Online Appendix for "Childhood Environment and Gender Gaps in Adulthood"

Raj Chetty, Stanford University and NBER
Nathaniel Hendren, Harvard University and NBER
Frina Lin, Stanford University
Jeremy Majerovitz, Harvard University
Benjamin Scuderi, Harvard University

January 2016

Appendix A. Panel Study of Income Dynamics: Sample and Variable Definitions

Appendix Figure 2 examines the relationship between parent income and children's employment rates by gender using the Panel Study of Income Dynamics (PSID).[1] To construct this sample, we use the family and individual files from the PSID, as well as a spine from the PSID which matches parents to children.

We define income as total labor income until 1993; after 1993, we define income as total labor income excluding farm and business income, which is the primary definition of labor income in the PSID after 1993. We compute the income of each parent as their average labor income from ages 30 to 50 (or whatever subset of those ages for which they are in the sample). We use a longer average in the PSID than we do in the tax data (where we average income over 5 years) to account for the greater degree of measurement error in the PSID, as discussed in Chetty et al. (2014). We define parent family income as the sum of the father's and mother's average income.

We define an indicator for the child's employment based on their response at age 30 to the employment status question. We consider a child employed if they answer that they are "working now" or "temporarily laid off" when asked about their employment status. A few children appear twice at age 30 due to the timing of the PSID survey; we assign these children a single observation where their employment is defined as the average of the two observed values of employment. Some children who are interviewed at age 30 are not asked the employment question (e.g., due to institutionalization); these children are coded as unemployed to match our definition in the tax data. In addition, employment in adulthood is not observed for many children due to attrition from the sample. We code these observations as missing and exclude them from our analysis.

Appendix B. Correlates of Gender Gaps: Permanent Residents' Outcomes vs. Causal Effects

This appendix discusses the analysis in Appendix Table 3, which correlates selected CZ characteristics with different estimates of gender gaps.

Motivation. In Figure 4 of the paper, we present correlations of gender differences in employment rates for children who grew up in different CZs with a range of characteristics of those CZs. These correlations with gender differences could reflect either the causal effect of growing up in CZ with certain characteristics or sorting of different types of families to such CZs. For instance, families who live in high-poverty neighborhoods might have different characteristics that lead to worse outcomes for boys relative to girls, which would lead to the observed correlation even if growing up in a

[1] We are grateful to Alex Bell for suggesting and helping us implement the comparison to the PSID.

high-poverty CZ does not have differential causal effects by gender. To address this problem, one must correlate characteristics such as poverty rates with estimates of the *causal effect* of growing up in each CZ on gender gaps.

Estimates of Causal Effects. Chetty and Hendren (2015) estimate gender-specific causal effects of spending an additional year growing up in each CZ by analyzing families who move across CZs. Their estimates are identified from differences in the ages at which children move to different CZs. Intuitively, if children who move from New York to Boston do better as adults the earlier they move, one can infer that growing up in Boston has a positive causal effect relative to New York. Building on this logic, Chetty and Hendren (2015) estimate a fixed effects model that identifies the causal effect of spending an additional year of childhood in each CZ on earnings in adulthood, separately for boys and girls. We multiply these annual exposure effect estimates by 20 to obtain an estimate of the causal effect of growing up in each CZ from birth.

Note that Chetty and Hendren (2015) do not compute causal effects on employment rates for children who grow up low-income families. Instead, they estimate effects on individual income ranks at age 26 for children whose parents are at the 25th percentile of the parental income distribution. Individual income is defined as the sum of individual W-2 wage earnings, UI benefits, SSDI payments, and half of household self-employment income (see Online Appendix A of Chetty et al. (2014) for details). We translate this income measure into percentile ranks (scaled from 0-100) by ranking children relative to other children in the same birth cohort.

Since estimates of causal effects for employment rates are not available, we first compare results for employment rates and individual income ranks in the sample of permanent residents (non-movers) and then turn to the causal effect estimates. In the interest of space, we focus on three of the strongest predictors of gender gaps in employment rates: segregation of poverty, the fraction of black residents in the CZ, and the fraction of single parents.

Results for Employment Rates at Age 30. In Panel A of Appendix Table 3, we regress gender differences (male - female) in employment rates at age 30 for children with parents in the bottom income quintile on the three CZ-level characteristics. Each covariate is standardized to have mean 0 and standard deviation of 1 in the full sample, so the regression coefficients can be interpreted as the effect of a 1 SD increase in the covariate on the male-female difference in employment rates (measured in percentage points). The sample used to estimate these regressions consists of children whose parents never moved across CZs, as in Figure 3.

Columns 1-3 report estimates from univariate regressions, replicating specifications reported

in Appendix Table 2. Boys who grow up in areas with greater segregation, a higher fraction of African Americans, or more single parents have lower employment rates than girls who grew up in the same areas. Column 4 includes all three of these variables in a multivariable regression. Both segregation and the fraction black continue to have a significant relationship with the gender difference in employment; however, the coefficient on the fraction of single parents falls and is no longer significantly different from zero. Column 5 adds state fixed effects and shows that both segregation and the fraction black remain significant predictors of the gender gap in employment across CZs within states.

Results for Income Ranks at Age 26. Next, we show that the pattern of correlations is similar when measuring children's outcomes using their income rank at age 26, as in Chetty and Hendren (2015), instead of employment rates at age 30. Panel B of Appendix Table 3 replicates the specifications in Panel A, changing the dependent variable to the gender difference in mean individual income rank at age 26 for children with parents at the 25th percentile of the national income distribution. These mean ranks are estimated as in Chetty et al. (2014) using linear regressions of children's percentile ranks on parent ranks.

