Internet Appendix to "E Pluribus, Pauciores (Out of Many, Fewer): Diversity and Birth Rates"

David H. Solomon (Boston College)

Internet Appendix

This appendix reports the data details, robustness checks and supplementary results referred to in the paper. Section 1 describes the data sources, the construction of the woman-by-year birth panel and the international censuses. Section 2 reports alternative constructions of the main regressions. Section 3 uses a placebo simulation to show that the results do not arise mechanically from the correlation between race and RaceShare. Section 4 reports the first stage of the instrumental variables regressions and weak-instrument-robust confidence sets. Section 5 reports longer-horizon and differenced versions of the time-series regressions. Section 6 reports the joint timing regressions. Section 7 reports the international results on the birth panel. Section 8 describes the census-matched versions of the main tables and reports the children-at-survey versions (Tables IA10 to IA16). The tables follow, with notes. Unless stated otherwise, standard errors are double clustered by year and state, coefficients are in the top row and t-statistics are below in parentheses, with *, **, and *** indicating statistical significance at the 10%, 5% and 1% level respectively. Table numbers in the paper refer to this version of the appendix (Tables IA1 to IA16).

1. Data

1.1 U.S. data – main sample

U.S. census data sources are taken from the IPUMS default samples for each year, namely: the 1% sample from 1850, 1860, 1870, 1880, 1900, 1910, 1920, 1930, 1940 and 1950 (the 1890 and 1960 samples are missing because they lack geographic information); the 1% metro fm1 sample from 1970; the 1% metro sample from 1980 and 1990; the 1% sample from 2000; the 10% sample from 2010; and the ACS surveys from 2003 to 2021 (earlier ACS vintages lack the necessary geographic information).

City information is available in 1850, 1860, 1870, 1880, 1900, 1910, 1920, 1930, 1940, 1950, 1980, 1990, 2000 and 2005-2021. County information is available in 1950, 1970, 1980, 1990, 2000 and 2005-2021.

Detailed metro area information is available in 1850, 1860, 1870, 1880, 1900, 1910, 1920, 1930, 1940, 1950, 1970, 1980, 1990, 2000, 2003 and 2005-2011. Because the 1850 census collects city and detailed metro area information, the combined-geography RaceShare measure is available from the first sample year.

1.2 Construction of the birth panel

The main tables use a woman-by-year panel of annual birth outcomes constructed from the number and ages of each woman's co-resident children at the time of survey, following Nickerson and Solomon (2024). I exclude group quarters, households with inmates or children-in-law, and households with no adult woman present (households with only an adult male and children lack information on the mother's age). Children in a household are assigned to the woman who is the household head or the head's spouse; if there are multiple women over 18 in a household, I split the women into separate observations and assign children to the corresponding mother. A child's age identifies the year of her birth. Each woman contributes one observation for every year from age 18 to her age at the survey (or 40, if earlier), with the dependent variable equal to 100 in the years a child was born and zero otherwise. Twins count as one birth event. As with the children-at-survey measure, only children still present in the household at the time of the survey are observed, so births to older women whose children have left home are undercounted; Table IA2 Panel D examines this directly.

Demographic variables are observed only at the time of survey and are merged to the woman's birth observations in all prior years. Area-level measures (RaceShare, RaceHerfindahl and the area traits) are matched to the year before the birth outcome, approximating the year of conception. Two matching conventions are used. Under the all-years convention, which the main tables use, a woman-year whose conception year is not a census or ACS year takes the measurement from the nearest subsequent census or ACS year for her area, carried back to all intervening years (up to 25 years, and in practice ten or fewer for most of the sample because of the decennial censuses and the annual ACS). Under the census-matched convention, used in Table 4 and Table 14, only woman-years whose conception year is itself a

measurement year are retained. The two conventions give the same conclusions in every table; the census-matched coefficients are equal or somewhat larger, so the all-years estimates are the more conservative. In both conventions, all area-level controls in a specification are taken from the same conception-year vintage as RaceShare. As in the children-at-survey sample, the implicit assumption is that the woman has not moved between the birth year and the survey; if women move, this introduces measurement error and likely attenuates the estimates.

Relative to the version of the panel used in an earlier draft (which entered only Table 4 and Table 14), two corrections were made and one convention was tightened. First, the earlier construction attached children to the household head only and then merged the woman's demographics by the head's person number, so households whose head was the husband were dropped without error; that panel covered 6.0 million women, 26% of them married, against 11.7 million women and 49% married in the children-at-survey sample. The corrected construction carries the woman's own person number and recovers the married-couple households, and the panel's marriage rates now match the survey sample. Second, the earlier expansion ran back to age 11 rather than 18, so 38% of woman-years had ages below 18, with near-zero birth rates that attenuated the coefficients; the corrected panel enforces ages 18-40 in the birth year. Third, all area-level controls are now taken from the conception-year vintage, where the earlier version mixed conception-year RaceShare with survey-year area controls. Every table that uses the panel was re-estimated on the corrected version; the corrections raise the main coefficients somewhat and leave every conclusion unchanged.

International censuses additionally provide a mother pointer (IPUMS MOMLOC) that identifies the mother of each child directly rather than through the head/spouse relationship. This makes it possible to measure how many children the head/spouse construction loses: children whose mother is present in the household but is neither the head nor the spouse (for example, an adult daughter's children coded as the head's grandchildren). For children under 18 the share is 2-7% in the United States, 3-11% in Brazil, 6-12% in Colombia, Costa Rica and Ecuador, 14-18% in Jamaica, 15-23% in South Africa and 22% in

Cuba (Table IA9). Re-estimating the international results with those children assigned to their mothers (Table IA8) changes no coefficient by more than its standard error.

1.3 International data

Table IA1 reports, for each country used in Table 8, the census samples, the number of observations, the control variables available and the racial classification collected. I measure race using whatever definitions each country collects.

2. Alternative Specifications for the Main Result

In this section I explore variations on the main specification. All regressions in Table IA2 use the birth panel under the all-years convention, as in Table 2.

Table IA2 Panel A constructs RaceShare using alternative definitions of race: i) omitting the Hispanic/Latino category, ii) using detailed race instead of broad race and omitting Hispanic/Latino, iii) using broad race with Hispanic/Latino as an interaction rather than a separate category, iv) using detailed race with the Hispanic/Latino interaction, v) using ancestry share instead of any race classification, and vi) using the base definition over the whole population, including those under eighteen. Each column includes area-by-year fixed effects and demographics interacted with state and year, corresponding to the last column of Table 2 Panel A. The coefficients are 1.169, 1.213, 1.343, 1.496, 2.146 and 1.315, with t-statistics between 4.60 and 6.42, so the effect is positive and significant at the 1% level under every definition.

Panel B constructs RaceShare at different geographical levels. Because geographic information varies across census years, limiting the analysis to one type of geography alters the range of years included. Recall that in the base case the levels are used sequentially by availability, so that county is used only if city information is missing. In columns 1 to 3 I use i) city, ii) county and iii) detailed metro area for all observations in each category, and in column 4 I reverse the priority order of city and county (city first,

then county, then metro area). Each of these columns includes the share's own geography interacted with year, together with demographics interacted with state and year. All four relationships are positive and significant: 1.542 (t-statistic of 4.41) for city, 2.178 (6.14) for county, 1.059 (2.54) for metro area and 2.162 (5.45) for the reversed ordering. Column 5 uses the state-level share, which includes every observation in the state, including those with missing finer geography; here state-by-year fixed effects would absorb the variable, so the column includes demographics interacted with state and year only. The state-level measure is significant only at the 10% level (0.933, t-statistic of 1.71), which suggests that it is the more local geography that drives the effect. Consistent with this, column 6 includes both the baseline measure and the state-level measure in the same regression: the local measure remains highly significant (1.595, t-statistic of 4.69) while the state-level measure is negative and insignificant (-1.456, t-statistic of -1.37).

Panel C explores different weighting schemes. The baseline regressions weight every woman-year equally, which draws more observations both from larger areas and from recent years. Column 1 repeats the baseline; column 2 weights every area-by-year combination equally regardless of the number of respondents; column 3 weights every year equally and, within the year, applies the census household weights. Columns 4 to 6 apply census household weights within the area when constructing RaceShare, with the same three observation-level weighting choices. All six specifications include area-by-year fixed effects and demographics interacted with state and year. The coefficients range from 1.321 to 1.441 with t-statistics between 4.99 and 5.81, so the weighting choice does not matter for the result.

Panel D examines the role of women's ages. The first concern is children leaving home: because only children living in the household are counted, a woman who had a child at 18 and is 40 at the survey may have a child who has left, and that birth is missed. Age, age-by-year and age-by-state fixed effects control for the average level of this omission, but RaceShare could be correlated with it in other ways. Columns 1 to 3 limit the sample to women aged 35 and under, 30 and under, and 25 and under at the time of survey, where the chance of a missing child is small (35 suffices to exclude most cases, as a woman who gave

birth at 18 would have a child aged 17 at the survey). The coefficients are 1.093 (t-statistic of 4.39), 0.675 (3.66) and 0.536 (3.07), compared with 1.315 in the corresponding column of Table 2 Panel A. The effect is therefore similar, though somewhat smaller, in samples where missing children are not an issue. That the coefficient continues to decline from the 35-and-under to the 25-and-under sample, where leaving home is already not a concern, suggests that the decline reflects the mechanics of young samples rather than data omissions: at young ages few women have children, so RaceShare cannot reduce births by as much.

