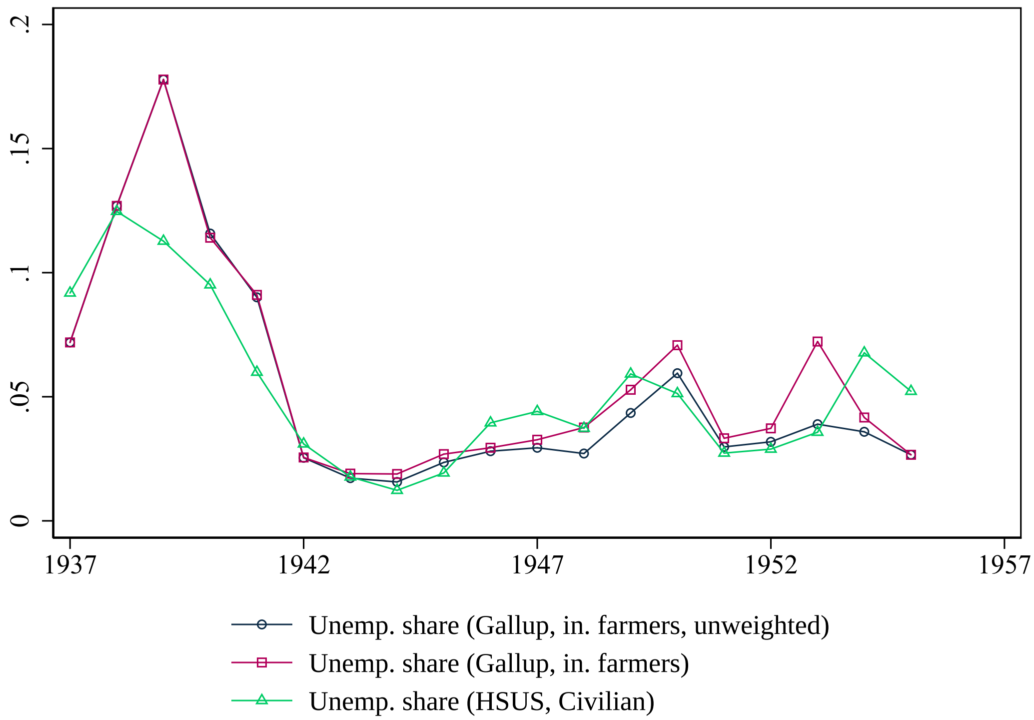
UNIONS AND INEQUALITY OVER THE TWENTIETH CENTURY:
APPENDIX MATERIALS

List of Appendix Figures

List of Appendix Tables

A. Supplementary Figures and Tables Noted in the Text

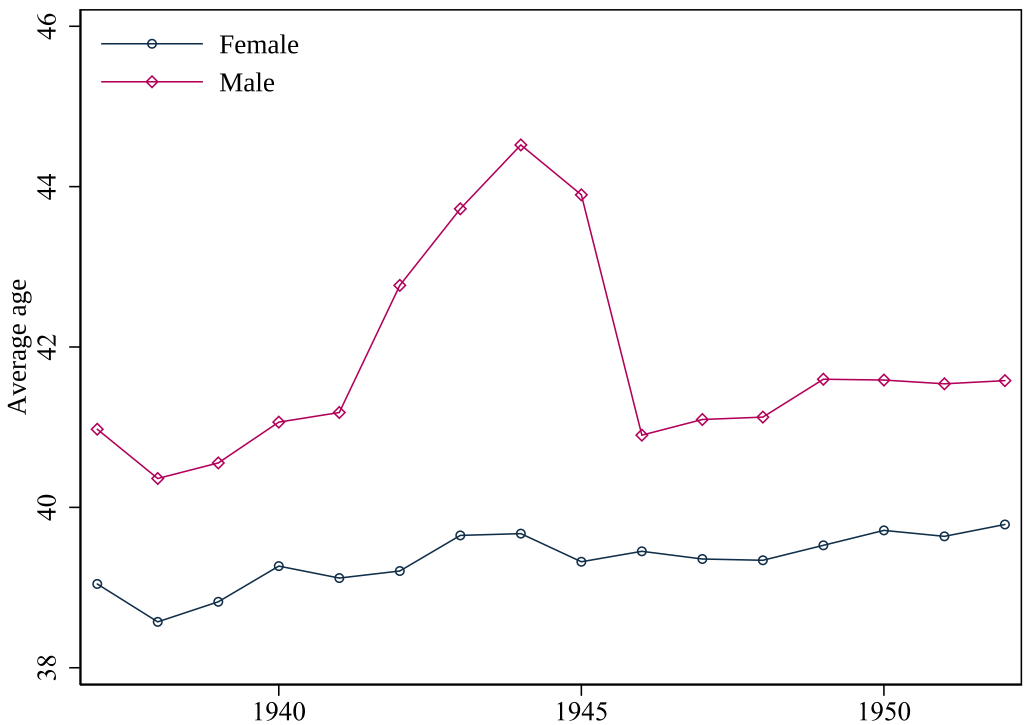
Appendix Figure A.1: Comparing unemployment rates in Gallup and the HSUS



Data sources: Gallup and Historical Statistics of the United States (HSUS)

Notes: Sample in Gallup includes farmers

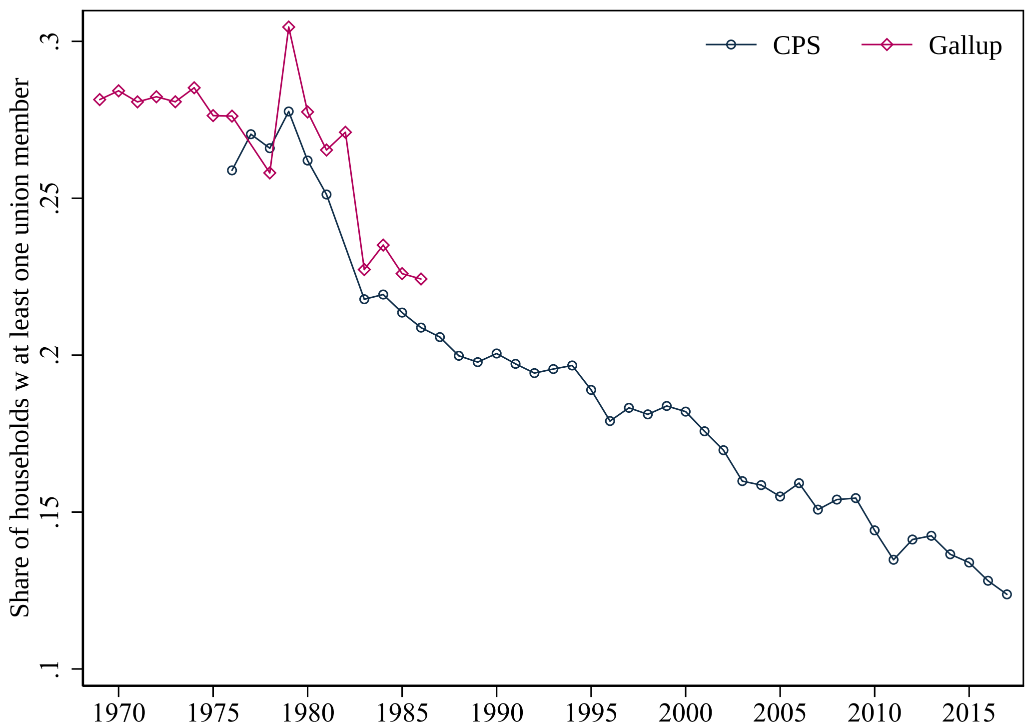
APPENDIX FIGURE A.2: AGE DISTRIBUTION IN GALLUP, BY GENDER, 1937-1952



Data sources: Gallup microdata.

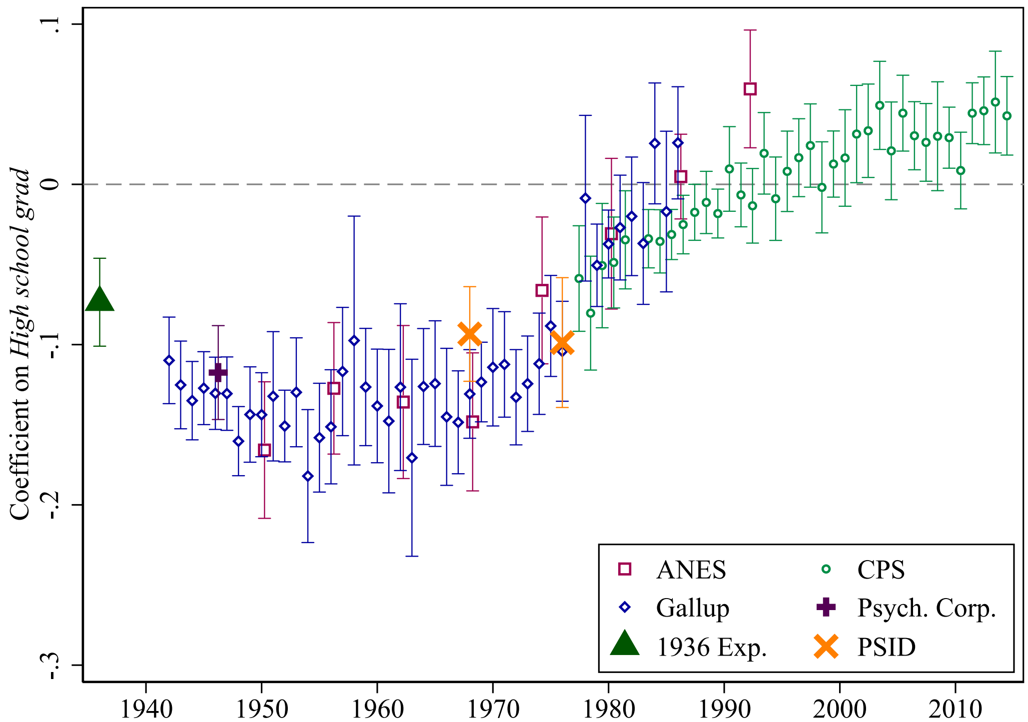
Notes: We show the large increase in the male average age in our Gallup dataset from 1942-1945 to demonstrate its ability to pick up high-frequency demographic changes (such as the deployment of young men overseas during World War II).

Appendix Figure A.3: Comparing household union density in Gallup and CPS, 1970–present



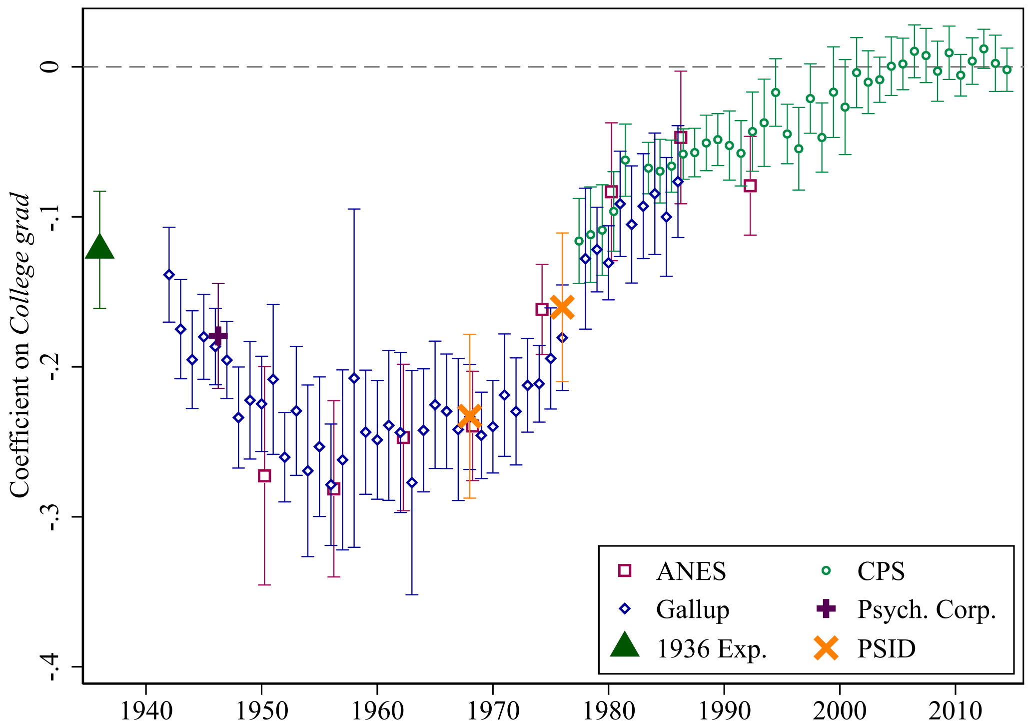
Data sources: Gallup and Current Population Survey

APPENDIX FIGURE A.4: SELECTION OF UNION HOUSEHOLDS BY HIGH-SCHOOL GRADUATION



Data sources: Gallup, 1937–1986. CPS, 1978–2016; BLS Expenditure Survey, 1936; ANES, 1952–1996, U.S. Psych. Corporation, 1946.
Notes: For each data source, we estimate, separately by year, household union status on a *High School Grad* dummy variable, state *s* and survey-date *t* fixed effects, age and its square, and gender. We plot in this graph the coefficients on *High School Grad* from each of these estimations. For the ANES, because the samples are smaller, we group surveys into six-year bins. Standard errors are clustered by state.

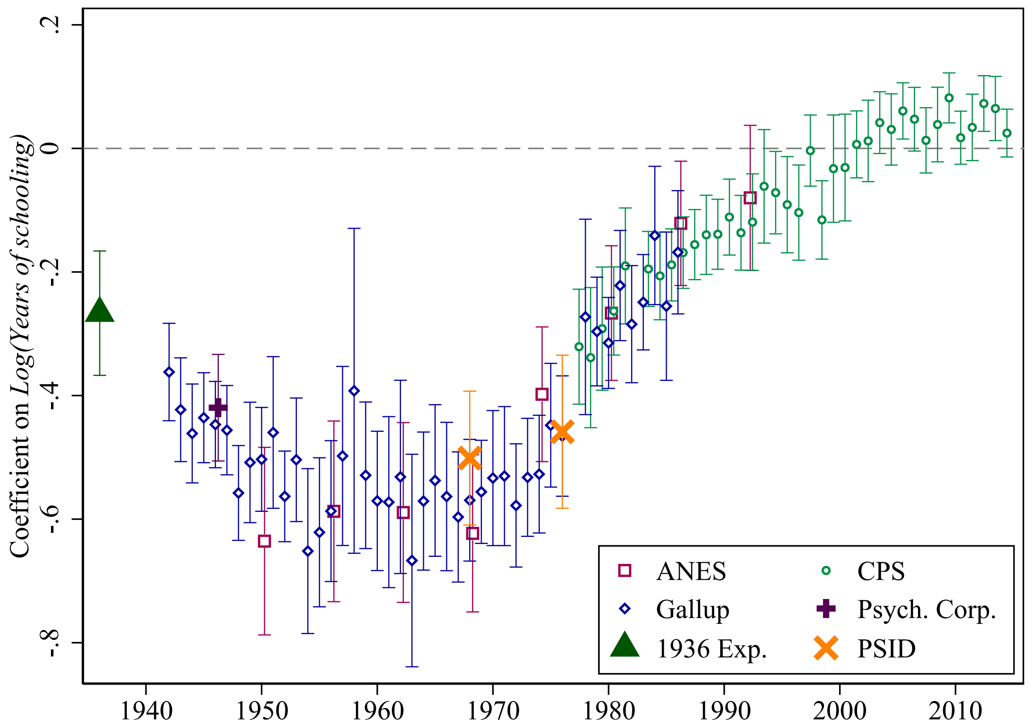
APPENDIX FIGURE A.5: SELECTION OF UNION HOUSEHOLDS BY COLLEGE GRADUATION



Data sources: Gallup, 1937–1986. CPS, 1978–2016; BLS Expenditure Survey, 1936; ANES, 1952–1996, U.S. Psych. Corporation, 1946.

Notes: For each data source, we estimate, separately by year, household union status on a *College Grad* dummy variable, state s and survey-date t fixed effects, age and its square, and gender. We plot in this graph the coefficients on *College Grad* from each of these estimations. For the ANES, because the samples are smaller, we group surveys into six-year bins. Standard errors are clustered by state.

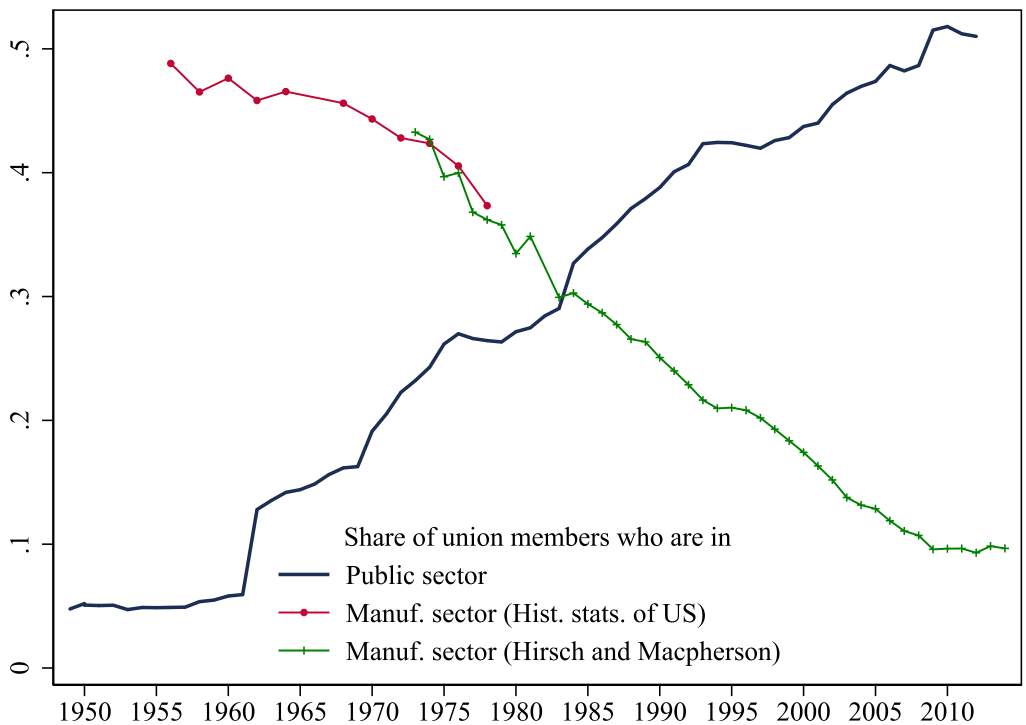
Appendix Figure A.6: Selection of union households by log years schooling



Data sources: Gallup, 1937–1986. CPS, 1978–2016; BLS Expenditure Survey, 1936; ANES, 1952–1996, U.S. Psych. Corporation, 1946.

Notes: For each data source, we estimate, separately by year, household union status on *Log Years Education*, state *s* and survey-date *t* fixed effects, age and its square, and gender. We plot in this graph the coefficients on *Log Years Education* from each of these estimations. For the ANES, because the samples are smaller, we group surveys into six-year bins. Standard errors are clustered by state.

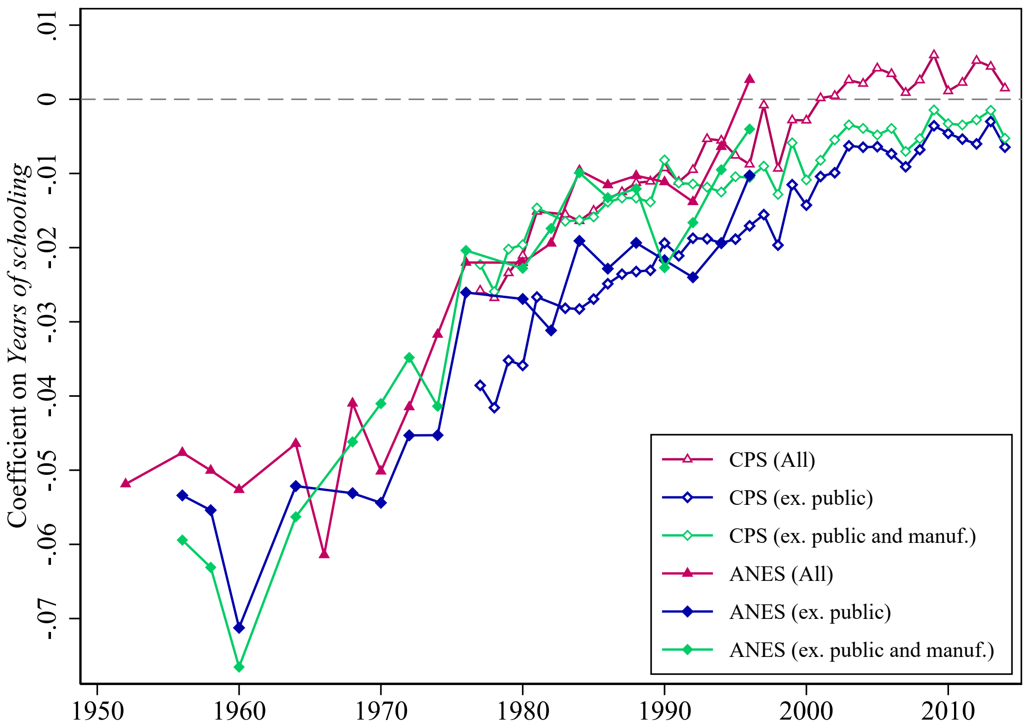
APPENDIX FIGURE A.7: SHARE OF UNION MEMBERS IN PUBLIC SECTOR AND MANUFACTURING



Data sources: For the public-sector series, we thank John Schmitt at EPI. The early manufacturing series is from the Historical Statistics of the United States. The later manufacturing series is from the CPS, calculated by Hirsch and Macpherson and posted on .

