Gender, Generation and Multiracial Identification in the United States

PAA 2019 Submission

Janet Xu
Department of Sociology
Princeton University

Aliya Saperstein
Department of Sociology
Stanford University

Ann Morning
Department of Sociology
New York University

Sarah Iverson
Department of Sociology
New York University

Short Abstract

Previous research finds that women are more likely to identify as multiracial than men. However, most studies offer a limited perspective on multiracial Americans by: 1) restricting attention to children of different-race parents (i.e. first-generation multiracials) and 2) measuring identification as checking two boxes on official forms such as censuses or surveys. Using the 2015 Pew Survey of Multiracial Adults, we offer a more comprehensive intersectional analysis that considers gender differences at various stages: who is aware of having multiracial ancestry, who identifies as mixed race in general, and who selects multiple races for self-identification. Generally, females are more likely to be *aware* of having multiracial ancestry, but these patterns differ by race. Furthermore, only *first-generation* females are more likely to choose more than one race for themselves. This offers a more nuanced picture of how gender interacts with other social processes to shape racial identification in the U.S.

Gender, Generation and Multiracial Identification in the United States

How individuals of mixed-race ancestry self-identify and racially "label" themselves is of great interest to demographers and scholars of race and ethnicity. Do mixed-race individuals identify with primarily one race, or do they see themselves as distinctly mixed race or multiracial, existing "betwixt and between" races? When multiple-race reporting is permitted (as has been the case for the U.S. Census since 2000), do they only report one race, or do they select two or more options? The pool of individuals who *can* report multiple races or identify as mixed-race is larger than those who *do* so on surveys and censuses. These patterns of self-identification have important implications for race relations, population forecasting, and public policy.

Recent scholarship on multiracial identification in the United States has paid careful attention to how these identification processes are patterned by gender and racial ancestry (Davenport 2016; Liebler 2016). However, like much research on mixed-race populations, these studies are limited to children of interracial unions – what Morning and Saperstein (2018) call "first-generation" multiracials. A narrow focus on these individuals not only ignores the self-identification patterns of the millions of Americans descending from interracial unions that took place before their parents' generation, but can also contribute to the misleading notion that American multiraciality is a relatively "new" phenomenon. Moreover, we have little empirical evidence that the gendered and racialized identification processes observed among first-generation mixed-race individuals generalize to the higher-generation, "submerged" multiracial population (Roth and Ivemark 2018).

To offer a more historically grounded, intersectional perspective on multiraciality, we use a new national dataset that gives us some insight on the distribution of knowledge about mixed-race ancestry to investigate differentials in *awareness* of multiraciality as well as differentials in *self-identification*. Research on multiracial identification has generally shown the association between multi-race reporting and social identities like gender as a single, undifferentiated relationship. For example, Davenport (2016) demonstrates that biracial women are more likely than biracial men to select more than two races when asked to self-identify. We contend, however, that there are (at least) two stages of multiracial identification at work with different gendered and generational patterns at each stage. First, individuals with mixed-race ancestry must be *aware* of their multiple-race ancestry. Second, and conditional on awareness, individuals choose to *identify* as multiracial (see Goldstein and Morning 2000). In this study, we also investigate two measures of racial self-identification: 1) selecting two or more races when asked to identify one's race or origin, and 2) whether one explicitly considers oneself "mixed race or multiracial."

We expect that, at both the awareness and self-identification stages, who ultimately gets counted as "multiracial" will vary by factors such as gender, multiracial generation, and one's specific racial ancestry combination. In the full paper, we also plan to explore two additional potential mechanisms that could explain observed differentials in propensity to report multiple races or identify as mixed race: the particular historical period in which the first interracial union in the family tree took place, and loss of contact with kin of a particular racial background. In doing so, we aim to offer the most comprehensive analysis of patterns of multiracial ancestry and identification in the U.S. to date.