The second concern is the opposite one: that women observed young have incomplete birth histories, so that delayed childbearing could be confused with lower fertility. Columns 4 to 8 restrict the sample to women aged exactly 40, 39, 38, 37 and 36 at the time of survey, whose histories from age 18 are close to complete. The coefficients are 1.292, 1.317, 1.446, 1.626 and 1.564, with t-statistics between 4.04 and 4.65, so in the samples where childbearing is most nearly complete the effect is, if anything, larger than in the full sample.

3. Effect of Correlations Between Race and RaceShare

One possible concern is that the higher coefficient on RaceShare after controlling for race (Table 2 Panel A, columns 1 and 2) is due to the correlation between RaceShare and race itself. Because race is categorical, I report correlations of RaceShare with individual race dummies: 0.829 with white, -0.382 with black, -0.102 with American Indian, -0.198 with Chinese, -0.088 with Japanese, -0.318 with other Asian or Pacific Islander, -0.083 with other, -0.212 with two races, -0.076 with three or more races, and -0.378 with Hispanic.

It is possible that these correlations drive the high significance when both variables are included, for instance through multicollinearity. To test this, I create a placebo that preserves the distributional properties of the variables and their correlation with race, but severs the link between a woman's own outcome and her own RaceShare. Within each race category, I randomly reassign the number of children

across women, so that each woman receives the outcome of another woman of her race (drawn without replacement). If the correlation structure between race and RaceShare, or any other mechanical feature of the RaceShare distribution, were producing the result, the placebo would reproduce it, because it has exactly the same correlation structure; if the placebo shows no effect, the result comes from the specific pairing of RaceShare with each woman's own outcome. I run 1,000 simulations with a fixed random seed and perform the regressions of Table 2 Panel A columns 1 and 2 on the children-at-survey measure (the measure of the earlier draft; a birth-panel version of the placebo would be straightforward to add).

In the simulated versions of the univariate specification, all 1,000 draws show negative effects of RaceShare, with a mean coefficient of -0.0456, a smallest value of -0.0482 and a largest of -0.0425. This confirms that, absent a true economic effect of RaceShare, the mechanical effect of the correlation with race is negative (whites have a high RaceShare and a low birth rate), whereas the true coefficient of 0.156 is positive and significant, consistent with a positive true effect in addition to the negative mechanical one. More importantly, in the specification with race controls, none of the simulations comes close to the actual result: the mean simulated coefficient is -0.00007, the smallest is -0.0094 and the largest is 0.0107, compared with a true value of 0.707. On average the correlation of race with the placebo induces no effect once race is controlled for, and the range of placebo coefficients is tiny next to the true value, which is roughly seventy times the largest of 1,000 draws. The correlation structure between race and RaceShare is therefore not an important driver of the results in Table 2, nor are other mechanical aspects of the distribution of RaceShare.

4. Instrumental Variables: First Stage and Weak-Instrument-Robust Inference

Table IA3 reports the first stage of the instrumental variables regressions of Table 3. I regress county-level RaceShare in the conception year on the immigration shock, keeping the fixed effects and controls of the corresponding column of Table 3, on the birth panel. The immigration shock is negative and significant in every column but the last: the coefficients run from -16.0 (t-statistic of -4.63) with no controls to -1.6 (t-statistic of -1.72) with area fixed effects. Because there is a single instrument, the first-

stage F-statistic is the square of the cluster-robust t-statistic: 21.4, 40.3, 7.9, 7.2, 8.3, 6.9 and 3.0 across the seven columns. The first stage is therefore strong in the first two columns and weak by conventional standards once state-by-year effects and demographics are included.

Because the first stage is weak in the controlled columns, Table IA3 also reports Anderson-Rubin 95% confidence sets for the RaceShare coefficient, which remain valid however weak the instrument. The set for a hypothesized effect b inverts the test that the instrument has no effect on the birth outcome net of b times RaceShare (a regression of the birth outcome minus b times RaceShare on the instrument and the column's controls); with one instrument the set is an interval when the first-stage F exceeds the squared critical value, and otherwise the complement of an interval, written "outside (a, b)", which excludes zero whenever zero lies between a and b. Because the immigration shocks are matched to five conception years, the year dimension has five clusters and the critical value is that of a t distribution with four degrees of freedom (2.78), the convention used for the t-statistics in Table 3 as well. A zero effect lies outside the Anderson-Rubin set in columns 1 to 6 ([1.05, 9.43], [7.87, 22.27], [14.82, 1668.8], outside (-491.6, 8.17), [10.20, 347.9] and outside (-446.3, 9.06)) and inside it in column 7 (outside (5.08, 21.65)). The pattern of significance in Table 3 therefore does not depend on the weak-instrument-sensitive 2SLS standard errors; the price of the weak first stage in columns 3 to 6 is that the sets are wide or unbounded on one side, so the magnitudes in those columns are not well determined.

5. Time Series: Longer Horizons and Differenced Regressions

The time-series regressions in Table 13 relate the U.S. total fertility rate since 1971 to the average of RaceShare in the prior year. Two supplementary sets of results are reported here.

5.1 Longer horizon time series regressions

While the TFR series begins in 1971, a rough approximation of birth rates can be constructed from the census back to 1850, using the average number of children of all female respondents aged 18-40. This is sensitive to the age profile of the women sampled, but it has the advantage of being available over a far

longer period. Table IA4 reports these regressions. In columns 1 to 4, all geographic measures of average RaceShare work as univariate predictors since 1850, with t-statistics above 6. R-squared values range from 0.571 to 0.796, and the predicted change as a fraction of the actual change is 85.8%, 70.4% and 83.6% for the combined, city and state measures respectively (county is an outlier at 3892%, partly because of the small number of observations). Columns 5 to 8 repeat the analysis using only decennial observations, to mitigate the influence of the numerous annual observations available since 2000; the results are consistent with those in the first four columns.

5.2 Differenced regressions

Both TFR and average RaceShare are strongly trending series, and level-on-level regressions of trending series can produce large R-squared values even when the underlying relationship is spurious (Granger and Newbold 1974). Table IA5 therefore presents first-differenced and ten-year-differenced versions of the Table 13 regressions. The timing follows Table 13: the dependent variable is the change in TFR from year t-1 to year t, and the independent variable is the change in average RaceShare from year t-2 to year t-1, so that the fertility change in year t is related to the change in diversity that had occurred by the time of the conception decisions. The macroeconomic controls are the corresponding lagged changes in inflation, unemployment and GDP growth. Columns 1 and 2 use OLS on the fifteen years with consecutive annual data (2007 onwards), without and with the macroeconomic controls; columns 3 and 4 use Newey-West standard errors with two lags on the same sample; column 5 uses the state-level measure of RaceShare; and column 6 regresses the ten-year change in TFR on the ten-year change in average RaceShare, lagged one year, on the eleven years where both are available.

Once the common trends are removed, the relationship between changes in RaceShare and changes in TFR is statistically indistinguishable from zero and negative in sign. The coefficient on the lagged change in RaceShare is -0.966 (t-statistic of -1.10) without controls and -1.205 (-1.14) with them, the Newey-West t-statistics are -0.99 and -1.09, the state-level version is -0.913 (-0.92), and the ten-year-difference coefficient is -0.202 (-0.14). With only fifteen annual observations, and with year-to-year changes in

average RaceShare largely reflecting sampling variation across ACS vintages, the power of these tests is limited and attenuation from measurement error is likely severe. The differenced results reinforce the interpretation adopted in the paper: the level regressions document low-frequency co-movement between the two aggregates, and the panel-based estimates of Section 4.13 are the appropriate basis for magnitude claims.

6. Joint Timing of Diversity Measures and Birth Outcomes

Table 4 shows that births respond to RaceShare measured in the conception year, with smaller or no response to RaceShare measured ten years before or ten years after. Table IA6 augments Table 4 by including the three windows jointly, in the subsample of woman-years where all three measures can be matched (which requires three separate census matches, leaving 2.5 million woman-years). Included separately, each window is positive and significant: 1.695 (t-statistic of 5.80) for the conception year, 1.582 (5.83) for ten years before and 1.835 (5.05) for ten years after, as the three measures are highly correlated. Included in pairs or all together, the measures are too collinear to separate: conception-year RaceShare retains its magnitude against the ten-years-before measure (1.563 against 0.137) but not against the ten-years-after measure, and no individual window is statistically distinguishable in column 6. The joint regressions therefore cannot identify the timing more finely than Table 4 does.

7. International Results on the Birth Panel

Table 8 in the main text uses the number of children at the time of survey for all twelve countries. This section reports the birth-panel version, which is feasible only where a race-bearing census precedes the births in question. Because international censuses are eight to sixteen years apart and a census's own race share is dated to the following conception year, every birth after a country's last prior race-bearing census is unmatched under the all-years convention. Countries whose race data begin with their most recent census (the United Kingdom, Zimbabwe, El Salvador and Colombia) therefore have no estimable woman-years, and Uruguay lacks fine geography in its earlier race-bearing census; Uruguay can be estimated with

the coarse-geography share, as Jamaica is in Table 8, in the first specification only (2.879, t-statistic of 1.02), because with a single prior census the share is constant within race and region and is absorbed by the fixed effects of the other two.