Notes: These series refer to union *members*, not households, as in much of the paper.

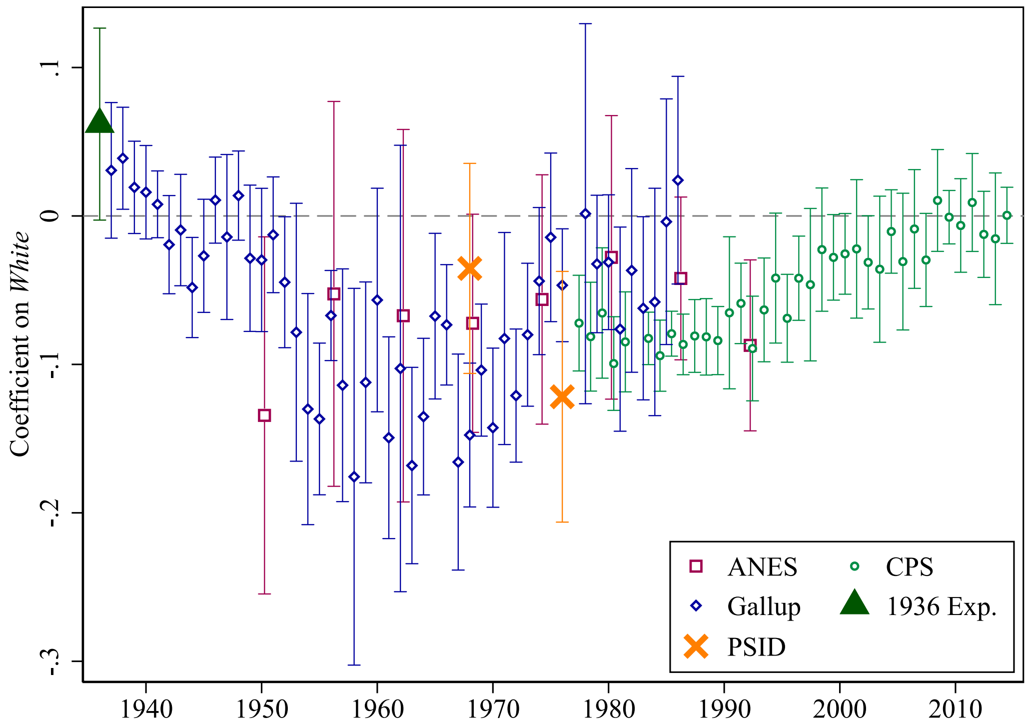
APPENDIX FIGURE A.8: SELECTION OF UNION HOUSEHOLDS BY EDUCATION IN THE ANES AND CPS (DROPPING HOUSEHOLDS WITH A PUBLIC- OR MANUFACTURING-SECTOR WORKER)



Data sources: CPS, 1978–2016; ANES, 1952–1996.

Notes: For each data source, we estimate, separately by year, household union status on a *Years of education* variable, state s and survey-date t fixed effects, age and its square, and gender. We plot in this graph the coefficients on *Years of education* from each of these estimations. For the ANES, because the samples are smaller, we group surveys into six-year bins. Note that we only include ANES and CPS in this graph, because other data sources do not allow us to identify industrial sectors of workers in the household.

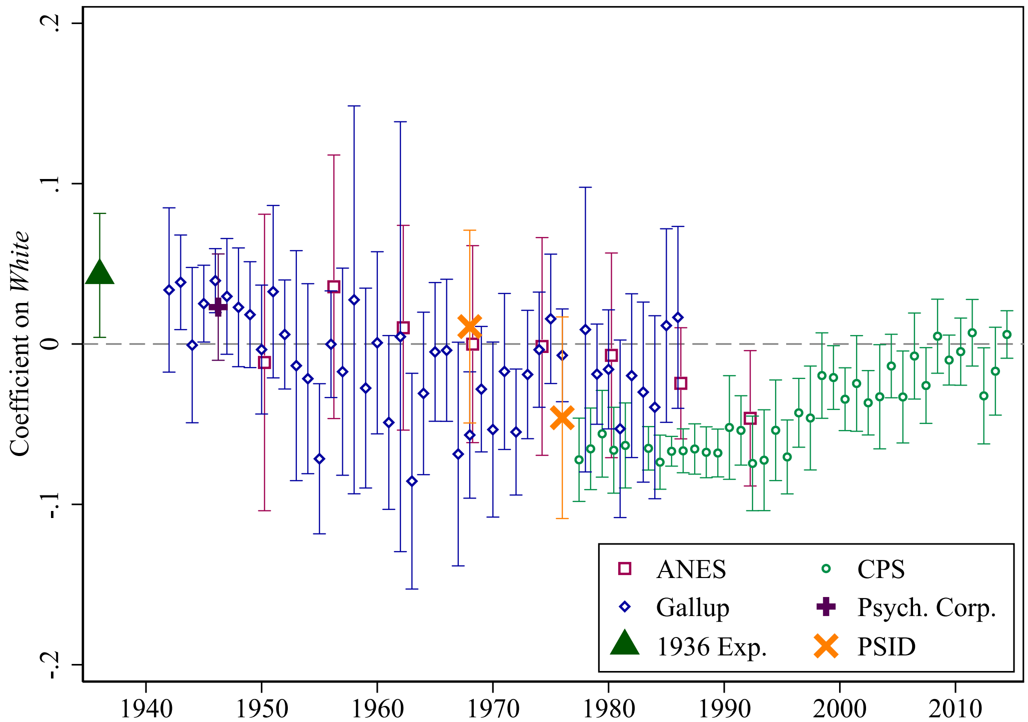
APPENDIX FIGURE A.9: SELECTION OF UNION HOUSEHOLDS BY RACE (DROPPING SOUTHERN STATES)



Data sources: Gallup data, 1937–1986; CPS, 1978–2016; BLS Expenditure Survey, 1936; ANES, 1952–1996. See Section II.B for a description of each data source.

Notes: For each data source, we estimate, separately by year, household union status on a *White* dummy variable, state s and survey-date t fixed effects, age and its square, and gender. We plot in this graph the coefficients on $White$ from each of these estimations. For the ANES, because the samples are smaller, we group surveys into six-year bins. Note that we cannot use the U.S. Psychological Corporation survey in this figure because, while it has state identifiers (thus we can thus control for state fixed effects), the codebook does *not* provide the state names that correspond to the codes (so we cannot drop the South). Confidence intervals are based on standard errors clustered by state.

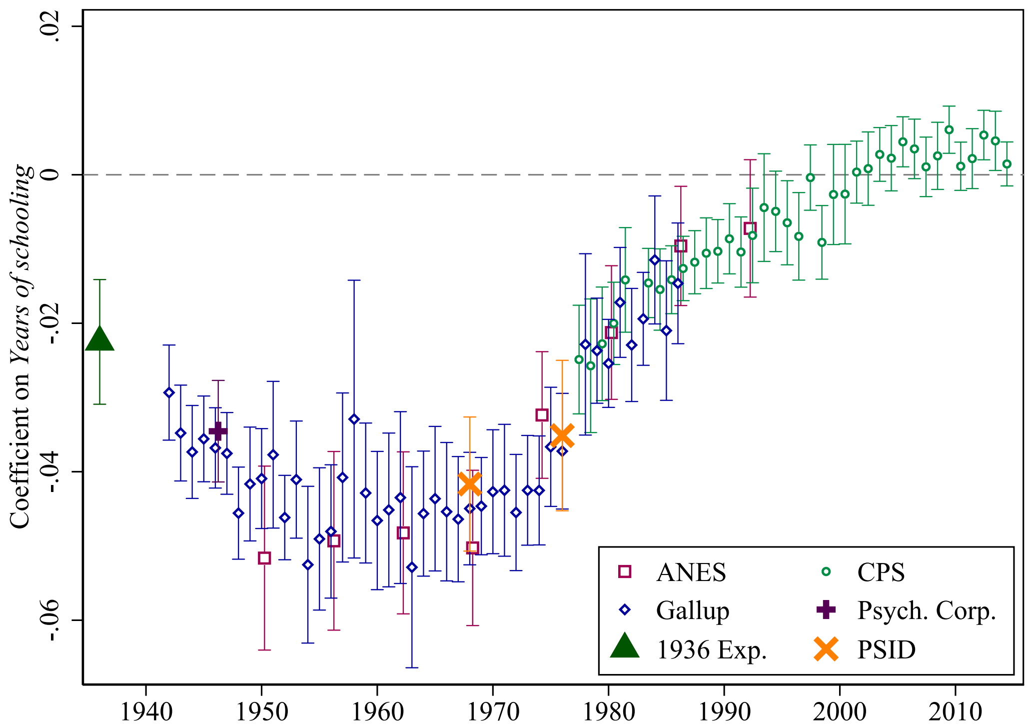
Appendix Figure A.10: Selection of union households by race (conditional on education)



Data sources: Gallup data, 1937–1986; CPS, 1978–2016; BLS Expenditure Survey, 1936; ANES, 1952–1996. See Section II.B for a description of each data source.

Notes: For each data source, we estimate, separately by year, household union status on a *White* dummy variable, state s and survey-date t fixed effects, age and its square, gender, *and* years of schooling. Otherwise, the analysis is identical to that in Figure IV. Note that conditioning on education means we lose data from 1937-1941, as the Gallup education question is not included in these surveys. Confidence intervals are based on standard errors clustered by state.

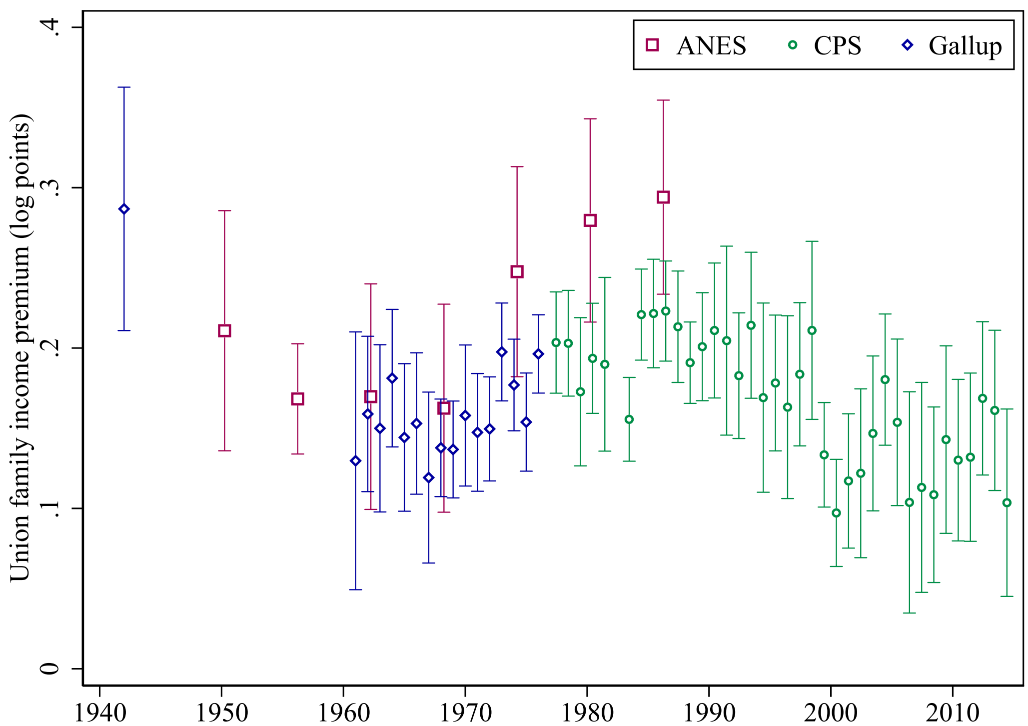
Appendix Figure A.11: Selection of union households by education (conditional on race)



Data sources: Gallup data, 1937–1986; CPS, 1978–2016; BLS Expenditure Survey, 1936; ANES, 1952–1996. See Section II.B for a description of each data source.

Notes: For each data source, we estimate, separately by year, household union status on years of schooling, state s and survey-date t fixed effects, age and its square, gender, *and* a *White* dummy variable. Otherwise, the analysis is identical to that in Figure III. Note that conditioning on education means we lose data from 1937-1941, as the Gallup education question is not included in these surveys. Confidence intervals are based on standard errors clustered by state.

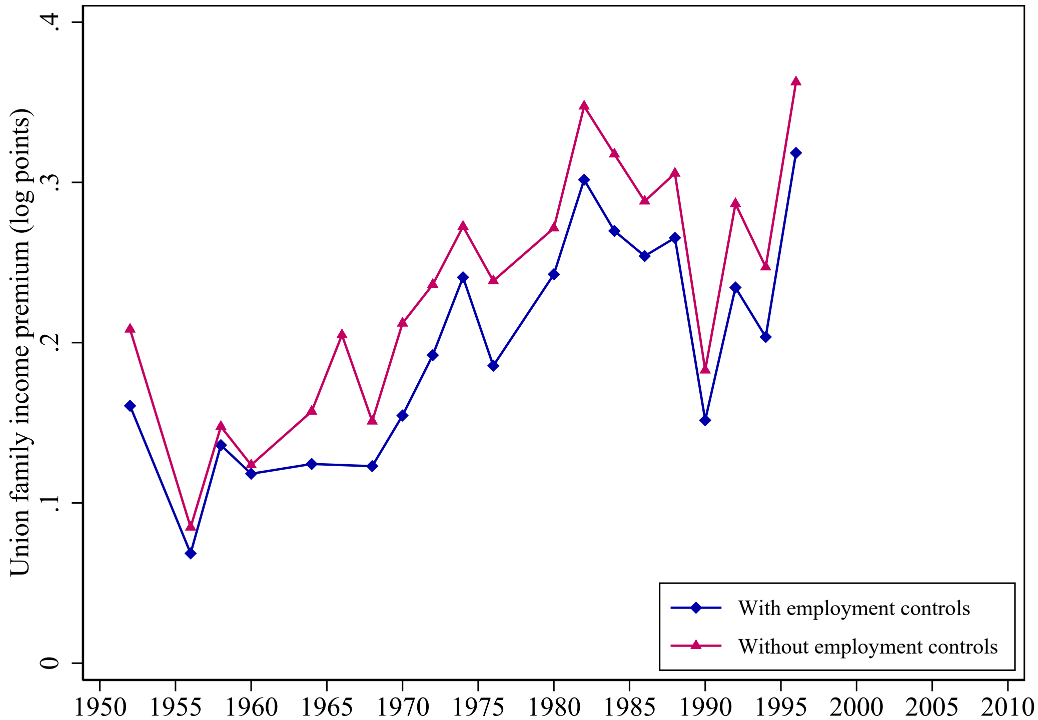
Appendix Figure A.12: Estimates of the union family income premium (including occupation controls when available)



Data sources: Gallup data, 1942, 1961–1976; CPS, 1978–2016; BLS Expenditure Survey, 1936; ANES, 1952–1996, U.S. Psych. Corporation, 1946. See Section II.B for a description of each data source. See Appendix C for details on family income variable construction.

Notes: Each plotted point comes from estimating equation (2), which regresses log family income on household union status and controls for age, gender, race, state and survey-date fixed effects and (in most cases) fixed effects for the occupation of the head. We cannot perfectly match occupation categories across regressions, which is why we relegate this graph to the appendix. For the ANES, because the samples are smaller, we group surveys into six-year bins. The plotted confidence intervals are based on standard errors clustered by state.

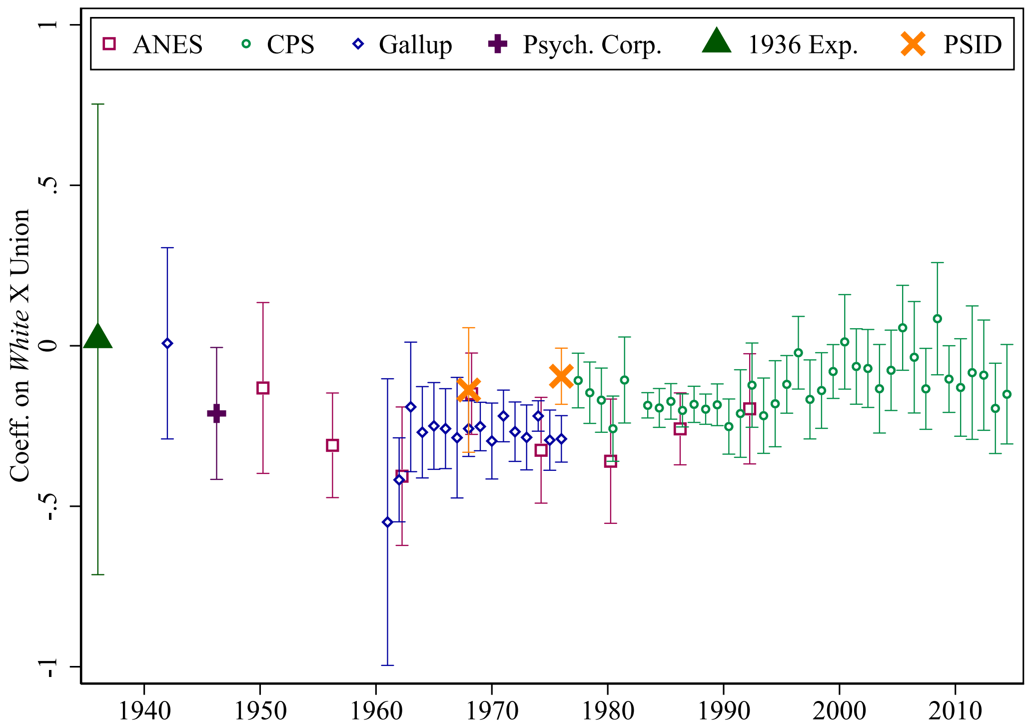
Appendix Figure A.13: Estimates of the union family income premium from ANES (with and without employment status controls)



Data sources: See Section II.B for a description of ANES data.

Notes: Each plotted point comes from estimating equation (2), which regresses log family income on the household union dummy and controls for age, gender, race, state and survey-date fixed effects. In addition, the first series includes an indicator for the household head being employed and a separate indicator for the respondent being employed. See Section IV.A for more detail.