Once again, we find that areas with greater segregation, a higher fraction of African Americans, and more single parents have better outcomes for women relative to men. For example, a one standard deviation increase in the segregation of poverty is associated with a 3 percentile reduction in men's expected rank relative to women who grow up in the same areas. In the multivariable regression in Column 4, both segregation and the fraction black have a significant relationship with the gender difference in individual income ranks. However, the coefficient on the fraction of single parents is positive but not statistically significantly different from zero, similar to the result for employment rates at age 30. Adding state fixed effects in Column 5 does not change these results. In short, the correlates of the gender gap in individual income ranks at age 26 are very similar to those observed for employment rates at age 30.

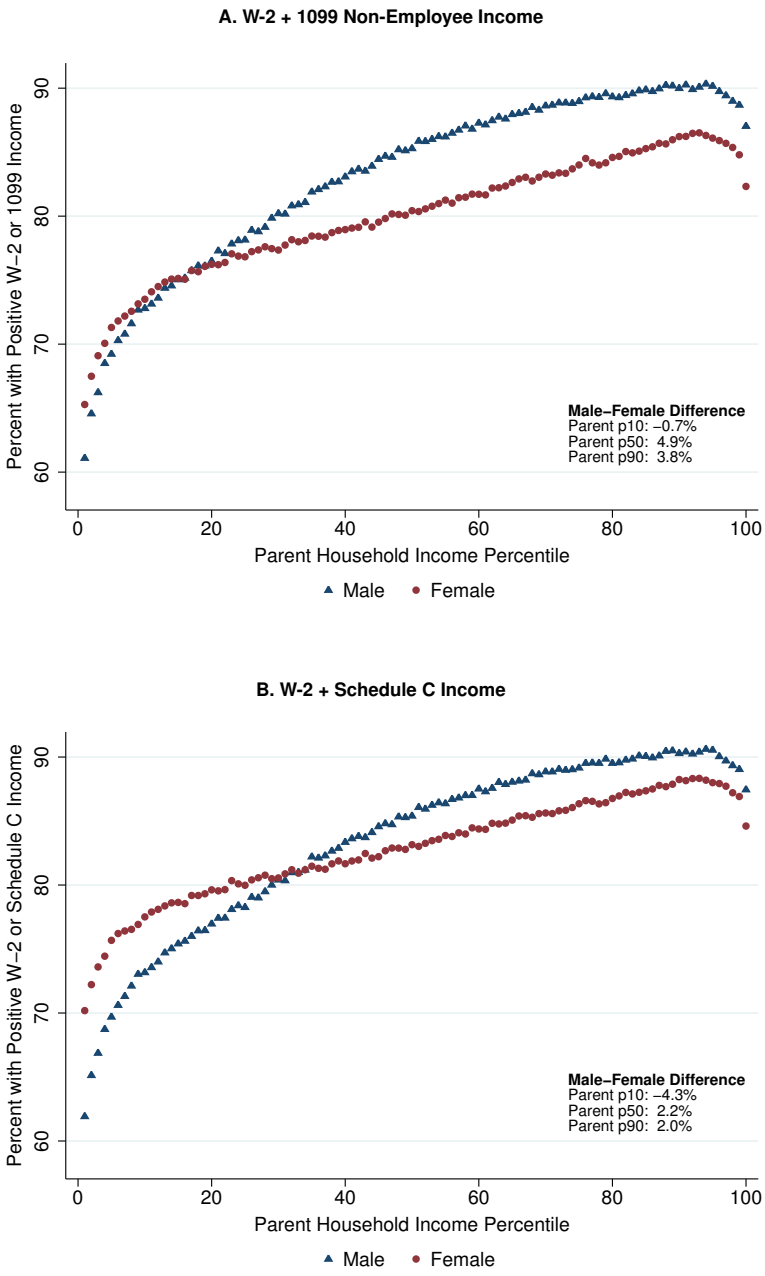
Results for Causal Effects on Income Ranks at Age 26. Finally, in Panel C, we turn to the correlates of the gender gap in causal effects estimated based on the outcomes of children whose families moved across CZs. We replicate the same specifications as in Panel B, changing the dependent variable to the gender difference in the causal exposure effect estimates reported in Chetty and Hendren (2015) for children with parents at the 25th percentile (multiplied by 20 to obtain the effect of 20 years of exposure).

Consistent with the pattern seen for the permanent residents in Panel B, we find that a one

standard deviation increase in the segregation of poverty is associated with a 2.5 percentile reduction in men's expected rank relative to women who grow up in the same areas. The coefficient on segregation remains quite stable even when controlling for the fraction of black residents and single parents (Column 4) and including state fixed effects (Column 5). Hence, there is robust evidence that growing up in an area with greater segregation produces worse economic outcomes for men relative to women in low-income families.

