Demographic Predictors of Voter Turnout in the 2020 U.S. Election

Cecily Wang-Munoz

Research Question

Understanding voter turnout is important for policymakers, election officials, and better voter outreach strategies, especially with the increasingly tense past presidential elections. In this study, I aimed to investigate **which demographic factors could best predict voter turnout in the 2020 U.S. election, with a particular focus on urban and rural dynamics**. Following my personal insight living in the very liberal city of Madison, WI, which is surrounded by a variety of rural and more conservative perspectives, I hypothesized that **counties with higher median income and predominantly urban populations will exhibit higher voter turnout rates.**

To analyze this question, I relied on county-level voter registration and turnout data disaggregated by race (White, Black, Hispanic, Asian, and Other), along with demographic data such as median household income and racial composition from the American Community Survey. I also incorporated urban-rural classifications based on federal urban-rural codes and geospatial shapefiles for mapping county-level turnout and demographic patterns.

Variables

Dependent Variable:
- **Voter Turnout Rate:** Proportion of registered voters who cast ballots, calculated as:
  Turnout Rate = Total Votes Cast/Total Registered Voters

Independent Variables:
- **Median Household Income** (scaled for interpretability).
- **Turnout by Race:** White, Black, Hispanic, Asian, and Other.
- **Urbanization:** Encoded as "Urban" for metro areas and "Rural" for non-metro areas.
- **Interaction Terms:** To see how race specific turnout patterns vary by urban-rural classification.

Methods & Regression Models

Data Cleaning and Processing:
- Merged voter turnout, demographic, and urban-rural datasets using fips identifiers.
- Scaled income and created binary urban-rural classifications.
- Removed missing/invalid data.

Regression Models:
- **Model 1:** Predict overall turnout using race-specific turnout rates, median income, and urbanization.

Turnout Rate = $\beta 0 + \beta 1$(Median Income Scaled) + $\beta 2$(Urban)+$\beta 3$(Race Turnout)+ϵ

- **Model 2:** Examine interaction effects between urbanization and race-specific turnout rates.

Turnout Rate = $\beta 0 + \beta 1$(Race Turnout × Urban) + $\beta 2$(Median Income Scaled) +ϵ

Results & Analysis

To see the factors influencing voter turnout in the 2020 U.S. election, I estimated two linear regression models: the first one predicting turnout rates using income, race-specific turnout rates, and urbanization, and the second incorporating interaction terms to evaluate how urbanization modifies racial turnout patterns. These models offer insights into the role of demographics and geography in shaping voter turnout.

Predicted vs. Actual Turnout

The figure below shows the predicted turnout rates from the interaction model against the observed turnout rates. The strong alignment along the regression line indicates that the model effectively predicts turnout (minimal residual errors). This suggests the predictors, including race-specific turnout and median income, can successfully capture turnout patterns across counties.