Data:

We use data from Pew Research Center's 2015 Survey of Multiracial Adults. To our knowledge, this is the first survey of multiracial adults drawn from a nationwide random sample. The survey was conducted online by GfK, which maintains a nationally representative panel of respondents (KnowledgePanel) recruited by random-digit dialing or address-based sampling. A screening survey of 22,719 Americans first asked for racial self-identification, and the races of parents, grandparents, and earlier ancestors. From this data, Pew identified a subsample of multiracial adults who were asked additional questions about their identity, family relationships, and social attitudes. To be included, respondents had to report multiple races for themselves, report that their parents were different races from each other and/or from themselves, report that their grandparents were different races from each other and/or themselves, *or* (in a more limited number of cases) report that their great- grandparents were different races from each other and/or themselves.[1] For more details of the study design see Pew Research Center (2015) and Patten (2015).

The structure of the survey allows us to treat self-identification as multiracial as the outcome of interest rather than the sole criteria for sample selection. Our analytic sample (N=2296) is limited to multiracial adults who completed the longer survey and do not have missing data on any of the race or ancestry measures.

Methods and Measures:

Most studies of multiraciality are limited to studying people who select two or more races for themselves. The Survey of Multiracial Adults also included a question about whether respondents consider themselves "mixed race or multiracial," regardless of whether or not they check two boxes on a typical census-style race question. This measure helps to capture someone who thinks of themselves as "mixed race or multiracial" but only selects one category for their self-identified race (such as Black). We expect the size of the population identifying as "mixed" in this more general sense to fall somewhere in between the population of people aware of their multiracial ancestry (i.e., everyone in our analytic sample) and the population estimated by counting only those who select two or more races on a survey. The relationship between identifying as mixed and other factors such as gender and multiracial generation may also differ. Thus, we consider both measures as dependent variables in our analysis.

Dependent Variables:

Selects 2 or more federally -designated racial or ethnic categories for self-identification: As part of the screening survey, respondents were asked, "What is your race or origin?" and could mark one or more of the following boxes: 1) White, 2) Hispanic, Latino, or Spanish origin, 3)Black or African American, 4) Asian or Asian-American, 5) American Indian, 6) Native Hawaiian or other Pacific Islander, and 7) some other race or origin. We count a multiracial response as anyone who selected two or more of the first six boxes.[2] In our sample, about one-third of adults who report multiracial ancestry also self-identify as multiracial by selecting two or more races.

---

[1] This approach does select on individuals who satisfy more presentist logics of multiracial ancestry. However, we find significant variation in self-identification even in this somewhat restricted sample.

[2] We err on the side of caution and do not include people who checked "Other" in either our assessment of multiracial ancestry or self-identification because respondents were only asked to specify what "Other" meant for themselves. Thus, if someone checked "Other" for both their mother and father, we could not determine if this

Considers self mixed-race or multiracial: In one of the first questions on the full multiracial survey, most respondents were asked, "Do you consider yourself to be mixed race or multiracial, that is more than one race, or not?" and could respond with "Yes" or "No". Respondents who previously identified as Hispanic, Latino or Spanish origin were given an alternate wording that included the phrases "mestizo" or "mulatto." We combined these two versions into a single indicator for identifying as "mixed." About 40 percent of our sample explicitly identifies as mixed or multiracial.

Key Independent Variables:

Generation: We distinguish between 1st generation and 2nd generation or higher multiracials. First-generation is operationalized as people who report their parents as both single-race but different races from each other, AND do not attribute additional races to grandparents, great-grandparents, and other ancestors. For more details on the coding of multiracial generation, see Morning and Saperstein (2018).

Gender: Most demographic characteristics in the Pew Survey of Multiracial Adults come from data previously collected by GfK. This includes information about whether the respondent identified as male or female, which we use to differentiate respondents by sex/gender in our analysis.[3]

Racial Ancestry Regimes: With six possible race responses across three generations, it is not possible for us to account for all possible combinations of racial ancestries reported in the survey. Instead, we highlight several key combinations that previous research has found to yield different patterns of multiracial identification (e.g., Gullickson and Morning 2011, Roth 2005), in part because of being subject to different historical norms of hypo- or hyperdescent:

- White/Indian (N = 1130): anyone who reports White and American Indian ancestry and no additional races; this serves as the reference category in our regression models.
- Any-Black (N = 660): anyone who reports some Black ancestry
- Non-Black Hispanic (N = 380): anyone with some Hispanic but not Black ancestry
- Non-Black, Non-Hispanic Asian (N = 126): anyone with some Asian ancestry who is not included in any previous category[4]

Preliminary Results:

In theory, there should be very little gender difference in estimates of the multiracial population, as we have no reason to think that interracial unions would give rise to more female descendants than male ones. Therefore, any significant departure from a 50:50 female-male population ratio signals a gendered process of either awareness or identification. However, imbalances in survey participation by race and gender also must be factored in to any such calculations.