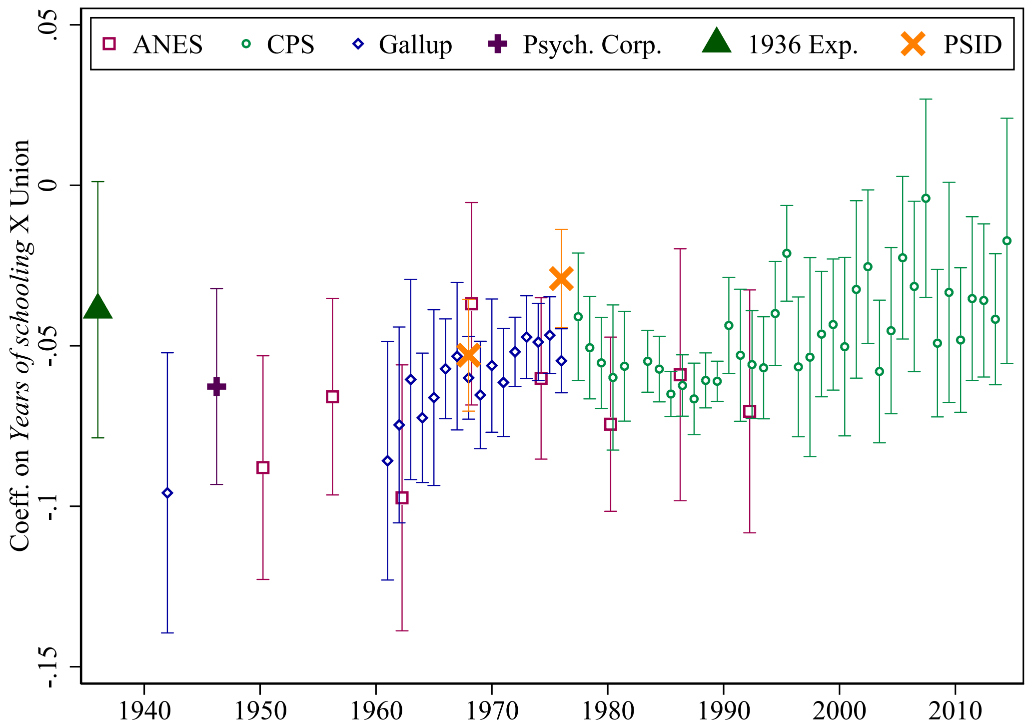
APPENDIX FIGURE A.14: UNION FAMILY INCOME PREMIUM BY RACE (CONDITIONAL ON *Yrs. schooling*×*Union*)



Data sources: Gallup data, 1942, 1961–1976; CPS, 1978–2016; BLS Expenditure Survey, 1936; ANES, 1952–1996; U.S. Psych. Corporation, 1946. See Section II.B for a description of each data source. See Appendix C for details on family income variable construction.

Notes: In this figure we estimate the differential union premium paid to white households, conditional on any differential premium by education of the respondent. This figure is identical to the union-premium-by-race analysis in Figure VII, except that we add *Years of schooling*$_h^R$ $\times Union_h$ to each estimating equation, where *Years of schooling*$_h^R$ is the years of schooling for the respondent from household h, and $Union_h$ is our usual household union measure. The plotted confidence intervals are based on standard errors clustered by state.

APPENDIX FIGURE A.15: UNION FAMILY INCOME PREMIUM BY EDUCATION (CONDITIONAL ON $White \times Union$)



Data sources: Gallup data, 1942, 1961–1976; CPS, 1978–2016; BLS Expenditure Survey, 1936; ANES, 1952–1996; U.S. Psych. Corporation, 1946. See Section II.B for a description of each dat

Notes: In this figure we estimate the differential union premium paid to more-educated households, conditional on any differential premium by race of the respondent. This figure is identical to the union-premium-by-education analysis in Figure VI, except that we add $White^R_h \times Union_h$ to each estimating equation, where $White^R_h$ is a dummy for the respondent from household h, and $Union_h$ is our usual household union measure. The plotted confidence intervals are based on standard errors clustered by state.

APPENDIX TABLE A.1: ESTIMATING FAMILY UNION INCOME PREMIUM AND REPORTING COEFFICIENTS ON ADDITIONAL COVARIATES, BY DATA SOURCE AND TIME PERIOD

	Dep't var: Logged family income					
	(1)	(2)	(3)	(4)	(5)	(6)
Union household	0.116***	0.259***	0.196***	0.160***	0.129***	0.246***
	[0.0239]	[0.0332]	[0.0337]	[0.0151]	[0.0212]	[0.0179]
Years of education		0.175***	0.146***	0.115***	0.122***	0.122***
		[0.00672]	[0.00681]	[0.00371]	[0.00624]	[0.00653]
Years of educ., household head	0.125***					
	[0.00770]					
White dummy	0.880***	0.461***	0.410***	0.443***	0.517***	0.326***
	[0.0477]	[0.0883]	[0.0317]	[0.0299]	[0.0656]	[0.0408]
Female		-0.109***	-0.203***	-0.0903***	-0.121***	-0.126***
		[0.0312]	[0.0195]	[0.00386]	[0.0200]	[0.0154]
Household head is female	0.0955***					
	[0.0261]					
Age	0.0749***	0.0521***	0.0682***	0.0698***	0.0640***	0.0740***
	[0.00824]	[0.0134]	[0.00515]	[0.00227]	[0.00407]	[0.00407]
Age squared, divided by 1000	-0.842***	-0.614***	-0.884***	-0.817***	-0.744***	-0.753***
	[0.0999]	[0.165]	[0.0625]	[0.0261]	[0.0518]	[0.0454]
Data source	Exp. survey	Gallup	U.S. Psych.	Gallup	ANES	ANES
Year(s) in sample	1936	1942-1942	1946	1961-1975	1952-1970	1972-1990
Observations	4976	2538	5415	171973	9212	12925

Data sources: Gallup data, 1942, 1961–1975; CPS, 1978–2016; BLS Expenditure Survey, 1936; ANES, 1952–1996, U.S. Psych. Corporation, 1946. See Section II.B for a description of each data source. See Appendix C for details on family income variable construction.

Notes: All regressions include state fixed effects and survey date fixed effects. We control for number of employed individuals in the household, except in the Gallup and U.S. Psych. data where this control is not available. Otherwise, all other samples include ages 21–64. Since the goal of the table is to show the coefficients from regressions run on the datasets least likely to be familiar to readers, we do not include the CPS. Standard errors in brackets, clustered by state. $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$

Appendix Table A.2: Estimating family union income premium using 1956-1960 ANES panel

	Dept. variable: Log family income				
	(1)	(2)	(3)	(4)	(5)
Union household	0.134***	0.114***	0.103*	0.0635	0.0692
	[0.0309]	[0.0283]	[0.0559]	[0.0809]	[0.0544]
Union household x Low-educ. respondent				0.0486	
				[0.106]	
Union household x Non-white respondent					0.249
					[0.209]
Added controls?	No	Yes	No	No	No
Respondent FE?	No	No	Yes	Yes	Yes
Observations	3303	3303	3303	3303	3303

Notes: All regressions include year fixed effects and a quadratic in age. Sample restricted to ages 18 to 65. Controls include race, sex, education and occupation fixed effects. "Low education" is high school degree or less. Standard errors in brackets, clustered by individual. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

APPENDIX TABLE A.3: HETEROGENEITY OF THE UNION PREMIUM

	Dept. variable: Log family income						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Union household	0.194***	0.202***	0.186***	0.191***	0.185***	0.170***	0.148***
	[0.00404]	[0.0233]	[0.0187]	[0.00749]	[0.00389]	[0.00354]	[0.00852]
Union x National unemp. rate		-0.117					
		[0.352]					
Union x National union density			0.0382				
			[0.0813]				
Union x Δ Ln(CPI)				0.0775			
				[0.148]			
Union x South					0.0562***	-0.0118*	-0.00312
					[0.00417]	[0.00641]	[0.00717]
Union x State ever RTW						0.0832***	0.0959***
						[0.00751]	[0.0100]
Union x State currently RTW							0.0214**
							[0.00879]
Observations	1,153,757	1,153,757	1,148,781	1,153,757	1,153,757	1,153,757	1,153,757

Notes: All regressions include state and survey-date fixed effects and number of employed individuals in household whenever available. *State ever RTW* is a state-level dummy indicating that a state passed a right-to-work law at some point during our sample period. *State currently RTW* is coded as one for any year after a state passes its first RTW law. Standard errors in brackets, clustered by year. $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$

APPENDIX TABLE A.4: PAID VACATION AS A FUNCTION OF UNION STATUS (GALLUP, 1949)

	Dep't var: Do you (or husband) get paid vacation?				
	(1)	(2)	(3)	(4)	(5)
Union household	0.223***	0.188***	0.323**	0.288	0.130***
	[0.0319]	[0.0292]	[0.129]	[0.222]	[0.0291]
White x Union household			-0.144		
			[0.130]		
Years educ. x Union household				-0.00904	
				[0.0194]	
Low-skill labor x Union					0.137***
					[0.0487]
Dept. var. mean	0.517	0.524	0.524	0.524	0.524
State FE?	Yes	Yes	Yes	Yes	Yes
Demographic controls?	Yes	Yes	Yes	Yes	Yes
Occupation FE?	No	Yes	Yes	Yes	Yes
Observations	1969	1911	1911	1911	1911

Notes: Data from Gallup, May 1949. Demographic controls include respondent's age and square, education (four fixed effects), gender, and race. When occupation controls are added, they refer to the head of the household. Low-skill occupation dummy denotes "unskilled and semi-skilled labor." Standard errors in brackets, clustered by state. $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$

APPENDIX TABLE A.5: EASE OF FINDING A JOB AS GOOD AS THE ONE YOU HAVE

	Dept. var: Would be easy to find a job as good as current one				
	Gallup (1939)			GSS (1977-2018)	
	(1)	(2)	(3)	(4)	(5)
Union household	-0.124***	-0.121***	-0.0943***	-0.0863***	-0.0766***
	[0.0275]	[0.0272]	[0.0310]	[0.00953]	[0.00960]
Mean, dept. var.	0.497	0.497	0.497	0.254	0.255
State FE	Yes	Yes	Yes	Reg.	Reg.
Demogr. controls	No	Yes	Yes	No	Yes
Educ. controls	No	No	Yes	No	Yes
Occup. controls	No	No	Yes	No	No
Observations	1978	1978	1978	12039	12019

Notes: The Gallup question reads: "If you lost your present job (business, farm), how hard do you think it would be for you to get another job (business, farm) just as good?" We code "impossible" and "quite hard" ("fairly hard" and "easy") as zero (one). Demographic controls include respondent's age and its square, education (four fixed effects), gender and race. Occupation controls refer to household head; low-skill occupation to "unskilled, semi-skilled labor." The GSS question reads: "About how easy would it be for you to find a job with another employer with approximately the same income and fringe benefits you now have?" We code "very easy" ("somewhat easy" and "not easy at all") as one (zero). All GSS regressions include year fixed effects. Demographic and education controls are as in Gallup. Standard errors are in brackets and clustered by state (region). $^{*} p < 0.1$, $^{**} p < 0.05$, $^{***} p < 0.01$

Appendix Table A.6: Covariance between union density and skill shares

	Annual regressions				State-year panel regs.	
	(1)	(2)	(3)	(4)	(5)	(6)
Skill share measure	-0.0828***	-0.0938***	-0.253***	-0.0742	-0.0208**	-0.00312
	[0.0201]	[0.0331]	[0.0453]	[0.0446]	[0.00917]	[0.0131]
Time polynomial?	None	Cubic	Quad.	Quartic	None	None
State FE?	N/A	N/A	N/A	N/A	No	Yes
Year FE?	No	No	No	No	No	Yes
Observations	56	56	56	56	1968	1968

Notes: This table shows how our union density variable and the skill-share measure (both used extensively in Section V) co-vary at different levels of aggregation as well as conditionally and unconditionally. See Section II for more information on the construction of the density variable. We follow Autor, Katz, and Kearney (2008) and Goldin and Katz (2008) in constructing skill-shares measures (see Appendix C for more information). $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

APPENDIX TABLE A.7: AGGREGATE COLL. PREMIUM, 90/10, 90/50 RATIOS AS FUNCTIONS OF DENSITY

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Panel A: College premium			
Skill share	-0.493***	-0.555***	-0.495***	-0.586***	-0.572***	-0.505***	-0.694***
	(0.089)	(0.089)	(0.075)	(0.100)	(0.125)	(0.121)	(0.115)
Gallup union density		-0.778**					
		(0.358)					
BLS union density			-1.162**				
			(0.435)				
Density (avg. of Gallup, BLS)				-1.090**	-1.115	-1.972***	-0.989*
				(0.477)	(0.693)	(0.449)	(0.499)
				Panel B: Log 90/10 ratio			
Skill share	0.028	-0.083	0.025	-0.158	0.179	0.245**	0.104
	(0.115)	(0.085)	(0.099)	(0.099)	(0.119)	(0.109)	(0.124)
Gallup union density		-1.407***					
		(0.379)					
BLS union density			-1.971***				
			(0.319)				
Density (avg. of Gallup, BLS)				-2.189***	-1.936***	-2.783***	-1.859***
				(0.415)	(0.688)	(0.451)	(0.547)
				Panel C: Log 90/50 ratio			
Skill share	-0.291***	-0.286***	-0.292***	-0.329***	-0.232***	-0.138	-0.229***
	(0.084)	(0.092)	(0.078)	(0.088)	(0.067)	(0.103)	(0.082)
Gallup union density		0.061					
		(0.279)					
BLS union density			-0.517*				
			(0.279)				
Density (avg. of Gallup, BLS)				-0.450	-0.489	-1.683***	-0.492
				(0.332)	(0.366)	(0.359)	(0.378)
				Panel D: Log 10/50 ratio			
Skill share	-0.319**	-0.204**	-0.317**	-0.172	-0.411***	-0.384***	-0.334***
	(0.136)	(0.099)	(0.139)	(0.125)	(0.121)	(0.123)	(0.116)
Gallup union density		1.468***					
		(0.307)					
BLS union density			1.454***				
			(0.401)				
Density (avg. of Gallup, BLS)				1.739***	1.447**	1.099**	1.368**
				(0.420)	(0.629)	(0.450)	(0.545)
Controls?	No	No	No	No	Yes	Yes	Yes
Time Polynomial	Cubic	Cubic	Cubic	Cubic	Cubic	Quadratic	Quartic
Observations	54	54	54	54	54	54	54

Notes: This table shows variants of the specifications estimated in cols. 1 and 2 (Panel A), cols. 3 and 4 (Panel B), cols. 5 and 6 (Panel C), and cols. 7 and 8 (Panel D) of Table II. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

APPENDIX TABLE A.8: AGGREGATE GINI, TOP-TEN, LABOR SHARE OF INCOME AS FUNCTIONS OF DENSITY

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Panel A: Gini coefficient			
Skill share	-0.006	-0.001	0.009	-0.006	-0.001	0.079***	-0.008
(interpolated)	(0.013)	(0.014)	(0.014)	(0.014)	(0.020)	(0.019)	(0.019)
Educ. share ratio	0.014***	0.009**	0.009**	0.006**	0.007**	0.020***	0.008***
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)
Gallup union density		-0.115***					
		(0.032)					
BLS union density			-0.120***				
			(0.035)				
Density (avg. of				-0.168***	-0.160***	-0.195***	-0.163***
Gallup, BLS)				(0.039)	(0.042)	(0.049)	(0.041)
Observations	65	65	65	65	65	65	65
				Panel B: Top-ten income share			
Skill share	-11.779***	-12.543***	-7.654**	-13.385***	-15.780***	-13.538***	-13.258**
(interpolated)	(3.644)	(3.496)	(2.967)	(3.190)	(5.483)	(4.798)	(5.743)
Educ. share ratio	2.196	0.779	-1.176	-1.430	-1.094	-1.075	0.359
	(2.296)	(2.329)	(1.371)	(1.443)	(1.588)	(1.594)	(1.779)
Gallup union density		-26.253**					
		(11.193)					
BLS union density			-66.186***				
			(13.841)				
Density (avg. of				-69.165***	-61.972***	-66.390***	-61.092***
Gallup, BLS)				(18.103)	(18.080)	(17.245)	(16.476)
Observations	75	75	75	75	75	75	75
				Panel C: Labor share of income			
Skill share	-7.399**	-6.715**	-10.687***	-6.398**	-5.247	7.881	-4.408
(interpolated)	(3.514)	(3.047)	(2.513)	(2.799)	(3.446)	(4.740)	(3.721)
Educ. share ratio	-3.241**	-1.973	-0.554	-0.980	-1.503	-1.388	-1.020
	(1.457)	(1.375)	(0.691)	(0.988)	(1.225)	(1.289)	(1.364)
Gallup union density		23.490***					
		(8.522)					
BLS union density			52.750***				
			(7.398)				
Density (avg. of				43.123***	39.434***	13.560	39.727***
Gallup, BLS)				(10.710)	(13.214)	(11.914)	(13.390)
Observations	75	75	75	75	75	75	75
Controls?	No	No	No	No	Yes	Yes	Yes
Time Polynomial	Cubic	Cubic	Cubic	Cubic	Cubic	Quadratic	Quartic

Notes: This table shows variants of the specifications estimated in cols. 9 and 10 (Panel A), cols. 11 and 12 (Panel B), and cols. 13 and 14 (Panel C) of Table II. $^{*} p < 0.1, ^{**} p < 0.05, ^{***} p < 0.01$

Appendix Table A.9: Skill premium, percentile ratios, and Gini coefficient as a function of state-year union density

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: College premium						
Household union share	-0.122**	-0.187	-0.169	-0.205	-0.214*	-0.195	-0.182
	[0.052]	[0.136]	[0.141]	[0.130]	[0.128]	[0.124]	[0.113]
	Panel B: Log 90/10 ratio						
Household union share	-0.227**	-0.345**	-0.291*	-0.293*	-0.307**	-0.251*	-0.197
	[0.098]	[0.168]	[0.160]	[0.155]	[0.149]	[0.136]	[0.125]
	Panel C: Log 90/50 ratio						
Household union share	-0.091*	-0.140	-0.118	-0.112	-0.122	-0.122	-0.097
	[0.048]	[0.088]	[0.088]	[0.088]	[0.086]	[0.079]	[0.084]
	Panel D: Log 10/50 ratio						
Household union share	0.135**	0.205*	0.173	0.181*	0.184*	0.129	0.100
	[0.063]	[0.113]	[0.106]	[0.104]	[0.102]	[0.105]	[0.100]
	Panel E: Gini coefficient						
Household union share	-0.035**	-0.055**	-0.041	-0.052**	-0.054**	-0.046**	-0.050**
	[0.016]	[0.027]	[0.027]	[0.023]	[0.022]	[0.022]	[0.025]
Observations	1,960	1,960	1,960	1,960	1,960	1,960	1,960
Min Year	1940	1940	1940	1940	1940	1940	1940
Max. Year	2014	2014	2014	2014	2014	2014	2014
SouthXyear FE	Yes	Yes	Yes	Yes	Yes	No	Yes
Split-Sample IV	No	Yes	Yes	Yes	Yes	Yes	Yes
Income covars.	No	No	Yes	Yes	Yes	Yes	Yes
Industry Shares	No	No	No	Yes	Yes	Yes	Yes
Policy covars.	No	No	No	No	Yes	Yes	Yes
RegionXyear FE	No	No	No	No	No	Yes	No
State-spec. quad.	No	No	No	No	No	No	Yes

Data sources: See notes to Table III.

Notes: IV estimates are from split-sample-IV regressions (see Section V.C for estimating equations). All regressions include state and year fixed effects; $South \times Year$ fixed effects; and state-year education controls (both from Gallup and CPS at the annual level, and interpolated from the IPUMS Census at the decade level). "Industry shares" controls for state-year share of employment in all one-digit industry categories. "State-spec. quad." indicates that state-specific quadratic time trends are included. "Policy covars." indicate that state-year minimum wage and a "policy liberalism" index (from Caughey and Warshaw, 2016) are included. Standard errors are clustered at the state level. $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$

APPENDIX TABLE A.10: STATE YEAR TOP-TEN INCOME SHARE, LABOR SHARE AS A FUNCTION OF UNION DENSITY

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Panel A: Top 10p. Income			
Household union share	-2.739**	-4.192**	-4.340**	-3.732**	-3.479**	-3.248**	-2.403**
	[1.125]	[1.917]	[1.704]	[1.788]	[1.693]	[1.614]	[1.075]
				Panel B: Labor share			
Household union share	3.656***	5.567***	6.018***	4.037**	3.972**	3.442*	1.090
	[1.198]	[1.870]	[2.010]	[1.906]	[1.789]	[1.857]	[1.029]
Observations	3,537	3,537	3,537	3,537	3,537	3,537	3,537
Min Year	1937	1937	1937	1937	1937	1937	1937
Max. Year	2014	2014	2014	2014	2014	2014	2014
SouthXyear FE	Yes	Yes	Yes	Yes	Yes	No	Yes
Split-Sample IV	No	Yes	Yes	Yes	Yes	Yes	Yes
Income covars.	No	No	Yes	Yes	Yes	Yes	Yes
Industry Shares	No	No	No	Yes	Yes	Yes	Yes
Policy covars.	No	No	No	No	Yes	Yes	Yes
RegionXyear FE	No	No	No	No	No	Yes	No
State-spec. quad.	No	No	No	No	No	No	Yes

Data sources: See notes to Table III.