Table IA7 reports the three specifications of Table 8 for the seven countries where the panel is feasible. Mozambique, South Africa and Ecuador are positive and significant in at least one specification, Costa Rica is negative and significant throughout, and Brazil, Cuba and Jamaica are insignificant. Wherever both versions are available they agree in sign with Table 8, with the panel version less precise. Table IA8 repeats the estimation with children whose mother is present but is neither the head nor the spouse assigned to their mother using the MOMLOC pointer; no coefficient changes by more than its standard error. Table IA9 reports the share of children recovered and lost by the head/spouse construction in each country's censuses.

8. Census-Matched and Children-at-Survey Versions of the Main Tables

Every table estimated on the birth panel was also estimated under the census-matched convention, keeping only woman-years whose conception year is a census or ACS year. The census-matched versions agree with the all-years versions in sign and significance in every specification with controls, with coefficients typically 10-45% larger; Table 14 reports the main ladder under this convention, and the remaining census-matched tables are available on request. Two exceptions are noted in the text: the census-matched immigration IV is degenerate because the five-year shocks intersect the matched years too thinly, and the census-matched historical-period columns of Table 6 are too thin to estimate before 1900.

Tables IA10 to IA16 report the main results on the original children-at-survey measure, which was the dependent variable of the previous draft: the number of children living with the woman at the time of survey, for women aged 18-40 in the same census and ACS samples. These are the children-at-survey versions of Tables 2, 3, 5, 6, 7, 11 and 12, with the same specifications, controls and clustering as the

corresponding birth-panel tables. The conclusions are the same in every respect. In Table IA10 Panel A, the coefficient on RaceShare is 0.156 (t-statistic of 2.78) with no controls, 0.707 (6.82) with race fixed effects, and between 0.152 and 0.432 across the remaining specifications, ending at 0.190 (5.77) with area-by-year fixed effects; a one-standard-deviation increase in RaceShare raises the number of children by 0.062 unconditionally, or 0.020 using the conditional standard deviation, in the last column. The race Herfindahl index (Panels B to D) behaves as in Table 2. Instrumenting county race share with immigration shocks (Table IA11) gives positive and significant coefficients in every column but the last, as in Table 3; the mobility subsamples (Table IA12), the historical periods (Table IA13), the race interactions (Table IA14), the trust measures (Table IA15) and the other-dimension shares (Table IA16) match the patterns of Tables 5, 6, 7, 11 and 12. Table IA15 Panel B uses the corrected county key described in Section 4.10 of the paper. Because the children-at-survey measure counts children accumulated over the woman's life rather than births in a given year, the coefficients are in different units from the panel tables and are not directly comparable in magnitude; the effect-size rows convert each to children per standard deviation of RaceShare.

Tables

Table IA1 – Summary Statistics and Control Availability for International Data

This table reports, for each country in Table 8, the census samples used, the number of observations of women aged 18-40, the control variables available in each sample, and the racial, ethnic or language classification used to construct RaceShare. Samples are IPUMS-International extracts.

Country	# Obs.	Mean # Children	# Samples	# Course Geography Units	# Fine Geography Units	Std Race Share	Control Variables	Races	Number	Pct
Brazil	22,877,029	1.646	6	25	2,040	0.242129	Age, Education, Employment, Marital Status, Race, Urban	White	10,232,595	54.92
								Black	1,202,607	6.46
								Indigenous	40,650	0.22
								Asian	126,174	0.68
								Brown	7,028,211	37.72
Colombia	2,215,042	1.542	4	22	438	0.252859	Age, Education, Employment, Marital Status, Race, Urban	White	561,681	82.77
								Black	74,113	10.92
								Indigenous	41,884	6.17
								Other	889	0.13
Costa Rica	231,878	1.499	4	7	55	0.226887	Age, Education, Employment, Marital Status, Race, Urban	White	138,357	93
								Black	2,232	1.5
								Indigenous	1,108	0.74
								Asian	156	0.1
								Chinese	154	0.1
								Mulatto	6,015	4.04
								Other	755	0.51
Cuba	383,755	0.970	2	14	137	0.262611	Age, Education, Employment, Marital Status, Race	White	242,150	63.1
								Black	35,451	9.24
								Mixed Race	106,154	27.66
Ecuador	909,119	1.601	5	14	79	0.291202	Age, Education, Employment, Marital Status, Race, Urban	White	39,427	8
								Black	7,513	1.52
								Afro-Ecuadorian	12,233	2.48
								Indigenous	30,938	6.28
								Mestizo	370,966	75.29
								Mulatto	11,909	2.42
								Other	1,753	0.36
								Montubio	17,969	3.65
El Salvador	201,637	1.499	2	14	103	0.261839	Age, Education, Employment, Marital Status, Race, Urban	White	14,437	13.32
								Black	117	0.11
								Indigenous	261	0.24
								Mestizo	92,930	85.75
								Other	623	0.57
Jamaica	121,582	1.404	3	14	N/A	0.2558	Age, Education, Employment, Marital Status, Race, Urban	White	229	0.2
								Black	103,095	88.49
								Chinese	213	0.18
								Indian	1,559	1.34
								Other Asian	13	0.01
								Mixed Race	11,304	9.7
								Other	98	0.08

Country	# Obs.	Mean # Children	# Samples	# Course Geography Units	# Fine Geography Units	Std Race Share	Control Variables	Races	Number	Pct
Mozambique	651,821	1.932	2	11	143	0.080113	Age, Education, Employment, Marital Status, Race, Urban	White	457	0.07
								Black	633,924	99.35
								Indian	489	0.08
								Pakistani	57	0.01
								Mixed Race	3,037	0.48
								Other	135	0.02
South Africa	3,141,423	1.042	5	5	19	0.270324	Age, Education, Employment, Marital Status, Race, Urban	White	155,624	6.39
								Black african	2,014,603	82.71
								Asian	53,757	2.21
								Coloured	208,378	8.56
								Other	3,288	0.13
United Kingdom	92,397	1.020	1	11	N/A	0.236817	Age, Employment, Marital Status, Race	White	86,347	93.45
								Black African	499	0.54
								Black Caribbean	1,084	1.17
								Other Black	350	0.38
								Chinese	388	0.42
								Indian	1,673	1.81
								Pakistani	820	0.89
								Bangladeshi	197	0.21
								Other Asian	504	0.55
								Other	535	0.58
Uruguay	282,446	1.229	6	19	67	0.262951	Age, Education, Employment, Marital Status, Race, Urban	White	80,648	88.89
								Black	3,433	3.78
								Indigenous	1,529	1.69
								Asian	183	0.2
								Mestizo	917	1.01
								Two or More Races	3,940	4.34
								Other	82	0.09
Zimbabwe	123,039	1.400	1	10	88	0.057442	Age, Education, Employment, Marital Status, Race, Urban	White	145	0.12
								Black	122,537	99.67
								Asian	84	0.07
								Mixed Race	170	0.14
								Other	8	0.01

Table IA2 – Alternative Constructions of the Diversity Variable

This table presents alternative versions of the main regressions in Table 2, on the woman-by-year birth panel under the all-years convention. The dependent variable is 100 times a dummy for giving birth in the year, for women aged 18-40. Panel A considers alternative definitions of race: i) omitting the Hispanic/Latino category, ii) detailed race, omitting the Hispanic/Latino category, iii) broad race interacted with Hispanic/Latino, iv) detailed race interacted with Hispanic/Latino, v) ancestry instead of race and ethnicity, and vi) the baseline measure (broad race plus a Hispanic category) for the whole population, including children under 18; all columns include area-by-year fixed effects and demographics interacted with state and year. Panel B varies the geography at which race share is measured: i) city only, ii) county only, iii) metro area only, iv) city, then county, then metro area, and v) state; column 6 includes the baseline measure (county, then city, then metro area) and the state measure together. Columns 1 to 4 include the share's geography interacted with year; all columns include demographics interacted with state and year. Panel C explores weighting: i) the unweighted baseline, ii) each area-by-year cell weighted equally, iii) each year weighted equally with census household weights within the year, and iv) to vi) the same three choices with census household weights used in constructing race share; all columns include area-by-year fixed effects and demographics interacted with state and year. Panel D restricts the sample by the woman's age at the time of survey: 35 and under, 30 and under, 25 and under, and each age from 40 to 36 separately. Standard errors are double clustered by year and state. Coefficients are in the top row, and t-statistics are below in parentheses, with *, **, and *** indicating statistical significance at the 10%, 5% and 1% level respectively.