---

indicated their parents shared the same "Other" race or if they had two different "Other" races. This suggests that, on this dimension, our count of people who report multiracial ancestry is conservative.

[3] Other controls included in our multivariate models, such as age, nativity, educational attainment, marital status, presence of children in the household, and place of residence also come from this profile data.

[4] We omit respondents who reported only White and Native Hawaiian or Pacific Islander ancestry (N = 13) from the analysis because there are too few cases to analyze statistically.

Table 1 shows counts and sex ratios for the full screening sample by our four racial regime categories.[5] Overall, the full sample sex ratio is 1.04, reflecting greater female participation in household-based surveys in the U.S. (Groves 2006). However, this imbalance ranges from 1.03 among people who identify as White to 1.62 among people who identify as Black (column 3). To calculate the sex ratio for respondents in the full sample who are aware that they have multiracial ancestry (column 6), we adjust for pre-existing imbalances in the population at risk – i.e, the population in each racial category that participated in the screening survey. For each row, we take the counts of female and male respondents who report two or more races in their family tree (columns 4 and 5), then divide them by the counts of female and male respondents in the full sample who share the same racial category (columns 1 and 2). We then divide the female proportion by the male proportion to reach our outcome of interest (column 6). We make analogous calculations in Table 2 to adjust our self-identification sex ratios for imbalances in our analytic sample – i.e., in the population that is aware they have multiracial ancestry.

We find departures from gender balance at both stages of awareness (Table 1) and self-identification (Table 2). Overall, female respondents are more likely to be aware of and report multiple races in their family trees, but this female skew differs by race and is driven by females who report white and/or Black ancestry (Table 1, column 6). Among people with non-Black Hispanic and non-Black, non-Hispanic Asian ancestry, the aware population skews slightly male, and awareness is gender-balanced among people who report American Indian ancestry.

In our analytic sample, all of whom are aware of multiracial ancestry, we find that gender imbalances further vary by generation and the measure of self-identification used (see Table 2). A larger proportion of females select two or more races than males, but only among first-generation adults. In higher generations, men are over-represented (column 6). Men also tend to be over-represented among people who consider themselves mixed race or multiracial more generally (column 4). This suggests that gender differences in awareness of multiracial ancestry cannot explain observed gender differences in multiracial identification. It also indicates that observed population compositions depend on how people choose to measure multiraciality.

The results hold in preliminary multivariate models. All else being equal, people whose multiracial ancestry dates to their grandparents or earlier ancestors are significantly less likely to choose two or more races, and generational attrition is even greater among female respondents (Table 3). Female and higher generation multiracial adults are also significantly less likely to think of themselves as "mixed" in a general sense (Table 4). Figures 1 and 2 present predicted probabilities calculated from the "Gender x Generation" models in Tables 3 and 4. Notably, people who report any Black ancestry are among the more likely to consider themselves mixed despite being the least likely to select two or more races for self-identification.

These patterns have been overlooked in previous studies in part because of a limited definition of who counts as "multiracial." Our intersectional analysis – which takes both gender and generation into account – reveals a more nuanced picture of the demography of U.S. multiracial adults that we hope can serve as a model for future research.

---

[5] In Table 1, the racial categories include anyone who self-identifies or reports any ancestry related to that category, including monoracial respondents with no other races in their family tree.

References

Davenport, Lauren D. 2016. "The Role of Gender, Class, and Religion in Biracial Americans' Racial Labeling Decisions." *American Sociological Review* 81 (1): 57–84. https://doi.org/10.1177/0003122415623286.

