 

Figure 2 displays the proportions of respondents who support autocratic leadership within each economic perception category, faceted by political leaning. Across all political groups, a higher proportion of those who perceive the economy as “good” express support for autocracy. Frequency tables for income, political party, race, gender, and economic perception are included in the appendix.

###### 

## Statistical Analysis

Table 1 below summarizes the multiple logistic regression model output for autocratic leadership support. The estimates and p-values will be interpreted in the following section. Log odds model: 

# Conclusion

The purpose of this study is to determine whether individuals’ social media reliance and perception of the economy as generally good or bad might predict support for autocratic leadership in the United States. In addition, an interaction between social media reliance and political ideology is also considered. There is very strong evidence that at least one variable in the model with social media preference index score, social media use frequency, economic perception, gender, race, political leaning, income, and the social media preference-political leaning interaction has predictive power of autocratic support (G = 141.3424, p-value < 0.001). However, the model overall is a poor fit, as evidenced by the relatively low Area Under Curve (AUC) value (AUC = 0.6589).

While the model overall is not very effective at predicting autocratic support, several predictors were statistically significant, including the intercept (z = -1.978, p-value = 0.0479), perceiving the economy as good (z = 4.009, p-value < 0.0001), being a woman (z = 2.100, p-value = 0.0357), being white (z = -5.475, p-value < 0.00001), being or leaning Republican (z = 2.371, p-value = 0.0177), being middle income (z = -6.309, p-value < 0.00001), and being upper income (z = -8.011, p-value < 0.00001). Social media preference index score, social media use frequency, being non-binary, being neither Republican nor Democrat, and the interaction between social media preference index score and political leaning were not statistically significant.

Based on the model, we can predict the following: When a male, non-white, lower income democrat’s social media preference index score is 0, they never use social media to get news information, and they perceive the economy as bad, their odds of supporting autocratic leadership are multiplied by 0.5636961. When all other variables are held constant, the odds of autocratic support are 1.56138 times higher when perceiving the economy as good compared to bad, which contradicts our expectations. White individuals odds of autocratic support are 0.5549762 times that of non-white individuals, indicating that whites are less likely to express autocratic support, again contradicting our expectations. Finally, middle and upper income individuals’ odds of autocratic support are respectively 0.4979827 and 0.2809496 times those for lower income individuals, indicating that being lower income increases the probability of autocratic support, when all other variables are held constant.

We hypothesized that perceiving the economy as “bad” has predictive power of support for autocratic leadership, increasing the odds of support. While economic perception was statistically significant, the odds of autocratic support were actually lower when perceiving the economy as bad compared to good, so our hypothesis was not supported. Second, we hypothesized that there is a significant interaction between political party and social media preference index score on the log-odds of the autocratic support; specifically, the increase in social media preference score was hypothesized to be associated with a greater increase in the odds of the outcome for Republicans than for other political parties. This hypothesis was not supported by sufficient evidence (z = -1.691, p-value = 0.0909).

The results largely disagree with past research and our hypotheses, except in several demographic areas. First, based on the results, macro-level economic insecurity is associated with lower odds of autocratic support. A preference for getting news and information primarily from social media is not associated with higher odds of autocratic support, and political ideology is not a significant mediating factor in the relationship. The frequency of social media use for news consumption is also not associated with increased odds of autocratic support. Interestingly, men and whites are less likely to express autocratic support, contradicting the scholarly consensus that autocratic support tends to be highest among these group. However, in agreement with previous research, autocratic support is more likely among Republicans and lower income individuals than other political and income groups. The inconsistency between our results and existing research may stem from several limitations of this study, and indicates a need for further research on this topic. First, certain explanatory variables related to social media behavior contained a high percentage of missing values which could have influenced the analysis. Second, the data was collected via a stratified sample with oversampling in underrepresented groups, then weighted. However, our analysis did not take weights into account, meaning some racial groups had a disproportionate influence over the analysis. These limitations prevent us from making confident inferences about the factors influencing autocratic support among Americans. However, the results provoke new questions and challenge existing assumptions about the nature of this critical and evolving issue.