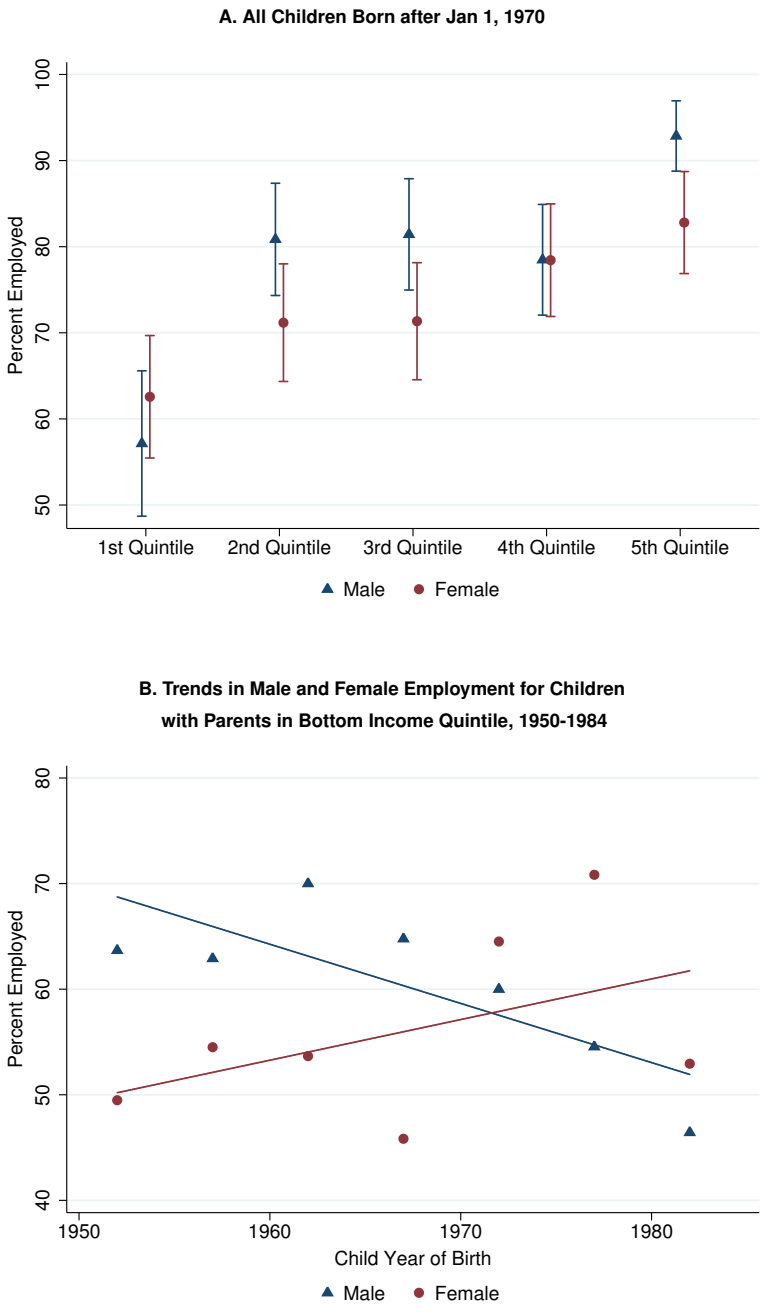
In contrast, using the causal estimates of the gender gap yields smaller coefficients on the fraction black and fraction single parents relative to the specifications for permanent residents in Panel B. This suggests that some of the correlation between permanent residents' outcomes with the fraction black and single parents is driven by sorting (selection effects) rather than a correlation with the causal effect of childhood exposure to such areas.

Appendix Figure 1. Employment Rates by Gender and Parent Income: Robustness Checks



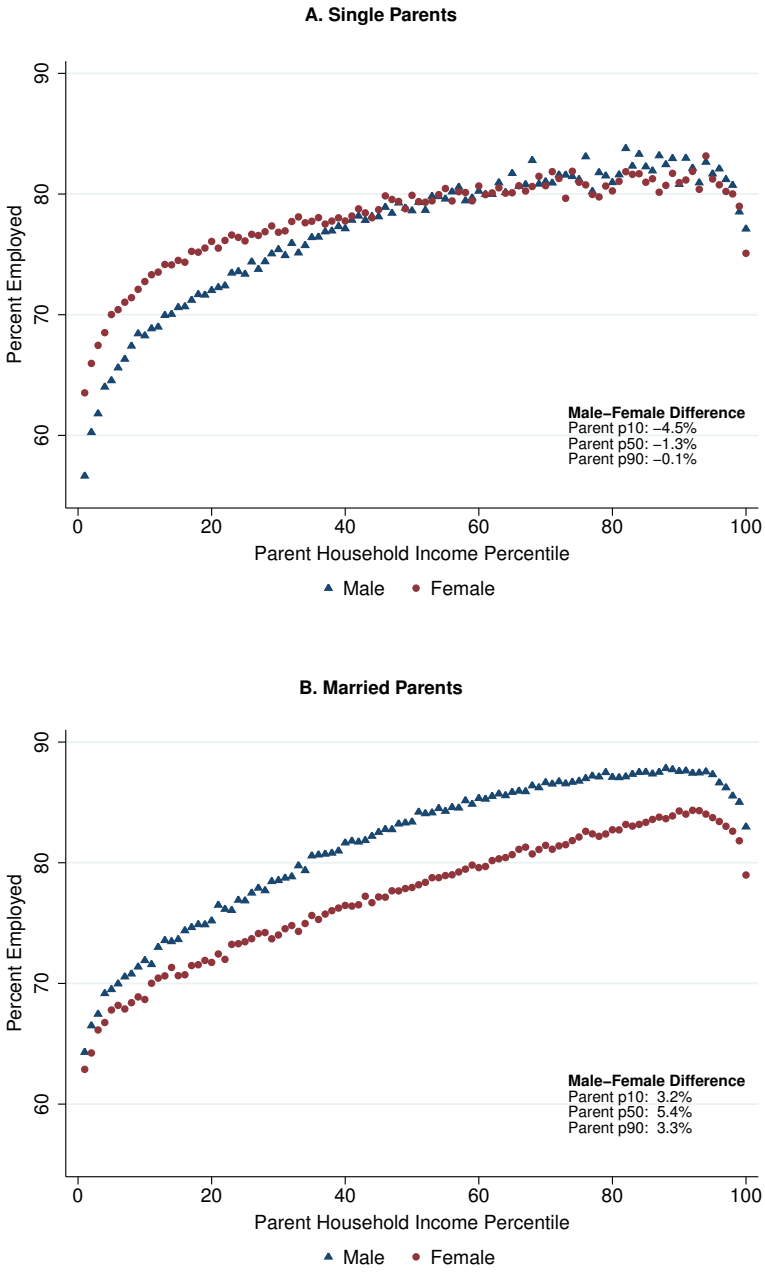
Notes: These figures replicate Figure 1 using alternative measures of employment. Panel A replicates Figure 1 defining employment as an indicator for having either a W-2 form or non-zero non-employee compensation in Form 1099 (box 7) in the tax year during which the child turns 30 years old. Non-employee compensation includes payments to independent contractors as well as other forms of income that may not be from employment, such as prize winnings. Panel B replicates Figure 1 defining employment as an indicator for having either a W-2 form or non-zero Schedule C income in the tax year during which the child turns 30 years old. The Schedule C form includes self-employment income, but only one form is filed for married spouses filing jointly. As a result, unemployed individuals whose spouses have self-employment income would be counted as employed in Panel B.