Goldstein, Joshua R., and Ann J. Morning. 2000. "The Multiple-Race Population of the United States: Issues and Estimates." *Proceedings of the National Academy of Sciences of the United States of America* 97 (11): 6230–35.

Groves, Robert M. "Nonresponse Rates and Nonresponse Bias in Household Surveys." *The Public Opinion Quarterly* 70, no. 5 (2006): 646–75.

Gullickson, Aaron, and Ann Morning. 2011. "Choosing Race: Multiracial Ancestry and Identification." *Social Science Research* 40 (2): 498–512. https://doi.org/10.1016/j.ssresearch.2010.12.010.

Liebler, Carolyn A. 2016. "On the Boundaries of Race: Identification of Mixed-Heritage Children in the United States, 1960 to 2010." *Sociology of Race and Ethnicity* 2 (4): 548–68. https://doi.org/10.1177/2332649216632546.

Morning, Ann, and Aliya Saperstein. 2018. "The Generational Locus of Multiraciality and Its Implications for Racial Self-Identification." *The ANNALS of the American Academy of Political and Social Science* 677 (1): 57–68. https://doi.org/10.1177/0002716218754774.

Patten, Eileen. 2015. "Who Is Multiracial? Depends on How You Ask." *Pew Research Center's Social & Demographic Trends Project* (blog). November 6, 2015. http://www.pewsocialtrends.org/2015/11/06/who-is-multiracial-depends-on-how-you-ask/.

Pew Research Center. 2015. "Multiracial in America." *Pew Research Center's Social & Demographic Trends Project*, 2015. http://www.pewsocialtrends.org/2015/06/11/multiracial-in-america/.

Roth, Wendy D. 2005. "The End of the One-Drop Rule? Labeling of Multiracial Children in Black Intermarriages." *Sociological Forum* 20 (1): 35–67. https://doi.org/10.1007/s11206-005-1897-0.

Roth, Wendy D., and Biorn Ivemark. 2018. "Genetic Options: The Impact of Genetic Ancestry Testing on Consumers' Racial and Ethnic Identities." *American Journal of Sociology* 124 (1): 150–84. https://doi.org/10.1086/697487.

Table 1: Sex Ratio by Racial Ancestry (Full Sample)						
	(1) Female Count	(2) Male Count	(3) Survey Participation Sex Ratio	(4) Aware of Multiraciality Female Count	(5) Aware of Multiraciality Male Count	(6) Aware Sex Ratio
Racial Ancestry						
Any Black	1101	681	1.62	645	345	1.16
Non-Black Hispanic	1456	1209	1.20	596	542	0.91
Non-Black/Hisp Asians	289	278	1.04	77	77	0.96
White-Indians	861	771	1.12	857	771	1.00
Any American Indian	1365	1054	1.30	1347	1036	1.00
Any White	9645	9336	1.03	1940	1653	1.14
Sample Total	**11591**	**11128**	**1.04**	**2320**	**1879**	**1.19**

Note: Racial ancestry categories are not mutually exclusive or exhaustive, so they do not necessarily sum to sample total. The survey participation sex ratio is the count of females divided by the count of males. The aware sex ratio is a ratio of proportions: we first calculate the number of females who report multiracial ancestry divided by the total number of females in the sample and the same for males. We then divide the female "aware" proportion by the male "aware" proportion. This adjusted for the initial sex ratio imbalance in survey participation. Sex ratios that favor females are shaded in gray.

Table 2: Sex Ratio in Multiracial Awareness and Self-Identification (Analytic Sample)

Racial Ancestry Regime and Multiracial Generation	Aware of multiracial ancestry		Considers self "mixed"		Selects 2+ races for self-id	
	(1) Count	(2) Sex Ratio	(3) Count	(4) Sex Ratio	(5) Count	(6) Sex Ratio
Any Black						
1st generation	47	1.76	35	0.96	24	1.70
2nd gen. or higher	613	1.95	281	0.97	164	0.91
Non-Black Hispanic						
1st gen	108	1.40	66	1.10	43	1.21
2nd gen+	272	1.18	143	0.69	137	0.66
Non-Black/Hisp Asians						
1st gen	70	1.00	63	0.97	49	1.04
2nd gen+	56	1.07	30	1.22	29	1.00
White-Indians						
1st gen	131	0.87	68	0.63	45	1.20
2nd gen+	999	1.24	237	0.70	327	0.75
Total	2296	1.36	923	0.84	818	0.82

Note: The aware sex ratios presented here (column 2) are ratios of counts. The self-identification sex ratios (columns 4 and 6) are ratios of proportions that account for the size of the population at risk of identifying as multiracial – in this case the counts of females and males with the given racial ancestry and of the given generation who are aware they have multiracial ancestry. Sex ratios that favor females are shaded in gray.