# References

Al Waroi, M. N. A. L., Riyanta, S., & Rustam, M. R. (2024). Populism and the erosion of democratic checks and balances: A systematic literature review across regions. International Journal of Multidisciplinary Research and Analysis, 7(10), 4728-4738. <https://doi.org/10.47191/ijmra/v7-i10-18>

Baturo, A., Kenny, P. D., & Balta, E. (2024). Leaders’ experience and the transition from populism to dictatorship. Democratization, 1–24. <https://doi.org/10.1080/13510347.2024.2391482>

Berlin, I., Hofstadter, R., MacRae, D., Schapiro, L., Seton-Watson, H., Touraine, A., Venturi, F., Walicki, A., & Worsley, P. (1968). To define populism. Government and Opposition, 3(2), 137–179. <http://www.jstor.org/stable/44481863>

Boese, V. A., Lundstedt, M., Morrison, K., Sato, Y., & Lindberg, S. I. (2022). State of the world 2021: Autocratization changing its nature? Democratization, 29(6), 983–1013. <https://doi.org/10.1080/13510347.2022.2069751>

Cambridge University Press & Assessment. (n.d.). Autocracy. In Cambridge dictionary. Retrieved March 16, 2025, from <https://dictionary.cambridge.org/us/dictionary/english/autocracy>

Cerda, A. (2025, February 11). Republicans think economy will improve over the next year, Democrats expect it to get worse. Pew Research Center. <https://www.pewresearch.org/short-reads/2025/02/11/republicans-think-economy-will-improve-over-the-next-year-democrats-expect-it-to-get-worse/>

Gidron, N., & Hall, P. A. (2017). The politics of social status: Economic and cultural roots of the populist right. The British Journal of Sociology, 68(S1). <https://doi.org/10.1111/1468-4446.12319>

Hutchens, M. J., Shaughnessy, B., & DuBosar, E. (2025). Populist hyperpartisans?: The interaction between partisan media exposure and populism in the 2020 US presidential election. Mass Communication and Society, 28(1), 51–75. <https://doi.org/10.1080/15205436.2023.2264265>

Mourão, R. R., Thorson, E., Chen, W., & Tham, S. M. (2018). Media repertoires and news trust during the early Trump administration. Journalism Studies, 19(13), 1945–1956. <https://doi.org/10.1080/1461670X.2018.1500492>

Pew Research Center. (2023). American Trends Panel Wave 124 [Data file and codebook]. Retrieved from <https://www.pewresearch.org/dataset/american-trends-panel-wave-124/.>

Watson, B., Law, S., & Osberg, L. (2022). Are populists insecure about themselves or about their country? Political attitudes and economic perceptions. Social Indicators Research, 159(2), 667–705. <https://doi.org/10.1007/s11205-021-02767-8>

# Appendix

#### Loading and cleaning the data

#read the data file  
amertrends <- read\_sav("data/data-raw/W124\_Mar23/ATP W124.sav")  
  
#load the codebook for the data to understand what each variable represents  
amertrends\_codebook <- read\_excel("data/data-raw/W124\_Mar23/ATP W124 Codebook.xlsx")  
  
#selecting a subset the variables for exploration   
amertrends\_subset <- amertrends |>   
 select(POLSYS\_AUTOC\_W124, ECON\_SIT\_W124, NEWSSOURCE\_e\_W124, NEWSUSE\_PLAT\_a\_W124, NEWSUSE\_PLAT\_b\_W124, NEWSUSE\_PLAT\_c\_W124, NEWSUSE\_PLAT\_d\_W124, NEWSUSE\_PLAT\_e\_W124, F\_AGECAT, F\_GENDER, F\_EDUCCAT, F\_RACECMB, F\_PARTYSUM\_FINAL, F\_INC\_TIER2)  
  
#change variable names to snake\_case  
amertrends\_subset <- janitor::clean\_names(amertrends\_subset)   
  
#converting 99 to NA (99 means "refused"/didn't answer that question so it must be treated as NA)  
amertrends\_subset <- amertrends\_subset |>   
 mutate(across(everything(), ~ replace(.x, .x == 99, NA)))  
  
#redefine column types to factor since they represent ordered categories/likert scale   
amertrends\_subset <- amertrends\_subset |>   
 mutate(across(where(is.numeric), as.factor))

#### Cleaning response variable

In the code chunk below, we create two versions of the binary response variable (autoc\_support): one with values 0 and 1 for the model, and one with descriptive labels for visualizations.