Predicted vs Actual Turnout Rate
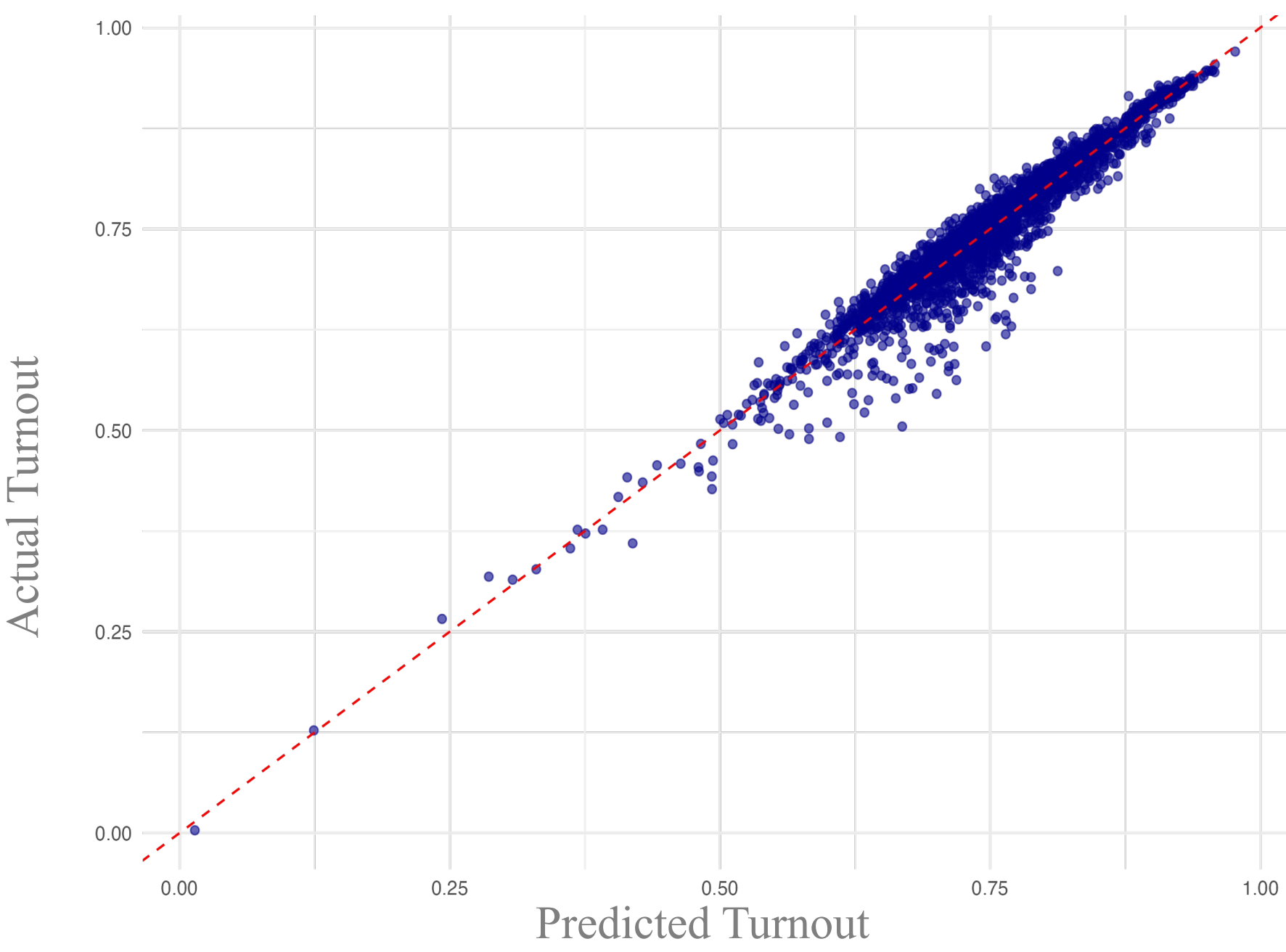
Regression Coefficients

The bar chart below presents the regression coefficients for the interaction model, along with their confidence intervals. White turnout emerges as the most significant predictor of overall turnout (β = 0.78, p < 0.001), followed by Hispanic turnout (β = 0.15, p < 0.001) and Asian turnout (β = 0.06, p < 0.001). While the interaction terms in the regression analysis were not statistically significant, this finding suggests that the effects of racial turnout predictors, such as White and Hispanic turnout, are consistent across urban and rural settings. This challenges assumptions about divergent voting behaviors in these regions and points to potential factors like shared socioeconomic or cultural influences.