Selecting 2+ Races

Table 3: Odds of Selecting 2+ Races

	Baseline Demographic	+ Racial Ancestry	+ Generation	Gender x Generation
	(1)	(2)	(3)	(4)
2nd Gen or higher			0.197***	0.335*
			(0.080)	(0.152)
Age*2ndGen			1.033***	1.029***
			(0.008)	(0.009)
Any Black		0.810	0.798	0.801
		(0.093)	(0.092)	(0.093)
Non-B Hispanic		2.073***	1.988***	1.970***
		(0.298)	(0.294)	(0.292)
Non-B Non-H Asian		2.923***	2.623***	2.642***
		(0.608)	(0.572)	(0.576)
Age	0.985***	0.988***	0.961***	0.964***
	(0.003)	(0.003)	(0.008)	(0.008)
Female	0.678***	0.705***	0.702***	1.218
	(0.063)	(0.066)	(0.066)	(0.290)
2nd Gen Female				0.522*
				(0.135)
Northeast	1.671***	1.634***	1.637***	1.643***
	(0.243)	(0.242)	(0.243)	(0.245)
Midwest	1.169	1.169	1.158	1.155
	(0.143)	(0.145)	(0.145)	(0.145)
West	1.729***	1.443**	1.435**	1.441**
	(0.194)	(0.168)	(0.168)	(0.168)
Non-Metro	1.251	1.276	1.266	1.268
	(0.162)	(0.169)	(0.168)	(0.169)
BA or higher	1.204*	1.169	1.165	1.174
	(0.112)	(0.111)	(0.111)	(0.112)
Married	0.925	0.884	0.895	0.897
	(0.089)	(0.088)	(0.090)	(0.090)
Foreign Born	0.796	0.595**	0.620*	0.617*
	(0.145)	(0.114)	(0.119)	(0.119)
Nativity Q Missing	1.109	1.357	1.392	1.363
	(0.614)	(0.764)	(0.781)	(0.768)
Spanish Version	1.333	0.947	0.964	0.968
	(0.367)	(0.274)	(0.280)	(0.282)
Children	0.856	0.834	0.835	0.836
	(0.101)	(0.100)	(0.101)	(0.102)
Intercept	1.151	0.962	3.867***	2.458*
	(0.223)	(0.201)	(1.586)	(1.089)
Observations	2253	2253	2253	2253
Log Likelihood	-1428	-1397	-1389	-1385
Akaike Inf. Crit.	2883	2826	2813	2809
Bayesian Inf. Crit.	2957	2918	2916	2918