Panel A - Different Race Definitions						
	Dependent variable is a dummy for a birth in that year (x100)					
Race Share (No Hispanic Category)	1.169***					
	(5.74)					
Race Share (Detailed Race, No Hispanic Category)		1.213***				
		(6.42)				
Race Share (Hispanic Interaction)			1.343***			
			(4.86)			
Race Share (Detailed Race, Hispanic Interaction)				1.496***		
				(5.53)		
Ancestry Share					2.146***	
					(4.60)	
Race Share (Including Under 18 Population)						1.315***
						(4.94)
Demographics*(State, Year) FE	Y	Y	Y	Y	Y	Y
Area * Year FE	Y	Y	Y	Y	Y	Y
Observations	56,745,423	56,252,690	56,684,159	55,965,727	47,623,220	56,747,886
R-squared	0.039	0.040	0.040	0.040	0.040	0.039

Panel B - Different Region Measures						
	Dependent variable is a dummy for a birth in that year (x100)					
Race Share (City)	1.542***					
	(4.41)					
Race Share (County)		2.178***				
		(6.14)				
Race Share (Metro Area)			1.059**			
			(2.54)			
Race Share (City, then County, then Metro)				2.162***		
				(5.45)		
Race Share (State)					0.933*	-1.456
					(1.71)	(-1.37)
Race Share (Baseline - County, then City, then Metro)						1.595***
						(4.69)
State-Year FE	Y	Y	Y	Y	N	N
Demographics*(State, Year) FE	Y	Y	Y	Y	Y	Y
Observations	13,893,798	49,594,291	27,844,896	57,063,498	94,711,660	13,893,798
R-squared	0.042	0.039	0.038	0.038	0.035	0.041

Panel C - Weighting						
	Dependent variable is a dummy for a birth in that year (x100)					
Race Share (Baseline - Unweighted)	1.441***	1.321***	1.385***			
	(5.53)	(4.99)	(5.29)			
Race Share (HH Weighted)				1.416***	1.334***	1.395***
				(5.81)	(5.37)	(5.70)
Sample Weighting	Area*Year (Indiv)	Year (Indiv)	Year (HH Weight)	Area*Year (Indiv)	Year (Indiv)	Year (HH Weight)
Demographics*(State, Year) FE	Y	Y	Y	Y	Y	Y
Area * Year FE	Y	Y	Y	Y	Y	Y
Observations	56,747,886	56,747,886	56,746,976	56,747,886	56,747,886	56,746,976
R-squared	0.037	0.039	0.039	0.037	0.039	0.039

Panel D - Female Age								
Female Age	<=35	<=30	<=25	40	39	38	37	36
	Dependent variable is a dummy for a birth in that year (x100)							
Race Share	1.093***	0.675***	0.536***	1.292***	1.317***	1.446***	1.626***	1.564***
	(4.39)	(3.66)	(3.07)	(4.53)	(4.04)	(4.13)	(4.65)	(4.51)
Demographics*(State, Year) FE	Y	Y	Y	Y	Y	Y	Y	Y
Area * Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Earliest Year	1980	1980	1980	1980	1980	1980	1980	1980
Observations	33,663,150	17,069,349	6,342,025	5,335,391	4,801,571	4,577,081	4,291,893	4,078,697
R-squared	0.051	0.065	0.078	0.040	0.040	0.040	0.040	0.041

Table IA3 – First Stage Regressions and Anderson-Rubin Confidence Sets for the Instrumental Variables Regressions

This table presents the first stage regressions for the instrumental variables regressions of Table 3, on the woman-by-year birth panel. The dependent variable is county-level race share in the conception year. The independent variable is the immigration shock, taken from Burchardi et al. (2024), based on predicted ancestry in a given county in 1975 and national immigration from that ancestry group over the following five-year period, scaled by county respondents. Each column includes the controls of the corresponding column of Table 3, which are defined in Table 2. The first-stage F-statistic is the square of the cluster-robust t-statistic on the single instrument. The Anderson-Rubin 95% confidence set for the race share coefficient in the corresponding column of Table 3 inverts the test that the instrument has no effect on the birth outcome net of a hypothesized race share effect b; with one instrument the set is an interval when the first-stage F exceeds the squared critical value, and otherwise the complement of an interval, written "outside (a, b)", which excludes zero whenever zero lies between a and b. Standard errors are clustered by year and state; because the immigration shocks are matched to five conception years, the year dimension has five clusters and the critical value is that of a t distribution with four degrees of freedom. Coefficients are in the top row, and t-statistics are below in parentheses, with *, **, and *** indicating statistical significance at the 10%, 5% and 1% level respectively.

	Dependent variable is RaceShare						
Immigration Shocks	-16.025***	-6.397***	-4.722**	-4.220*	-3.974**	-1.845*	-1.595
	(-4.63)	(-6.35)	(-2.81)	(-2.69)	(-2.83)	(-2.63)	(-1.72)
Race	N	Y	Y	N	N	N	N
State-Year FE	N	N	Y	Y	Y	Y	Y
Demographics FE	N	N	N	Y	N	N	N
Demographics*(State, Year) FE	N	N	N	N	Y	Y	Y
Area Type	N	N	N	N	Y	N	N
Area Traits	N	N	N	N	Y	N	N
Area Traits * Year FE	N	N	N	N	N	Y	Y
Area Type * Population FE *Year	N	N	N	N	N	Y	Y
Area FE	N	N	N	N	N	N	Y
Earliest Year	1850	1850	1850	1980	1980	1980	1980
First-stage F (t^2 on instrument)	21.4	40.33	7.91	7.25	8.28	6.89	2.97
Anderson-Rubin 95% confidence set (from Table 3)	[1.049, 9.434]	[7.867, 22.274]	[14.816, 1668.774]	(-inf, -491.644] or [8.173, +inf)	[10.201, 347.938]	(-inf, -446.324] or [9.055, +inf)	(-inf, 5.081] or [21.649,
Observations	5,760,379	5,760,379	5,760,379	5,760,379	5,732,523	5,732,523	5,732,523
R-squared	0.058	0.745	0.772	0.776	0.876	0.887	0.914

Table IA4 – Long Horizon Time Series Effects of Diversity on Fertility

This table examines how time-series changes in the average local level of diversity across the US are associated with changes in US birth rates over the long run. The dependent variable is the unadjusted average number of children of all women aged 18-40 in the survey year, from 1850 onwards. The independent variable is the average across all respondents of race share, measured using combined geography (county, then city if county is unavailable, then detailed metro area if both are unavailable), city only, county only, or state, in the year before the survey. Columns 1 to 4 use all available survey years; columns 5 to 8 use only decennial census observations. "Full Sample Change" is the change in the dependent variable over the period in question, "Predicted Change" is the regression coefficient multiplied by the change in the independent variable from the first sample year to the last, and "Fraction of Change Explained" is the ratio of the two. Coefficients are in the top row, and t-statistics are below in parentheses, with *, **, and *** indicating statistical significance at the 10%, 5% and 1% level respectively.

	Dependent variable is Number of children							
Race Share (Base Combined)	0.903***				0.706**			
	(6.96)				(2.56)			
Race Share (City Only)		0.672***				0.527**		
		(6.68)				(2.87)		
Race Share (County Only)			2.020***				1.764**	
			(8.77)				(3.58)	
Race Share (State)				0.839***				0.656**
				(6.76)				(2.54)
Full Sample Change	0.496	0.496	0.018	0.496	0.517	0.517	0.039	0.517
Predicted Change	0.451	0.373	0.692	0.440	0.344	0.292	0.582	0.334
Fraction of Change Explained	0.909	0.751	38.897	0.887	0.665	0.564	15.027	0.646
First Year	1850	1850	1950	1850	1850	1850	1850	1850
Years Included	All	All	All	All	Decades	Decades	Decades	Decades
Method	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Observations	32	30	22	32	16	15	7	16
R-squared	0.617	0.615	0.794	0.604	0.318	0.388	0.719	0.315

Table IA5 – Time Series Changes in Diversity and Fertility: Differenced Regressions

This table presents differenced versions of the time series regressions in Table 13, to remove the influence of common trends in the two series. The dependent variable is the annual change in the U.S. total fertility rate from year t-1 to year t, from the St Louis Fed FRED database. The main independent variable is the change in the average race share across all respondents from year t-2 to year t-1, so that the change in diversity is lagged one year relative to the change in fertility, as the level of race share is in Table 13; the average is measured using combined geography as in Table 13, or state-level geography in column 5. Macroeconomic controls are the corresponding lagged changes in inflation, unemployment, and GDP growth. Columns 1 and 2 use OLS on all years where consecutive annual data is available (in practice, 2007 onwards); columns 3 and 4 use Newey-West standard errors with two lags on the same sample. Column 6 regresses the ten-year change in TFR on the ten-year change in average race share, lagged one year, using all years where both are available. Coefficients are in the top row, and t-statistics are below in parentheses, with *, **, and *** indicating statistical significance at the 10%, 5% and 1% level respectively.

Dependent variable	Annual change in US TFR					Ten-year change in US TFR
Change in avg Race Share, combined geography (lagged)	-0.966	-1.205	-0.966	-1.205		
	(-1.10)	(-1.14)	(-0.99)	(-1.09)		
Change in avg Race Share, state (lagged)					-0.913	
					(-0.92)	
10-year change in avg Race Share (lagged)						-0.202
						(-0.14)
Change in inflation (lagged)		0.001		0.001	0.000	
		(0.11)		(0.16)	(0.02)	
Change in unemployment (lagged)		-0.004		-0.004	-0.004	
		(-0.59)		(-0.57)	(-0.59)	
Change in GDP growth (lagged)		-0.249		-0.249	-0.316	
		(-0.62)		(-0.63)	(-0.74)	
Race share geography	Combined	Combined	Combined	Combined	State	Combined
Macroeconomic controls (inflation, unemployment, GDP growth)	N	Y	N	Y	Y	N
Newey-West standard errors (two lags)	N	N	Y	Y	N	N
Observations	15	15	15	15	15	11
R-squared	0.085	0.146			0.109	0.002

Table IA6 – Joint Timing of Diversity Measures and Birth Outcomes

This table augments Table 4 by including race share measured ten years before conception, in the conception year, and ten years after conception jointly. The sample is woman-by-year birth observations under the census-matched convention of Table 14, restricted to woman-years where all three measures can be matched. Columns 1 to 3 include each measure separately; columns 4 and 5 include pairs; column 6 includes all three. All specifications include area-by-year fixed effects and demographics interacted with state and year. Standard errors are double clustered by year and state. Coefficients are in the top row, and t-statistics are below in parentheses, with *, **, and *** indicating statistical significance at the 10%, 5% and 1% level respectively.