Notes: IV estimates are from split-sample-IV regressions (see Section V.C for estimating equations). All regressions include state and year fixed effects; $South \times Year$ fixed effects; and state-year education controls (both from Gallup and CPS at the annual level, and interpolated from the IPUMS Census at the decade level). "Industry shares" controls for state-year share of employment in all one-digit industry categories. "State-spec. quad." indicates that state-specific quadratic time trends are included. "Policy covars." indicate that state-year minimum wage and a "policy liberalism" index (from Caughey and Warshaw, 2016) are included. Standard errors are clustered at the state level. $^{*}p<0.1,^{**}p<0.05,^{***}p<0.01$

Appendix Table A.11: Log state-year income per capita as a function of union density

	Dep't var: Log state-year income per capita					
	(1)	(2)	(3)	(4)	(5)	(6)
Household union share	0.112***	0.170***	0.138**	0.141**	0.032	-0.010
	[0.034]	[0.059]	[0.066]	[0.064]	[0.059]	[0.038]
Observations	3,537	3,537	3,537	3,537	3,537	3,537
Min Year	1937	1937	1937	1937	1937	1937
Max. Year	2014	2014	2014	2014	2014	2014
SouthXyear FE	Yes	Yes	Yes	Yes	No	Yes
Split-Sample IV	No	Yes	Yes	Yes	Yes	Yes
Industry Shares	No	No	Yes	Yes	Yes	Yes
Policy covars.	No	No	No	Yes	Yes	Yes
RegionXyear FE	No	No	No	No	Yes	No
State-spec. quad.	No	No	No	No	No	Yes

Data sources: Details on Log State Net Income/Capita data construction are in Appendix H
Notes: IV estimates are from split-sample-IV regressions (see Section V.C for estimating equations). All regressions include state and year fixed effects; $South \times Year$ fixed effects; and state-year education controls (both from Gallup and CPS at the annual level, and interpolated from the IPUMS Census at the decade level). "Industry shares" controls for state-year share of employment in all one-digit industry categories. "State-spec. quad." indicates that state-specific quadratic time trends are included. "Policy covars." indicate that state-year minimum wage and a "policy liberalism" index (from Caughey and Warshaw, 2016) are included. Standard errors are clustered at the state level. $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$

B. Background on Gallup and other historical data sources

B.1. Brief history of Gallup and other historical polling data

One of the main contributions of the paper is the introduction of newly available household-level data that include information on union membership. We draw much of these data from public opinion polls, which have recently been posted online by the Roper Center at Cornell.[46]

Polling has a long history in American life. The earliest systematic polls were conducted by magazines, in particular *Literary Digest*, which would include a returnable postcard with opinion questions to conduct "straw polls" on the issues of the day (Igo, 2007).[47] Beginning in the late 1930s, George Gallup, Elmo Roper, and Archibald Crossley began importing techniques from market research into the domain of public opinion polling.

Gallup established the American Institute of Public Opinion (AIPO) and set out to conduct nationwide surveys of American opinions on a range of social and political issues.[48] Gallup was scrupulously non-partisan, never running polls on behalf of a particular party. AIPO also devoted considerable efforts to develop neutral, easy to understand question wording. By 1940, about eight million people had read Gallup's tri-weekly polling report, *America Speaks!* which was syndicated in newspapers. Gallup and other pollsters made money by selling their results to businesses for consumer research and newspapers for public opinion.

B.2. Evolution of Gallup's sampling methodology

B.2.1. Gallup methodology before 1950 Before 1950, Gallup used so-called "quota-based" sampling. Survey-takers had to fill quotas for each pre-determined strata thought to capture distinct political views. Enumerators were given both hard (e.g., gender, must have one-third female) and soft (e.g., age, "get a good spread") quotas, but within each quota, interviewers had a lot of discretion. As Berinsky (2006a) notes, "interviewers preferred to work in safer areas and tended to question ap-

46. See .

47. The Earliest *Literary Digest* poll we could find referenced was a poll to predict the outcome of the 1916 presidential election.

48. Similar organizations were formed at roughly the same time: Roper's company was steadily employed by Fortune magazine starting in 1935, Henry Cantril started the Organization of Public Opinion Research (OPOR) in 1940, and the University of Chicago's National Opinion Research Center (NORC) was founded in 1941.

proachable respondents," which likely led to Gallup over-sampling, within each quota strata, more prosperous and well-off respondents.[49]

Gallup once noted that the "the voting public...is the universe of the opinion researcher," suggesting his aim was to be representative of *voters*, which implies substantial underrepresentation of certain segments of the population. Presumably because the South had low turnout (given many of its elections during this time did not even manage a Republican challenger), it was under-sampled. Southern blacks were differentially underrepresented among Southerners, consistent with their near total disenfranchisement during this period. Gallup purposely over-sampled men because of a belief that women merely adopted their husbands' opinions on Election Day.[50]

Documentation for Gallup surveys prior to 1950 describe the sampling procedure as follows:

> Prior to 1950, the samples for all Gallup surveys, excluding special surveys, were a combination of what is known as a purposive design for the selection of cities, towns, and rural areas, and the quota method for the selection of individuals within such selected areas. The first step in obtaining the sample was to draw a national sample of places (cities, towns, and rural areas). These were distributed by six regions and five or six city size, urban rural groups or strata in proportion to the distribution of the population of voting age by these regional-city size strata. The distribution of cases between the non-south and south, however, was on the basis of the vote in presidential elections. Within each region the sample of such places was drawn separately for each of the larger states and for groups of smaller states. The places were selected to provide broad geographic distribution within states and at the same time in combination to be politically representative of the state or group of states in terms of

49. Berinsky, 2006a provides great detail on Gallup's quota-based sampling procedures, from which we draw much of the information in this subsection. Consistent with discretion within the quota-based sampling leading to oversampling of the well-to-do, Gallup over-predicts the Republican vote share in 1940 and 1944, though in both cases he still correctly predicts Roosevelt victories. In 1948, this over-sampling of Republican voters leads him to incorrectly call the election.

50. It is worth noting that any oversampling of men is not a substantial problem for our purposes since we are interested in measures of union status and income at the *household* level. Since most men and women are combined in households, particularly in the earlier years, reports of "any union members in the household" and "household income" should not be affected by whether the surveyed individual in the household was male or female.

> three previous elections. Specifically they were selected so that in combination they matched the state vote for three previous elections within small tolerances. Great emphasis was placed on election data as a control in the era from 1935 to 1950. Within the civil divisions in the sample, respondents were selected on the basis of age, sex and socioeconomic quotas. Otherwise, interviewers were given considerable latitude within the sample areas, being permitted to draw their cases from households and from persons on the street anywhere in the community.

B.2.2. Gallup methodology after 1950 From 1950 onward, Gallup uses modern-day probabilistic sampling procedures. Weights are often provided, but their documentation is not consistent. As a result, in our analyses of the Gallup data we use weights that we generate from the Census, as detailed in Appendix B.5.

The following excerpt is taken from post-1950 Gallup survey documentation on sampling:

> All Gallup polls since 1950, excluding certain special surveys, have been based on a national probability sample of interviewing areas. Refinements in the sample design have been introduced at various points in time since then. However, over this period the design in its essentials has conformed to the current procedure, as follows:
>
> 1. The United States is divided into seven size-of-community strata: cities of population 1,000,000 and over; 250,000 to 999,999; and 50,000 to 249,999; with the urbanized areas of all these cities forming a single stratum; cities of 2,500 to 49,999; rural villages; and farm or open country rural areas.
>
> 2. Within each of these strata, the population is further divided into seven regions: New England, Middle Atlantic, East Central, West Central, South, Mountain, and Pacific Coast.
>
> 3. Within each size-of-community and regional stratum the population is arrayed in geographic order and zoned into equal-sized groups of sampling units.
>
> 4. In each zone, pairs of localities are selected with probability of selection proportional to the size of each locality's population–producing two replicated samples of localities.

5. Within selected cities for which population data are reported by census tracts or enumeration districts, these sample subdivisions are drawn with probability of selection proportional to the size of the population.

6. For other cities, minor civil divisions, and rural areas in the sample for which population data are not reported by census tracts or enumeration districts, small, definable geographic areas are drawn, with the probability of selection proportional to size where available data permit; otherwise with equal probability.

7. Within each subdivision selected for which block statistics are available, a block or block cluster is drawn with probability of selection proportional to the number of dwelling units.

8. In cities and towns for which block statistics are not available, blocks are drawn at random, that is, with equal probability.

9. In subdivisions that are rural or open country in character, segments approximately equal in size of population are delineated and drawn with equal probability.

10. In each cluster of blocks and each segment so selected, a randomly selected starting point is designated on the interviewer's map of the area. Starting at this point, interviewers are required to follow a given direction in the selection of households, taking households in sequence, until their assigned number of interviews has been completed.

11. Within each occupied dwelling unit or household reached, the interviewer asks to speak to the youngest man 18 or older at home, or if no man is at home, the oldest woman 18 or older. This method of selection within the household has been developed empirically to produce an age distribution by men and women separately which compares closely with the age distribution of the population. It increases the probability of selecting younger men, who are at home relatively infrequently, and the probability of reaching older women in the household who tend to be under-represented unless given a disproportionate chance of being drawn from among those at home. The method of selection among those at home within the household

is not strictly random, but it is systematic and objective and eliminates interviewer judgement in the selection process.

12. Interviewing is conducted at times when adults are most likely to be at home, which means on weekends or if on weekdays, after 4:00 p.m. for women and after 6:00 p.m. for men.

13. Allowance for persons not at home is made by a "times-at-home" weighting procedure rather than by "call-backs." this procedure is a standard method for reducing the sample bias that would otherwise result from underrepresentation of persons who are difficult to find at home.

14. The pre-stratification by regions is routinely supplemented by fitting each obtained sample to the latest available census bureau estimates of the regional distribution of the population. Also, minor adjustments of the sample are made by educational attainment (by men and women separately), based on the annual estimates of the census bureau derived from their current population survey. The sampling procedure described is designed to produce an approximation of the adult civilian population living in the United States, except for those persons in institutions such as hospitals.

Note that not until the 1980s does Gallup switch from face-to-face interviews to phone interviews. For this period we make use of the much larger CPS data instead of Gallup, so the vast majority of our Gallup data comes from face-to-face interviews.

B.3. The Gallup union question

The typical Gallup union question is "Are you (or is your husband) a member of a labor union?", with the choices most often being: "neither," "yes, I am," "yes, he is," "yes, both are." In 1959, "husband" changes to "husband/wife." In some years, however, the question does *not* ask which member or members of the household is or are in a union, so we cannot, for example, always measure individual union status. We harmonize these questions to form a measure of *household* union status, where we code a household as union if either household head or spouse is a union member. While technically the implied unit of observation is *couple* (or individual if the respondent is not part of a couple), we will generally refer to this measure as *household union status*. Importantly, Gallup asks this question of *all* respondents, not skipping those in, say, agricultural occupations or who are unemployed.

B.4. Weighting the Gallup data

To construct weights, we use post-stratification methods (i.e., cell-weighting). Specifically, we weight observations in the Gallup data so that the annual proportions of education-race-region cells in Gallup match the corresponding proportions in U.S. Census data. The process involves several steps: First, we construct comparable measures of education (less than high school, high school graduate, some college, college graduate), race (white, non-white), and region (South, non-South) in both Gallup and Census data. Second, we construct annual proportions of each education-race-region cell for each dataset. In the Census data, we apply representative household weights and linearly interpolate values for intercensal years to best approximate the "true" annual proportions of each cell. Third, we generate cell-specific weights w_{ct} by applying the following formula:

$$w_{ct} = \frac{\pi^C_{ct}}{\pi^G_{ct}} \tag{B.1}$$

where c denotes a particular education-race-region cell (e.g., white Southerners with a college degree), and π^C_{ct} and π^G_{ct} denote annual cell proportions for Census and Gallup, respectively. Finally, we let $w_{it} = w_{ct}$ for each respondent i in year t corresponding to cell c in the Gallup data and re-normalize so $\sum_i^{N_t} \frac{w_{it}}{N_t} = 1$ for each year t.

We repeat the procedure above for several alternative cell definitions (e.g., education-race-age-state, age-gender-region). Our preferred weights use education-race-region cells because we find this definition makes our sample as representative as possible without compromising comparability across surveys or creating excessively small or "empty" cells.[51] For surveys without education data, we use race-region weights.

B.5. Comparing Gallup to Census Microdata

We begin with Gallup data from 1950 onward, returning shortly to earlier data. Table B.1 compares Gallup data to 1950–1980 Census data. To summarize how the *actual* (unweighted) Gallup observations compare to the full U.S. adult population, we compare unweighted Gallup data to Census IPUMS tabulations. Given Gallup's well-documented under-sampling of the South, we show results separately for Southern and non-Southern states.

51. For a more thorough discussion of post-stratification weighting, including optimal cell "fineness," see Berinsky, 2006b

In 1950, Gallup exhibits some under-sampling of the South, but, by 1960, this bias had disappeared. From 1950 to 1960, Gallup under-sampled blacks in both the South and the Non-South. This bias continued in the South through 1970, to a smaller degree. These biases reflect the substantial disenfranchisement of blacks, particularly in the South during this period. Age and gender appear representative in Gallup in both regions in each decade.

Gallup respondents outside the South are more educated than their Census counterparts, with the largest gap being a high school completion difference of around 8.5 percentage points in 1950 and 1960. In the South, except for 1950, Gallup and IPUMS show similar levels of education. Gallup Southern respondents have higher high school completion rates than those in the Census in 1950, as Gallup was still under-sampling Southern blacks in that year. In Appendix D we show some of our key results with the Gallup data both unweighted and weighted to match Census characteristics, but Appendix Table B.1 gives some sense of how much "work" the weights must do.

Appendix Table B.2 looks separately at 1940, given that Gallup's sampling procedures were quite different during its earlier years. In fact, in 1940, very few Gallup surveys ask about education (the summary statistics we present for that variable are based on only 5,767 observations), so in this table we include occupation categories as supplemental proxies for socio-economic status. The first column shows, again, unweighted Gallup data. Col. (2) presents summary statistics for all adults in the 1940 IPUMS. Perhaps the most striking discrepancy is gender: consistent with their stated methodology at the time, Gallup over-samples men. Col. (3) adjusts the Census sampling so that men are sampled at the Gallup frequencies and also down-weights large households (since Gallup only interviews one person per household). Comparing col. (1) versus (3) shows, as expected, that Gallup significantly under-samples the South.

Consistent with concerns about Gallup over-sampling the affluent, Gallup respondents in 1940 are substantially more educated than their Census counterparts. Unfortunately, given that only in 1942 does Gallup begin to regularly include an education question, the Gallup sample for which we have an education measure in 1940 is quite small (about 5,700 individuals, relative to over 150,000 for the other Gallup variables in 1940). Given the small education sample in 1940, we use occupational categories to further explore socio-economic status in Gallup versus the 1940 Census. Gallup and IPUMS use different occupation categories—Gallup's are much coarser and unfortunately IPUMS categories do not completely nest Gallup categories—so comparisons are not straightforward. Consistent with the concerns cited earlier that Gallup over-sampled the well-to-do, Gallup respondents appear to have slightly higher-status occupations relative to their Census counterparts, with "white-collar" workers significantly overrepresented.

For the most part, these patterns hold when we drop Southern states from both samples (the final two columns of Table B.2). Importantly, outside of the South, Gallup appears to sample blacks in proportion to their population, even in the very early years of its existence. Also, outside the South, Gallup appears to accurately sample the remaining six regions of the US.[52]

In general, we show results with Gallup data using weights to match (interpolated) Census IPUMS summary statistics, even though the need for weights is not obvious after 1960. From 1937 until 1941, we weight so that Gallup matched the IPUMS in terms of $White \times South$ cells, given that the summary statistics show that Gallup sampling along these dimensions appears suspect in the early years. Beginning in 1942 (the first year in which Gallup surveys ask the union and education questions in the same survey) we weight by $White \times Education \times South$, where $Education \in$ {No high school degree, HS degree, Some college, College graduate}, thus giving us $2 \times 4 \times 2 = 16$ cells on which to match. In practice, however, our results are very similar with and without weights.

References

Berinsky, Adam J (2006a). "American public opinion in the 1930s and 1940s: The analysis of quota-controlled sample survey data". *International Journal of Public Opinion Quarterly* 70.4, pp. 499–529.

— (2006b). "American public opinion in the 1930s and 1940s: The analysis of quota-controlled sample survey data". *International Journal of Public Opinion Quarterly* 70.4, pp. 499–529.

Igo, Sarah Elizabeth (2007). *The averaged American: Surveys, citizens, and the making of a mass public*. Harvard University Press.

52. We use Gallup-defined geographic regions in this table.

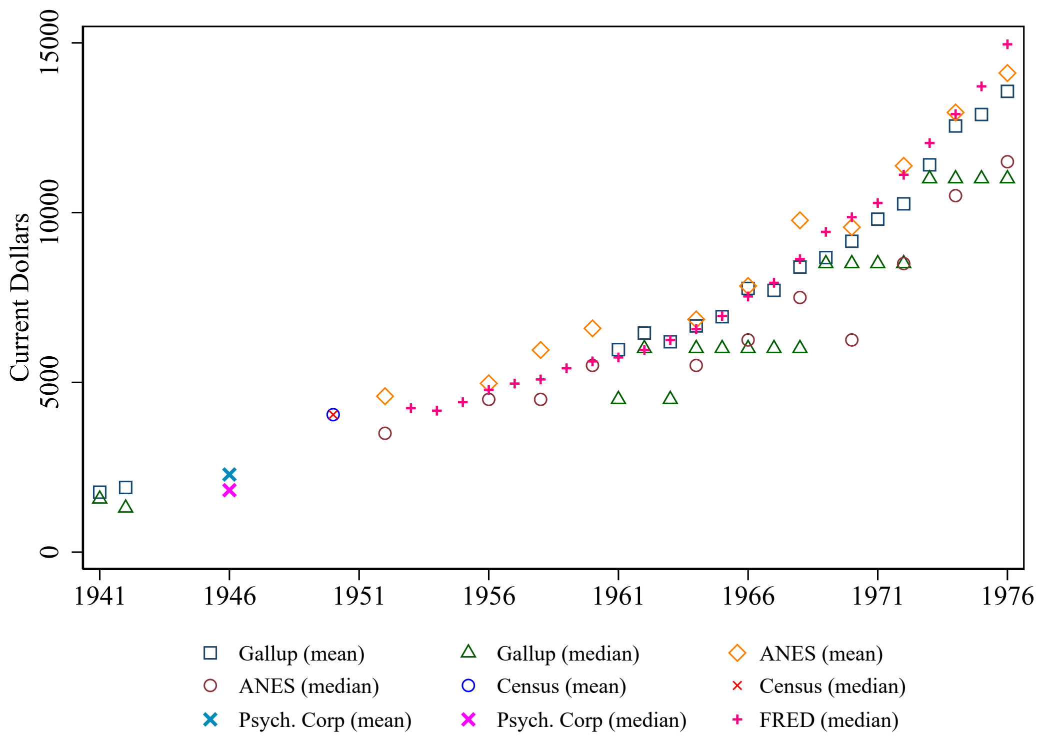
APPENDIX TABLE B.1: COMPARING GALLUP AND IPUMS, 1950–1980

	1950		1960		1970		1980	
	Census	Gallup	Census	Gallup	Census	Gallup	Census	Gallup
South Share	0.258	0.133	0.260	0.254	0.270	0.262	0.295	0.264
—South								
Female	0.530	0.516	0.532	0.539	0.528	0.515	0.521	0.508
Age	39.48	40.90	41.18	42.99	41.12	41.97	39.84	41.45
Black	0.205	0.0759	0.183	0.137	0.159	0.124	0.159	0.157
HS grad.	0.280	0.405	0.387	0.376	0.513	0.565	0.674	0.703
—Non-South								
Female	0.523	0.508	0.520	0.527	0.523	0.514	0.517	0.510
Age	40.64	40.44	41.68	41.71	41.33	41.47	39.97	40.63
Black	0.0506	0.0479	0.0638	0.0577	0.0742	0.0616	0.0816	0.0880
HS grad.	0.418	0.502	0.491	0.578	0.634	0.712	0.768	0.814
Observ.	250519	91682	4488254	23620	2023944	75911	6186033	59138

Data sources: Gallup surveys and 1950–1980 IPUMS.

Notes: We use the Gallup definition of the "South": all eleven states of the former Confederacy plus Oklahoma. All Census results use IPUMS person weights.

Appendix Figure B.1: Household income measures in our historical survey data compared to official statistics



Data sources: See Section II for a description of each of our historical data sources. The 1950 data points come from the Census and from 1953 onward from Federal Reserve Economic Data (FRED).