Appendix Figure 2. Employment Rates by Gender and Parent Income in the Panel Study of Income Dynamics



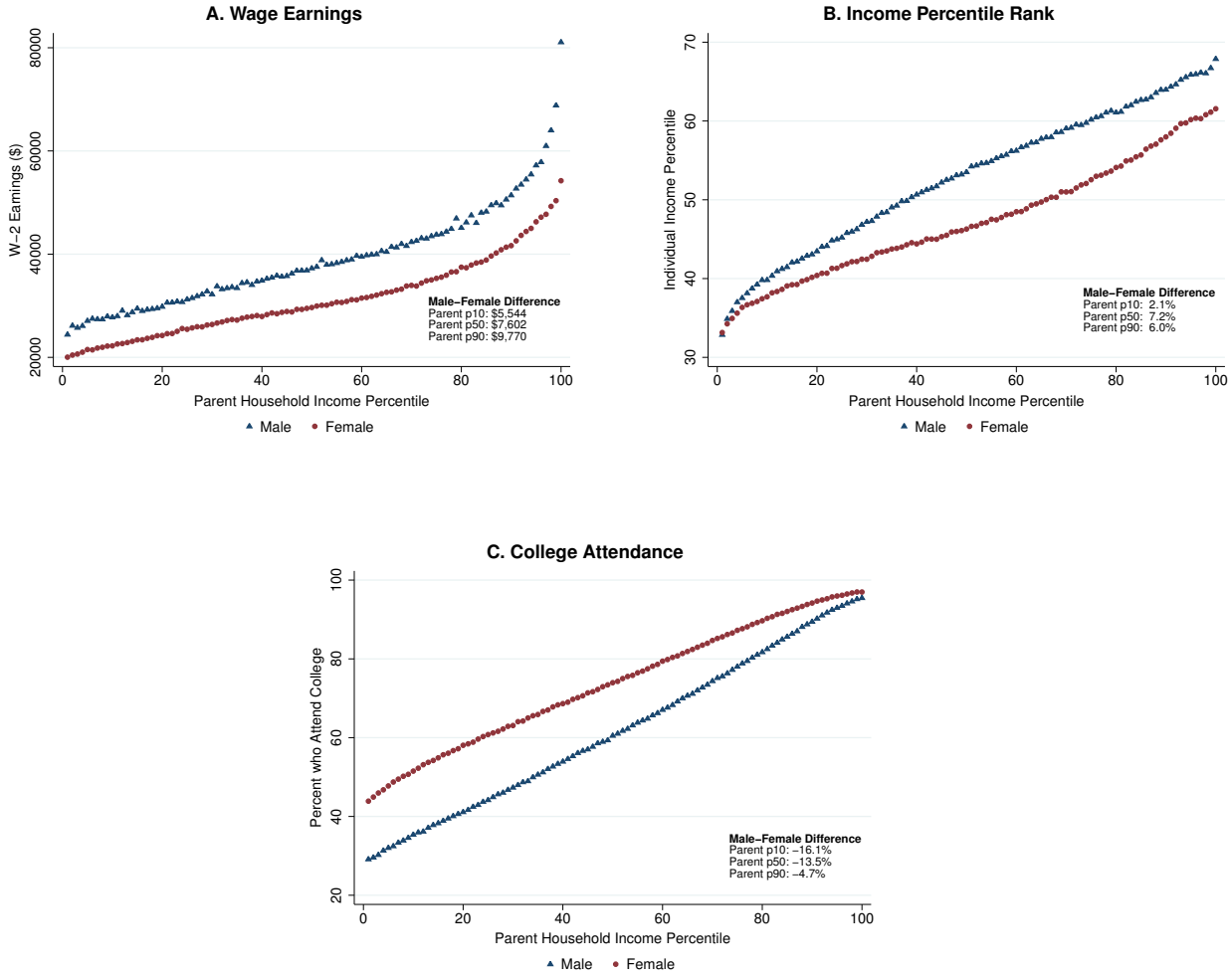
Notes: This figure shows employment patterns for children in the PSID by gender and parental income. The sample consists of all children in the PSID observed at age 30 who have non-missing parental income measures. We define employment as being currently employed or temporarily laid off (see Appendix A for details). We define parental income as mean labor income averaged over all years when the parents are between the ages 30 and 50 for which we observe them in the sample. Panel A pools all cohorts from 1970-1984 and reports employment rates by parental income quintile (defining quintiles within each birth cohort) for males and females at age 30, along with 95% confidence intervals for these estimates. Panel B reports estimates of employment rates for men and women with parents in the bottom income quintile by birth cohort (in five-year bins), along with linear trend estimates.