	Dependent variable is a dummy for a birth in that year (x100)					
Race Share (Conception Year)	1.695**			1.563*	-0.106	-0.382
	(5.80)			(2.68)	(-0.10)	(-0.27)
Race Share (Ten Years Before)		1.582**		0.137		0.267
		(5.83)		(0.26)		(0.49)
Race Share (Ten Years After)			1.835**		1.937	1.957
			(5.05)		(1.54)	(1.55)
Demographics*(State, Year) FE	Y	Y	Y	Y	Y	Y
Area * Year FE	Y	Y	Y	Y	Y	Y
Sample	Census-matched years	Census-matched years	Census-matched years	Census-matched years	Census-matched years	Census-matched years
Observations	2,520,360	2,520,360	2,520,360	2,520,360	2,520,360	2,520,360
R-squared	0.040	0.040	0.040	0.040	0.040	0.040

Table IA7 – International Results on the Birth Panel

This table repeats the specifications of Table 8 on the woman-by-year birth panel described in Section 1.2, for the countries where it is feasible. The dependent variable equals 100 if the woman gave birth in that year and zero otherwise; the independent variable is the race share of her local area in the year before the birth, taken from the nearest subsequent race-bearing census of that country (the all-years convention of Table 2). Women are aged 18-40 in the birth year. Children are assigned to the woman who is the household head or the head's spouse, as in the US tables. Specifications, fixed effects and clustering are as in Table 8, with the survey-sample interactions replaced by birth-year interactions. Countries whose race data begin with their most recent census (the United Kingdom, Zimbabwe, El Salvador and Colombia) have no woman-years matched to a prior race-bearing census and are omitted; Uruguay uses the coarse-geography share, which is absorbed by the fixed effects in the second and third specifications. Coefficients are in the top row, and t-statistics are below in parentheses, with *, **, and *** indicating statistical significance at the 10%, 5% and 1% level respectively.

Panel A - Africa						
Country	Mozambique			South Africa		
	Dependent variable is a dummy for a birth in that year (x100)					
Race Share (fine geography)	7.566**	4.812*	-0.229	0.475	1.345	1.219**
	(2.03)	(1.91)	(-0.08)	(1.05)	(1.63)	(2.60)
Year FE	Y	Y	Y	Y	Y	Y
State-Year FE	Y	Y	N	Y	Y	N
Demographics FE	Y	N	N	Y	N	N
Demographics*Year FE	N	Y	Y	N	Y	Y
Demographics*State FE	N	Y	Y	N	Y	Y
Ln Population Density	Y	N	N	Y	N	N
Ln Population Density*Year	N	Y	Y	N	Y	Y
Local Region*Year FE	N	N	Y	N	N	Y
Race Share Level	Local	Local	Local	Local	Local	Local
Clustering	Local	Local	Local	Local	Local	Local
Observations	1,048,591	1,048,586	1,048,586	5,750,637	5,750,637	5,750,637
R-squared	0.023	0.026	0.029	0.022	0.024	0.024

Panel B - Central America and Caribbean									
Country	Costa Rica			Cuba			Jamaica		
	Dependent variable is a dummy for a birth in that year (x100)								
Race Share	-2.656***	-2.677**	-2.861***	0.124	0.161	0.748	2.842	2.160	-3.921
	(-2.95)	(-2.58)	(-5.70)	(0.56)	(0.29)	(1.27)	(0.75)	(0.53)	(-0.58)
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
State-Year FE	Y	Y	N	Y	Y	N	N	N	Y
Demographics FE	Y	N	N	Y	N	N	Y	N	N
Demographics*Year FE	N	Y	Y	N	Y	Y	N	Y	Y
Demographics*State FE	N	Y	Y	N	Y	Y	N	N	N
Ln Population Density	Y	N	N	Y	N	N	Y	N	N
Ln Population Density*Year	N	Y	Y	N	Y	Y	N	Y	Y
Local Region*Year FE	N	N	Y	N	N	Y	N	N	N
Race share geography	Fine	Fine	Fine	Fine	Fine	Fine	Coarse	Coarse	Coarse
Clustering	Local	Local	Local	Local	Local	Local	State	State	State
Observations	206,702	206,702	206,702	620,428	620,428	620,428	125,426	125,424	125,424
R-squared	0.027	0.030	0.033	0.009	0.010	0.013	0.013	0.016	0.017

Panel C - South America									
Country	Brazil			Ecuador			Uruguay		
	Dependent variable is a dummy for a birth in that year (x100)								
Race Share	-0.091	-0.080	0.070	2.684**	2.121	-0.624	2.879	absorbed	absorbed
	(-0.98)	(-0.69)	(0.87)	(2.22)	(1.66)	(-0.79)	(1.02)		
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
State-Year FE	Y	Y	N	Y	Y	N	Y	Y	N
Demographics FE	Y	N	N	Y	N	N	Y	N	N
Demographics*Year FE	N	Y	Y	N	Y	Y	N	Y	Y
Demographics*State FE	N	Y	Y	N	Y	Y	N	Y	Y
Ln Population Density	Y	N	N	Y	N	N	Y	N	N
Ln Population Density*Year	N	Y	Y	N	Y	Y	N	Y	Y
Local Region*Year FE	N	N	Y	N	N	Y	N	N	Y
Race share geography	Fine	Fine	Fine	Fine	Fine	Fine	Coarse	Coarse	Coarse
Clustering	Local	Local	Local	Local	Local	Local	Local	Local	Local
Observations	38,918,369	38,918,369	38,918,369	896,617	896,617	896,617	364,328	364,326	364,326
R-squared	0.036	0.041	0.044	0.031	0.033	0.035	0.030	0.036	0.039

Table IA8 – International Results Assigning Children to Mothers Using the MOMLOC Pointer

As Table IA7, but children whose mother is present in the household without being the head or the head's spouse are assigned to their mother using the IPUMS-International MOMLOC pointer, forming a separate mother-and-children unit; the head/spouse construction of Table IA7 and of the US tables loses these children. Samples without a pointer (Colombia 1964, Costa Rica 1963, Ecuador 1962, Suriname 2004 and the United Kingdom 2001) use the head/spouse construction. Coefficients are in the top row, and t-statistics are below in parentheses, with *, **, and *** indicating statistical significance at the 10%, 5% and 1% level respectively.

Panel A - Africa						
Country	Mozambique			South Africa		
	Dependent variable is a dummy for a birth in that year (x100)					
Race Share (fine geography)	7.554**	5.241**	0.227	0.692*	0.691	1.108***
	(2.04)	(2.08)	(0.09)	(1.88)	(0.81)	(2.90)
Year FE	Y	Y	Y	Y	Y	Y
State-Year FE	Y	Y	N	Y	Y	N
Demographics FE	Y	N	N	Y	N	N
Demographics*Year FE	N	Y	Y	N	Y	Y
Demographics*State FE	N	Y	Y	N	Y	Y
Ln Population Density	Y	N	N	Y	N	N
Ln Population Density*Year	N	Y	Y	N	Y	Y
Local Region*Year FE	N	N	Y	N	N	Y
Race Share Level	Local	Local	Local	Local	Local	Local
Clustering	Local	Local	Local	Local	Local	Local
Observations	1,126,781	1,126,776	1,126,776	7,138,965	7,138,965	7,138,965
R-squared	0.024	0.026	0.029	0.013	0.014	0.015

Panel B - Central America and Caribbean									
Country	Costa Rica			Cuba			Jamaica		
	Dependent variable is a dummy for a birth in that year (x100)								
Race Share	-2.613***	-2.643***	-1.851***	0.058	-0.066	0.757	2.535	0.382	-4.621
	(-3.25)	(-2.89)	(-4.99)	(0.33)	(-0.13)	(1.47)	(0.85)	(0.11)	(-0.96)
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
State-Year FE	Y	Y	N	Y	Y	N	N	N	Y
Demographics FE	Y	N	N	Y	N	N	Y	N	N
Demographics*Year FE	N	Y	Y	N	Y	Y	N	Y	Y
Demographics*State FE	N	Y	Y	N	Y	Y	N	N	N
Ln Population Density	Y	N	N	Y	N	N	Y	N	N
Ln Population Density*Year	N	Y	Y	N	Y	Y	N	Y	Y
Local Region*Year FE	N	N	Y	N	N	Y	N	N	N
Race share geography	Fine	Fine	Fine	Fine	Fine	Fine	Coarse	Coarse	Coarse
Clustering	Local	Local	Local	Local	Local	Local	State	State	State
Observations	225,297	225,297	225,297	764,982	764,982	764,982	145,501	145,498	145,498
R-squared	0.025	0.028	0.031	0.009	0.010	0.012	0.014	0.016	0.018