Notes: As our historical data sources are unfamiliar and non-standard sources of household income, we compare them to official government statistics. Beginning in the 1970s, we use the more standard CPS and thus do not show comparisons.

APPENDIX TABLE B.2: COMPARING GALLUP AND IPUMS IN 1940

	Gallup	Census	Census	Gallup	Census
—Demographics					
Black	0.0291	0.0895	0.0906	0.0325	0.0357
Female	0.338	0.505	0.344	0.341	0.343
Age	40.46	39.61	40.06	40.41	40.55
HS Graduate	0.493	0.278	0.266	0.494	0.290
College Graduate	0.0720	0.0472	0.0499	0.0709	0.0543
—Geography					
Northeast	0.0836	0.0660	0.0629	0.0947	0.0854
Mid Atlantic	0.261	0.253	0.241	0.295	0.327
East Central	0.208	0.187	0.186	0.236	0.252
West Central	0.177	0.127	0.129	0.200	0.175
South	0.117	0.258	0.263	0	0
Rocky Mountain	0.0752	0.0284	0.0308	0.0851	0.0418
Pacific Coast	0.0783	0.0754	0.0818	0.0887	0.111
—Occupation					
Farmer	0.213	0.156	0.159	0.188	0.109
Professional	0.0792	0.113	0.122	0.0808	0.129
Propietors, managers, officials	0.0105	0.0928	0.0875	0.0108	0.0933
Clerks (white collar)	0.299	0.0535	0.0539	0.306	0.0609
Skilled workmen and foremen	0.0926	.	.	0.0970	.
Unskilled or semi-skilled labor	0.194	.	.	0.204	.
Sales workers	.	0.0462	0.0457	.	0.0499
Craftsmen	.	0.142	0.139	.	0.153
Operatives	.	0.146	0.147	.	0.159
Service workers (priv. HH)	.	0.0103	0.0105	.	0.00626
Other service workers	.	0.0477	0.0468	.	0.0508
Laborers	.	0.0932	0.0973	.	0.0944
No answer, N/A, etc.	0.111	0.0999	0.0920	0.113	0.0949
Gender/HH adj?	No	No	Yes	No	No
Ex. S/SW?	No	No	No	Yes	Yes
Observations	144996	736832	736832	127995	544375

Data sources: Gallup surveys and 1940 IPUMS.

Notes: The Gallup sample size varies substantially by variable during this period. For the col. (1) sample, all demographics except for education and all geographic variables have a sample size around 159,000 (with small variations due to missing observations). The occupation codes have a sample size of roughly 21,000. The high school completion indicator has a sample size of 5,700. In col. (4) each sample size is roughly twelve percent smaller. “HH / gender adjustment” underweights women and people in large households in the IPUMS to better match Gallup sampling (which only sampled one person per household and had a target female share of one-third). “Ex S/SW" excludes Southern and Southwestern states (all eleven states of the former Confederacy plus Oklahoma). Note that occupation categories are coarser in Gallup than in the Census (but unfortunately, Gallup categories do not nest Census categories). We do our best to match occupation across these different categorizations. All Census results use IPUMS person weights.

APPENDIX TABLE B.3: SUMMARY STATISTICS FROM SUPPLEMENTARY DATA SETS

	(1) ANES	(2) BLS exp. dataset	(3) U.S. Psych. Corp.	(4) NORC	(5) ANES panel
Union household	0.240	0.116	0.172	0.274	0.284
Female	0.596	0.507	0.496	0.514	0.538
White	0.848	0.819	0.890	0.903	0.906
Age	41.35	40.98	42.13	39.84	41.72
HS graduate	0.738	0.363	0.442	0.403	0.532
South	0.288	0.271	0.208	.	0.239
Log fam. inc.	10.73	10.07	10.11	7.913	8.511
Sample period	1952-2012	1936	1946	1950	1956-1960
Observations	32475	5517	5665	1106	3783

Notes: See Section II.B and Appendix B for details on the data sources.

C. Sample selection and construction of key variables

C.1. Sample Selection

To construct our main Gallup sample, we apply the following selection criteria to the population of recorded Gallup survey respondents from years 1937 through 1987. First, we eliminate respondents to surveys in which the union membership question was not asked. Second, we remove any respondents younger than 21 or older than 64 (we cap at 65 to focus on the working-age population, and only halfway through our sample period did Gallup begin to include 18-20 year olds and we wish to have a consistent sampling rule throughout the entire period). Third, we remove respondents who live in Alaska, Hawaii, or Washington DC (again, Gallup did not include these respondents at the beginning of our sample period). For the state-year analyses we also exclude Idaho because the state identifiers are often miscoded as Hawaii.

Our CPS sample is taken from the May supplements in years 1976 to 1981, the Merged Outgoing Rotation Groups in years 1983 to 1989, and the Annual Social and Economic Supplement in years 1990 to 2015. Note that the CPS did not ask about union status in 1982. Since the CPS contains information for all individuals within a household, to make the CPS comparable with Gallup, we restrict our CPS sample to one randomly selected observation from each household, which we refer to as the "designated" respondent.[53] For state-year measures our CPS-based series begins in 1977, as individual state-of-residence identifiers are not available before that time. We exclude designated respondents in armed forces. Additionally, we exclude Alaska, DC and Hawaii from all analyses, and Idaho from the state-year analysis to make it comparable with the Gallup sample.

C.2. Variable Construction and Trends in Inequality Measures

Union Density In both Gallup and CPS, union density is calculated as the number of households with at least one reported union member divided by the total number of households. The Gallup sample is limited to respondents aged 21-65 whereas the CPS sample is limited to "designated" respondents aged 18-65.

Family Income Our Gallup measure of family income covers years 1942 and 1961 through 1976. Gallup family income is derived from the responses to survey ques-

53. The exception to this is Appendix Figure D.5, which examine the robustness of our premium estimates to using all observations within a household.

tions of the following form: "Which best represents the total annual income, before taxes, of all the members of your immediate family living in your household?" Responses are coded into income bins which vary across surveys. We construct a harmonized income measure by calculating the midpoint of each interior binned response. For top and bottom bins, we estimate implied midpoints from a fitted Pareto distribution as in Von Hippel, Hunter, and Drown (2017). Our CPS measure of family income is taken from the May and March supplements in years 1978 through 2015. This measures combines all reported income from household members 15 years and older. To construct this variable in early CPS years (May and March before 1990), we use the family income variable, which is binned into 12 categories. For the following years (CPS March only) we use the continuous family income variable, which reports the total income for the respondent's family. To make the continuous variable comparable with the binned variable of earlier years, we recode it into bins matching those of the ANES income variable in the corresponding year.

College premium, college high school share ratio, wage ratios The college wage premium, college high school share ratio, and the 90-10, 90-50, and 10-50 wage ratios are calculated using a sample of 18 to 65 year-old full-time, full-year wage and salary workers who make at least one-half of the minimum wage and who have 0-48 years of potential experience in the March CPS (1964-2019 for the time series analysis and 1977-2019 for the state-year analysis) and the 1940-1970 Census.[54] Unemployed and NILF respondents are excluded from the analysis.

In the time-series analysis, we calculate changes in each measure between 1940-50, 1950-60, and 1960-70 in the Census data and append these changes to the measure from 1964-2019 (or 1977-2019 in the state-year analysis) calculated from the March CPS.

The **college-high school share ratio** is calculated in terms of efficiency units following the methodology outlined in Autor, Katz, and Kearney, 2008. Workers are divided into cells based on two sexes, five education categories (high school drop outs, high school graduates, some college, college graduates, greater than college), and years of experience (ten-year bins for the state-year analysis).

For each cell in each year we calculate the weighted sum of weeks worked by all individuals in the cell using the individual weights from the data. This comprises the "quantity" of labor supplied. To translate this into efficiency units of labor supply

54. We follow Autor, Katz, and Kearney, 2008 and calculate years of potential experience as age minus assigned years of schooling minus six, rounded down to the nearest integer value.

we also calculate the "price" of each week of labor in a particular cell. The "price" of labor corresponds to weighted average of log real weekly income in each cell, normalized by a reference wage (the wage of male high school graduates with the highest category of experience cell in our data, which is 40-48 in the main sample but 30-40 in the backwards projection to 1930 described below), and averaged over the entire period. The efficiency units of labor supplied by each cell is the product of the "quantity" and "price" of labor.

The total efficiency units of labor supplied in a given year is calculated by summing across cells. We calculate aggregate college-equivalent labor supply as the share of total efficiency units of labor supplied by college or college-plus workers plus half of the share of labor supplied by workers with some college. The college-high school share ratio is the natural logarithm of the ratio of college-equivalent to non-college-equivalent labor supply shares in each year.

As the 1930 census does not ask years of schooling, we construct the 1929 college-high school share ratios by projecting backwards from cohorts in 1940, using their state of residence in 1935. We use the efficiency units in 1940 aggregated across 34-64 age groups, which are the cohorts that would be 24-54 in 1930. The correlation between these age groups in 1940 is 0.885 and 0.883 updated by migration, which validates the backward projection for that year.

The **college wage premium** is calculated following the methodology outlined in Autor, Katz, and Kearney, 2008 and Goldin and Katz, 2008. The premium is the fixed weighted average of the premium earned by college graduates vs high school graduates and more than college educated workers vs high school graduates. These premiums are estimated by regressing the log real hourly earnings on a set of five education dummies, a full-time dummy, a female dummy, a non-white dummy, a set of three geographic division dummies, a quartic in experience and the interaction of female with both non-white and the quartic in experience. The weights are the relative employment shares of college and more than college educated workers in 1980.

Weights are calculated as follows:

$$CollegeShare = \frac{\text{Number of workers with exactly college education}}{\text{Number of workers with exactly college or more than college education}}$$

$$MoreThanCollegeShare = \frac{\text{Number of workers with more than college education}}{\text{Number of workers with exactly college or more than college education}}$$

The **90-10, 90-50, and 10-50 wage ratios** are calculated as the difference in the Xth and Yth percentile of log real weekly earnings among men in our sample.

Gini coefficient For the aggregate time-series analysis, the Gini coefficient is taken from Kopczuk, Saez, and Song, 2010. For the state-year analysis, we estimate the Gini coefficient from a sample of 18 to 65 year-old workers who are not self-employed, have non-allocated income, and have 0-48 years of potential experience in the March CPS (1977-2019) and the 1940-1970 Census. We append changes in the Gini coefficient between 1940-50, 1950-60, and 1960-70 in the Census data to the coefficient in 1977-2019 calculated from the March CPS.

Appendix Figure C.2 shows the time-series plots of our various measures of inequality, confirming that they all broadly tend to exhibit U-shapes over the 20th century.

Manufacturing Employment We estimate major industry employment shares from 1910 to 2015 by combining data from the Census, BLS State and Area Employment, Hours and Earnings series, and ACS. Although the BLS is our preferred data source, it is only available between 1939 and 2001. Furthermore, not every state-industry pair has data beginning in 1939, and for some pairs data starts as late as 1982. We therefore supplement the BLS series with Census data from 1910 to 1980 and ACS data from 2001 to 2015.

For each dataset, state, and year we group calculate the share of employed individuals that work within each major industry: mining, construction, manufacturing, transportation, trade, finance, services, and government. We group 1950 census industry codes in the Census and ACS to match these BLS industries.[55] To combine the Census and ACS with the BLS, we append changes in the Census and ACS to the BLS series in its first and last year, respectively.

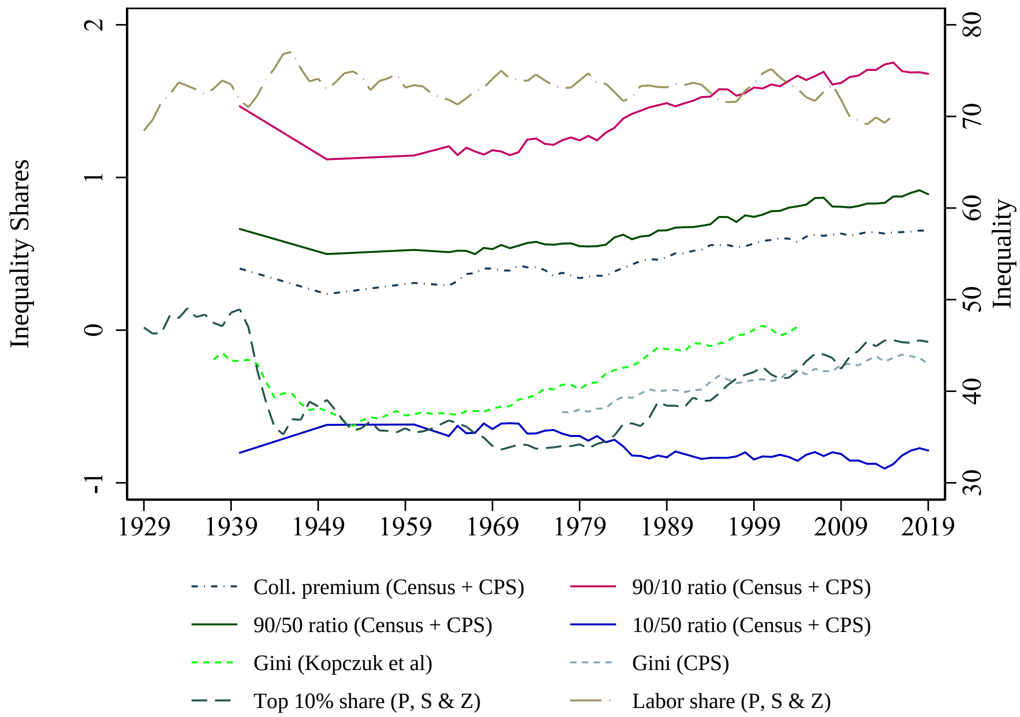
References

Autor, David H, Lawrence F Katz, and Melissa S Kearney (2008). "Trends in US wage inequality: Revising the revisionists". *The Review of economics and statistics* 90.2, pp. 300–323.

Goldin, Claudia Dale and Lawrence F Katz (2008). *The race between education and technology*. Harvard University Press.

55. Mining corresponds to 1950 census industry codes 206-239, construction to 246, manufacturing to 306-499, transportation to 506-598, trade to 606-699, finance to 716-756, services to 806-899, and government to 906-976.

APPENDIX FIGURE C.2: MEASURES OF INEQUALITY OVER THE 20TH CENTURY



Data sources: The college wage premium, the 90-10, 90-50, and 10-50 log wage ratios are calculated using a sample of 18 to 65 year-old full-time, full-year wage and salary workers who make at least one-half of the minimum wage and who have 0-48 years of potential experience in the March CPS (1964-2019 for the time series analysis and 1977-2019 for the state-year analysis) and the 1940-1970 Census. The labor share and top ten share of income are from Piketty, Saez, and Zucman (2018). The Gini coefficient for all workers is from Kopczuk, Saez, and Song (2010), while the CPS Gini is calculated using 18 to 65 year-old workers who are not self-employed, have non-allocated income, and have 0-48 years of potential experience in the March CPS (1977-2019). See text of section C.2 for details and sources of measures.

Kopczuk, Wojciech, Emmanuel Saez, and Jae Song (2010). "Earnings inequality and mobility in the United States: evidence from social security data since 1937". *The Quarterly Journal of Economics* 125.1, pp. 91–128.

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman (2018). "Distributional national accounts: methods and estimates for the United States". *The Quarterly Journal of Economics* 133.2, pp. 553–609.

Von Hippel, Paul T, David J Hunter, and McKalie Drown (2017). "Better estimates from binned income data: Interpolated CDFs and mean-matching". *Sociological Science* 4, pp. 641–655.

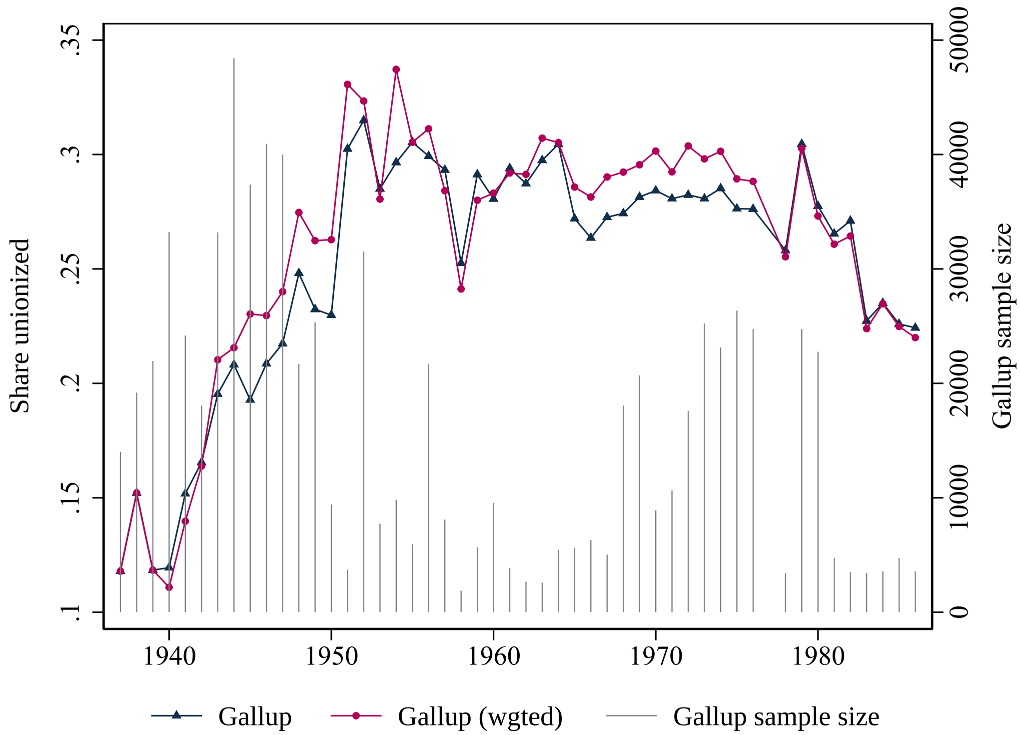
D. Main results using various weighting schemes and individual- instead of household-level union membership

As described in Section II and Appendix B, two issues in the Gallup data complicate comparisons with the CPS and other standard data sources. First, especially in its first few decades, Gallup polls over-sampled the well-off and under-sampled all Southerners but particularly black Southerners. Second, we cannot always infer individual-level union membership in the Gallup and other historical survey data, so instead we mostly use a household-level measure (i.e., is anyone in the household a union member).

An obvious concern is that some of the trends in the size of the union premium or selection into union that we document over our long sample period are in fact artifacts of these aspects of Gallup's data. For example, changes in selection into union *households* might reflect changes in assortative mating and not union membership *per se*.

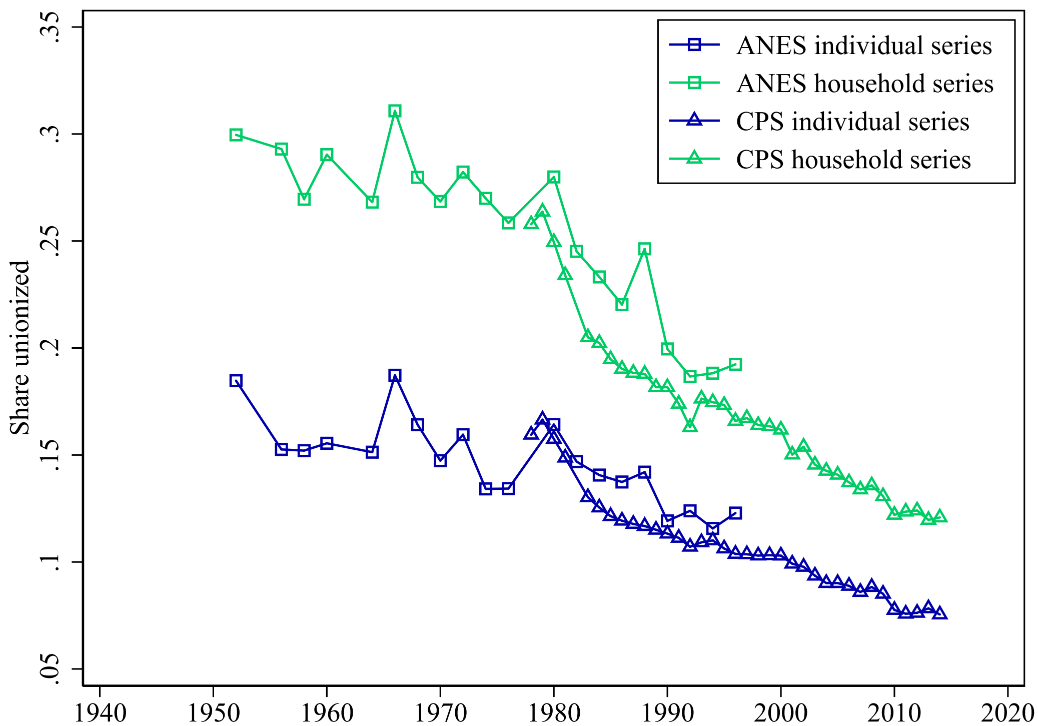
In this appendix, we reproduce, when possible, some of our main results (a) under various weighting schemes and (b) using an individual- instead of household-based measure of union membership. We also show some results for men only, as in the early years union membership was almost entirely male. Thus, for this subsample the household membership will closely proxy individual membership.

