Race and the Costs of Prosocial Behavior[*]

Alexis H. Villacis[†]

December 4, 2025

Abstract

This paper examines racial differences in costly prosocial behavior using a nationally representative sample of U.S. adults recruited through Prolific. Respondents completed measures of trust, willingness to share, reciprocity, and a hypothetical donation task. We document pronounced racial disparities in behaviors that require vulnerability or material sacrifice, but not in low-cost prosocial actions. Multiracial respondents donate substantially less and report markedly lower trust than White respondents, even after adjusting for demographic, socioeconomic, and political covariates. These patterns are robust to alternative functional forms, including log-transformed, median, and upper-tail quantile regressions, and do not vary with income interactions. Black respondents exhibit higher generosity in mean and median specifications but do not differ from Whites in the upper tail of the donation distribution. Asian and Other respondents show no consistent behavioral differences. In contrast, sharing and reciprocity are uniformly high across all racial groups. The results suggest that costly prosociality is shaped by group-specific expectations about reciprocity and institutional reliability, consistent with insights from behavioral economics and stratification economics. The findings underscore the importance of distinguishing between types of prosocial behavior and highlight mechanisms through which racial classification structures cooperative decision-making in the United States.

Keywords: prosocial behavior; trust; generosity; racial classification; multiracial identity; behavioral economics; stratification economics; cooperation; inequality.
JEL Codes: C90, D91, D64, Z13, I31.

[*]The authors received no external funding or financial support for this research, and report no financial relationships or other conflicts of interest. All errors are our own.

[†]**Corresponding author.** Assistant Professor, Department of Agricultural, Environmental, and Development Economics, The Ohio State University, Columbus, OH 43210. Corresponding author email: Villacis.9@osu.edu.

1 Introduction

Prosocial behavior is often described as the quiet machinery of social life, a set of everyday actions such as trust, generosity, and reciprocity through which societies sustain cooperation. Economists typically model these domains as reflections of altruism, warm–glow motivations, social–image concerns, or reciprocal preferences (Andreoni, 1990; Fehr and Gächter, 2000; Bénabou and Tirole, 2006). While these frameworks emphasize heterogeneity in motives, they generally treat such heterogeneity as individual-specific rather than systematically patterned across groups.

Identity economics demonstrates that social identity shapes economic decisions by altering norms, expectations, and perceived payoffs (Akerlof and Kranton, 2000; Bénabou and Tirole, 2011). Stratification economics extends this logic by arguing that group-based differences reflect structural position rather than inherent preferences (Darity Jr, 2022). Empirical evidence of racial discrimination in labor markets and institutions further supports the idea that expectations about fairness are shaped by social position (Bertrand and Mullainathan, 2004). Recent work shows that decisions involving generosity, cooperation, and other costly forms of prosociality are sensitive to institutional reliability, social expectations, and perceived vulnerability (Gneezy, Keenan and Gneezy, 2014; DellaVigna, List and Malmendier, 2012). These insights suggest that prosocial behavior may be stratified along group-based lines, reflecting systematic differences in expectations, exposure, and vulnerability rather than idiosyncratic variation.

A substantial empirical literature demonstrates that trust varies sharply across racial groups and is shaped by historical experience, social environments, and institutional reliability. Low trust among marginalized groups has been linked to discrimination, historical trauma, and neighborhood composition (Alesina and La Ferrara, 2002; Nunn and Wantchekon, 2011; Guiso, Sapienza and Zingales, 2008). Sociological studies further show that ethnoracial diversity can reduce generalized trust and channel cooperation toward narrower in-group ties (Abascal and Baldassarri, 2015; Baldassarri and Abascal, 2020). Generosity and support for redistribution exhibit similarly patterned racial dynamics. Individuals rely on racialized beliefs about deservingness when evaluating charitable giving or welfare programs (Fong and Luttmer, 2011; Luttmer, 2001); even private giving responds to the perceived social identity of recipients (Bordalo et al., 2016).

Despite these findings, most economic research continues to treat race as a fixed monoracial classification, largely overlooking the rapid demographic rise and distinct social positioning of multiracial Americans. Multiracial individuals often experience inconsistent

racial labeling, pressure to conform to monoracial categories, and social invalidation (Miniear and Atkin, 2023; Zamora and Padilla, 2024). Their racial self-classification is also context-sensitive, varying with social environments and labor market conditions (Jackson, 2024). These structural and interpersonal dynamics imply that multiracial classification is not simply a demographic attribute but a social position with implications for beliefs, expectations, and prosocial decisions.

If prosocial behavior is shaped by social position, then racial classifications may correspond to systematic differences in how individuals engage in costly forms of cooperation. Stratification economics emphasizes that group-based inequality arises from structural positioning and inherited group status rather than from differences in preferences or individual capabilities (Darity Jr, 2022). This perspective implies that prosocial behaviors requiring vulnerability or sacrifice may reveal patterns of stratification that are not apparent in lower-cost behaviors. The position of multiracial Americans within the U.S. racial hierarchy, as a group often situated at the margins of monoracial categories, makes them an especially informative population for examining these dynamics.

To investigate how racial classifications correspond to patterns of prosocial behavior, we fielded a behavioral study with adult participants recruited from Prolific. Respondents completed hypothetical monetary donation tasks, a battery of trust items measuring willingness to assume benign intent, and measures of willingness to share and reciprocate favors. These four prosocial domains differ in the costs they impose. Donation and trust require exposure to potential disappointment or loss, while sharing and reciprocity impose minimal personal risk. This distinction aligns with theories of warm-glow giving, reciprocal preferences, and social-image considerations that highlight the role of psychological or material costs in defining prosocial actions (Andreoni, 1990; Fehr and Gächter, 2000; Bénabou and Tirole, 2006).

The results reveal a clear and asymmetric pattern. Individuals classified as multiracial donate substantially less than White, Black, Asian, or other respondents, and they exhibit the lowest levels of trust, even after adjusting for income, education, gender, age, political affiliation, employment status, and economic sentiment. Black respondents also report lower trust, consistent with prior work, but their donation behavior does not differ significantly from other groups once covariates are included. White and Asian respondents are statistically similar across all prosocial measures. In contrast, willingness to share and reciprocity exhibit no meaningful variation across racial classifications. Racial differences emerge only in domains where prosocial behavior is costly or risky, and they disappear entirely in domains requiring minimal personal sacrifice.