Appendix Figure 3. Employment Rates by Gender and Parent Income: Single vs. Married Parents



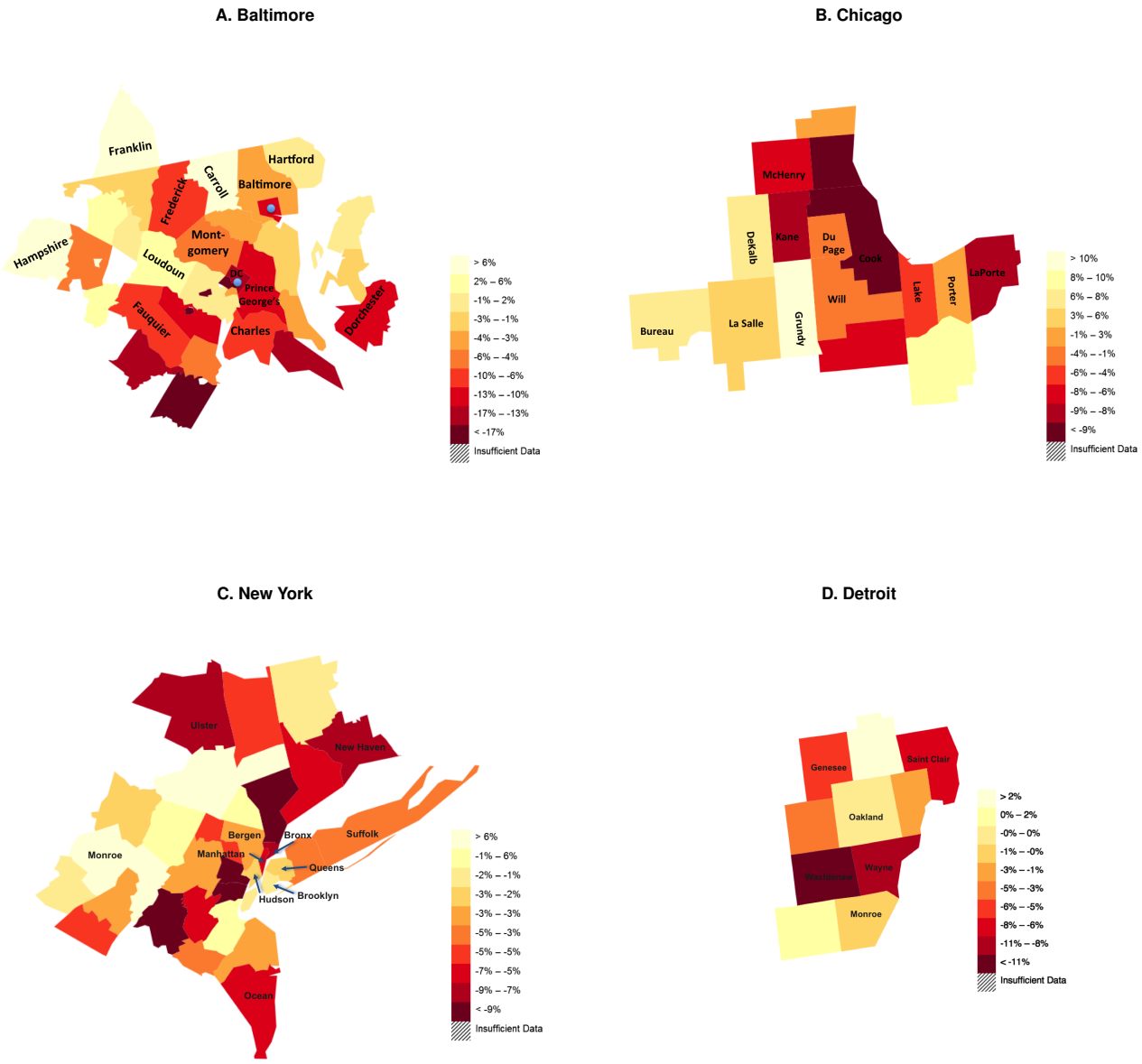
Notes: This figure replicates Figure 1, splitting the sample into children with single (Panel A) and married (Panel B) parents. Parent income ranks are the same as those used in Figure 1, pooling across married and un-married parents. Parental marital status is defined as the parents' marital status on the tax form on which the child is first claimed as a dependent.

Appendix Figure 4. Other Outcomes by Gender and Parent Income Percentile



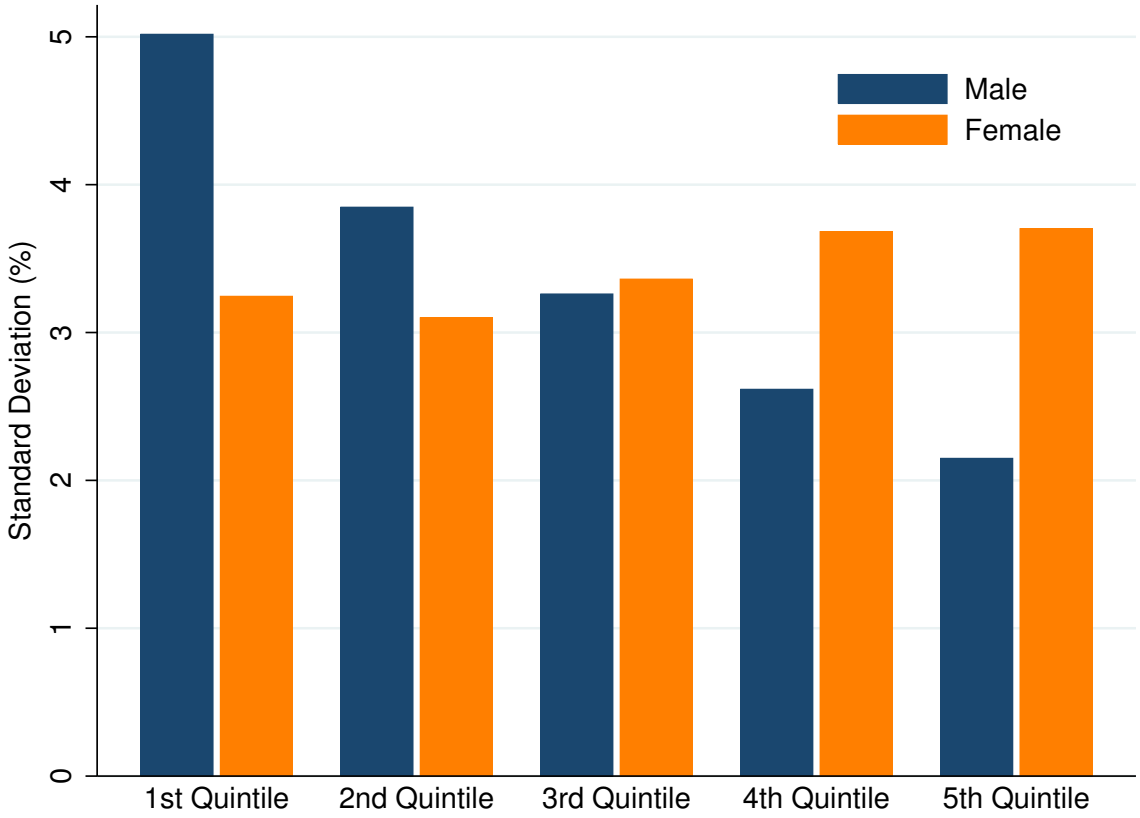
Notes: These figures replicate Figure 1 using three different outcomes for children: wage earnings (Panel A), individual income rank (Panel B), and college attendance rates (Panel C). Wage earnings is defined as the sum of all earnings on W-2s in the tax year during which the child turns age 30. Following Chetty et al. (2014), individual income is defined as the sum of wage earnings, unemployment income, disability income, and other (e.g. self employment) income from Schedule C (self-employment income is divided by 2 for married spouses). We translate this income measure into percentile ranks by ranking children relative to other children in the same birth cohort. College attendance is defined as an indicator for having a 1098-T form filed by a higher education institution on one's behalf between ages 18-23. Panels A and B use the 1980-82 birth cohorts, for whom we observe outcomes at age 30 in our dataset (which ends in 2012). Panel C uses the 1981-89 birth cohorts, for whom we observe outcomes at ages 18-23.