#make the response autoc variable binary/logical so it will work in the logistic model  
#change column type to numeric  
amertrends\_subset <- amertrends\_subset |>   
 mutate(autoc\_support = fct\_collapse(polsys\_autoc\_w124,  
 `1` = c("1", "2"),  
 `0` = c("3", "4"))) |>   
 drop\_na(autoc\_support)  
  
#make another version with descriptive labels for frequency table  
amertrends\_subset <- amertrends\_subset |>   
 mutate(autoc\_support\_cat = fct\_collapse(polsys\_autoc\_w124,  
 `Supports` = c("1", "2"),  
 `Does not support` = c("3", "4"))) |>   
 drop\_na(autoc\_support\_cat)

#### Cleaning explanatory variables

In the code chunk below, we recode the explanatory variables and create a new, clean data csv.

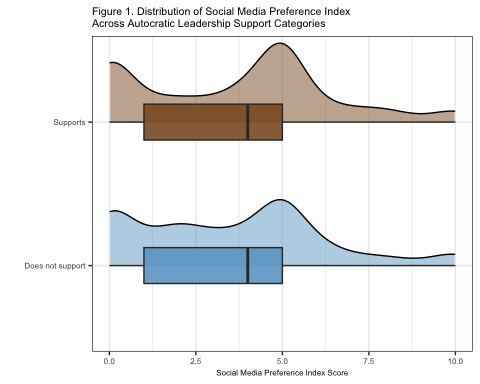
#defining the levels of key explanatory variables  
  
#step 1: economic perception  
amertrends\_subset <- amertrends\_subset|>   
 drop\_na(econ\_sit\_w124) |>   
 mutate(econ\_sit\_w124 = fct\_recode(econ\_sit\_w124,  
 "very good" = "1",  
 "somewhat good" = "2",  
 "somewhat bad" = "3",  
 "very bad" = "4"))   
  
#since there are considerably fewer observations at the extremes, collapse econ\_sit\_w124 into 2 categories: good and bad  
amertrends\_subset <- amertrends\_subset |>   
 mutate(econ\_percep = fct\_collapse(econ\_sit\_w124,  
 good = c("very good", "somewhat good"),  
 bad = c("very bad", "somewhat bad")))  
  
#step 2: creating social media reliance index variable  
  
#re-coding newsuse variables  
#newsuse\_plat\_a basic facts  
amertrends\_subset <- amertrends\_subset |>   
 drop\_na(newsuse\_plat\_a\_w124) |>   
 mutate(sm\_pref\_basicfacts = fct\_collapse(newsuse\_plat\_a\_w124,  
 `2` = c("2"),  
 `1` = c("3"),  
 `0` = c("1", "4")))   
#newsuse\_plat\_b opinions  
amertrends\_subset <- amertrends\_subset |>   
 drop\_na(newsuse\_plat\_b\_w124) |>   
 mutate(sm\_pref\_opinions = fct\_collapse(newsuse\_plat\_b\_w124,  
 `2` = c("2"),  
 `1` = c("3"),  
 `0` = c("1", "4")))   
#newsuse\_plat\_c in depth info  
amertrends\_subset <- amertrends\_subset |>   
 drop\_na(newsuse\_plat\_c\_w124) |>   
 mutate(sm\_pref\_indepth = fct\_collapse(newsuse\_plat\_c\_w124,  
 `2` = c("2"),  
 `1` = c("3"),  
 `0` = c("1", "4")))   
#news\_plat\_d up to date info  
amertrends\_subset <- amertrends\_subset |>   
 drop\_na(newsuse\_plat\_d\_w124) |>   
 mutate(sm\_pref\_uptodate = fct\_collapse(newsuse\_plat\_d\_w124,  
 `2` = c("2"),  
 `1` = c("3"),  
 `0` = c("1", "4")))   
#news\_plat\_e info that impacts the respondent  
amertrends\_subset <- amertrends\_subset |>   
 drop\_na(newsuse\_plat\_e\_w124) |>   
 mutate(sm\_pref\_impactsr = fct\_collapse(newsuse\_plat\_e\_w124,  
 `2` = c("2"),  
 `1` = c("3"),  
 `0` = c("1", "4")))   
  
#create the index variable for social media preference  
amertrends\_subset <- amertrends\_subset |>   
 drop\_na(sm\_pref\_basicfacts, sm\_pref\_opinions, sm\_pref\_indepth, sm\_pref\_uptodate, sm\_pref\_impactsr) |>   
 rowwise() |>   
 mutate(sm\_pref\_index = sum(as.numeric(as.character(c\_across(starts\_with("sm\_pref\_")))))) |>  
 ungroup()   
  