These findings extend behavioral economics by showing that prosocial preferences are not homogeneous across racial groups and that costly forms of prosociality are stratified. Where canonical models assume universal psychological foundations, our results indicate that racial group classifications correspond to systematic differences in generosity and trust. The study also deepens the literature on race and social cooperation by replicating established findings on racial disparities in trust (Alesina and La Ferrara, 2002; Abascal and Baldassarri, 2015) and by demonstrating that multiracial respondents face even steeper trust and generosity constraints. In addition, the results highlight that multiracial classification captures a set of structural and social experiences that translate into distinct behavioral outcomes, an insight unavailable when race is reduced to monoracial categories.

The distinction between costly and low-cost prosociality is central: costly actions depend on expectations about reciprocity and institutional fairness, whereas low-cost actions are governed more by social norms. Costly behaviors such as donation and trust require exposure to disappointment, reputational loss, or exploitation, making them especially sensitive to beliefs about reciprocity, fairness, and institutional reliability (Fudenberg and Levine, 2006; Gneezy, Keenan and Gneezy, 2014). Low–cost behaviors such as sharing and norm–driven reciprocity, in contrast, are more strongly governed by stable moral norms and impose minimal vulnerability. If racial classification shapes expectations through structural position (Darity Jr, 2022; Bertrand and Mullainathan, 2004), then group differences should appear most clearly in costly prosocial domains. The growing multiracial population, whose social position is often ambiguous or contested, provides a theoretically rich case for testing these predictions.

The remainder of the paper is organized as follows. Section 2 reviews related literature and develops the conceptual framework that links racial classification, social position, and costly prosocial decisions. Section 3 describes the data, sampling design, and measurement strategy. Section 4 outlines the empirical approach, including identification assumptions and robustness procedures. Section 5 presents the main results for donation, trust, sharing, and reciprocity. Section 6 reports additional robustness tests and evaluates potential mechanisms. Section 7 concludes with implications for behavioral theory, stratification economics, and policy.

2 Related Literature and Conceptual Framework

The behavioral economics literature provides a foundational account of why individuals engage in prosocial behavior. Early models conceptualized giving as pure altruism, but subsequent work demonstrated that individuals derive private psychological benefits from giving, including warm–glow and reputation concerns (Andreoni, 1990; Bénabou and Tirole, 2006). Experimental evidence shows systematic heterogeneity in giving across individuals (Fisman, Kariv and Markovits, 2007), challenging the notion that prosocial motives are purely idiosyncratic. Yet this literature rarely examines whether such heterogeneity is structured by group-based social position.

A parallel literature examines how reciprocity shapes cooperation. Positive and negative reciprocity, defined as rewarding kind actions and punishing unkind ones, have been shown to exert powerful influences on behavior even in anonymous or one-shot interactions (Fehr and Gächter, 2000). This evidence refutes the view that self-interest alone governs cooperative behavior and demonstrates that strong reciprocity norms are widespread. These theories imply that individuals respond differently to social interactions that entail vulnerability or potential exploitation, an insight that bears directly on the distinction between costly and low-cost prosocial behaviors.

Research on trust, diversity, and cooperation demonstrates that prosociality is shaped by lived environments and historical legacies. Trust varies sharply across racial groups in the United States and is influenced by discrimination, institutional reliability, and inherited historical experiences (Alesina and La Ferrara, 2002; Nunn and Wantchekon, 2011; Algan and Cahuc, 2010). Experimental and observational studies show that trust governs economic exchange, willingness to transact, and perceptions of institutional fairness (Guiso, Sapienza and Zingales, 2008). Sociological work further shows that ethnoracial diversity tends to reduce generalized trust and redirect cooperation to more circumscribed networks (Abascal and Baldassarri, 2015; Baldassarri and Abascal, 2020).

Racial differences in generosity and redistribution have also been extensively documented.

Preferences for redistribution depend not only on income and ideology but also on beliefs about recipient deservingness and industriousness (Fong and Luttmer, 2011; Luttmer, 2001). Racialized stereotypes systematically shape judgments of fairness and worthiness, producing divergent generosity toward in-group and out-group recipients (Bordalo et al., 2016). These mechanisms help explain persistent racial differences in support for redistribution and charitable giving.

Despite extensive work on race, cooperation, and redistribution, economic research rarely distinguishes between monoracial and multiracial populations. Yet racial classification in the United States is more fluid and context dependent than is often assumed. Sociological research demonstrates that multiracial individuals navigate inconsistent racial labeling, monoracial expectations, and experiences of identity invalidation (Minniear and Atkin, 2023; Zamora and Padilla, 2024). Quantitative work using census and survey data shows that changes in measurement practices have contributed to the rapid growth in multiracial identification (Ventura and Flores, 2025) and that racial self-classification among multiracial respondents is responsive to demographic and labor market conditions (Jackson, 2024). These findings indicate that multiracial classification reflects not only individual identity but also a structurally shaped social position—one characterized by inconsistent categorization, contested group boundaries, and context-dependent treatment. Although we do not observe internal identity processes directly, these literatures provide a substantive basis for treating multiracial classification as a meaningful construct when analyzing behavioral inequality.

The theoretical framework of stratification economics provides a unifying perspective on these patterns. Stratification economics rejects explanations of group inequality that rely on individual preferences or cultural deficiencies and instead emphasizes structural positioning, group boundaries, and intergroup relations as the primary mechanisms shaping economic outcomes (Darity Jr, 2022). From this viewpoint, behavioral differences across racial classifications are not expressions of inherent preferences but reflections of lived experiences, expectations, and constraints embedded in group-based hierarchies. Prosocial behaviors that entail risk or vulnerability are particularly likely to reflect these structural conditions, since they require individuals to make judgments about whether others will reciprocate, behave fairly, or refrain from exploitation. If some groups face more uncertain or hostile social environments, or if they occupy more marginal positions within the racial hierarchy, they may be less willing to take these risks.