Appendix Figure D.1: Union share of households in the Gallup data (weighted vs. unweighted)



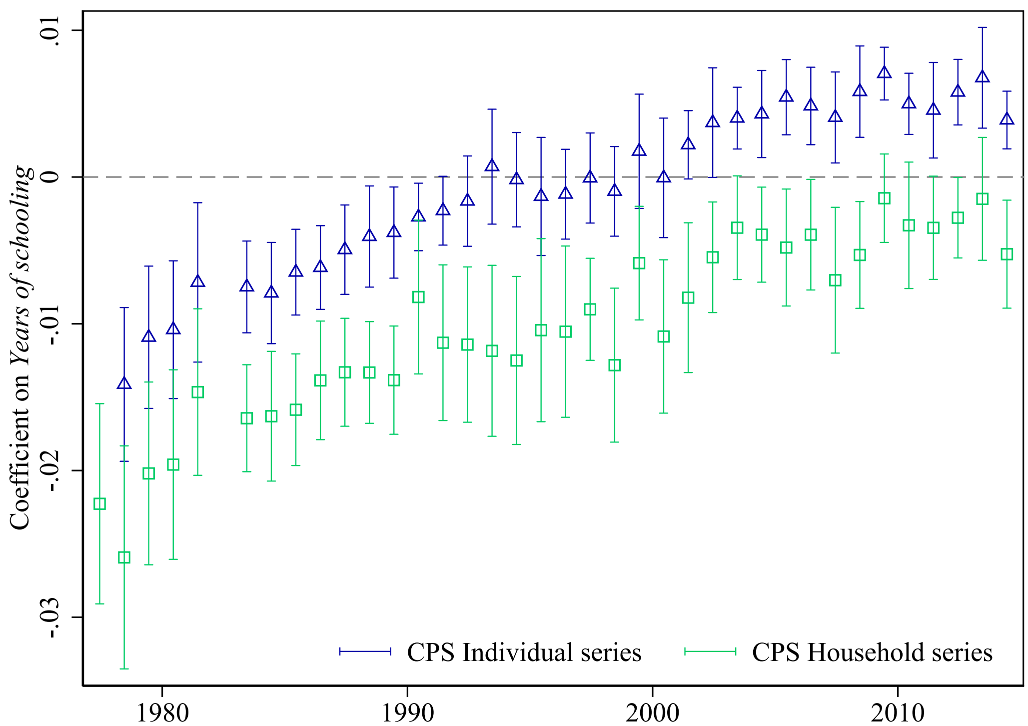
Data sources: Gallup. See Section II.B and Appendix B for more details on data and weight construction.

APPENDIX FIGURE D.2: COMPARING INDIVIDUAL VERSUS HOUSEHOLD UNION DENSITY IN CPS AND ANES, 1952–PRESENT



Data sources: Current Population Survey and American National Election Survey

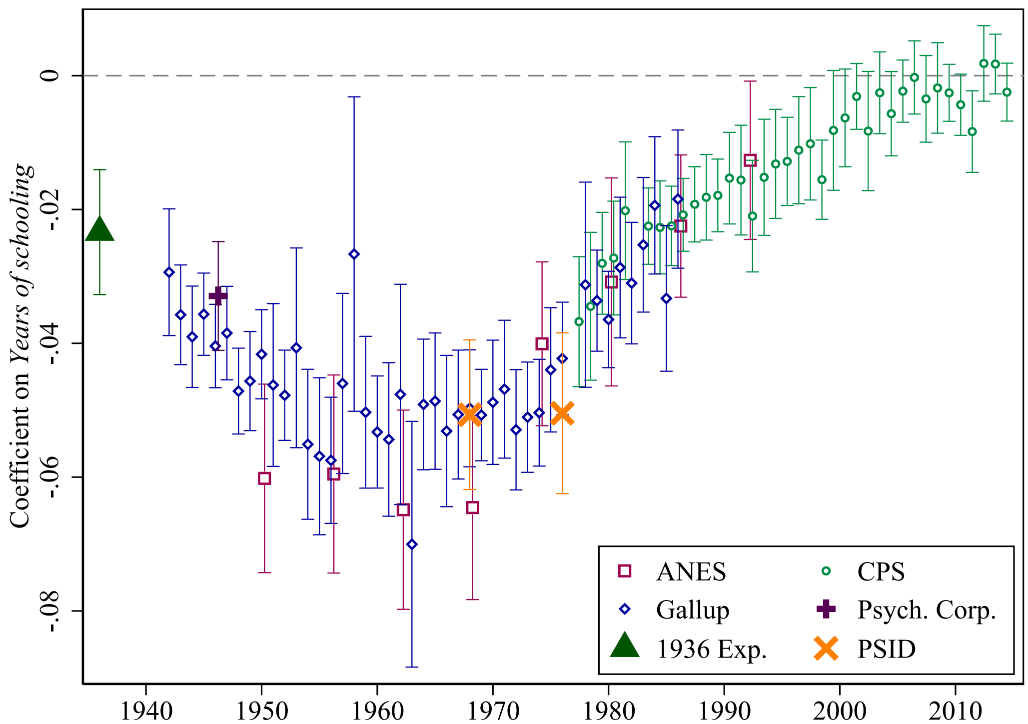
APPENDIX FIGURE D.3: SELECTION INTO UNIONS BY YEARS OF SCHOLLING IN THE CPS, INDIVIDUAL AND HOUSEHOLD MEASURES



Data sources: Current Population Survey.

Notes: The "household series" replicates the CPS analysis in Figure III (i.e., regresses, separately by year, a household union dummy on years of schooling, gender and state fixed effects, plotting the coefficient on years of schooling. The "individual series" substitutes *individual union membership* as the outcome variable instead of the household union dummy.

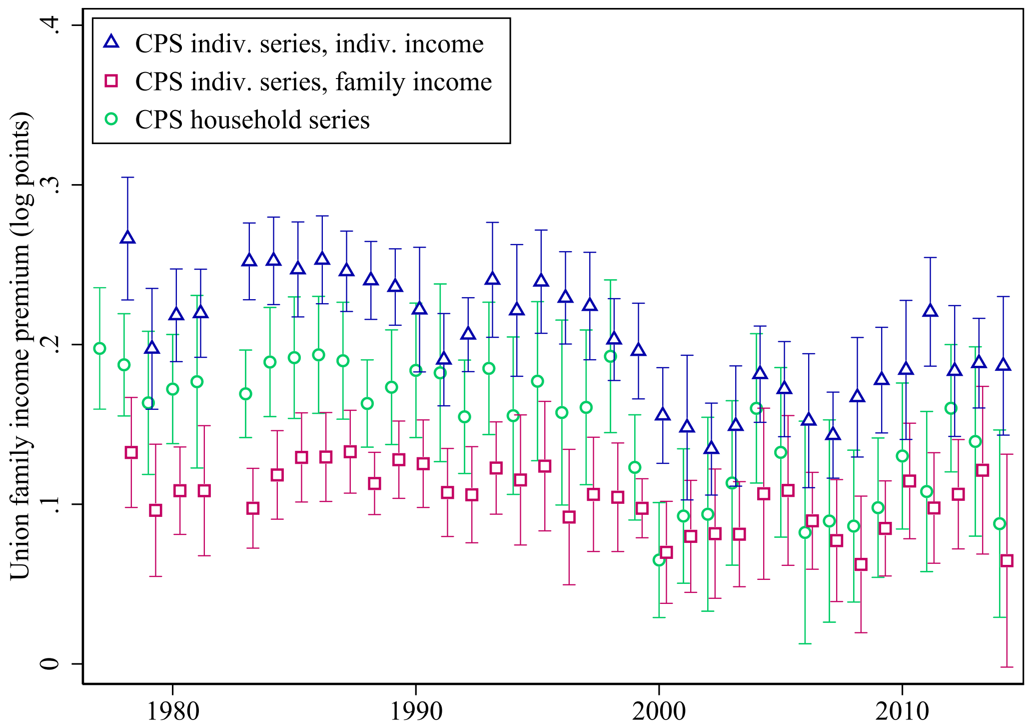
APPENDIX FIGURE D.4: SELECTION INTO UNIONS BY EDUCATION, MALE SURVEY RESPONDENTS ONLY



Data sources: Gallup data, 1937–1986; CPS, 1978–2016; BLS Expenditure Survey, 1936; ANES; 1952–1996; U.S. Psych. Corporation, 1946. See Section II.B for a description of each data source.

Notes: We regress household union status for male respondents only on $Years of education$, state s and survey-date t fixed effects, age and its square, and gender. (The notes to Figure VI describe how we impute years of schooling if the survey source only gives us categories of educational attainment.) We estimate this equation separately by survey source and by year. The figure plots the coefficient on $Years of education$. For the ANES, because the samples are smaller, we group surveys into six-year bins. The plotted confidence intervals are based on standard errors clustered by state.

APPENDIX FIGURE D.5: COMPARING UNION FAMILY AND INDIVIDUAL PREMIUM IN THE CPS



Data sources: CPS, 1978–2016. See Appendix C for details on CPS individual and family income variable construction.

Notes: Each plotted point comes from estimating equation (2), which regresses log family income on a union dummy and controls for age, gender, race, and state fixed effects. Occupation controls are not included. For each series, we estimate a separate regression for each year. The first series regresses log *individual* earnings on *individual*-level union membership. The second series regresses log *family* income on *individual*-union membership. The third series regresses log *family* income on whether the individual has a union member in the *household* (whether or not the individual himself is in a union) and is the concept we use in most of the paper. The plotted confidence intervals are based on standard errors clustered by state.

APPENDIX TABLE D.1: GALLUP SELECTION RESULTS THROUGH 1950, ROBUSTNESS TO WEIGHTS

	Dependent variable: Union household			
	(1)	(2)	(3)	(4)
yrsed	-0.0394***	-0.0386***	-0.0369***	-0.0307***
	[0.00309]	[0.00274]	[0.00299]	[0.00266]
Dept. var. mean	0.257	0.258	0.258	0.195
Weighting scheme	Baseline	None	White x Sth	Schickler
Observations	600744	610126	610126	62085

Data sources: See Section III and Appendix B for details.

Notes: All regressions include state and survey-date fixed effects. Respondents are include ages 21–64. Baseline weights are those we use throughout the paper (weights to make Gallup match interpolated Census cells for $White \times South \times Education\,categories$ (16 cells)). $White \times Sth$ are analogous, but match only on those four cells. Raking weights are constructed by matching yearly marginal mean population shares by $Black$, $Female$, and $Region$ to interpolated census shares. See Deville, Särndal, and Sautory, 1993 for more details. "Schickler weights" are taken from Schickler and Caughey, 2011 and match on $Black$ and whether a residence has as phone. They are only available through 1945. Standard errors in brackets, clustered by state. $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$

E. Existing Measures of Union Density Pre-Dating the Current Population Survey

The CPS first asks respondents their union status in 1973, and then only in selected months until 1983 from which time information on union status was collected each month in the CPS as part of the outgoing rotation group supplement. Before this survey, the primary sources for union density are the BLS and Troy/NBER historical time series mentioned in the introduction. The data underlying these calculations are union reports of membership and dues revenue when available, and a variety of other sources when not available. Neither of these data sources ever used representative samples of individual workers to calculate union density.

In general, the data derived from union reports likely become more accurate by the 1960s. Post-1959 the BLS collected mandatory financial reports from unions as a condition of the Labor-Management Reporting and Disclosure (Landrum-Griffin) Act, and Troy and Sheflin (1985) incorporate these data into their estimates of union density. Beginning in 1964, the BLS disaggregates union membership counts by state, and Hirsch, Macpherson, and Vroman (2001) splice these reports together with the CPS to form state-year union density panel beginning in 1964 and continuing through today.[56]

Before the 1960s, however, union data were far less standardized. In the remainder of this section, we detail the methodology of the two most widely used data sources on aggregate union density: the BLS and Troy series.

E.1. The BLS Estimate of Early Union Density

The BLS series is based on union-reported membership figures starting in the late 1940s. Prior to 1948, the methodology for calculating union membership does not appear standardized. For example, the 1945 Monthly Labor Report notes as its sources: “This study is based on an analysis of approximately 15,000 employer-union agreements as well as employment, union membership, and *other data available to the Bureau of Labor Statistics* [emphasis ours]" (Bureau of Labor Statistics, 1945)[57]

56. Freeman et al. (1998) constructs a time-series of union density from 1880 to 1995, splicing together the official series from the BLS with series constructed from the CPS. Freeman reports alternative series constructed by other scholars (Troy (1965), Troy and Sheflin (1985),Wolman (1924), and Galenson (1960)) in the Appendix to his paper.

57. For example, one alternative source the BLS used was convention representation formulas. “Convention formulas” specified the number of seats, as a function of membership, each union would have at the umbrella organization convention. By inverting this formula

It is obviously hard to verify information from unspecified "sources available to the BLS" but even in instances where the BLS can rely on union membership reports, concerns arise. A key issue is that unions had important incentives to overstate their membership and until the late 1950s faced no penalty for doing so. In the early and mid-1930s, the main umbrella organization for local unions was the American Federation of Labor (AFL). They were often charged with over-stating their membership, presumably to inflate their political influence. For example, a 1934 *New York Times* story casts doubt on the AFL's claim to represent over six million workers, noting that "complete and authoritative data are lacking" and that the figures provided by the AFL "are not regarded as accurate."[58] Individual unions also had an incentive to inflate the numbers they reported to the AFL. For example, the number of seats each union would receive at the annual convention was based on a formula to which membership was the main input.

If anything, these incentives to over-report likely grew after 1937, when the Committee on Industrial Organization broke away from the AFL to form a rival umbrella organization, the Congress of Industrial Organizations (CIO). Both federations of labor, the AFL and CIO, now competed for local unions to join their umbrella organizations, as well as for sympathies of government officials, tasks that were aided by a public perception that the federation was large and growing. Based on our read of *New York Times* articles on unions in the late 1930s and early 1940s, one of the most common if not the most common topic is the conflict between the two federations.[59] Individual unions still had incentives to compete for influence within their given federation, and thus inflate membership.

Membership inflation became such an issue that the federations themselves may not have known how many actual members they had. In fact, the CIO commissioned an *internal* investigation into membership inflation, conducted by then-United Steelworkers of America president Philip Murray. Murray's 1942 report con-

and using the convention records, rough estimates of union membership could be formed.

58. See, *"Organized Labor is Put at 6,700,000"*, New York Times, May 1935. reporting that "For one thing, complete and authoritative data are lacking, and this is especially true during times of depression, when some unions drop unemployed workers from the rolls and exempt them from paying dues. . . . . The [AFL] reported an average membership of 2,609,011 for the year ended Aug. 31, 1934. These official figures, which are not regarded as an accurate measure of the movement, are far below the peak figure of 4,078,740 for 1920."

59. As just one example, a 1938 *NYT* headline and subtitles read: "Green Says Lewis Falsified Report; A.F.L. Head Alleges Statement on C.I.O. membership is an 'Amazing Inflation; Questions Income Data," referring to AFL head William Green and CIO head John Lewis, respectively.

cluded that actual CIO membership was less than fifty percent of the official number the federation was reporting (Galenson, 1960).

E.2. The Troy Estimates of Early Union Density

In his NBER volumes estimating union density, Troy is well aware of the problems documented above with the BLS estimates. For this reason, he defines membership as "dues-paying members" and proceeds to estimate union membership using unions' financial reports where available, presumably under the assumption that financial reports were less biased than membership reports. For each union, he divides aggregate union dues revenue by average full-time member dues to recover an estimate of union membership. While Troy is cognizant of the limitations of his data and methodology, he believes the biases are largely *understating* union membership (e.g. some groups, such as veterans, pay lower than average or no dues).

But union financial reports, like membership reports, are also not verified until the late 1950s. Nor is it obvious that union revenue data are not similarly inflated (in fact, the AFL accused the CIO of lying about their income data, as we mention in footnote 59). Moreover, revenue data are largely incomplete for the 1930s and 1940s. For example, in his 1940 estimates, Troy (1965) notes that the sources for 54.4% of his total is *not* in fact from financial reports, but instead an "Other" category, which includes personal correspondence with unions, asking their membership.[60] As such, for these early years, the Troy data in fact appears to face the same issue with membership-inflation as does the BLS data.[61]

In addition, Troy imputes the membership of many CIO unions in the late 1930s and 1940s by assigning them the membership of their AFL counterpart in the same sector.[62] This procedure likely over-states CIO membership, given that the AFL was believed to be twice as large as the CIO during this period (we also find this 2:1 ratio

60. "Other" is down to 10% by 1960 (Troy (1965)).

61. Troy (1965) also only presents validation exercises for his post-1950 data, comparing reported measurement with that inferred from dues receipts for the Chemical and Rubber Workers in 1953, leaving it open whether the BLS or Troy (or neither) is correct for the pre-1950 series.

62. From Troy (1965) [pp. A53]: "The average membership per local industrial union is arbitrarily estimated to be 300, and this figure is multiplied each year by the number of such unions reported by the ClO. The estimate of an average membership of 300 is deemed a fair one since the average membership of the local trade and federal labor unions of the AFL, a class of unions similar to the local industrial unions of the CIO, varies from a low of 82 in 1937 to a high of 193 in 1948."

in our Gallup data), though obviously that average ratio may vary by sector.

In summary, while a likely improvement over the BLS series, it is difficult to believe that Troy's estimates (or Troy and Sheflin (1985)) are without extensive mismeasurement. Given the limitations of the existing pre-CPS data on union density, in the next section we introduce a new source: Gallup and other opinion surveys.

E.3. Other pre-CPS state-year measures of union density

The only sources of state-year data on union density prior to the CPS we are aware of are measures created by Hirsch, Macpherson, and Vroman (2001) from BLS reports (which begin disaggregating union membership regionally, often by state, in 1964) from 1964-1977, and measures created by Troy and Sheflin (1985) for the years 1939 and 1956. Our Gallup measure is quite highly correlated (correlation = .724) with the existing Hirsch-Macpherson measures (individual union density as a fraction of non-farm employment) for the 1964-1986 years, which are where there is overlap. This correlation increases to .75 when we restriction attention to the CPS years with state identifiers (1978-1986).

The historical Troy measures for 1939 and 1953 are constructed from even more fragmentary records than the annual series we discuss above (as many union reports did not disaggregate either revenue or membership by state). Nevertheless our Gallup measures are also correlated with these data in both cross-sections and changes (1939 correlation = 0.78, 1953 correlation = 0.75, correlation in changes =0.5).

Finally, to test for pre-trends in our IV design, we make use of the 1929 Handbook of American Trade Unions, which reports the number of locals for each union by state. We then take the national membership of each union and apportion it to states in 1929 based on the share of locals in that state to form a proxy for the number of members of a given union in a given state, and then sum across unions to get a state-level measure of union membership in 1929. Cohen, Malloy, and Nguyen (2016) construct a similar measure and validate it for a number of states.

References

Bureau of Labor Statistics (1945). *Extent of Collective Bargaining and Union Status, January 1945*. Tech. rep. Bulletin # 829.

Cohen, Lauren, Christopher J Malloy, and Quoc Nguyen (2016). "The impact of forced migration on modern cities: Evidence from 1930s crop failures". *Available at SSRN 2767564*.

Freeman, Richard B et al. (1998). “Spurts in Union Growth: Defining Moments and Social Processes”. *NBER Chapters*, pp. 265–296.
Galenson, Walter (1960). *The CIO challenge to the AFL: a history of the American labor movement, 1935-1941*. Harvard University Press.
Hirsch, Barry T, David A Macpherson, and Wayne G Vroman (2001). “Estimates of union density by state”. *Monthly Labor Review* 124.7, pp. 51–55.
Troy, Leo (1965). *Trade Union Membership, 1897–1962*. NBER.
Troy, Leo and Neil Sheflin (1985). “Union Sourcebook: Membership, Finances, Structure, Directory”. *West Orange, NJ: Industrial Relations Data and Information Services*.
Wolman, Leo (1924). *The Growth of American Trade Unions, 1880-1923*. NBER, pp. 163–170.