**Regression Coefficients with Confidence Intervals**
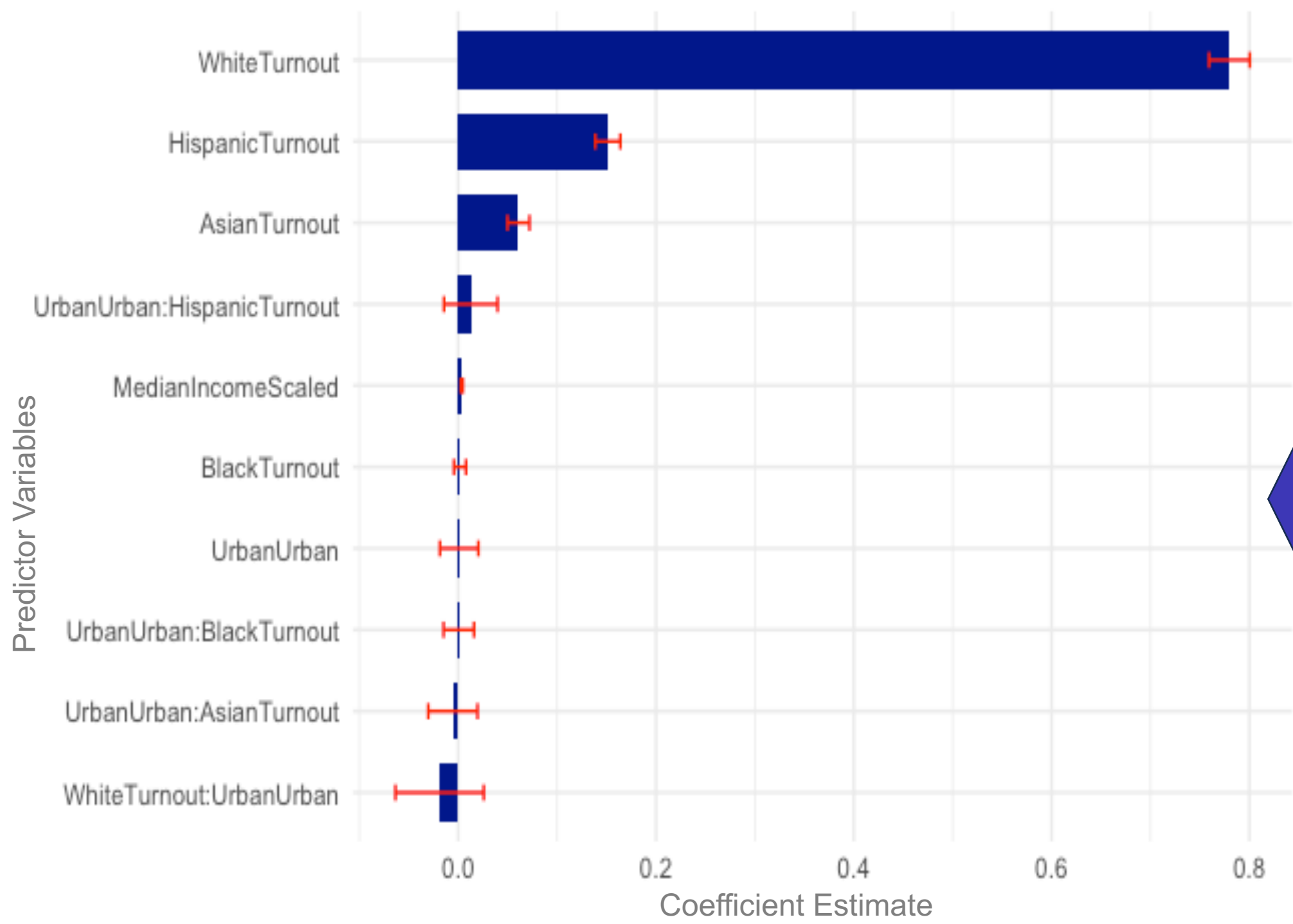
Regression Results

	Dependent variable:	
	Turnout Rate	
	(1)	(2)
White Turnout	0.775***	0.780***
	(0.009)	(0.010)
Black Turnout	0.002	0.002
	(0.003)	(0.003)
Hispanic Turnout	0.154***	0.151***
	(0.006)	(0.006)
Asian Turnout	0.060***	0.061***
	(0.005)	(0.006)
Median Income (Scaled)	0.004***	0.004***
	(0.001)	(0.001)
Urban		-0.019
		(0.023)
UrbanUrban:BlackTurnout		0.001
		(0.008)
UrbanUrban:HispanicTurnout		0.013
		(0.014)
UrbanUrban:AsianTurnout		-0.005
		(0.013)
UrbanUrban	-0.009***	0.001
	(0.001)	(0.010)
Constant	0.014***	0.012**
	(0.005)	(0.005)