This conceptual framework predicts that costly prosocial behaviors such as monetary donation and trust will exhibit greater racial stratification than low-cost behaviors such as sharing and reciprocity. Costly behaviors require exposure to potential loss or disappointment, which magnifies the role of expectations about others' fairness and reliability. Low-cost behaviors, in contrast, are more strongly governed by norms and social expectations and may therefore be less sensitive to group-based differences in social position. Multiracial individuals occupy a distinctive social location that may heighten these dynamics. Their classification is often contested, their group boundaries are ambiguous, and

their social experiences frequently involve inconsistent racial treatment. These conditions may shape expectations about vulnerability, reciprocity, and fairness in ways that affect costly prosocial behaviors.

In summary, prior research documents that prosociality is shaped by both individual motives and social environments, that trust and generosity vary systematically across racial groups, and that multiracial classification corresponds to distinctive structural and interpersonal experiences. Stratification economics provides a framework for understanding how these factors intersect. Building on these literatures, we examine whether and how racial classifications correspond to differences in prosocial behavior, with particular attention to the contrast between costly and low-cost forms of prosociality. This framework generates clear empirical predictions that guide the analyses that follow.

3 Data

We fielded a preregistered survey experiment with 1,000 adult participants recruited through Prolific Academic. Respondents were required to reside in the United States and were quota-matched to the U.S. population with respect to age, gender, and ethnicity. This design ensured broad demographic representation while allowing precise estimation of racial differences in prosocial outcomes. The first 6 percent of respondents were recruited without quotas in order to confirm that recruitment and survey delivery proceeded without technical issues. All responses were collected in Qualtrics. One duplicate Prolific ID was removed, producing a final analytic sample of $N = 999$.

The questionnaire measured four prosocial outcomes. The monetary donation measure followed a standard hypothetical-windfall paradigm often used in behavioral economics and psychology to assess generosity. Respondents were asked: "Imagine the following situation: Today you unexpectedly received $1,000. How much of this amount would you donate to a good cause?" They could enter any amount between 0 and 1000. Although hypothetical, such measures reliably capture variation in generosity and correlate with real-stakes giving in prior research.

Trust was measured using a single item assessing respondents' willingness to assume benign intent in others. Respondents were asked: "How well does the following statement describe you as a person? 'I assume that people may have only the best intentions.'" Responses were recorded on a scale from 0 to 10, where 0 indicated "does not describe me at all" and 10 indicated "describes me perfectly."

Willingness to share was measured with the item: "How willing are you to give to

good causes without expecting anything in return?" Respondents answered using a 0 to 10 scale, where 0 indicated being "completely unwilling to share" and 10 indicated "very willing to share."

Reciprocity was measured with the statement: "When someone does me a favor, I am willing to return it." Respondents rated how well this statement described them on a scale from 0 to 10, where 0 indicated "does not describe me at all" and 10 indicated "describes me perfectly."

TABLE 1: Summary Statistics by Racial Classification

Variable	White	Black	Asian	Multiracial	Other
N	716	118	56	70	39
Prosocial Outcomes					
Assume best intentions (0–10)	5.303	4.661	5.393	4.171	4.513
Willingness to share (0–10)	7.311	7.941	6.679	7.043	7.154
Reciprocity (0–10)	8.844	8.610	8.732	8.814	8.615
Donation amount ($)	170.97	208.22	154.11	116.63	139.23
Demographics					
Age	47.25	43.09	42.29	48.33	40.59
Income category (1–5)	3.161	3.246	3.500	2.871	3.179
Female (0/1)	0.541	0.636	0.571	0.586	0.436
Bachelor's or higher (0/1)	0.507	0.644	0.786	0.414	0.462
Employment Status					
Full-time (0/1)	0.429	0.551	0.482	0.429	0.462
Part-time (0/1)	0.151	0.144	0.143	0.157	0.103
Unemployed (0/1)	0.094	0.102	0.071	0.100	0.103
Self-employed (0/1)	0.112	0.093	0.179	0.143	0.103
Retired (0/1)	0.138	0.059	0.071	0.057	0.051
Economic Opinion					
Terrible (0/1)	0.166	0.246	0.107	0.257	0.256
Poor (0/1)	0.355	0.305	0.500	0.400	0.436
Fair (0/1)	0.346	0.356	0.286	0.286	0.282
Good (0/1)	0.117	0.093	0.089	0.029	0.026
Excellent (0/1)	0.015	0.000	0.018	0.029	0.000
Political Affiliation					
Democrat (0/1)	0.398	0.466	0.536	0.357	0.308
Republican (0/1)	0.332	0.271	0.196	0.257	0.205
Independent (0/1)	0.258	0.254	0.250	0.343	0.410
Other (0/1)	0.011	0.008	0.018	0.043	0.077

We also collected detailed demographic and socioeconomic information. Respondents

TABLE 2: Correlation Matrix of Prosocial Outcomes

	Trust	Sharing	Reciprocity	Donation
Trust	1.00	0.20	0.15	0.18
Sharing	0.20	1.00	0.34	0.31
Reciprocity	0.15	0.34	1.00	0.05
Donation	0.18	0.31	0.05	1.00

reported their annual family income by selecting one of five ordered categories: (1) less than $10,000 to $24,999, (2) $25,000 to $49,999, (3) $50,000 to $74,999, (4) $75,000 to $99,999, and (5) $100,000 or more. This ordered income variable is included in all regression specifications. Political affiliation was recorded using Prolific's standard measure, which distinguishes Republicans, Democrats, Independents, and those selecting Other. Additional covariates include age, gender, education, employment status, and subjective evaluations of the national economy.

Race was measured using Prolific's self-classification item, which allows respondents to select multiple racial categories. To construct analytically meaningful groups, we coded respondents as "White," "Black," or "Asian" when they selected only that single category. Respondents who selected more than one racial category were coded as "Multiracial," and those selecting a single category outside the three major groups (for example, American Indian or Alaska Native, Native Hawaiian or Pacific Islander, Other) were coded as "Other."

Table 1 reports the distribution of racial groups in the sample as well as descriptive characteristics by race. Table 2 reports pairwise correlations among the four prosocial outcomes. All data and replication code will be made publicly available at the Open Science Framework upon publication.