F. DISTRIBUTIONAL DECOMPOSITION APPENDIX

Re-weighting Let households' selection into unions be given by $u(X,\epsilon)$ in reality and $u^C(X,\epsilon)$ under some counterfactual, C. The true income distribution, F_Y, is observed, but the counterfactual, F_Y^C, must be estimated. Using Bayes rule, we find that

$$\begin{aligned} F_Y^C &= \int\int F_{Y|X,u}\, dF_{u^C|X}\, dF_X \\ &= \int\int F_{Y|X,u}\, dF_{u|X}\Psi(u,X)\, dF_X \\ &= \int\int F_{Y|X,u}\Psi(u,X)\, dF_{u,X}, \end{aligned} \tag{F.1}$$

where $\Psi(u,X)$ is reweighting factor given by

$$\Psi(u,X) \equiv u * \frac{\Pr(u^C=1|X)}{\Pr(u=1|X)} + (1-u)\frac{\Pr(u^C=0|X)}{\Pr(u=0|X)}. \tag{F.2}$$

Equation F.1 illustrates how the counterfactual income distribution relates to the observed income distribution, allowing us to simulate the former by reweighting on observables in the latter. As Equation F.2 shows, the nature of this reweighting depends not only on $\Pr(u=1|X)$, which we estimate using predicted values from logistic regressions of observed union status, but also on $\Pr(u^C=1|X)$, which depends on the counterfactual in question. In our case we will consider setting a within-year counterfactual where $\hat{\Pr}(u^C=1|X)=0$, effectively deunionizing the income distribution by reweighting union members to have the same income distribution as the non-union members with the same X. We will also consider an over-time counterfactual where $\hat{\Pr}(u^C = 1|X) = \hat{\Pr}(u^{t_B} = 1|X)$, where u^B indicates union membership in a base year t_B.

Decomposing the Total Union Effect Unions can contribute to changes in inequality through two channels: first, changes in union membership over time; and second, changes to the union-non-union wage structure. For each time period, we further decompose the total union component into these respective "unionization" and "union wage" effects by considering an alternative counterfactual. For each time period t_B to t, we reweight year-t households to unionize as they would in year t_B:

$$\Pr(u^{C_B} = 1|X, t) \equiv \Pr(u = 1|X, t_B). \tag{F.3}$$

Applying this counterfactual to Equation F.2 allows us to generate weights by predicting year-t households' union status with year-t_B estimates of union-selection.[63] Applying these weights to year-t households allows us to separate Equation 4 into its respective subcomponents:

$$\begin{aligned}\Delta^U = &\underbrace{\left[\mathrm{Gini}(F_{Y_t}) - \mathrm{Gini}(F^{C_B}_{Y_t})\right]}_{\text{Unionization Effect}} \\ &+ \underbrace{\left(\left[\mathrm{Gini}\left(F^{C_B}_{Y_t}\right) - \mathrm{Gini}\left(F^{C_0}_{Y_t}\right)\right] - \left[\mathrm{Gini}(F_{Y_{t_B}}) - \mathrm{Gini}\left(F^{C_0}_{Y_{t_B}}\right)\right]\right)}_{\text{Union Wage Effect}},\end{aligned} \tag{F.4}$$

Ideally, we could compare the results of our decomposition to a similar exercise conducted using 1951 Palmer survey data by Callaway and Collins (2018), but they report all of their effects in percentile ratios. We are limited by only having binned income data in the years closest to 1951, so our percentile ratios are unstable. Therefore, we elect to use the Gini coefficient instead. Nonetheless, our results are qualitatively consistent with theirs: union members are negatively selected, and the union premium is larger for otherwise lower-wage workers. We can infer from these results that unions exercised a considerable compressing effect. In Callaway and Collins (2018), the reduction in inequality amounts to 16-24 percent across per-

63. The union selection equation in the base year is estimated using logistic regression of household union membership against education, race, a quadratic in respondent age, and state fixed effects. When 1936 is the base year, we replace state fixed effects with region fixed effects, as incomplete coverage in the 1936 Expenditure Survey means many states' fixed effects cannot be identified in that year.

centile ratios in their 1951 urban wage-earners sample, while we observe a 5 and 7.6 percent decrease in the Gini coefficient in household income in 1947 and 1960, respectively.

Incorporating spillover effects One limitation of the standard DFL reweighting procedure is that it uses observed non-union wages to simulate de-unionization, assuming that changes in unionization have no spillover effects. To relax this assumption, we adopt the distributional-regression strategy developed by Fortin, Lemieux, and Lloyd (2018). Specifically, we model the year-t likelihood of household income falling between quantiles k and $k+1$ for each of twenty-five income quantiles:

$$p_k\left(X_{it}, U_{sjt}, y_k\right) \equiv \Pr(y_k \leq Y_{it} < y_{k+1} | X_{it}, U_{sjt}) \text{ for k = 1,...,K,} \tag{F.5}$$

where Y_{it} denotes realized household income, y_k denotes income at the kth quantile, X_{it} denotes household demographics (including union status), and U_{sjt} denotes the share of unionized workers in state s and industry j at year t.[64] $p_k(\cdot)$ is estimated separately for union and non-union households using a heteroskedastic-robust ordered probit model:

$$\Pr\left(Y_{it} \geq y_k | X_{it}, U_{sjt}\right) = \Phi\left(X_{it}\beta + y_k X_{it}\lambda + \sum_{m=0}^{4}\left[y_k^m U_{sjt}\phi_m\right] - c_k\right). \tag{F.6}$$

We then construct a spillover reweighting factor, ξ_{ik}, which captures the change in the likelihood of falling into income bin k one would experience if their state(-industry) union share were at some counterfactual level U_{sjt}^C:

$$\xi_{ik} = \frac{\hat{\Pr}(y_k \leq Y_{it} < y_{k+1} | X_{it}, U_{sjt}^C)}{\hat{\Pr}(y_k \leq Y_{it} < y_{k+1} | X_{it}, U_{sjt})} = \frac{\hat{p}_k\left(X_{it}, U_{sjt}^C, y_k\right)}{\hat{p}_k\left(X_{it}, U_{sjt}, y_k\right)} \tag{F.7}$$

We then generate predicted probabilities for each household in year-t using true and counterfactual union densities in their state or state-industry. For the "within-year" impact of spillovers shown in the dashed lines of Figures F.1b, F.1c, and F.1d, these counterfactual union shares are simply set zero, $U_{sjt}^{C_0} = 0$. For the spillover-adjusted

64. Because we lack panel data on households' industries prior to 1977, we use state union shares rather than state-industry union shares in earlier years. Similarly, incomplete state coverage and absence of year variation prevents us from estimating any spillover effects prior to the 1960s.

unionization components of the decompositions reported in Table F.2, we generate predictions using state- or state-industry-level unionization rates from the *base* year, $U_{sjt}^{C_0} = U_{sjt_B}$.[65] Finally, we adjust the counterfactual income distributions from Section V.A by simply multiplying a given household's union-selection weighting factor, Ψ_i, by the spillover weight $\hat{\xi}_{ik}$ corresponding to the income bin k_i in which it falls. The result is an income distribution that looks as though individuals unionized as they did in year t_B *and* received the spillover benefits of year-t_B unions.

Appendix Table F.2 shows the results of the decomposition, with and without spillovers. The effects of unions are again large for the 1936-1968 period, but are small for the recent period. This result, as well as the relatively small effect of unions on household income inequality in the recent period is in contrast with DiNardo, Fortin, and Lemieux (1996) and Fortin, Lemieux, and Lloyd (2018) who find both larger effects of unionization and larger effects of spillovers in the recent period. As Appendix Table F.3 shows, the difference is primarily due to the inequality concept and population being used, rather than the differences in the selection equation. We use household income inequality, while DiNardo, Fortin, and Lemieux (1996) use individual earnings inequality, and often focus on men. Changing household composition, female labor force participation, and wealth inequality are just some of the forces affecting household income inquality that would be missed in simply looking at individual male earnings. The divergence between household and individual inequality changes is smaller in the early part of our sample than the latter part: the top 10% measured by individual income in Piketty, Saez, and Zucman (2018) between 1936 and 1968 is 13.2, while it is 12.4 when measured in tax units (which are closer to our notion of households), while the change in top 10% between 1968 and 2014 is 8.6 when measured at the individual level while it is 12.4 when measured at the tax unit level.

65. For year-t households in states or state-industry pairs not represented in the base year, we predict their counterfactual union shares using predictions from a regression of union shares against a quadratic time trend and state-specific linear time trends. When year-t includes industry information, we include industry-specific time trends in the regression and interpolate early state-industry shares using industry-level density estimates from Troy (1965) reweighted by employment shares from IPUMS, following Collins and Niemesh (2019).

APPENDIX TABLE F.1: YEARLY UNION IMPACT AND UNION DENSITY: $\theta_{Gini} \equiv$ GINI - CF GINI

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	θ_{Gini}	θ_{Gini}	θ_{Gini}	θ_{Gini}	Gini	Gini	Gini
Union Density	-0.0875***	-0.0663***	-0.0891***	-0.0761**	-0.304***	-0.118***	-0.105***
	(0.00829)	(0.0176)	(0.0237)	(0.0227)	(0.0678)	(0.0198)	(0.0192)
College Share				0.0426*			0.0440*
				(0.0206)			(0.0166)
CF Gini						0.864***	0.863***
						(0.0317)	(0.0349)
Linear Time Trend?	No	Yes	Yes	Yes	Yes	Yes	Yes
Quadratic Time Trend?	No	No	Yes	Yes	Yes	Yes	Yes
R-squared	0.767	0.791	0.802	0.816	0.948	0.997	0.998
N	57	57	57	57	57	57	57

Note: This table reports OLS estimates of the marginal response of the Gini coefficient to historical changes in union density, adjusting for observable changes in the population via the counterfactual-weighting procedure described in Section V.A. Columns 1-2 report coefficients from an OLS regression of yearly union impact, $\nu(F_{Yt}) - \nu\left(\hat{F}_{Y_{nt}}\right)$, against the yearly unionization rate. Columns 4 and 5 report coefficients from alternative specifications, which put $\nu\left(F^C_{Y^n_t}\right)$ on the right-hand side. Robust standard errors are shown in parentheses. $^{*}p = 0.1, ^{**}p = 0.05, ^{***}p = 0.01$.

Appendix Table F.2: The Impact of Unionization with and without Spillovers

	Time Period	*Total Change in Statistic*	*Unionization Component*	
			no spillovers	w/spillovers
	(1)	(2)	(3)	(4)
Panel A: Gini	1936 to 1968	-0.0526	-0.0149	-0.0188
			(28.37)	(35.74)
	1968 to 2014	0.144	0.00587	0.00723
			(4.075)	(5.016)
Panel B: 90/10	1936 to 1968	-0.188	-0.0980	-0.135
			(52.17)	(71.83)
	1968 to 2014	0.817	0.0494	0.0417
			(6.041)	(5.097)
Panel C: 90/50	1936 to 1968	-0.102	-0.0328	-0.0455
			(31.99)	(44.45)
	1968 to 2014	0.360	0.0281	0.0258
			(7.818)	(7.183)
Panel D: 10/50	1936 to 1968	0.0855	0.0653	0.0895
			(76.33)	(104.6)
	1968 to 2014	-0.458	-0.0213	-0.0158
			(4.644)	(3.457)

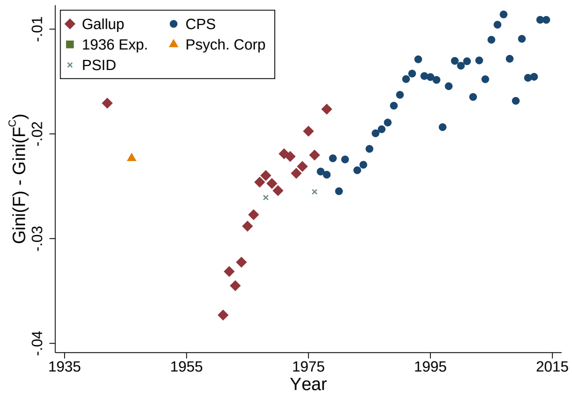
Note: This table reports the union-related components of decompositions of changes in Gini coefficient over time with and without spillovers estimated as described in Appendix F. Each row represents a separate decomposition. Column 1 specifies the beginning and end years of the decomposition. Column 2 reports the total change in computed Gini coefficient. Column 3 reports the change in Gini attributable to changes in union versus non-union incomes. Column 4 reports the change in Gini attributable to changes in the conditional unionization rate. Column 5 reports the total effect of both union wage changes and unionization (Column 3 + Column 4). Numbers in parentheses report components as a percentage of total change in Gini coefficient.

Appendix Table F.3: Decomposition of Change in Gini (CPS) from Individual to Household Measure

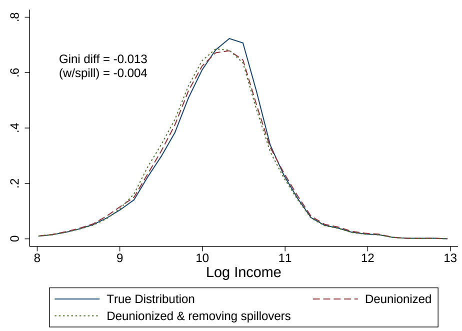
	Time Period	*Total Change in Statistic*	*Change Attributable to:* Δ Union Wages	Δ Unionization	Total Union Component
	(1)	(2)	(3)	(4)	(5)
Individual Union Status and Earnings, Men Only	1979 to 2014	0.0925	0.00566	0.00928	0.0149
			(6.117)	(10.03)	(16.15)
	1981 to 1988	0.0176	0.000816	0.00506	0.00587
			(4.631)	(28.67)	(33.30)
	1990 to 2014	0.0268	0.00311	0.00467	0.00778
			(11.60)	(17.41)	(29.01)
Individual Union Status and Earnings, Men and Women	1979 to 2014	0.0590	0.00536	0.00533	0.0107
			(9.091)	(9.032)	(18.12)
	1981 to 1988	0.00890	0.000462	0.00359	0.00405
			(5.191)	(40.37)	(45.56)
	1990 to 2014	0.0209	0.00361	0.00313	0.00673
			(17.23)	(14.95)	(32.18)
HH Union Status and Individual Earnings, Men and Women	1979 to 2014	0.0590	0.00141	0.00482	0.00623
			(2.391)	(8.174)	(10.57)
	1981 to 1988	0.00890	-0.000507	0.00264	0.00213
			(-5.698)	(29.67)	(23.97)
	1990 to 2014	0.0209	0.00410	0.00302	0.00711
			(19.58)	(14.41)	(33.99)
Household Union Status and Income	1979 to 2014	0.102	0.00300	0.00842	0.0114
			(2.929)	(8.223)	(11.15)
	1981 to 1988	0.0476	-0.00327	0.00463	0.00136
			(-6.880)	(9.729)	(2.850)
	1990 to 2014	0.0730	0.00372	0.00264	0.00636
			(5.090)	(3.612)	(8.703)

Note: This table reports the contribution of unions to inequality in different CPS samples, showing how the population, income, and union measure affect the decomposition. The top row shows the results for just individual male workers, with unionization and earnings measured at the individual level. Row 2 adds women. Row 3 changes the definition of union to be the household measure we use in the main text, but keeps earnings measured at the individual level. Row 4 then changes the measure to be household income, and changes the population to be households rather than individuals.

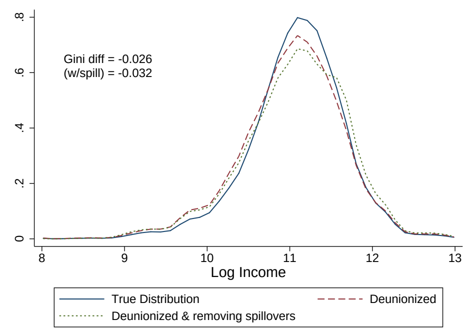
APPENDIX FIGURE F.1: INCOME DISTRIBUTIONS: TRUE VS. NO-UNIONS COUNTERFACTUAL



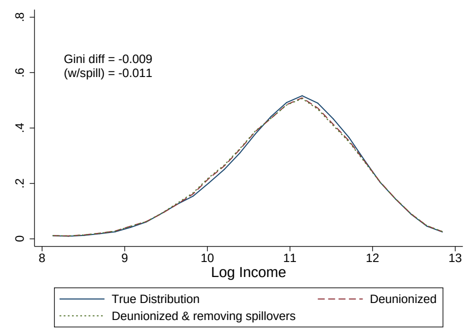
(A) YEARLY UNION IMPACT (ASSUMING NO SPILLOVERS)



(B) 1936



(C) 1968 (PSID)



(D) 2014 (CPS)

Note: This figure compares the observed population (F_Y) and the counterfactual population without unions (F_{Y_n}) in selected years. Panel F.1a plots yearly differences in true and counterfactual Gini coefficients. Panels F.1b, F.1c, and F.1d plot kernel-density estimates of true and counterfactual log-income distributions for selected years. Spillovers are estimated using state-year-industry level union density in the CPS and state-year union density in the other samples, imputed where necessary. Income is denominated in 2014 dollars using CPI.

References

Callaway, Brantly and William J Collins (2018). "Unions, workers, and wages at the peak of the American labor movement". *Explorations in Economic History* 68, pp. 95–118.

Collins, William J and Gregory T Niemesh (2019). "Unions and the Great Compression of wage inequality in the US at mid-century: evidence from local labour markets". *The Economic History Review* 72.2, pp. 691–715.

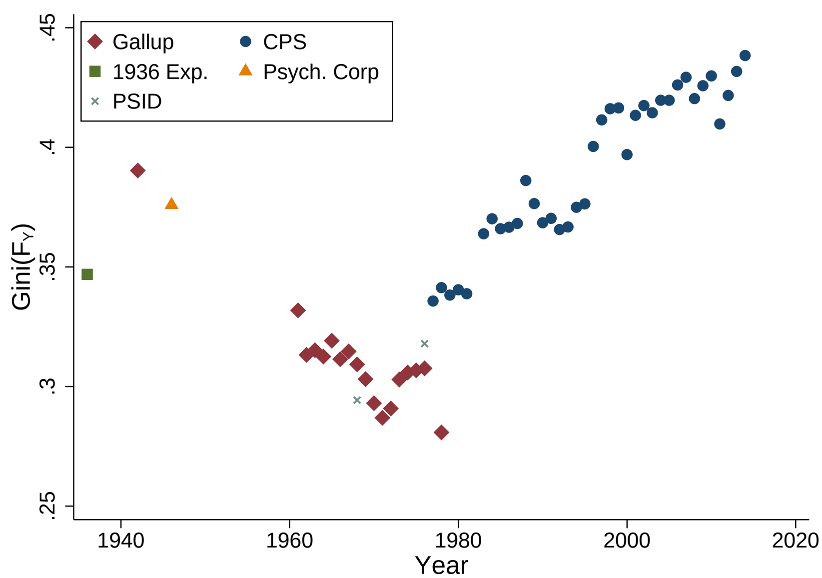
DiNardo, John, Nicole M Fortin, and Thomas Lemieux (1996). "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach". *Econometrica* 64.5, pp. 1001–1044.

Fortin, Nicole, Thomas Lemieux, and Neil Lloyd (2018). *Labor Market Institutions and the Distribution of Wages: The Role of Spillover Effects*. Tech. rep. Working Paper.

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman (2018). "Distributional national accounts: methods and estimates for the United States". *The Quarterly Journal of Economics* 133.2, pp. 553–609.

Troy, Leo (1965). *Trade Union Membership, 1897–1962*. NBER.

Appendix Figure F.2: Gini Coefficient In Survey Data Over Time



Note: This figure reports the Gini coefficient in each year, computed using Gallup, ANES, and CPS data.

G. Detailed IV analysis

As we demonstrate in Section V of the main text, there is a robust, negative relationship between union density and a variety of inequality measures, both at the aggregate time-series level (Section V.B) and at the state-year level (Section V.C). In this Appendix, we provide a more detailed treatment of the IV analysis summarized in Section V.D of the main paper. We focus on two key policy shocks that take place in the 1930s and 1940s, both of which had large but differential effects across states, allowing for identification of the effects of changes in state-level union density on changes in state-level measures of inequality. We begin by presenting historical details on the relevance of the two policy changes and qualitative evidence on the plausibility of the exclusion restriction. We then present our first-stage and 2SLS estimates, followed by a variety of econometric checks on our identification assumptions.