Appendix Figure 5. Gender Differences in Employment Rates by County for Children with Parents in Bottom Income Quintile



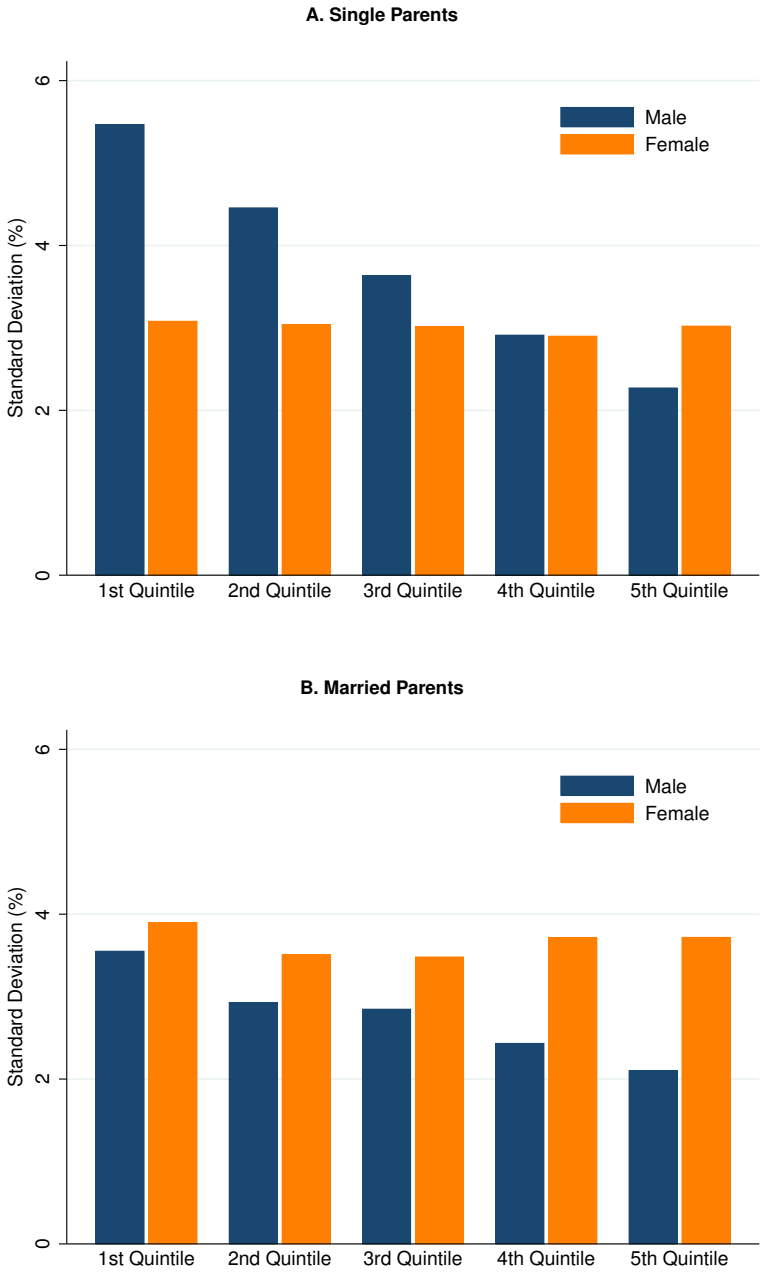
Notes: These maps show the difference in employment rates at age 30 (males - females) by the county where children grew up (i.e., the county where their parents lived). The sample consists of children in the 1980-82 birth cohorts whose parents stay in the same county throughout our sample window, 1996-2012, and who are in the bottom quintile of the national income distribution. Darker colors represent areas where men have lower employment rates than women. We show maps for counties in four combined statistical areas (CSAs): Baltimore, Chicago, New York, and Detroit.

Appendix Figure 6. Standard Deviations of Employment Rates Across CZs by Gender and Parent Income Quintile



Notes: This figure plots the standard deviation of employment rates across CZs for men and women by parent income quintile. The sample consists of children in the 1980-82 birth cohorts whose parents stay in the same CZ throughout our sample window, 1996-2012. Parents are grouped into quintiles based on their ranks in the national income distribution relative to other parents with children in the same birth cohort. Each value plotted in the figure is constructed in three steps. First, we compute the total variance of the employment rates across CZs for each gender and parent income quintile, weighting by the population of the CZ in the 2000 Census. Second, we calculate the portion of this variance that is due to estimation error (the noise variance) by squaring the standard errors of the employment rate estimates in each CZ and taking the (population-weighted) mean of these values. Finally, we subtract the noise variance from the total variance of employment rates across CZs to obtain estimates of the signal variance. The figure plots the square root of these values (the signal standard deviations) in each group.