4 Empirical Strategy

Our empirical strategy is designed to provide a transparent and theoretically grounded assessment of whether and how prosocial behavior varies across racial classifications. The central aim is descriptive and diagnostic rather than causal. Racial classification is not manipulable in the sense required for causal identification, and our goal is therefore not to estimate causal effects of race but to measure systematic contrasts across racial groups that are informative about underlying patterns of stratification. This approach aligns with the theoretical framework outlined in Section 2, in which prosocial behavior is shaped by structural position and group-based historical experiences.

Let Y_i denote a prosocial outcome for respondent i. We estimate variants of the following baseline specification:

$$Y_i = \beta_0 + \beta_1 \text{Black}_i + \beta_2 \text{Asian}_i + \beta_3 \text{Multiracial}_i + \beta_4 \text{Other}_i + \mathbf{X}_i'\boldsymbol{\gamma} + \varepsilon_i. \tag{1}$$

The primary coefficients of interest are β_1 through β_4, which capture differences in prosocial outcomes relative to White respondents, the omitted category. The vector $\mathbf{X}_i$ includes age, gender, education, family income category, employment status, political affiliation, and subjective economic opinion. All covariates enter the model flexibly through sets of indicator variables except for age, which is treated as continuous. This adjustment set captures major demographic and socioeconomic correlates of prosocial tendencies documented in prior work.

This specification is estimated using ordinary least squares with heteroskedasticity-consistent (HC1) standard errors. Linear models provide coefficients that are directly interpretable as mean differences relative to the White category, and they facilitate comparisons across multiple outcomes. Although trust, sharing, and reciprocity are measured on 0 to 10 scales, prior research shows that linear models perform well for approximately continuous attitudinal outcomes and yield coefficients that are substantively meaningful and easily comparable.

The donation measure exhibits substantial right skew, with a long tail generated by a small number of large hypothetical donations. To address this skewness, we estimate models using the transformation $\log(1 + \text{donation})$, which reduces sensitivity to outliers and yields a distribution closer to normality. We also estimate quantile regressions for the donation outcome, focusing on the conditional median. Quantile regression allows for heterogeneity across the distribution of the dependent variable and provides robustness to heavy tails and non-normal error structures.

The racial classification indicators merit special discussion. Following standard practice in stratification research, we treat racial groups as mutually exclusive but socially constructed categories. Respondents selecting more than one category are coded as Multiracial, while all remaining single-category responses outside the three largest monoracial groups (White, Black, Asian) are coded as Other. This operationalization does not attempt to measure subjective racial identity or its salience, consistent with our emphasis on social classification rather than self-understood identity. Instead, it captures the categorical classification that respondents themselves report and that structures social interactions and institutional treatment in the United States. The contrasts estimated in Equation (1) there-

fore represent differences across racial classifications that are consistent with the conceptual framework presented in Section 2, which emphasizes the importance of group-based social position.

Our identification strategy is descriptive and rests on two assumptions. First, conditional on $\mathbf{X}_i$, the estimated coefficients reflect differences across racial classifications that are not driven by demographic or socioeconomic composition. Second, the specification captures the relevant forms of heterogeneity that are necessary to interpret the comparisons as contrasts in conditional means rather than artifacts of omitted variable bias. Because we do not aim to estimate causal effects of race, omitted variable bias does not pose the same inferential challenge it would in a causal design. Instead, the estimates quantify the patterns associated with racial classification after adjusting for major observable characteristics. In other words, our estimates characterize conditional associations, not causal effects, and are interpreted as descriptive contrasts across classification groups.

We assess the robustness of our findings through several additional specifications. First, we estimate models that include interactions between racial classification and income category to allow for heterogeneous income gradients across groups. Second, we examine the sensitivity of the results to alternative codings of education and economic opinion. Third, we estimate models using the untransformed donation amount and confirm that the main patterns are unchanged. Fourth, we re-estimate the models using only monoracial respondents to verify that the estimated Multiracial coefficients are not driven by misclassification or marginal cases. Across all robustness checks, the direction, magnitude, and significance of the estimated coefficients for Multiracial respondents remain stable.

Taken together, this empirical strategy provides a consistent framework for measuring contrasts in prosocial behavior across racial classifications. By estimating comparable models across costly and low-cost prosocial behaviors, the analysis leverages the conceptual distinction between these domains to evaluate whether stratification is more pronounced in actions that require vulnerability or sacrifice. This approach yields estimates that are straightforward to interpret and that speak directly to the theoretical predictions derived from stratification economics.

5 Results

This section presents the associations between racial classification and prosocial behavior. We begin with hypothetical donation amounts, then turn to trust, willingness to share,

and reciprocity. All models adjust for age, gender, education, family income, employment status, political affiliation, and subjective economic opinion. White respondents serve as the reference category throughout.

5.1 Donation

We first examine differences in the hypothetical donation measure. Column (1) of Table 3 reports the ordinary least squares estimates using the donation amount in levels. Multiracial respondents donate substantially less than White respondents, with a coefficient of -51.54 dollars ($p = 0.003$). Black respondents donate modestly more than White respondents, and the estimate is statistically significant at the 5% level (38.78, $p = 0.042$). Asian and Other respondents do not differ significantly from Whites.

TABLE 3: Race Coefficients Across Prosocial Outcomes

	Donation Amount	Trust	Sharing	Reciprocity
Asian	-16.42	0.017	-0.697^{*}	-0.006
	(25.12)	(0.296)	(0.355)	(0.197)
Black	38.78^{**}	-0.500^{**}	0.659^{***}	-0.151
	(19.06)	(0.252)	(0.216)	(0.210)
Multiracial	-51.54^{***}	-0.878^{***}	-0.205	-0.040
	(17.17)	(0.307)	(0.346)	(0.218)
Other	-12.64	-0.453	0.054	-0.179
	(29.34)	(0.300)	(0.412)	(0.246)
Additional Controls	Yes	Yes	Yes	Yes
Observations	999	999	999	999
R-squared	0.051	0.119	0.073	0.054

Notes: Each column reports coefficients from an OLS regression of the prosocial outcome on racial classifications. White is the omitted category. All models adjust for age, gender, education, income, party identification, employment status, and subjective economic opinion. Robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 1 displays adjusted predictions for donation amounts by race. Holding covariates at their modal or median values, Multiracial respondents are predicted to donate markedly less than members of any other racial group, reinforcing the results from the linear model.