Panel C - South America									
Country	Brazil			Ecuador			Uruguay		
	Dependent variable is a dummy for a birth in that year (x100)								
Race Share	-0.100	-0.085	0.063	2.459**	1.921	-0.744	2.735	absorbed	absorbed
	(-1.12)	(-0.75)	(0.81)	(2.13)	(1.56)	(-1.02)	(1.07)		
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
State-Year FE	Y	Y	N	Y	Y	N	Y	Y	N
Demographics FE	Y	N	N	Y	N	N	Y	N	N
Demographics*Year FE	N	Y	Y	N	Y	Y	N	Y	Y
Demographics*State FE	N	Y	Y	N	Y	Y	N	Y	Y
Ln Population Density	Y	N	N	Y	N	N	Y	N	N
Ln Population Density*Year	N	Y	Y	N	Y	Y	N	Y	Y
Local Region*Year FE	N	N	Y	N	N	Y	N	N	Y
Race share geography	Fine	Fine	Fine	Fine	Fine	Fine	Coarse	Coarse	Coarse
Clustering	Local	Local	Local	Local	Local	Local	Local	Local	Local
Observations	41,931,431	41,931,431	41,931,431	996,093	996,093	996,093	399,014	399,012	399,012
R-squared	0.034	0.039	0.042	0.030	0.032	0.033	0.027	0.032	0.035

Table IA9 – Children Whose Mother Is Present but Is Not the Household Head or Spouse

For children under 18 in each international census, this table reports the share with a valid mother pointer (MOMLOC), the share whose pointed-to mother is the household head or the head's spouse (and who are therefore recovered by the construction used in the US tables), and the share whose mother is present in the household but is neither (and who are therefore lost by that construction). Ranges are across census years within the country. Colombia 1964, Costa Rica 1963, Ecuador 1962, Suriname 2004 and the United Kingdom 2001 have no pointer and are excluded from the ranges.

Country	Census years	Pointer available	Mother is head/spouse (recovered)	Mother present, not head/spouse (lost)
United States	1960-2015	91-96%	84-92%	2-7%
Brazil	1960-2010	90-93%	80-89%	3-11%
Colombia	1973-2005	83-88%	75-76%	8-12%
Costa Rica	1973-2011	91-94%	82-86%	6-11%
Cuba	2002-2012	86-88%	63-65%	22-23%
Ecuador	1974-2010	84-88%	75-80%	8-11%
El Salvador	1992-2007	78-85%	69-71%	9-14%
Jamaica	1982-2001	72-79%	58-61%	14-18%
Mozambique	1997-2007	81-83%	75-77%	6-7%
Puerto Rico	1970-2010	88-93%	77-87%	5-12%
South Africa	1996-2016	70-76%	50-59%	15-23%
Uruguay	1963-2011	87-94%	80-84%	6-10%
Zimbabwe	2012	66%	58%	9%
United Kingdom	1991	97%	95%	2%

Table IA10 – Number of Children and Racial Diversity (Children-at-Survey Measure)

This table reports the children-at-survey version of Table 2. Data is from the U.S. decennial census from 1850 to 2000, and the American Community Survey from 2001 to 2021. Observations are of women aged 18-40 at the time of survey, and the dependent variable is the number of children living with the woman at the time of survey. In Panel A, the independent variable is Race Share, the fraction of the local area population with the same race as the woman, where race is the nine census broad racial groups and a tenth category for Hispanic/Latino, measured in the survey year. Local area is measured as city (if present), then county (if city is missing), then metro area (if both city and county are missing). Fixed effects are included for race, state, year, state-by-year, demographics (age, marital status, education, race, employment status, income decile, and citizenship), demographics by state and year, area type (county, city or metro), deciles of population within that area type, deciles of population by year, area (measured at the same level as the race share), and area-by-year. Area parametric controls are average income decile, average age, fraction employed, and a z-score for the fraction of residents who moved in the last one or five years (based on data availability), also interacted with year fixed effects. The earliest year of data given the controls is noted. The effect of a one standard deviation change in race share on the number of children is reported both for an unconditional standard deviation and for a conditional standard deviation (the standard deviation of the residual from regressing race share on the fixed effects in the regression). In Panel B, the race share variable is replaced with a race Herfindahl index for that area and year. In Panel C, both the race share and the race Herfindahl index are included. In Panel D, the Herfindahl index is computed only among races other than the respondent's own. Standard errors are double clustered by year and state. Coefficients are in the top row, and t-statistics are below in parentheses, with *, **, and *** indicating statistical significance at the 10%, 5% and 1% level respectively.

Panel A - Race Share											
	Dependent variable is the number of children living with the woman at the time of survey										
Race Share	0.156***	0.707***	0.432***	0.310***	0.241***	0.290***	0.198***	0.152***	0.158***	0.188***	0.190***
	(2.78)	(6.82)	(6.87)	(7.30)	(8.54)	(7.07)	(6.54)	(4.77)	(5.08)	(5.78)	(5.77)
Effect of 1 σ change (unconditional)	0.051	0.230	0.141	0.101	0.078	0.094	0.064	0.049	0.051	0.061	0.062
Effect of 1 σ change (conditional)	0.051	0.123	0.066	0.049	0.038	0.037	0.025	0.018	0.019	0.020	0.020
Race	N	Y	Y	N	N	N	N	N	N	N	N
State	N	N	Y	Y	N	N	N	N	N	N	N
Year	N	N	Y	Y	N	N	N	N	N	N	N
State-Year FE	N	N	N	N	Y	Y	Y	Y	Y	Y	N
Demographics FE	N	N	N	Y	Y	N	N	N	N	N	N
Demographics*(State, Year) FE	N	N	N	N	N	Y	Y	Y	Y	Y	Y
Area Type	N	N	N	N	N	N	Y	N	N	N	N
Area Traits	N	N	N	N	N	N	Y	Y	N	N	N
Area Traits * Year FE	N	N	N	N	N	N	N	N	Y	Y	N
Area Type * Population FE	N	N	N	N	N	N	N	Y	N	N	N
Area Type * Population FE *Year	N	N	N	N	N	N	N	N	Y	Y	N
Area FE	N	N	N	N	N	N	N	N	N	Y	N
Area * Year FE	N	N	N	N	N	N	N	N	N	N	Y
Earliest Year	1850	1850	1850	1980	1980	1980	1980	1980	1980	1980	1980
Observations	7,099,506	7,099,506	7,099,505	5,925,639	5,925,639	5,925,627	5,925,627	5,925,627	5,925,627	5,925,627	5,925,627
R-squared	0.002	0.017	0.041	0.317	0.394	0.402	0.404	0.404	0.405	0.406	0.407

Panel B - Race Herfindahl							
	Dependent variable is the number of children living with the woman at the time of survey						
Race Herfindahl	0.592***	0.821***	0.666***	0.348***	0.208***	0.114***	0.041***
	(5.40)	(7.91)	(15.20)	(15.83)	(8.96)	(3.69)	(4.74)
Effect of 1 σ change (unconditional)	0.126	0.175	0.142	0.074	0.044	0.024	0.009
Effect of 1 σ change (conditional)	0.126	0.157	0.084	0.046	0.023	0.011	0.001
Race	N	Y	Y	N	N	N	N
State-Year FE	N	N	Y	Y	Y	Y	Y
Demographics FE	N	N	N	Y	N	N	N
Demographics*(State, Year) FE	N	N	N	N	Y	Y	Y
Area Type	N	N	N	N	Y	N	N
Area Traits	N	N	N	N	Y	N	N
Area Traits * Year FE	N	N	N	N	N	Y	Y
Area Type * Population FE *Year	N	N	N	N	N	Y	Y
Area FE	N	N	N	N	N	N	Y
Earliest Year	1850	1850	1980	1980	1980	1980	1980
Observations	7,099,506	7,099,506	7,099,505	5,925,639	5,925,627	5,925,627	5,925,627
R-squared	0.009	0.022	0.043	0.395	0.404	0.405	0.406

Panel C - Race Share and Race Herfindahl							
	Dependent variable is the number of children living with the woman at the time of survey						
Race Share	-0.168***	0.255***	0.227***	0.137***	0.148***	0.151***	0.190***
	(-5.27)	(4.54)	(4.65)	(4.47)	(4.07)	(4.43)	(5.75)
Race Herfindahl	0.759***	0.681***	0.538***	0.271***	0.119***	0.021	-0.071***
	(6.20)	(6.35)	(11.77)	(9.81)	(3.78)	(0.59)	(-3.49)
Race	N	Y	Y	N	N	N	N
State-Year FE	N	N	Y	Y	Y	Y	Y
Demographics FE	N	N	N	Y	N	N	N
Demographics*(State, Year) FE	N	N	N	N	Y	Y	Y
Area Type	N	N	N	N	Y	N	N
Area Traits	N	N	N	N	Y	N	N
Area Traits * Year FE	N	N	N	N	N	Y	Y
Area Type * Population FE *Year	N	N	N	N	N	Y	Y
Area FE	N	N	N	N	N	N	Y
Earliest Year	1850	1850	1980	1980	1980	1980	1980
Observations	7,099,506	7,099,506	7,099,505	5,925,639	5,925,627	5,925,627	5,925,627
R-squared	0.010	0.023	0.043	0.395	0.404	0.405	0.406