#step 3: social media use frequency  
  
#re-level newssource\_plat\_e: how frequently they get their news from social media  
#make numeric for regression purposes (treat as a score: higher = more frequently)  
amertrends\_subset <- amertrends\_subset |>   
 mutate(sm\_freq = fct\_recode(newssource\_e\_w124,   
 "5" = "1",  
 "4" = "2",  
 "3" = "3",  
 "2" = "4",  
 "1" = "5"))   
#another version with descriptive labels for freq table  
amertrends\_subset <- amertrends\_subset |>   
 mutate(sm\_freq\_cat = fct\_recode(newssource\_e\_w124,   
 "Extremely often" = "5",  
 "Fairly often" = "4",  
 "Sometimes" = "3",  
 "Rarely" = "2",  
 "Never" = "1"))   
  
#step 4: demographics   
#re-code age to more descriptive values  
amertrends\_subset <- amertrends\_subset |>   
 mutate(age = fct\_recode(f\_agecat,  
 "18-29" = "1",  
 "30-49" = "2",  
 "50-64" = "3",  
 "65+" = "4"))   
  
#recode race to binary, set reference level to non-white  
amertrends\_subset <- amertrends\_subset |>   
 mutate(race\_1 = fct\_collapse(f\_racecmb,  
 "Non-White" = c("2", "3", "4", "5"),  
 "White" = "1")) |>   
 mutate(race\_1 = fct\_relevel(race\_1,  
 "White", after = 1))   
  
#recode party lean, set reference level to democrat  
amertrends\_subset <- amertrends\_subset |>   
 mutate(pol\_lean = fct\_recode(f\_partysum\_final,  
 "Republican, or leans Republican" = "1",  
 "Democrat, or leans Democrat" = "2",  
 "Other" = "9"),  
 pol\_lean = fct\_relevel(pol\_lean,   
 "Republican, or leans Republican", after = 1))  
  
#recode income  
amertrends\_subset <- amertrends\_subset |>   
 mutate(income = fct\_recode(f\_inc\_tier2,  
 "Lower income" = "1",  
 "Middle income" = "2",  
 "Upper income" = "3"))  
  
#drop NAs to ensure later codes work (especially ROC curve codes)  
amertrends\_subset <- amertrends\_subset |>   
 drop\_na(autoc\_support, sm\_pref\_index, sm\_freq, econ\_percep, age, race\_1, pol\_lean, income)  
  
#recode gender  
amertrends\_subset <- amertrends\_subset |>   
 mutate(gender = fct\_recode(f\_gender,  
 "Man" = "1",  
 "Woman" = "2",  
 "Other" = "3"))  
  
  
write\_csv(amertrends\_subset, "data/data-clean/amertrends\_clean.csv")

#### Figures

In the code chunk below, we create Figures 1 and 2.

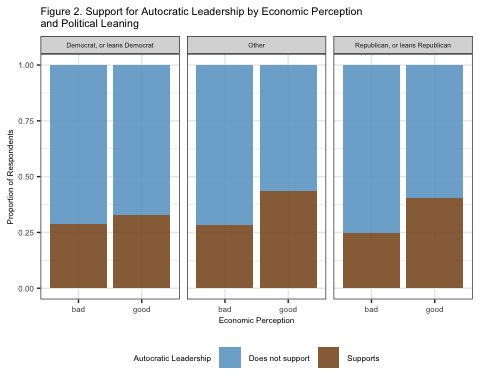
#boxplot + ridgelines for sm\_pref\_index by autoc\_support categories  
amertrends\_clean |>   
 ggplot(aes(x = sm\_pref\_index,   
 y = autoc\_support\_cat,   
 fill = autoc\_support\_cat)) +  
 geom\_density\_ridges(scale = 0.55,   
 alpha = 0.5) +  
 geom\_boxplot(aes(fill = autoc\_support\_cat),  
 width = 0.25,  
 alpha = 0.9) +  
 theme\_bw() +  
 scale\_fill\_manual(values = c("skyblue3", "tan4")) +  
 labs(x = "Social Media Preference Index Score",  
 y = "",  
 title = "Figure 1. Distribution of Social Media Preference Index \nAcross Autocratic Leadership Support Categories") +  
 guides(fill = "none") +  
 xlim(0,10) +  
 theme(legend.text = element\_text(size = 6),  
 legend.title = element\_text(size = 6),  
 axis.text.x = element\_text(size = 6),  
 axis.title.x = element\_text(size = 6),  
 axis.text.y = element\_text(size = 6),  
 plot.title = element\_text(size = 8))