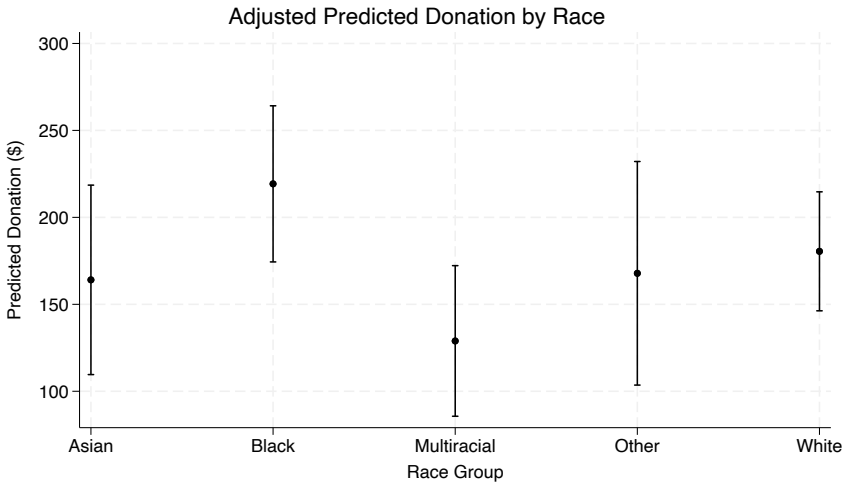


FIGURE 1: Adjusted predicted donation amount by racial classification. Predictions are evaluated at the median age and modal values of gender, education, income, and party identification. Error bars represent 95 percent confidence intervals.

5.2 Trust

We next examine trust, measured as willingness to assume that others have the best intentions. Column (2) of Table 3 reports the estimates. Black respondents report significantly lower trust relative to White respondents (-0.500, $p = 0.047$), while Multiracial respondents show an even larger trust deficit (-0.878, $p = 0.004$).[1] These differences remain stable across alternative specifications and are among the strongest in the analysis. Asian respondents do not differ from Whites, and the coefficient for the Other category, although negative, is not statistically distinguishable from zero.

The predictive margins shown in Figure 2 highlight these differences. Multiracial respondents exhibit the lowest expected trust scores, followed by Black respondents, with White and Asian respondents showing comparatively higher values. The size and robustness of the Multiracial trust deficit align with the theoretical expectations described in Section 2.

5.3 Willingness to Share

We then examine willingness to share resources with good causes without expecting anything in return. Column (3) of Table 3 presents the results. In contrast to donation and

[1] On a 0–10 scale, these deficits correspond to approximately 5–9 percent of the full range and 12–18 percent of the sample standard deviation.

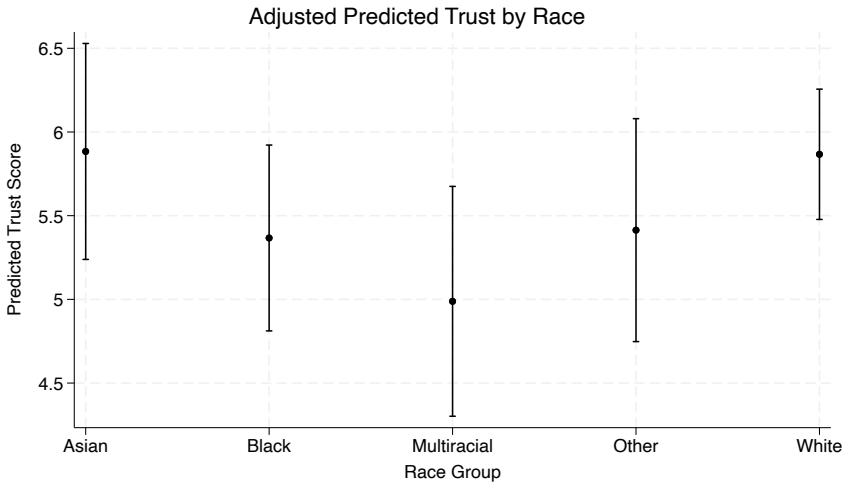


FIGURE 2: Adjusted predicted trust scores by racial classification. Predictions are evaluated at the median age and modal values of gender, education, income, and party identification. Error bars represent 95 percent confidence intervals.

trust, racial differences in sharing are limited. Multiracial respondents do not differ significantly from Whites (-0.205, $p = 0.555$). Asian respondents display a marginally negative association (-0.697, $p = 0.050$), and Black respondents show a positive and statistically significant association (0.659, $p = 0.002$). Respondents in the Other category are not statistically distinguishable from Whites.

The predicted values in Figure 3 indicate that willingness to share is high across all racial groups, with average scores between seven and nine on the ten-point scale. Black respondents report somewhat higher willingness to share than other groups, but the overall range is narrow and does not display the steep stratification observed in trust or, to a lesser degree, in donation behavior.

5.4 Reciprocity

Finally, Column (4) of Table 3 reports estimates for reciprocity. Consistent with the idea that reciprocity is a widely shared social norm, none of the racial coefficients are statistically significant, and the magnitudes are small. Multiracial, Black, Asian, and Other respondents all report reciprocity levels similar to those of Whites. Figure 4 shows that predicted reciprocity values are tightly clustered near the upper end of the 0-10 scale, with little variation across groups.