Panel D - Race Share and Race Herfindahl Across Other Races							
	Dependent variable is the number of children living with the woman at the time of survey						
Race Share	0.247***	0.696***	0.527***	0.321***	0.239***	0.177***	0.195***
	(4.45)	(7.10)	(6.79)	(7.48)	(7.11)	(5.34)	(5.39)
Other Race Herfindahl	0.704***	0.578***	0.315***	0.242***	0.104***	0.043**	0.013
	(6.49)	(5.43)	(3.44)	(4.85)	(4.03)	(2.22)	(0.69)
Race	N	Y	Y	N	N	N	N
State-Year FE	N	N	Y	Y	Y	Y	Y
Demographics FE	N	N	N	Y	N	N	N
Demographics*(State, Year) FE	N	N	N	N	Y	Y	Y
Area Type	N	N	N	N	Y	N	N
Area Traits	N	N	N	N	Y	N	N
Area Traits * Year FE	N	N	N	N	N	Y	Y
Area Type * Population FE *Year	N	N	N	N	N	Y	Y
Area FE	N	N	N	N	N	N	Y
Earliest Year	1850	1850	1980	1980	1980	1980	1980
Observations	7,055,592	7,055,592	7,055,591	5,925,639	5,925,627	5,925,627	5,925,627
R-squared	0.014	0.025	0.042	0.395	0.404	0.405	0.406

Table IA11 – Instrumenting Race Share with Immigration Shocks (Children-at-Survey Measure)

This table reports the children-at-survey version of Table 3. County-level race share in the survey year is instrumented with prior years' county-level immigration shocks, and the instrumented race share is related to the number of children living with the woman at the time of survey. Immigration shocks are taken from Burchardi et al. (2024), based on predicted ancestry in a given county in 1975 and national immigration from that ancestry group over a given five-year period. All other control variables are defined in Table IA10. Coefficients are in the top row, and t-statistics are below in parentheses, with *, **, and *** indicating statistical significance at the 10%, 5% and 1% level respectively.

	Dependent variable is the number of children living with the woman at the time of survey						
County Race Share, Instrumented	0.206**	0.919**	2.417***	1.375***	1.381***	1.754**	0.243
With Immigration Shocks	(2.12)	(4.42)	(6.74)	(5.87)	(6.98)	(3.34)	(0.93)
Effect of 1 σ change (unconditional)	0.064	0.286	0.751	0.427	0.429	0.545	0.076
Effect of 1 σ change (conditional)	0.064	0.151	0.359	0.203	0.153	0.183	0.022
Race	N	Y	Y	N	N	N	N
State-Year FE	N	N	Y	Y	Y	Y	Y
Demographics FE	N	N	N	Y	N	N	N
Demographics*(State, Year) FE	N	N	N	N	Y	Y	Y
Area Type	N	N	N	N	Y	N	N
Area Traits	N	N	N	N	Y	N	N
Area Traits * Year FE	N	N	N	N	N	Y	Y
Area Type * Population FE *Year	N	N	N	N	N	Y	Y
Area FE	N	N	N	N	N	N	Y
Earliest Year	1850	1850	1850	1980	1980	1980	1980
Observations	1,026,653	1,026,653	1,026,653	1,026,653	1,026,625	1,026,625	1,026,625
R-squared	-0.005	0.002	-0.054	-0.029	-0.016	-0.029	0.000

Table IA12 – Mobility (Children-at-Survey Measure)

This table reports the children-at-survey version of Table 5. The regressions of Table IA10 are run on subsamples of women less likely to have relocated: i) those living in the state they were born in, ii) those who have not moved in the past year, iii) those who have not moved in the past five years, and combinations of these. The dependent variable is the number of children living with the woman at the time of survey. Standard errors are double clustered by year and state. Coefficients are in the top row, and t-statistics are below in parentheses, with *, **, and *** indicating statistical significance at the 10%, 5% and 1% level respectively.

Selection	Living in State of Birth	No Move in Last Year	No Move in Last Five Years	No Move in Last One or Five Years	Any of Previous
	Dependent variable is the number of children living with the woman at the time of survey				
Race Share	0.154***	0.195***	0.118***	0.186***	0.183***
	(4.79)	(5.05)	(3.79)	(5.03)	(5.31)
Demographics*(State, Year) FE	Y	Y	Y	Y	Y
Area * Year FE	Y	Y	Y	Y	Y
Clustering	State, Year		State	State, Year	
Observations	3,187,811	3,843,774	421,774	4,265,600	5,067,659
R-squared	0.407	0.408	0.458	0.414	0.409

Table IA13 – Different Time Periods (Children-at-Survey Measure)

This table reports the children-at-survey version of Table 6: how racial diversity is associated with the number of children living with the woman at the time of survey in different periods of US history. Specifications from Table IA10 are run separately for i) 1850-1860, ii) 1870-1890, iii) 1900-1940, iv) 1950-1970, v) 1980-1990, and vi) 2000-2021. Panel A includes state-by-year fixed effects as well as age and race each interacted with state and year. Panel B also includes area-by-year fixed effects. The effect of a one unconditional standard deviation change in race share within each period is reported. Standard errors are double clustered by year and state. Coefficients are in the top row, and t-statistics are below in parentheses, with *, **, and *** indicating statistical significance at the 10%, 5% and 1% level respectively.

Panel A - No Area*Year Fixed Effects						
Period	1850-1860	1870-1890	1900-1940	1950-1970	1980-1990	2000-2021
	Dependent variable is the number of children living with the woman at the time of survey					
Race Share	2.565**	0.856**	0.947***	1.049***	0.546***	0.511***
	(2.69)	(2.57)	(3.16)	(3.88)	(6.34)	(5.90)
Effect of 1 σ change (unconditional)	0.451	0.180	0.226	0.308	0.171	0.157
State-Year FE	Y	Y	Y	Y	Y	Y
(Age, Race)*(State, Year) FE	Y	Y	Y	Y	Y	Y
Area * Year FE	N	N	N	N	N	N
Observations	18,701	101,526	459,410	594,085	650,416	5,275,201
R-squared	0.311	0.258	0.184	0.283	0.262	0.271

Panel B - With Area*Year Fixed Effects						
Period	1850-1860	1870-1890	1900-1940	1950-1970	1980-1990	2000-2021
	Dependent variable is the number of children living with the woman at the time of survey					
Race Share	-1.031	0.028	0.078	0.333*	0.327***	0.310***
	(-1.30)	(0.18)	(0.76)	(1.83)	(5.27)	(5.43)
Effect of 1 σ change (unconditional)	-0.181	0.006	0.019	0.098	0.103	0.096
State-Year FE	Y	Y	Y	Y	Y	Y
(Age, Race)*(State, Year) FE	Y	Y	Y	Y	Y	Y
Area * Year FE	Y	Y	Y	Y	Y	Y
Observations	18,701	101,524	459,405	594,085	650,416	5,275,201
R-squared	0.319	0.273	0.201	0.291	0.274	0.285

Table IA14 – Effects by Race (Children-at-Survey Measure)

This table reports the children-at-survey version of Table 7: how the effect of racial diversity on the number of children living with the woman at the time of survey varies with the race of the woman. Observations are women aged 18-40 in US census surveys from 1850 to 2021. Specifications from Table IA10 are run with interactions between the baseline race share variable and the ten racial and ethnic groups considered (from census categories): white, black, Native American, Chinese, Japanese, Asian/Pacific Islander, other, two races, three or more races, and Hispanic. Controls are the same as in Table IA10. Standard errors are double clustered by year and state. Coefficients are in the top row, and t-statistics are below in parentheses, with *, **, and *** indicating statistical significance at the 10%, 5% and 1% level respectively.

	Dependent variable is the number of children living with the woman at the time of survey						
Race Share * White	1.005***	0.706***	0.334***	0.418***	0.256***	0.256***	0.181***
	(7.44)	(7.61)	(8.93)	(9.95)	(6.15)	(5.94)	(3.68)
Race Share * Black	0.069	0.091	0.117*	0.051*	0.114**	0.120***	0.256***
	(0.64)	(1.05)	(1.97)	(1.78)	(2.74)	(3.13)	(4.08)
Race Share * Native American	0.432***	0.395***	0.657***	0.420***	0.084	0.104	0.250*
	(3.21)	(4.02)	(8.30)	(3.93)	(0.85)	(1.01)	(1.84)
Race Share * Chinese	-0.513**	-0.026	-0.086	-0.085	0.667***	0.702***	0.972***
	(-2.18)	(-0.12)	(-0.87)	(-0.44)	(4.08)	(4.20)	(7.88)
Race Share * Japanese	0.575**	-0.874***	0.146*	-0.055	0.603	0.659	1.078**
	(2.58)	(-5.39)	(1.94)	(-0.13)	(1.29)	(1.41)	(2.17)
Race Share * Asian / Pacific Islander	-0.252	-0.164	0.041	-0.104	0.171	0.178	0.157
	(-0.88)	(-1.09)	(0.27)	(-0.77)	(1.32)	(1.39)	(0.90)
Race Share * Other	0.546	14.466***	4.166	2.384	2.965	2.871	4.903*
	(0.17)	(2.92)	(1.34)	(0.80)	(0.98)	(0.95)	(1.83)
Race Share * Two Races	0.543	0.825	0.185	1.538*	2.632***	2.807***	1.975***
	(0.46)	(1.57)	(0.43)	(2.08)	(4.39)	(4.64)	(4.38)
Race Share * Three or More Races	2.512***	1.064	1.423***	-0.466	-0.709	-0.708	-1.003
	(5.42)	(1.68)	(4.57)	(-0.55)	(-0.74)	(-0.72)	(-1.13)
Race Share * Hispanic	0.115	0.084	0.132*	0.131	-0.020	-0.004	0.151
	(1.29)	(1.02)	(1.93)	(1.66)	(-0.36)	(-0.07)	(1.57)
Race	Y	Y	N	N	N	N	N
State-Year FE	N	Y	Y	Y	Y	Y	N
Demographics FE	N	N	Y	N	N	N	N
Demographics*(State, Year) FE	N	N	N	Y	Y	Y	Y
Area Type * Population FE	N	N	N	Y	Y	N	N
Area Type * Population FE *Year	N	N	N	N	N	Y	N
Area Traits	N	N	N	N	Y	N	N
Area Traits * (State, Year) FE	N	N	N	N	N	Y	N
Area FE	N	N	N	N	N	N	N
Area * Year FE	N	N	N	N	N	N	Y
Earliest Year	1850	1850	1850	1980	1980	1980	1980
Observations	7,099,506	7,099,505	5,925,639	5,925,627	5,925,627	5,925,627	5,925,627
R-squared	0.020	0.042	0.395	0.403	0.405	0.405	0.407