Appendix Figure 7. Standard Deviations of Employment Rates Across CZs by Gender and Parent Income Quintile: Single vs. Married Parents



Notes: This figure replicates Appendix Figure 6, splitting the sample into children with single (Panel A) and married (Panel B) parents. Parent income ranks are the same as those used in Appendix Figure 6, pooling across married and un-married parents. Parental marital status is defined as the parents' marital status on the tax form on which the child is first claimed as a dependent.

APPENDIX TABLE 1

Gender Differences in Employment Rates for Children with Parents in the Bottom Quintile: Top and Bottom 10 CZs

Rank	CZ	Difference (1)	Male (2)	Female (3)	Rank	CZ	Difference (1)	Male (2)	Female (3)
1	Salt Lake City, UT	9.8	78.9	69.1	91	Milwaukee, WI	-9.2	65.0	74.2
2	Bakersfield, CA	7.3	76.8	69.5	92	Dallas, TX	-9.4	64.7	74.1
3	El Paso, TX	7.2	81.8	74.6	93	Washington DC, DC	-9.7	66.6	76.3
4	Brownsville, TX	5.8	82.6	76.8	94	St. Louis, MO	-11.0	65.0	76.0
5	Erie, PA	4.1	75.6	71.5	95	Atlanta, GA	-11.1	59.3	70.4
6	Eugene, OR	4.0	69.0	65.0	96	Virginia Beach, VA	-11.6	65.0	76.6
7	Canton, OH	3.7	69.0	65.3	97	Charlotte, NC	-12.4	60.1	72.5
8	Reading, PA	3.2	73.7	70.5	98	Raleigh, NC	-13.6	59.9	73.5
9	Spokane, WA	2.5	70.3	67.8	99	Memphis, TN	-15.3	59.2	74.5
10	Syracuse, NY	2.4	74.2	71.8	100	Richmond, VA	-16.0	62.3	78.3

Notes: This table presents the male-female difference in employment rates as well as the levels of male and female employment rates for selected CZs in which children grew up. The commuting zones displayed are the top ten and bottom ten commuting zones by male-female employment rate difference, among the 100 most populous commuting zones in the U.S. The sample consists of children in the 1980-82 birth cohorts whose parents stay in the same CZ throughout our sample window, 1996-2012, and who are in the bottom quintile of the national income distribution.

APPENDIX TABLE 2

Correlations between Gender Differences in Employment Rates for Children with Parents in Bottom Income Quintile and CZ Characteristics

		Regression Coefficients						Correlations		
		Full Sample (1)		Married Parents (2)		Single Parents (3)		Full Sample (4)	Married Parents (5)	Single Parents (6)
		Coeff	(s.e.)	Coeff	(s.e.)	Coeff	(s.e.)	Corr	Corr	Corr
Segregation and Poverty	Fraction Black Residents	-3.848	(0.388)	-1.884	(0.373)	-4.102	(0.404)	-0.699	-0.368	-0.700
	Racial Segregation	-2.845	(0.340)	-1.364	(0.357)	-3.068	(0.304)	-0.517	-0.267	-0.524
	Segregation of Poverty	-2.820	(0.350)	-1.759	(0.325)	-2.933	(0.336)	-0.512	-0.344	-0.501
	Fraction < 15 Mins to Work	2.820	(0.349)	1.709	(0.302)	2.982	(0.328)	0.512	0.334	0.509
Income Distribution	Mean Household Income	-2.259	(0.408)	-1.752	(0.273)	-2.233	(0.441)	-0.410	-0.343	-0.381
	Gini Coefficient (Bottom 99%)	-2.081	(0.590)	-1.403	(0.521)	-2.220	(0.540)	-0.378	-0.274	-0.379
	Top 1% Income Share	-1.452	(0.478)	-1.200	(0.522)	-1.513	(0.434)	-0.264	-0.235	-0.258
K-12 Education	Student-Teacher Ratio	0.405	(0.395)	0.035	(0.470)	0.441	(0.452)	0.073	0.007	0.076
	Test Scores (Income Adjusted)	0.696	(0.681)	0.748	(0.251)	0.578	(0.770)	0.126	0.146	0.099
	High School Dropout Rate (Income Adjusted)	-1.606	(0.421)	-0.519	(0.326)	-1.697	(0.432)	-0.281	-0.097	-0.276
Social Capital	Social Capital Index (Rupasingha and Goetz 2008)	0.843	(0.447)	0.413	(0.300)	0.977	(0.510)	0.153	0.081	0.167
	Fraction Religious	0.657	(0.525)	0.453	(0.451)	0.634	(0.551)	0.119	0.089	0.108
	Violent Crime Rate	-1.049	(0.818)	-0.793	(0.436)	-1.157	(0.945)	-0.188	-0.153	-0.195
Family Structure	Fraction of Children with Single Mothers	-2.714	(0.458)	-1.317	(0.506)	-2.932	(0.388)	-0.493	-0.258	-0.501
	Fraction of Adults Divorced	0.204	(0.479)	0.374	(0.396)	0.216	(0.521)	0.037	0.073	0.037
	Fraction of Adults Married	2.415	(0.493)	1.620	(0.461)	2.503	(0.449)	0.439	0.317	0.427
Tax	Local Tax Rate	-1.208	(0.521)	-0.823	(0.467)	-1.313	(0.480)	-0.219	-0.161	-0.224
	State EITC Exposure	0.117	(0.261)	-0.240	(0.288)	0.260	(0.304)	0.021	-0.047	0.044
	Tax Progressivity	0.356	(0.348)	-0.524	(0.215)	0.599	(0.366)	0.065	-0.103	0.102
College	Number of Colleges per Capita	1.157	(0.297)	0.817	(0.230)	1.190	(0.325)	0.214	0.163	0.207
	Mean College Tuition	-1.148	(0.449)	-0.768	(0.337)	-1.237	(0.474)	-0.212	-0.153	-0.215
	College Graduation Rate (Income Adjusted)	0.770	(0.353)	0.400	(0.292)	0.734	(0.373)	0.142	0.080	0.128
Local Labor Market	Fraction Working in Manufacturing	0.576	(0.434)	0.590	(0.369)	0.578	(0.454)	0.105	0.115	0.099
	Growth in Chinese Imports 1990-2000 (Autor and Dorn 2013)	0.755	(0.344)	0.075	(0.306)	1.061	(0.361)	0.137	0.015	0.181
	Teenage (14-16) LFP Rate	0.502	(0.486)	0.220	(0.310)	0.630	(0.564)	0.091	0.043	0.108
Migration	Migration Inflow Rate	-0.091	(0.461)	-0.095	(0.356)	-0.061	(0.521)	-0.017	-0.019	-0.010
	Migration Outflow Rate	-0.341	(0.470)	-0.534	(0.345)	-0.281	(0.502)	-0.062	-0.104	-0.048
	Fraction of Foreign Born Residents	-0.539	(0.505)	-0.907	(0.300)	-0.493	(0.592)	-0.098	-0.177	-0.084