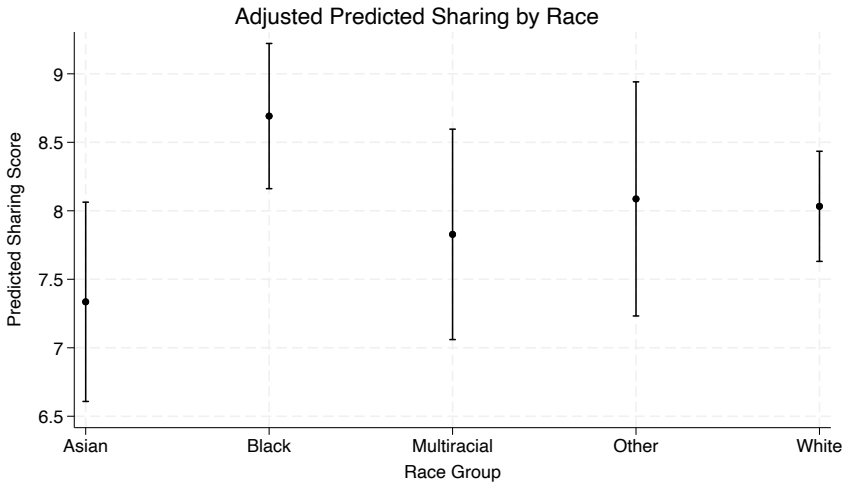


FIGURE 3: Adjusted predicted willingness to share by racial classification. Predictions are evaluated at the median age and modal values of the covariates. Error bars represent 95 percent confidence intervals.

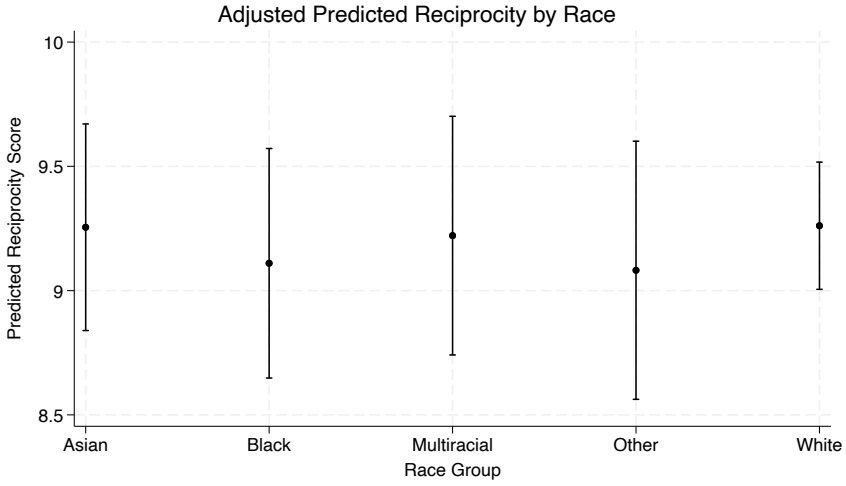


FIGURE 4: Adjusted predicted reciprocity scores by racial classification. Predictions are evaluated at the median age and modal values of the covariates. Error bars represent 95 percent confidence intervals.

5.5 Summary and Discussion

The results across the four prosocial domains reveal a consistent and theoretically informative pattern. Racial differences emerge most clearly in prosocial behaviors that require vulnerability or material sacrifice, namely donation and trust, and are far less pronounced in low-cost behaviors such as willingness to share or reciprocity. This asymmetry provides insight into the mechanisms through which racial classification shapes prosocial decision-making.

Multiracial respondents display a distinctive profile. They donate substantially less than White respondents and report markedly lower trust. The magnitude of the Multiracial trust deficit is among the strongest in the analysis. Although we do not observe subjective identity processes directly, prior research documents that multiracial classification is associated with inconsistent racial labeling and social invalidation. These experiences may generate uncertainty about how others will behave in interpersonal or institutional contexts. The reduced willingness to assume benign intent and to engage in costly prosocial action may therefore reflect expectations shaped by a more ambiguous or unstable social position within the racial hierarchy.

Black respondents also report lower trust than Whites, although their donation levels do not differ significantly once controls are included. This pattern suggests that trust and donation may be governed by related but distinct evaluative processes. Trust directly involves expectations about the intentions of others, whereas donation may be shaped more by personal norms, internal obligations, or altruistic motivations that are less sensitive to these expectations.

In contrast, willingness to share and reciprocity do not follow the same racial pattern. These behaviors impose little risk and rely more heavily on established social norms or moral obligations. Reciprocity in particular is a widely endorsed norm and shows virtually no variation across racial groups. Willingness to share is uniformly high across the sample, with small differences across racial classifications that do not resemble the disparities observed in donation or trust. These findings indicate that group-based differences in prosociality are concentrated in domains where the perceived costs of vulnerability are greater.

The combined pattern supports the theoretical argument that stratification operates most strongly in behaviors that require individuals to form expectations about the likelihood of fair or reciprocal treatment. Structural position and group-based historical experience influence these expectations. For respondents classified as Multiracial, who often oc-

cupy ambiguous or contested positions in the racial schema, the perceived costs of entering situations that require trust or sacrifice may be particularly high. For Black respondents, the historically documented consequences of discrimination and institutional unreliability correspond to lower trust but do not translate uniformly across all costly prosocial behaviors.

Overall, the results do not suggest broad differences in altruistic sentiment. Instead, they imply that the willingness to undertake prosocial actions that involve vulnerability is shaped by racialized expectations formed through structural experience. This pattern is consistent with the theoretical framework of stratification economics, which emphasizes the role of group-based social position rather than inherent preferences in shaping economic and social behavior.

6 Additional Robustness Tests

This section evaluates whether the main findings are sensitive to alternative functional forms of the donation outcome and to heterogeneity across income categories. Across all robustness checks, the central results remain unchanged: Multiracial respondents exhibit lower donation levels in the mean and upper tail of the distribution,[2] while Black respondents give more in the median and mean specifications but not in the upper tail. Asian and Other respondents do not differ significantly from Whites in any specification.

6.1 Alternative Functional Forms

Given the skewed distribution of donation amounts, we estimate several alternative specifications. First, the log transformation $\log(1 + \text{donation})$ is used to downweight extreme values. Results are reported in Column (1) of Table 4. The Multiracial coefficient remains negative but statistically insignificant, while the Black coefficient is positive and significant (0.672, $p < 0.001$). Asian and Other respondents again do not differ from Whites. These results parallel the patterns observed in the level specification and demonstrate that our conclusions are not driven by extreme donations.

Second, we estimate a median (0.5) quantile regression to assess robustness at the center of the donation distribution.[3] Results are reported in Column (2) of Table 4. In this specification, Multiracial, Asian, and Other respondents do not differ significantly from Whites. Black respondents, however, donate significantly more than Whites at the median

[2] The upper tail refers to relatively high self-reported donation amounts within the 0–1000 range.