Table IA15 – Trust and Number of Children (Children-at-Survey Measure)

This table reports the children-at-survey version of Table 11: how local measures of trust affect the relationship between racial isolation and the number of children living with the woman at the time of survey. In Panel A, the generalized state-level trust measures in 2006 from Putnam (2007) are compared with state-level race share measures (columns 1-6) and baseline local area measures (columns 7-12), for all years, the years 2001 to 2010 only, or 2006 only. In Panel B, the county-level measures of social capital from Chetty et al. (2022), namely the volunteering rate, friendship clustering and economic connectedness, are compared with county-only measures of race share, for all years, 2011 to 2021, or 2021 only; counties are matched on the five-digit state and county FIPS code. Standard errors are double clustered by year and state. Coefficients are in the top row, and t-statistics are below in parentheses, with *, **, and *** indicating statistical significance at the 10%, 5% and 1% level respectively.

Panel A - Social Capital Survey Direct Trust Measure												
Years	All		2001-2010		2006		All		2001-2010		2006	
	Dependent variable is the number of children living with the woman at the time of survey											
Race Share (State)	0.362***	0.286***	0.357***	0.271***	0.332***	0.246***						
	(7.42)	(7.07)	(6.48)	(5.84)	(5.75)	(4.76)						
State Level Trust		0.431***		0.467***		0.452***		0.325**		0.351***		0.362***
		(4.26)		(4.71)		(4.31)		(2.72)		(2.94)		(2.96)
Race Share (Base)							0.294***	0.274***	0.284***	0.262***	0.280***	0.259***
							(9.52)	(9.75)	(7.75)	(7.57)	(8.61)	(8.47)
Demographics*Year FE	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y	N
Observations	9,112,239	9,112,239	3,138,415	3,138,415	411,956	411,956	5,925,639	5,873,004	1,945,868	1,934,113	320,926	319,237
R-squared	0.385	0.386	0.371	0.371	0.373	0.374	0.395	0.396	0.383	0.384	0.380	0.381

Panel B - Meta Social Capital Metrics						
Years	All		2011-2021		2021	
	Dependent variable is the number of children living with the woman at the time of survey					
Race Share (County)	0.407***	0.309***	0.420***	0.314***	0.450***	0.320***
	(6.73)	(5.42)	(6.01)	(4.87)	(6.47)	(5.64)
Volunteering Rate		-1.604***		-1.687***		-1.878***
		(-4.09)		(-4.40)		(-3.83)
Friendship Clustering		8.582***		8.926***		10.373***
		(8.97)		(8.95)		(9.15)
Economic Connectedness		-0.106		-0.071		-0.040
		(-1.11)		(-0.74)		(-0.41)
State-Year FE	Y	Y	Y	Y	Y	Y
Demographics*(State, Year) FE	Y	Y	Y	Y	Y	Y
Observations	4,971,530	4,971,530	2,957,743	2,957,743	281,284	281,284
R-squared	0.327	0.329	0.327	0.329	0.321	0.324

Table IA16 – Similarity in Other Variables and Number of Children (Children-at-Survey Measure)

This table reports the children-at-survey version of Table 12: how diversity in other demographic variables is associated with the number of children living with the woman at the time of survey. For each demographic variable, the independent variable is the fraction of residents in the local area who share the same value of the trait as the woman, measured in the survey year. Local area is taken as county, then city if county is missing, then detailed metro area if both city and county are missing. Panel A examines education, income and age. Education is a dummy variable for the highest level of schooling (e.g., high school, college, graduate degree). Income is deciles of income across the US in the year in question. Age is the fraction of the population that is between two years younger and ten years older than the woman. Panel B examines country of birth and citizenship. Country of birth is a dummy for whether the person was born in the US, and citizenship is a dummy for whether the person is a US citizen. Panel C includes these variables together, and also reports the effect of a one-standard-deviation unconditional change in each variable for each specification. All other control variables are defined in Table IA10. Standard errors are double clustered by year and state. Coefficients are in the top row, and t-statistics are below in parentheses, with *, **, and *** indicating statistical significance at the 10%, 5% and 1% level respectively.

Panel A - Education, Income, Age												
	Dependent variable is the number of children living with the woman at the time of survey											
Education Share	0.097*	0.171**	-0.021	-0.022								
	(1.89)	(2.56)	(-0.62)	(-0.65)								
Income Decile Share					0.562***	0.745***	0.717***	0.725***				
					(6.93)	(9.93)	(10.70)	(10.89)				
Age (-2,+10) Share									-0.839***	-0.603***	0.406**	0.432***
									(-5.64)	(-5.22)	(2.81)	(2.96)
State-Year FE	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y	N
Demographics FE	Y	N	N	N	Y	N	N	N	Y	N	N	N
Demographics*(State, Year) FE	N	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y
Area Type * Population FE *Year	N	N	Y	N	N	N	Y	N	N	N	Y	N
Area Traits * Year FE	N	N	Y	N	N	N	Y	N	N	N	Y	N
Area * Year FE	N	N	N	Y	N	N	N	Y	N	N	N	Y
Observations	5,925,639	5,925,627	5,925,627	5,925,627	5,925,639	5,925,627	5,925,627	5,925,627	5,925,639	5,925,627	5,925,627	5,925,627
R-squared	0.393	0.401	0.406	0.407	0.394	0.402	0.406	0.407	0.394	0.402	0.406	0.407

Panel B - Country of Birth, Citizenship								
	Dependent variable is the number of children living with the woman at the time of survey							
US Born Share	0.160***	0.187***	-0.019	-0.020				
	(4.50)	(4.62)	(-0.93)	(-0.98)				
Citizenship Share					0.364***	0.418***	0.053	0.051
					(11.61)	(8.08)	(1.39)	(1.35)
State-Year FE	Y	Y	Y	N	Y	Y	Y	N
Demographics FE	Y	N	N	N	Y	N	N	N
Demographics*(State, Year) FE	N	Y	Y	Y	N	Y	Y	Y
Area Type * Population FE *Year	N	N	Y	N	N	N	Y	N
Area Traits * (State, Year) FE	N	N	Y	N	N	N	Y	N
Area * Year FE	N	N	N	Y	N	N	N	Y
Observations	5,925,639	5,925,627	5,925,627	5,925,627	5,925,639	5,925,627	5,925,627	5,925,627
R-squared	0.394	0.402	0.406	0.407	0.394	0.402	0.406	0.407

Panel C - Other Variables in Combination				
	Dependent variable is the number of children living with the woman at the time of survey			
Race Share	0.191***	0.220***	0.180***	0.182***
	(9.94)	(7.19)	(5.81)	(5.84)
Education Share	0.010	0.012	-0.096***	-0.097***
	(0.23)	(0.25)	(-3.29)	(-3.37)
Income Decile Share	0.515***	0.652***	0.638***	0.645***
	(7.50)	(10.80)	(10.33)	(10.59)
Age (within 10 years) Share	-0.651***	-0.366***	0.347**	0.371**
	(-3.80)	(-3.05)	(2.42)	(2.56)
US Born Share	-0.403***	-0.346***	-0.163***	-0.164***
	(-8.66)	(-5.74)	(-5.35)	(-5.10)
Citizenship Share	0.687***	0.636***	0.180***	0.182***
	(11.93)	(11.11)	(4.06)	(3.82)
Effect of 1 σ change (unconditional), Race Share	0.062	0.072	0.059	0.059
Effect of 1 σ change (unconditional), Education Share	0.001	0.002	-0.014	-0.014
Effect of 1 σ change (unconditional), Income Decile Share	0.020	0.025	0.024	0.024
Effect of 1 σ change (unconditional), Age (within 10 years) Share	-0.035	-0.020	0.018	0.020
Effect of 1 σ change (unconditional), US Born Share	-0.100	-0.086	-0.040	-0.041
Effect of 1 σ change (unconditional), Citizenship Share	0.213	0.197	0.056	0.057
State-Year FE	Y	Y	Y	N
Demographics FE	Y	N	N	N
Demographics*(State, Year) FE	N	Y	Y	Y
Area Type * Population FE *Year	N	N	Y	N
Area Traits * (State, Year) FE	N	N	Y	N
Area * Year FE	N	N	N	Y
Observations	5,925,639	5,925,627	5,925,627	5,925,627
R-squared	0.396	0.403	0.407	0.408