## Picking joint bandwidth of 0.63



#ggsave("Figure\_1.png", width = 4, height = 2.75)  
  
#summary stats for each category   
amertrends\_clean |>   
 group\_by(autoc\_support\_cat) |>   
 summarize(`mean sm\_pref\_index` = mean(sm\_pref\_index),  
 sd = sd(sm\_pref\_index))

amertrends\_clean |>   
 ggplot() +  
 geom\_bar(aes(x = econ\_percep,  
 fill = autoc\_support\_cat),  
 position = "fill",  
 alpha = 0.9) +  
 facet\_wrap(~pol\_lean, ncol = 3) +  
 theme\_bw() +  
 theme(legend.position = "bottom",  
 legend.text = element\_text(size = 6),  
 legend.title = element\_text(size = 6),  
 axis.text.x = element\_text(size = 6),  
 axis.title.x = element\_text(size = 6),  
 axis.title.y = element\_text(size = 6),  
 axis.text.y = element\_text(size = 6),  
 plot.title = element\_text(size = 8),  
 strip.text = element\_text(size = 5)) +  
 scale\_fill\_manual(values = c("skyblue3", "tan4")) +  
 labs(x = "Economic Perception",  
 y = "Proportion of Respondents",  
 fill = "Autocratic Leadership",  
 title = "Figure 2. Support for Autocratic Leadership by Economic Perception \nand Political Leaning")



ggsave("Figure\_2.png", width = 4, height = 2.75)

#### Building version 1 of the model

In the code chunk below, we create the first model with all demographic and theory-driven variables under consideration. We then check the ROC curve.

#creating the first model using all theory-driven variables AND all demographic/control variables  
autoc\_lm\_v1 <- glm(autoc\_support ~ sm\_pref\_index + sm\_freq + factor(econ\_percep) + factor(age) + factor(gender) + factor(race\_1) + factor(pol\_lean) + factor(pol\_lean)\*sm\_pref\_index + factor(income), family = binomial(link = "logit"), data = amertrends\_clean)  
summary(autoc\_lm\_v1)  
  
#ROC curve for v1  
#predicted probabilities  
predicted\_probs\_v1 <- predict(autoc\_lm\_v1, type = "response")  
#actual outcomes  
actual\_outcomes\_v1 <- autoc\_lm\_v1$y  
#ROC plot and AUC  
roc\_v1 <- roc(actual\_outcomes\_v1, predicted\_probs\_v1)  
plot(roc\_v1, legacy.axes = TRUE, main = "ROC Curve for autoc\_lm\_v1",   
 col = "blue", xlab = "False Positive Rate (1 - Specificity)",   
 ylab = "True Positive Rate (Sensitivity)")  
auc(roc\_v1)  
  
#not a great model, so we will try adjusting which demographic variables are included to see if that improves things

#### Building version 2 of the model

In the code chunk below, we use stepwise selection to find the “best” model. Because the variable selection is partly theory-driven, we make the smaller model include the theory-driven variables to ensure they are included in the final model, rather than setting the lower model to a null model. The process results in a model that includes every predictor except age. We then check the ROC curve and find that it is not much better or worse than the previous model. Given the stepwise results and the fact that all age categories were insignificant per the Wald-z tests, we decide to proceed with the model without age. There appears to be collinearity between the interaction term sm\_pref\_index\*pol\_lean and pol\_lean.

#stepwise selection to determine best demographic variables for the model  
#fit the lower and upper models  
#lower includes the theory-driven variables that must be included in the model regardless of their effect on AIC  
small\_v1<- glm(autoc\_support ~ sm\_pref\_index + sm\_freq + factor(econ\_percep) + factor(pol\_lean) + factor(pol\_lean)\*sm\_pref\_index , family = binomial(link = "logit"), data = amertrends\_clean)  
#big has optional demographic variables  
big\_v1<- glm(autoc\_support ~ sm\_pref\_index + sm\_freq + factor(econ\_percep) + factor(age) + factor(gender) + factor(race\_1) + factor(pol\_lean) + factor(pol\_lean)\*sm\_pref\_index + factor(income), family = binomial(link = "logit"), data = amertrends\_clean)  
#stepwise selection based on AIC  
autoc\_lm\_v2 <- step(big\_v1, scope = list(lower = small\_v1, upper = big\_v1), direction = "both")  
#final selected model (based on AIC) includes race, political leaning, and income as controls (age was removed)  
  