[3] All quantile regressions include the same full set of covariates as in the baseline OLS models.

(40.40, $p = 0.008$). This result complements the OLS findings by indicating that the higher generosity of Black respondents appears not only in the transformed mean but also at the median.

Third, to examine whether racial differences emerge among high-donation respondents, we estimate a 0.9 quantile regression. Results are reported in Column (3) of Table 4. Here the Multiracial deficit becomes particularly pronounced: Multiracial respondents give substantially less than Whites in the upper tail (-167.61, $p = 0.010$). No other racial group shows significant differences at the 0.9 quantile. This result confirms that the lower donation levels among Multiracial respondents are especially salient among high-potential donors, consistent with the interpretation that donation is a costly prosocial behavior shaped by expectations of reciprocity and vulnerability.

Across all specifications, the direction and interpretation of the racial contrasts are stable: the Multiracial donation deficit persists in both mean and upper-tail behavior, the Black donation premium is evident at the mean and median but not at the upper tail, and Asian and Other respondents do not exhibit systematic differences.

TABLE 4: Robustness of Donation Outcomes Across Alternative Specifications

	log(1 + Donation) OLS	Median Donation (q = 0.5)	Donation (q = 0.9)
Asian	-0.159	-14.18	3.76
	(0.275)	(21.29)	(72.23)
Black	0.672^{***}	40.40^{***}	25.82
	(0.162)	(15.20)	(51.59)
Multiracial	-0.415	-3.39	-167.61^{***}
	(0.278)	(19.23)	(65.26)
Other	0.258	-3.13	-18.78
	(0.318)	(25.19)	(85.49)
Additional controls	Yes	Yes	Yes
Observations	999	999	999
R-squared / Pseudo R2	0.086	0.005	0.043

Notes: Models include age, gender, education, income category, party identification, employment status, and subjective economic opinion. log(1 + donation) estimated by OLS with robust standard errors. Median and 0.9 quantile regressions estimated via the interior point algorithm.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6.2 Income Interactions and Socioeconomic Heterogeneity

To assess whether the racial differences in donation behavior are driven by differential income gradients across groups, we estimate models that interact racial classification with income categories. The interaction effects are uniformly small in magnitude and statistically insignificant. None of the Multiracial $\times$ income terms approach significance, and the main Multiracial coefficient remains large, negative, and unchanged relative to the baseline log specification. Similarly, the Black $\times$ income terms are small and nonsignificant, indicating that the positive donation tendencies among Black respondents at the mean and median are not driven by higher marginal effects of income.

TABLE 5: Race $\times$ Income Interactions in Donation Behavior (Compact Summary)

	Largest Interaction Coefficient (Coefficient with SE)	Joint Test of Interactions (p-value)
Asian	1.684 (1.098)	0.31
Black	-0.854 (0.550)	0.44
Multiracial	0.683 (1.039)	0.89
Other	-1.674^{*} (0.892)	0.14
Observations	999	
R-squared	0.097	

Notes: Compact summary of interaction effects between racial classification and income categories from the log(1 + donation) specification with full covariates. Column 1 reports the interaction term with the largest magnitude for each race group, along with its robust standard error. Column 2 reports the p-value from a joint significance test (Wald test) for all income interactions for that race group. None of the interaction sets are statistically significant at conventional levels.
* $p < 0.10$.

The absence of meaningful interaction terms suggests that the racial contrasts in donation behavior are not explained by socioeconomic heterogeneity in income gradients. Instead, the differences reflect stable contrasts across racial classifications after conditioning on demographic and socioeconomic covariates. This reinforces the conclusion that racial classification is associated with donation behavior through mechanisms other than income differences alone.

Overall, these robustness checks confirm that the main findings are not sensitive to functional form, outlier influence, or differential income slopes. The patterns identified in the primary analyses reflect stable and substantively meaningful contrasts in prosocial behavior across racial classifications.

7 Conclusion

The findings of this study reveal systematic racial differences in costly prosocial behavior that are robust across model specifications and insensitive to alternative functional forms and socioeconomic interactions. Multiracial respondents consistently donate less and report lower trust than White respondents, even after adjusting for a wide set of demographic, socioeconomic, and political covariates. These disparities persist in log-transformed models and are especially pronounced in the upper tail of the donation distribution. Black respondents exhibit higher donation levels in the mean and median specifications, though this pattern does not carry into the upper tail. Asian and Other respondents show no consistent differences relative to Whites across any of the costly prosocial behaviors examined. In contrast, willingness to share and reciprocity are high across all groups and display minimal variation.

These patterns align with the conceptual framework discussed earlier in the paper. From a behavioral economics perspective, costly prosocial actions such as donation and trust require forming beliefs about the likelihood of reciprocity or exploitation. These expectations are shaped by prior experience, structural position, and institutional reliability (Guiso, Sapienza and Zingales, 2008; Fudenberg and Levine, 2006). Experiments show that generosity and participation in cooperative activities are highly sensitive to perceived motives of others and contextual cues about fairness and reliability (DellaVigna, List and Malmendier, 2012; Gneezy, Keenan and Gneezy, 2014). Thus, if racial classifications correspond to different lived experiences with institutions or social evaluation, group-based differences in costly prosociality are to be expected. Differences in trust and generosity across racial classifications may therefore reflect group-specific beliefs about the likelihood of reciprocity or institutional fairness. The particularly large trust deficit among Multiracial respondents, combined with their lower donation levels, is consistent with the idea that ambiguous or inconsistent social classification generates uncertainty and increases the perceived cost of vulnerability.

Our findings are consistent with stratification economics, which emphasizes the role of structural positioning and historical group boundaries in shaping economic expectations and behavior (Darity Jr, 2022). They also resonate with identity-based economic models showing that ambiguity or instability in social classification can influence expectations, norms, and willingness to engage in cooperative behavior (Akerlof and Kranton, 2000; Bénabou and Tirole, 2011). The results here suggest that the Multiracial donation and trust deficits are not simply a function of income, education, or political preferences. Instead,

these behavioral differences reflect positional inequalities rooted in social meanings and institutional patterns of treatment. The behavior of Black respondents, who report lower trust but equal or higher donations in some specifications, further illustrates the complexity of group-based adaptation within stratified environments.