Observations	2,943	2,943
R²	0.926	0.926
Adjusted R²	0.926	0.926
Note:	*p<0.1; **p<0.05; ***p<0.01	

Turnout and Median Income

Next, the scatterplot comparison below highlights the relationship between median income (scaled) and turnout rates, with graphs showing the trends for urban and rural counties. A positive relationship is evident in both settings, although turnout is generally higher in rural areas. We can also note that the difference in slopes suggests that income disparities may have a more pronounced impact in urban contexts.

Turnout Rate vs Median Income
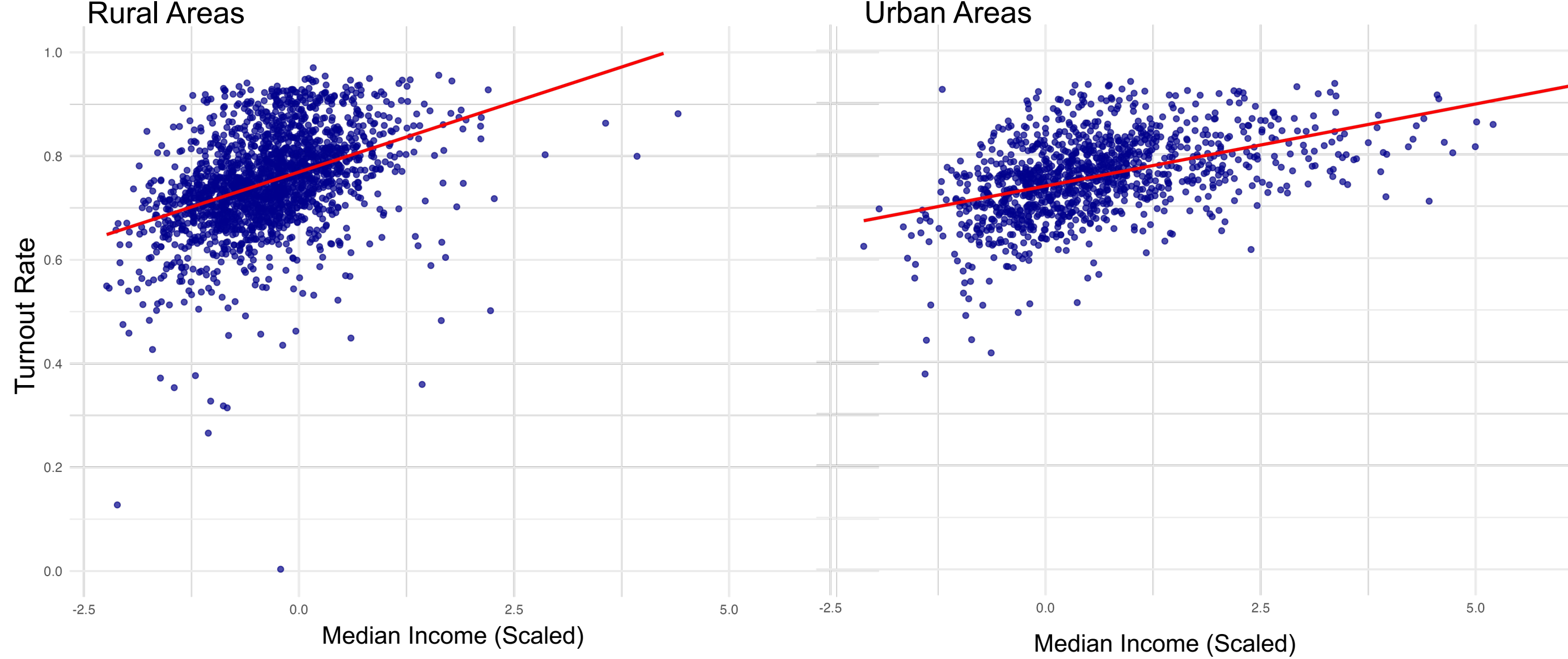
Urban vs. Rural Dynamics