Notes: This table presents correlations of gender differences in employment rates (males - females) for children with parents in the bottom income quintile with a set of CZ-level characteristics constructed in Chetty et al. (2014). We consider nine categories of covariates: race and segregation, properties of the income distribution, K-12 education, social capital, family structure, local tax policies, college education, labor market conditions, and migration rates. See Appendix G of Chetty et al. (2014) for definitions and sources of each variable. Column (1) presents the coefficient from a univariate regression of the gender difference in employment rates (male-female) on the covariate using the full sample. Column (2) restricts the sample to children of married parents, defined using the marital status of the parent when he or she first claims the child as a dependent, while Column (3) restricts the sample to children of single parents. Each covariate is standardized to have mean 0 and standard deviation of 1 in the full sample, so the regression coefficients can be interpreted as the effect of a 1 SD increase in the covariate on the male-female difference in employment rates (measured in percentage points). Standard errors, reported in parentheses alongside each estimate, are clustered by state. Columns (4)-(6) report the correlations corresponding to the regression coefficients in Columns (1)-(3) by dividing the regression coefficient by the standard deviation of the dependent variable in the relevant sample. All regressions and correlations are weighted by CZ population in the 2000 Census.

APPENDIX TABLE 3

Correlates of Gender Gaps in Permanent Residents' Outcomes vs. Causal Effects based on Movers

	Panel A: Male-Female Difference in Employment Rates				
	(1)	(2)	(3)	(4)	(5)
Segregation of Poverty	-2.823			-1.620	-1.948
	(0.350)			(0.323)	(0.197)
% Black		-3.848		-3.552	-3.335
		(0.388)		(0.536)	(0.563)
% Single Mothers			-2.716	0.404	0.526
			(0.459)	(0.666)	(0.413)
State FE					X
	Panel B: Male-Female Difference in Mean Income Rank				
	(1)	(2)	(3)	(4)	(5)
Segregation of Poverty	-3.030			-2.485	-2.231
	(0.231)			(0.246)	(0.186)
% Black		-2.389		-1.311	-1.820
		(0.442)		(0.410)	(0.449)
% Single Mothers			-1.890	-0.217	0.288
			(0.441)	(0.516)	(0.391)
State FE					X
	Panel C: Male-Female Difference in Causal Effects on Income Rank using Movers				
	(1)	(2)	(3)	(4)	(5)
Segregation of Poverty	-2.536			-2.464	-2.780
	(0.562)			(0.576)	(0.556)
% Black		-1.070		-0.452	1.389
		(0.560)		(0.777)	(1.326)
% Single Mothers			-0.647	0.350	-0.300
			(0.485)	(0.743)	(0.866)
State FE					X

Notes: This table presents further evidence on the relationship between three of the CZ-level characteristics in Appendix Table 2 (segregation of poverty, fraction black, and fraction single mothers) and gender gaps in adulthood. In Panel A, we regress gender differences (male - female) in employment rates at age 30 for children with parents in the bottom income quintile on these characteristics. Each covariate is standardized to have mean 0 and standard deviation of 1 in the full sample, so the regression coefficients can be interpreted as the effect of a 1 SD increase in the covariate on the male-female difference in employment rates (measured in percentage points). Columns 1-3 report estimates from univariate regressions, replicating specifications reported in Appendix Table 2. Column 4 reports estimates from a multivariable regression with three of the strongest predictors of gender differences. Column 5 replicates column 4 including state fixed effects. The sample in Panel A consists of children whose parents never moved across CZs, as in Figure 3. Panel B replicates Panel A, changing the dependent variable to the gender difference in the mean individual income rank at age 26 for children with parents at the 25th percentile of the national income distribution. Individual income is defined as the sum of wage earnings, unemployment income, disability income, and other (e.g. self employment) income from Schedule C (self-employment income is divided by 2 for married spouses). We translate this income measure into percentile ranks (scaled from 0-100) by ranking children relative to other children in the same birth cohort. Panel C replicates Panel B, replacing the dependent variable with the gender difference in the causal effects of each CZ on mean individual income rank at age 26 for children with parents at the 25th percentile. These causal effect estimates are taken from Chetty and Hendren's (2015) analysis of movers. Standard errors, reported in parentheses, are clustered by state. See Appendix B for further details on the specifications in this table.