These findings have implications for both policy design and the interpretation of behavioral evidence. Interventions intended to promote civic participation, charitable giving, or cooperative engagement should recognize that individuals differ not in baseline prosocial motivation but in the expected costs of exposure to risk. Policies that assume uniform behavioral responses may overlook important heterogeneity in how different groups experience vulnerability. For Multiracial populations in particular, the heightened uncertainty associated with social evaluation and institutional engagement may warrant tailored outreach strategies that reduce perceived risks and increase institutional reliability.

The study also suggests caution in interpreting racial differences in trust or generosity as evidence of deficits in social capital. For some groups, lower trust or lower donation levels may be adaptive responses to environments characterized by inconsistent treatment or greater exposure to discrimination. Addressing underlying structural inequalities, rather than attempting to modify individual preferences or civic attitudes, may yield more durable improvements in prosocial engagement.

At the same time, the results highlight the importance of distinguishing between types of prosocial behavior. Low-cost prosocial actions such as sharing and reciprocity do not reveal the disparities that emerge in costly forms of cooperation. Focusing on these low-cost measures alone may obscure meaningful inequalities in behaviors that matter for resource allocation, institutional functioning, and collective action.

Future research could build on this study by linking behavioral measures to contextual indicators of institutional quality, discrimination, and local social environments. Longitudinal data on Multiracial respondents would be particularly valuable for understanding how experiences of social validation or invalidation affect trust formation and economic decision making over time. In addition, experimental interventions that manipulate the perceived reliability of partners or institutions may shed light on the specific mechanisms through which racial classification shapes costly prosociality.

In conclusion, the evidence presented here shows that costly prosocial behavior varies across racial classifications in ways that reflect both behavioral beliefs and structural context. The patterns observed among Multiracial and Black respondents highlight the intertwined influence of expectations, social position, and vulnerability. By integrating behavioral and stratification perspectives, this study demonstrates that racial disparities in

cooperation and generosity are not simply a matter of preferences, but arise from the environments in which individuals evaluate the risks of helping others. These insights highlight the importance of designing policies and institutions that reduce uncertainty, enhance fairness, and recognize the social foundations of economic behavior, particularly for individuals situated in ambiguous or structurally marginalized racial positions.

References

Abascal, Maria, and Delia Baldassarri. 2015. "Love Thy Neighbor? Ethnoracial Diversity and Trust Reexamined." *American Journal of Sociology*, 121(3): 722–782.

Akerlof, George A., and Rachel E. Kranton. 2000. "Economics and Identity." *Quarterly Journal of Economics*, 115(3): 715–753.

Alesina, Alberto, and Eliana La Ferrara. 2002. "Who Trusts Others?" *Journal of Public Economics*, 85(2): 207–234.

Algan, Yann, and Pierre Cahuc. 2010. "The Economic Effects of Trust." *American Economic Review*, 101(3): 2060–2092.

Andreoni, James. 1990. "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving." *The Economic Journal*, 100(401): 464–477.

Baldassarri, Delia, and Maria Abascal. 2020. "Diversity and Prosocial Behavior." *Science*, 369(6508): 1183–1187.

Bénabou, Roland, and Jean Tirole. 2011. "Identity, Morals, and Taboos: Beliefs as Assets." *American Economic Review*, 101(3): 322–326.

Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review*, 94(4): 991–1013.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. 2016. "Stereotypes." *Quarterly Journal of Economics*, 131(4): 1753–1794.

Bénabou, Roland, and Jean Tirole. 2006. "Incentives and Prosocial Behavior." *American Economic Review*, 96(5): 1652–1678.

Darity Jr, William A. 2022. "Position and Possessions: Stratification Economics and Intergroup Inequality." *Journal of Economic Literature*, 60(2): 400–426.

DellaVigna, Stefano, John A. List, and Ulrike Malmendier. 2012. "Testing for Altruism and Social Pressure in Charitable Giving." *Quarterly Journal of Economics*, 127(1): 1–56.

Fehr, Ernst, and Simon Gächter. 2000. "Fairness and Retaliation: The Economics of Reciprocity." *Journal of Economic Perspectives*, 14(3): 159–181.

Fisman, Raymond, Shachar Kariv, and Daniel Markovits. 2007. "Individual Preferences for Giving." *American Economic Review*, 97(5): 1858–1876.

Fong, Christina M, and Erzo FP Luttmer. 2011. "Do Fairness and Race Matter in Generosity? Evidence from a Nationally Representative Charity Experiment." *Journal of Public Economics*, 95(5–6): 372–394.

Fudenberg, Drew, and David K. Levine. 2006. "A Dual-Self Model of Impulse Control." *American Economic Review*, 96(5): 1449–1476.

Gneezy, Uri, Elizabeth A. Keenan, and Ayelet Gneezy. 2014. "Avoiding Overhead Aversion in Charity." *Science*, 346(6209): 632–635.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2008. "Trust and Social Capital." *Quarterly Journal of Economics*, 122(3): 1095–1131.

Jackson, Osborne. 2024. "The Impact of Market Factors on Racial Identity: Evidence from Multiracial Survey Respondents." *Federal Reserve Bank of Boston Working Paper*, 1(24-13).

Luttmer, Erzo FP. 2001. "Group Loyalty and the Taste for Redistribution." *Journal of Political Economy*, 109(3): 500–528.

Minniear, Mackensie, and Annabelle L Atkin. 2023. "Exploring Multiracial Identity, Demographics, and the First Period Identity Crisis: The Role of the 2020 United States Census in Promoting Monocentric Norms." *Journal of Applied Communication Research*, 51(1): 37–54.

Nunn, Nathan, and Leonard Wantchekon. 2011. "The Slave Trade and the Origins of Mistrust in Africa." *American Economic Review*, 101(7): 3221–3252.

Ventura, Ilana M, and René D Flores. 2025. "The "Rise" of Multiracials? Examining the Growth in Multiracial Identification in the 2020 U.S. Census." *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 11(1): 44–64.

Zamora, Tatiana I, and Amado M Padilla. 2024. "Making Sense of Conflicting Messages of Multiracial Identity: A Systematic Review." *Frontiers in Psychology*, 15: 1307624.