Note: *p<0.05; **p<0.01; ***p<0.001

Explicitly Considers Self Mixed Race or Multiracial

Table 4: Odds of Identifying as Mixed or MR

	Baseline Demographic	+ Racial Ancestry	+ Generation	Gender x Generation
	(1)	(2)	(3)	(4)
2nd Gen or higher			0.057***	0.044***
			(0.027)	(0.023)
Age*2ndGen			1.037***	1.039***
			(0.009)	(0.009)
Any Black		2.519***	2.670***	2.665***
		(0.280)	(0.304)	(0.303)
Non-B Hispanic		3.338***	2.740***	2.761***
		(0.486)	(0.414)	(0.418)
Non-B Non-H Asian		6.953***	4.415***	4.435***
		(1.572)	(1.050)	(1.057)
Age	0.986***	0.991**	0.960***	0.958***
	(0.003)	(0.003)	(0.008)	(0.008)
Female	0.674***	0.652***	0.653***	0.495**
	(0.061)	(0.062)	(0.063)	(0.129)
2nd Gen Female				1.379
				(0.385)
Northeast	1.310	1.216	1.244	1.244
	(0.189)	(0.182)	(0.189)	(0.189)
Midwest	0.934	0.976	0.993	0.994
	(0.113)	(0.122)	(0.127)	(0.127)
West	1.517***	1.363**	1.377**	1.376**
	(0.167)	(0.159)	(0.164)	(0.164)
Non-Metro	0.960	1.196	1.207	1.206
	(0.125)	(0.163)	(0.168)	(0.167)
BA or higher	0.911	0.840	0.843	0.840
	(0.084)	(0.081)	(0.083)	(0.082)
Married	0.674***	0.779*	0.793*	0.793*
	(0.064)	(0.078)	(0.081)	(0.080)
Foreign Born	1.674**	1.081	1.199	1.198
	(0.291)	(0.198)	(0.222)	(0.222)
Nativity Q Missing	1.286	1.418	1.648	1.651
	(0.826)	(0.919)	(1.068)	(1.070)
Spanish Version	0.842	0.641	0.754	0.751
	(0.223)	(0.179)	(0.213)	(0.212)
Children	0.919	0.848	0.838	0.837
	(0.107)	(0.103)	(0.105)	(0.104)
Intercept	1.815**	0.851	9.937***	12.617***
	(0.349)	(0.179)	(4.667)	(6.529)
Observations	2243	2243	2243	2243
Log Likelihood	-1463	-1388	-1348	-1347
Akaike Inf. Crit.	2953	2808	2732	2732
Bayesian Inf. Crit.	3027	2899	2835	2841