#model parameter estimates and significance testing  
summary(autoc\_lm\_v2)  
  
#ROC curve for v2  
#predicted probabilities  
predicted\_probs\_v2 <- predict(autoc\_lm\_v2, type = "response")  
#actual outcomes  
actual\_outcomes\_v2 <- autoc\_lm\_v2$y  
#ROC  
roc\_v2 <- roc(actual\_outcomes\_v2, predicted\_probs\_v2)  
#plot the curve  
plot(roc\_v2, legacy.axes = TRUE, main = "ROC Curve for autoc\_lm\_v2",   
 col = "blue", xlab = "False Positive Rate (1 - Specificity)",   
 ylab = "True Positive Rate (Sensitivity)")  
auc(roc\_v2)  
#AUC gets slighter smaller but the difference is negligible  
#decided to proceed with model v2 (sans age) because it isn't a significant predictor anyway and does not add much value to the model  
  
#assumptions: check for collinearity  
#some collinearity between the interaction sm\_pref\_index\*pol\_lean and pol\_lean, but the rest are fine  
car::vif(autoc\_lm\_v2)  
  
#check significance overall  
G <- autoc\_lm\_v2$null.deviance - autoc\_lm\_v2$deviance  
G  
df <- autoc\_lm\_v2$df.null - autoc\_lm\_v2$df.residual  
1 - pchisq(G, df)

#### Creating model version 2 summary table

In the code chunk below, we create the summary table for the final mode. It doesn’t appear in the rendered word document unfortunately, so we include a screenshot in the text instead.

#### Frequency tables

Table 2. Frequency table for gender categories.

## x <character>   
## # total N=2065 valid N=2064 mean=1.80 sd=0.97  
##   
## Value | N | Raw % | Valid % | Cum. %  
## ---------------------------------------  
## Man | 1225 | 59.32 | 59.35 | 59.35  
## Other | 23 | 1.11 | 1.11 | 60.47  
## Woman | 816 | 39.52 | 39.53 | 100.00  
## <NA> | 1 | 0.05 | <NA> | <NA>

Table 3. Frequency table for income categories.

## x <character>   
## # total N=2065 valid N=2065 mean=1.89 sd=0.70  
##   
## Value | N | Raw % | Valid % | Cum. %  
## -----------------------------------------------  
## Lower income | 632 | 30.61 | 30.61 | 30.61  
## Middle income | 1028 | 49.78 | 49.78 | 80.39  
## Upper income | 405 | 19.61 | 19.61 | 100.00  
## <NA> | 0 | 0.00 | <NA> | <NA>

Table 4. Frequency table for race categories.

## x <character>   
## # total N=2065 valid N=2065 mean=1.55 sd=0.50  
##   
## Value | N | Raw % | Valid % | Cum. %  
## -------------------------------------------  
## Non-White | 935 | 45.28 | 45.28 | 45.28  
## White | 1130 | 54.72 | 54.72 | 100.00  
## <NA> | 0 | 0.00 | <NA> | <NA>

Table 5. Frequency table for economic perception categories.

## x <character>   
## # total N=2065 valid N=2065 mean=1.29 sd=0.46  
##   
## Value | N | Raw % | Valid % | Cum. %  
## ---------------------------------------  
## bad | 1458 | 70.61 | 70.61 | 70.61  
## good | 607 | 29.39 | 29.39 | 100.00  
## <NA> | 0 | 0.00 | <NA> | <NA>

Table 6. Frequency table for political leaning categories.

## x <character>   
## # total N=2065 valid N=2065 mean=1.85 sd=0.97  
##   
## Value | N | Raw % | Valid % | Cum. %  
## -----------------------------------------------------------------  
## Democrat, or leans Democrat | 1152 | 55.79 | 55.79 | 55.79  
## Other | 69 | 3.34 | 3.34 | 59.13  
## Republican, or leans Republican | 844 | 40.87 | 40.87 | 100.00  
## <NA> | 0 | 0.00 | <NA> | <NA>