Looking further into the urban-rural dynamic, the first map shows overall turnout rates by county, revealing higher turnout in the Midwest and Northeast. The second map disaggregates turnout by urban and rural classifications, showing us that urban counties display lower turnout compared to rural counties, further highlighting the importance of addressing barriers to participation in urban areas.

Voter Turnout by County
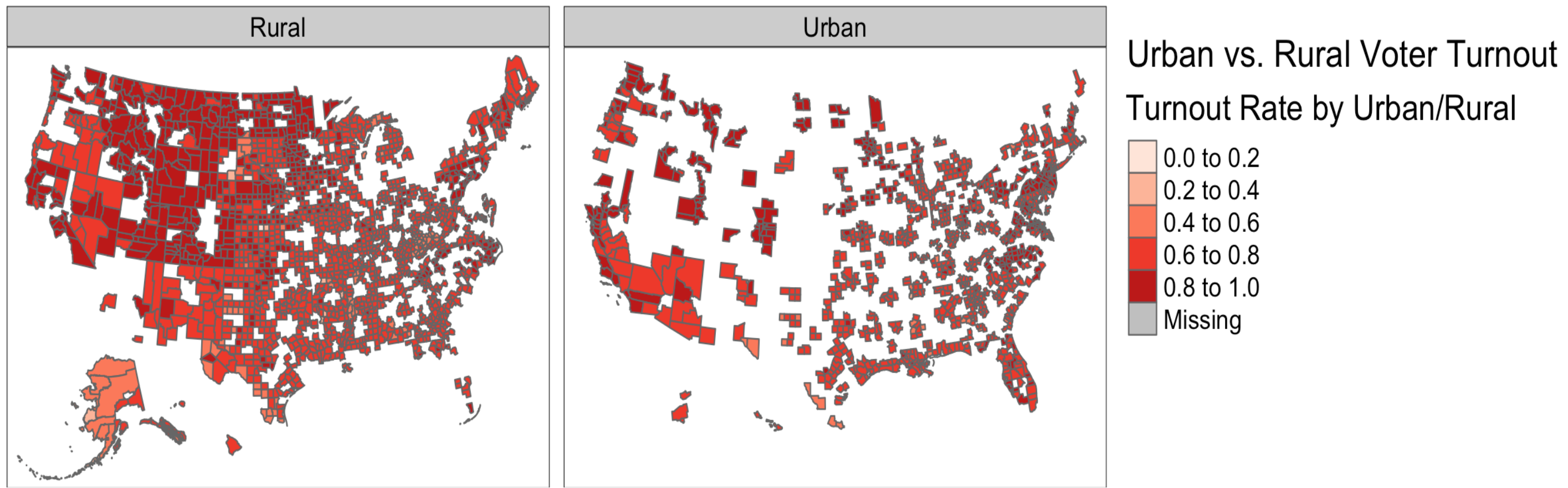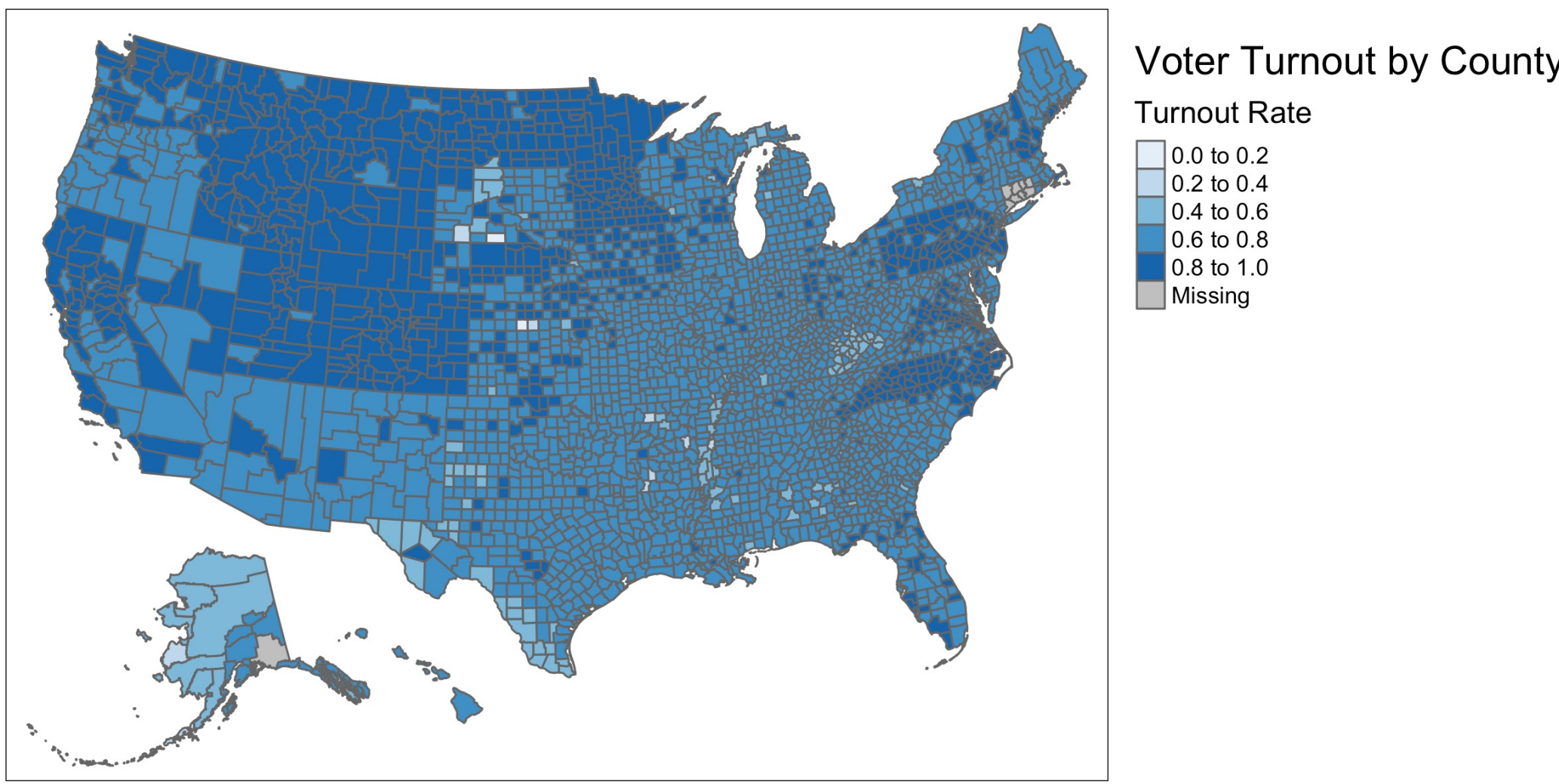
Conclusions

I found out I was actually wrong about my hypothesis. Higher income does not lead to increased voter turnout, neither do more urban counties. In fact, rural counties, despite economic and logistical barriers, demonstrated higher turnout rates compared to urban counties. I did find that race-specific turnout rates and median income are strong predictors of voter turnout at the county level during the 2020 election. White turnout consistently emerged as the most dominant factor driving overall turnout, followed by Hispanic turnout. In the end, these findings show the persistent role of geographic and demographic disparities in shaping electoral participation.