Note: *p<0.05; **p<0.01; ***p<0.001

Figure 1: Probability of Selecting 2+ Races

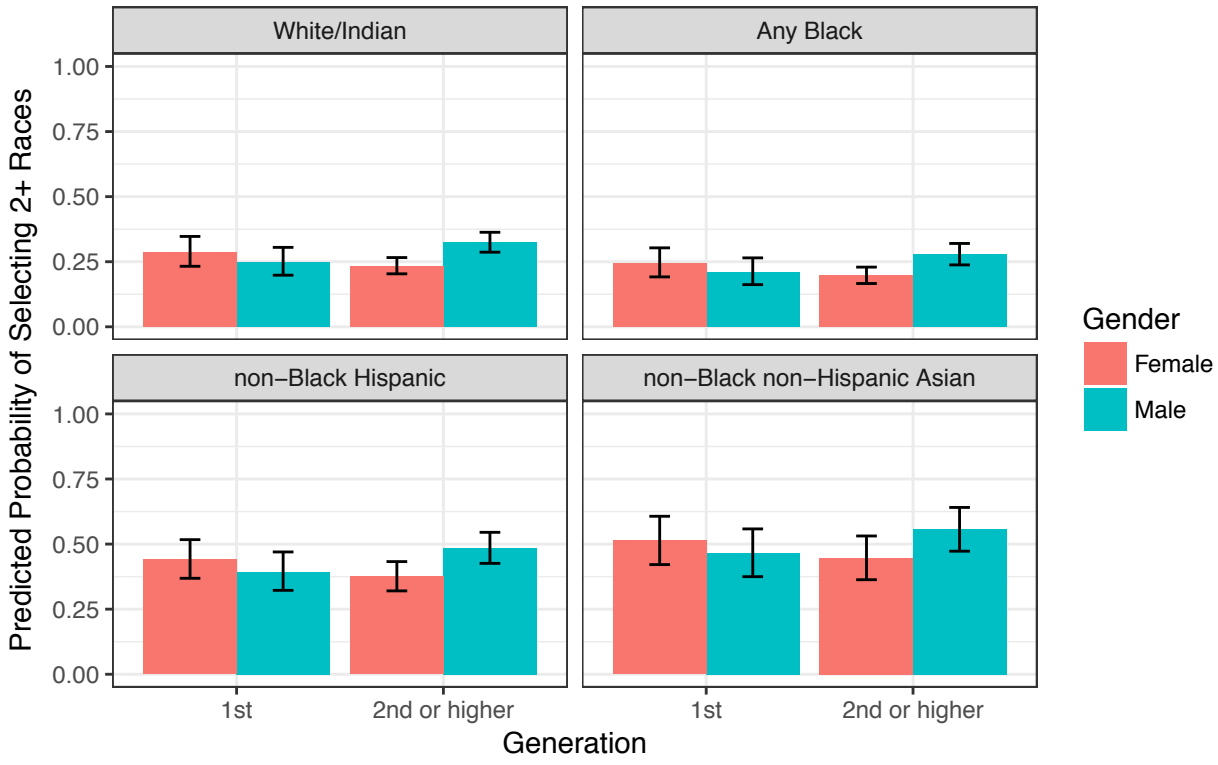
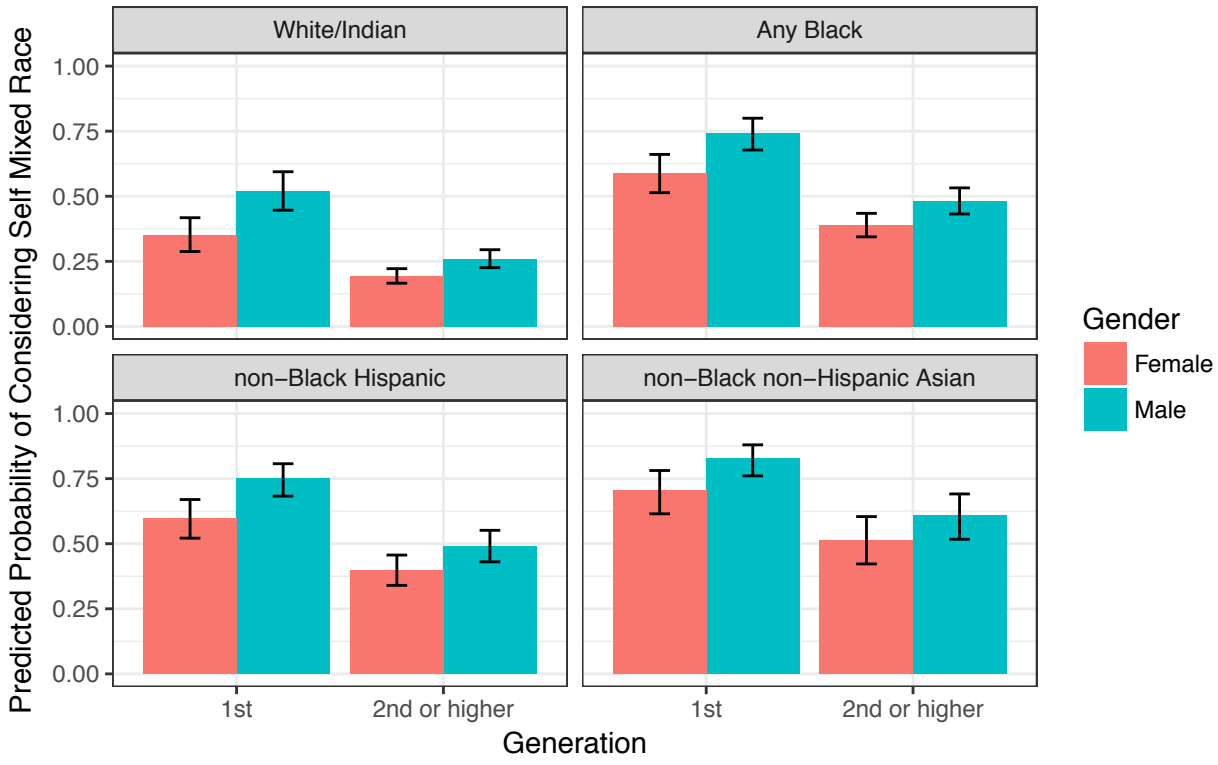


Figure 2: Probability of Considering Self Mixed Race



Predicted probabilities were generated from "Gender x Generation" models with varying generation, and racial ancestry regime, and gender values. Controls are held at median for numeric variables and mode for categorical variables, so these estimates are for a 52-year-old, non-rural, native-born, Southern, English-speaking married adult with no BA and no children at home.