G.1. Two policy shocks that increased union density

We make use of two historical policies that together spurred a substantial increase in union density over a short, roughly ten-year period of time. First, we use the legalization of union organizing itself via the 1935 National Labor Relations Act (the NLRA, or "Wagner Act") and the 5-4 Supreme Court decision that upheld its constitutionality in 1937. As we will show, these events are associated with a modest increase in strike activity but a much larger increase in the probability of a strike's success, as well as a large increase in union members via the Act's establishment of a union recognition process via the National Labor Relations Board (NLRB). *We construct our "Wagner Act" shock as follows: new union members, by state, added from 1935 to 1938 via NLRB elections and successful recognition strikes, divided by 1930 state population.*[66] Appendix Figure G.1 shows a map of U.S. states, grouped by level of the "Wagner Act" shock.

66. For the NLRB elections data, we thank Ethan Kaplan. The strikes data come from The Labor Fact Book, a publication of *Labor Research Associates (LRA)*, which was a labor journal that operated from the 1930s through the 1960s. *The Labor Fact Books* only record *large* strikes, but unlike BLS strike measures they allow us to tabulate successful union recognition strikes by *state*, obviously crucial to our state-year analysis. Where multiple states are listed we assign them equally, but have also experimented with allocation based on share manufacturing. Note that BLS reports also records much of this information (whether a strike is for union recognition or some other goal, the strike's outcome, the state, etc.), but all in *separate* tables, and thus constructing cross-tabulations by state is not possible.

In the midst of this new legal opportunity for union organizing, Germany invaded Poland in September of 1939, marking the start of World War II in Europe. By the spring of 1940, the war created enormous U.S. government demand for military equipment to aid the Allied cause. Between 1940 and 1945, the federal government mobilized much of the country's industrial capacity for war production, spending $340 billion on national defense (or over three times the nominal GDP in 1940).[67] Because the war coincided with unprecedented union power, important concessions were made to labor in exchange for its cooperation. First, Roosevelt announced in 1940 that only firms that were NLRA-compliant would receive defense contracts from the National Defense Advisory Commission.[68] Second, when the US enters the war after Pearl Harbor as a military combatant, the newly established National War Labor Board (NWLB) imposes automatic enrollment and maintenance-of-membership at any firm receiving war-related production orders: if the firm was unionized, then any new worker would be default-enrolled into the union upon starting a job and would be maintained as a union member. He would only have a 15-day window to dis-enroll, but "few workers took the initiative to withdraw from the union in their first hectic weeks on the job."[69] Third, the NWLB allowed unions to have dues automatically deducted from members' paychecks ("dues checkoff"), eliminating the onerous practice of nagging members in-person for late dues and creating for the first time a steady source of revenue for unions. As we show in Appendix Figure G.5, unions managed to organize other "superstar" firms of the time during the war. Further, the new union members brought in under these policies were disproportionately low education and, as we show in Appendix Figure G.6, Black, who have larger union premia than the average union member.

Given the thumb the government put on the scale in unions' favor in war-related industries, we posit that the more defense contracts a state received during the war, the more union density grew. *We construct our "war spending shock" as follows: total 1940-1945 military spending by state, divided by state population.*[70] Appendix Figure G.2 shows a map of U.S. states, grouped by level of the "war-spending" shock. The map is quite similar to Appendix Figure G.1, and indeed, as we show more

67. See Brunet (2018).

68. This convinced even the staunchly anti-union Henry Ford to recognize the United Auto Workers (UAW) in 1941, lest he lose out on these enormous defense contracts. See chapter six of Loomis (2018).

69. See Lichtenstein (2003), Kindle Location 1415.

70. We use newly digitized war-era military supply contract data to construct per capita 1940-1945 war spending for each state. This measure is in 1942 dollars. We are very grateful to Gillian Brunet for sharing these data.

directly in Appendix Figure G.3, the two policy shocks are highly correlated across states.

G.2. *First-stage relationship between the policy shocks and union density*

G.2.1. Results in changes We take two approaches to documenting the first-stage relationship and other results. First, we examine results in changes within state. We take care to avoid years during the Second World War itself because war-specific institutions (most obviously wage controls, which were not fully lifted until 1946) could have a direct effect on inequality. We are also limited by data availability, particularly that of the endogenous variable, state-year union density, which we only have in 1929 and then from 1937 onward. These constraints lead us to estimate the following first-stage equation:

(G.1)

$$\begin{aligned}Union_{st} - Union_{s,t-9} =& \beta_1 \textit{Wagner shock}_s \times \mathbb{I}_t^{t=1938} + \beta_2 \textit{War-spending shock}_s \times \mathbb{I}_t^{t=1947} \\ &+ \gamma_1 \textit{Wagner shock}_s + \gamma_2 \textit{War-spending shock}_s \\ &+ \lambda_{r(s)t} + \gamma\left(log\left(\frac{N_{st}^{Col}}{N_{st}^{HS}}\right) - log\left(\frac{N_{s,t-9}^{Col}}{N_{s,t-9}^{HS}}\right)\right) \\ &+ \eta \mathbb{X}_{st} + e_{st},\end{aligned}$$

where the outcome variable is a *nine-year* change in union density in state s, $Wagner\, shock_s$ is the per capita number of new members added via NLRB elections and recognition strikes from 1935-1938 in state s, $\mathbb{I}_t^{t=1938}$ is an indicator variable for year $t = 1938$ (so, an interaction term that turns on for the 1929-1938 interval), $\mathbb{I}_t^{t=1947}$ is an indicator for year $t = 1947$ (so, for the 1938-1947 interval), $\lambda_{r(s)t}$ are Census region-by-year fixed effects, and $\mathbb{X}_{st}$ are other controls that we vary to probe robustness. Using nine-year intervals may seem odd, but it is done intentionally. Our data constraints (i.e. missing state-level union density from 1930-1936) plus our desire to avoid any year with war-related wage controls means that intervals included in this regression are 1929-1938, 1938-1947, 1947-1956, and so on until the end of our sample in 2014. The nine-year intervals allow us to skirt the wage-control period (which ends in 1946) and make use of our only year of pre-Wagner state density data, 1929.

Appendix Table G.1 shows the results of estimating equation (G.1). Col. (1) is our preferred specification and shows that the two interaction terms substantially shift upward union density in the appropriate window (i.e., the Wagner-Act shock during the 1929-1937 window and the war-spending shock during the 1938-1947 window). Importantly, the main effects of the Wagner and war-spending variables are not significant, meaning that outside of the specific windows captured by the interaction terms, Wagner and war-spending states are not predisposed to union-density growth. The associated F-statistic is also well above the rule-of-thumb cut-off value.

The rest of the table examines robustness. Col. (2) adds state fixed effects. Since the regression is in changes, adding state fixed effects is analogous to adding state-specific trends in an in-levels regression. Col. (3) weights the state-year observations by 1930 state population, and col. (4) drops Michigan (the outlier for both policy shocks), all with minimal effect on the coefficients of interest. Col. (5) adds interactions of each policy shock with the "wrong" window to the col. 1 regression—a demanding specification check given the high correlation between the two variables, as shown earlier in Appendix Figure G.3. While the standard errors on the variables of interest increase somewhat, the point-estimates are quite stable. Moreover, the coefficients on the "wrong" interactions are insignificant: the effect of the Wagner shock is only significant in the earlier window and that of the war shock only significant in the later window.

Nonetheless, it is clear from the changes in the coefficients and the fall in the F-statistic between columns (3) and (5) that the two shocks are highly correlated. Appendix Figure G.3 shows a scatter plot of our two policy shock variables. The figure shows, as expected, that Michigan (the birthplace of the modern U.S. labor movement in the 1930s and the "Arsenal of Democracy" during the war) is an outlier for both of the shocks. More generally, the two shocks have a correlation of 0.7, and so we pool the two shocks into a single state-level shock variable. Using this single instrument and interacting it with the two treatment windows gives similar results, as shown in column (6) of Table G.1.

G.2.2. Results in levels The second approach we take is more graphical and non-parametric: we simply regress state-year union density (in levels) on the pooled policy shock variable, separately in each year. Instead of using nine-year intervals to avoid the war and specifying in which windows we expect the see effects, we plot the relationship in each year and observe whether the changes emerge in the periods we predict.

In particular, we estimate:

$$Union_{st} = \sum_{y \in 1929, 1937 \ldots 2014} \beta_y IV_s \mathbb{I}_t^{t=y} + \lambda_{r(s)t} + \mathbb{X}_{st}\gamma + e_{st}, \tag{G.2}$$

where $Union_{st}$ is state-year density, IV_s is the time-invariant pooled policy shock variable for state s, $\mathbb{I}_t^{t=y}$ is an indicator variable for when year t is equal to year y, $\lambda_{r(s)t}$ is a vector of $region \times year$ fixed effects, and $\mathbb{X}_{st}$ is a vector of covariates that we vary to probe robustness, but always includes log skill shares $log\left(\frac{N_{st}^{Col}}{N_{st}^{HS}}\right)$. y is summed over all years for which we have a state-year union density estimate (i.e. 1929 and then 1937 onward). In our baseline estimation, $\mathbb{X}_{st}$ is omitted, and thus equation (G.2) is equivalent to regressing union density on the pooled IV and region fixed effects separately by each year of the sample period, and then plotting the resulting β_y values.

As the results are in levels, our hypothesis makes predictions about the *changes* in the relationship between union density and the pooled IV variable. We argue that the only time union density should exhibit a sustained change in its relationship with the IV is during the treatment period (1935 to the end of the war), and that the relationship should *increase*. We are agnostic as to the sign of the density-IV correlation before the treatment period, but we expect that it should increase from this level on during the treatment period.

The results from the baseline estimation are shown in Figure G.4. We only have pre-period data for 1929, but we see a large increase from 1929 to 1937. Unfortunately we cannot show the precise timing due to lack of data. The coefficient in 1929 is close to zero, showing that before the treatment period, states about to be hit by our policy shock variables were not historically union friendly. From 1937 onward we have annual data, and the relationship between the IV and union density increases steadily during the remainder of the treatment period. Afterwards, we see no sustained increase but also no back-sliding, suggesting that the states hit by the policy shock variables retain (relative to other states) greater density levels even after the war ends.

G.3. Are the policy shocks plausibly exogenous?

Appendix Table G.1 and Figure G.7 show that our shocks appear to have a strong, first-stage effect on union density, but of course they do not speak to whether the shocks provide a valid experiment. In arguing that these policy shocks provide quasi-exogenous variation in union density, we never claim that they hit a random subset of states. Indeed, states with larger IV values (i.e., those that gained more union members via strikes and elections in the mid and late 1930s as well as received more dollars per capita of government war contracts) were different in important ways from other states. Table G.2 uses the 1920 Census to examine what state-level characteristics predict the pooled IV variable. By far the strongest predictor is the manufacturing share of employment in the state. Not only is pre-period manufacturing a key predictor of the IV, but the manufacturing sector is key to the first-stage of the IV as well, as we are arguing that the government taking over manufacturing production during World War II was the driving mechanism for why war spending increased union density in a state. For these reasons, we will give special attention to the potential confound of manufacturing in Section G.5.1.

The rest of this section provides historical context for the two policy shocks, which helps establish their validity as sources of identification.

G.3.1. The "Wagner shock" The historical consensus, both from contemporaneous accounts as well as more modern assessments, argues that the decision of the federal government to no longer intervene on the side of employers—*not* a sudden increase in union demand among workers—led to the historic gains in density immediately after the Wagner Act's passage. Employers had considerable latitude, both legal and extra-legal, in combating unions before Wagner. Firms put down strikes and other organizing activity with an array of raw paramilitary power and espionage, and if needed, military assistance from the state. White, 2016 describes the weapons the major steel companies stockpiled to deter or put down organizing activity: "[T]he major steel companies had evolved potent systems of labor repression that included political and legal resources as well as extensive police forces and stockpiles of armaments....massive arsenal[s] of firearms and gas weapons." Henry Ford not only commanded a "brutal private army", but also paid an espionage force of over 1,000 employees to spy on fellow workers and report back any hints of organizing activity.[71]

A final recourse for firms was the power of the state. Prior to the NLRA, the coercive powers of the American government, at all levels, were regularly used against organized labor, with military deployments and judicial repression commonplace (Naidu and Yuchtman, 2018). Riker (1979) documents that the most frequent domestic use of the national military in the nineteenth century was to put down labor unrest. As late as summer 1934 the national guard was called in to put down major strikes in Toledo and Minneapolis, as well as a general strike of West-coast dockworkers lead by the Teamsters. In all cases the national guard succeeded after pitched street battles.

The Wagner Act legally protected collective actions such as picketing and strikes, bypassed judicial injunctions, and mandated resources for independent enforcement

71. The "private army" quote is from Loomis (2018, p. 122), and Lichtenstein (1995) discusses anti-union espionage at Ford.

of organizing rights. It was this policy shift, not an increase in union organizing, that led to the sudden gains in the second half of the 1930s. Writing about the 1937 Flint sit-down strike (which led to GM's official recognition of the UAW), Lichtenstein, 1995 notes that: "The UAW victory was possible not so much because of the vast outpouring of union sentiment among autoworkers, but because General Motors was temporarily denied recourse to the police power of the state." Taking a more modern perspective, Loomis (2018) agrees: "[T]he government played a critical role in determining Flint's outcome. Ten years earlier, with the stridently anti-union Calvin Coolidge as president, the outcome would likely have turned out very different, no matter what the Flint strikers did."

We provide two pieces of evidence on strikes in support of historians' contention that organizing successes immediately after Wagner's passage did not stem from an increase in grass-roots organizing activity, but rather a top-down change in the rules government used to referee management-labor relations. We treat strikes as a proxy for labor activism and mobilization. First, zooming in on the period immediately before and after the Wagner Act passes, we show in Appendix Figure G.8 that strike activity increases only modestly upon passage of the Wagner Act. We also show in Appendix Figure G.5 that strike activity increases only modestly upon passage of the Wagner Act. Although leaders in the CIO urged their colleagues to "seize the once-in-a-lifetime organizing opportunities so evident in the mid-1930s" (Lichtenstein, 2003)[72], strike activity only rises by twenty percent. Nor do their goals change remarkably, as there is only a modest uptick (15 percent) in the share of strikes for which union recognition is a key goal.

The most dramatic change is the share of strikes that are *successful*, which increases from just over twenty percent to forty percent. This time-series evidence supports the conclusion of White (2016) that "poverty and resentments alone did not undermine the open shop. The surge of unionization was influenced by the arrival from above of a new political economy premised on greater regulation of industrial production by the federal government."

Appendix Figure G.8 only speaks to national time-series evidence; it is possible that organizing activity shifted toward union-friendly states after Wagner, in violation of our identification story. By constrast, Appendix Figure G.9 shows that the relationship between our Wagner variable and state strike activity is roughly constant since 1914 (the first year of state-level strikes data). Essentially, the same

72. As further evidence that the modest increase in organizing was likely endogenous to the NLRA, the CIO, with its unprecedented focus on organizing industrial workers, was only formed as a committee within the AFL six months after the NLRA's passage.

states were striking before and after the Wagner Act, but only met with success after its passage.

This steady relationship supports the reading that the geographic variation in post-NLRA density gains can be modeled as arising from (a) *constant* differences in latent union demand at the state level interacted with (b) a national policy shock in 1935 that allowed that demand to translate into density gains. Latent union demand likely comes from industrial structure (such as high fixed-cost capital investments and product market power enabling workers to capture rents) or cultural and ideological differences across states. Political scientists and sociologists (Davis, 1999; Eidlin, 2018; Goldfield, 1989) who study the period emphasize the role of persistent communities and networks of highly ideological labor activists pushing for strikes and other forms of collective action even when success was impossible. Appendix Figure G.9 supports these arguments. If, as we claim, the geographic variation in post-Wagner gains in density are explained by the interaction between longstanding differences in demand for unions in certain localities and a shift in the federal government's position on the legality of organizing, then it should be possible to construct an alternative IV using *earlier* episodes of union demand interacted with the treatment period. We perform this exercise in Section G.5.2.

G.3.2. The "war shock" While we will perform extensive robustness tests later in this Appendix, here we provide evidence from existing work that per capita war spending is plausibly exogenous to other factors that could shape inequality.

Brunet (2018), whose war-spending data we in fact use to construct our war-shock measure, shows that war spending had only a modest state-level fiscal multiplier (0.25 to 0.3).[73] She conducts a battery of tests showing that war spending was independent of a variety of other state-level changes during World War II. For example, she shows that war spending was not correlated with increases in government employment, nor was it targeted to places with more available labor (e.g., those states with lower pre-war employment levels). These results foreshadow the success of our robustness checks in Section G.5, in that flexibly controlling for a variety of state-level characteristics typically has little effect on our main results.

Furthermore, the war contracts did not radically change the geography of American industry; contracts favored existing manufacturing firms and their subcontractors. As we will show in Section G.5.1, any differential increase in manufac-

73. This result echoes Fishback and Cullen (2013), who find that war spending at the county-level led to some modest population growth, but limited if any sustained per capita economic growth.

turing employment correlated with the IV was extremely short-lived (disappearing by 1946), and states that received more war contracts do not subsequently show faster growth in manufacturing employment after the war ends. Much of war production involved *conversion* of existing factories, and as such not substantially the expanding overall manufacturing share of employment. Yet, even in states that built new factories to accommodate the demands of war production, such as those in the South, manufacturing employment rapidly returned to baseline and did not gain a solid foothold until decades later (Jaworski, 2017).

Finally, Rhode, Snyder Jr, and Strumpf (2017) show that during the war, defense contracts were free of the usual political considerations. They find that the electoral importance of a state did not predict the volume of its war contracts, perhaps because contracts were drawn up directly by military, not Congressional or White House, agencies.[74]

G.4. Main IV results

G.4.1. Results in changes We begin with the two-stage-least-squares (2SLS) analogue of our first-stage results in Table G.1, with $Wagner\,shock_s \times \mathbb{I}_t^{t=1938}$ and $War\,shock_s \times \mathbb{I}_t^{t=1947}$ as the two excluded instruments.

The first six columns of Appendix Table G.3 show results when the top-ten income share is the outcome, following the same specifications as in Appendix Table G.1. Our preferred estimate in col. (1) suggests that a ten-percentage-point increase in state union density decreases the top-ten share by roughly 6.2 percentage points, with the point-estimates from other specifications ranging from 3.6 to 8.1 percentage points.

The remaining six columns of Appendix Table G.3 show analogous results using the state labor-share as the outcome, with our preferred estimate indicating a 3.6 percentage point increase from a ten percentage-point increase in density. The remaining specifications cluster quite tightly around this baseline result.

For completeness, Appendix Table G.4 shows the corresponding reduced form specifications. Reassuringly, both instruments have independently significant effects on both labor share and top ten share in most specifications that include the two together, with the war shock having a larger reduced form effect than the Wagner Act shock.

74. In his memoirs, Donald Nelson, the chairman of the War Production Board, frequently emphasizes the importance of ensuring that production orders came directly from the military and were free of interference from civilian authorities. See Nelson (1946).

G.4.2. Results in levels As we did with the first-stage results, we also show annual results in levels. Again, predictions in this setting map to *changes* in the relationship between the pooled IV variable and the inequality outcomes. The only time when the relationship between the IV and our inequality outcomes should change is during the treatment period. One advantage of this approach over the 2SLS regressions is that we do not need to observe union density to plot the reduced-form relationship between our inequality outcomes and the pooled IV variable. We can thus look further back in time in the reduced form than we can in the first-stage.

The first series of Appendix Figure G.10 shows the relationship between the pooled IV and the top-ten income share from 1917 onward, using the same specification as we showed for the first-stage relationship in Appendix Figure G.7. The figure shows that in the pre-period, the pooled IV is associated with a *higher* share of income going to the richest ten percent, meaning states that would soon be hit by our pro-union policy shocks were not historically more egalitarian (in fact, the opposite), at least by this measure. While noisy, this positive pre-period relationship can generally be distinguished from zero each year and is largely unchanged until the mid- to late-1930s. It then begins a dramatic and sustained decline. By the start of the war in Europe, the sign of the relationship has flipped. The relationship slowly recovers some of its magnitude over the rest of the sample period, but the changes cannot be distinguished from zero in any of these years. The shape of the relationship between the pooled IV and the top-ten share echoes the results from Appendix Table G.3: the only period of sustained decrease in the relationship between top-ten inequality and the IV is during the treatment period.

The first series of Appendix Figure G.11 is the labor-share analogue of this analysis. It tells a very similar story, though data limitations shorten the pre-period relative to that of state top-ten inequality. In the early 1930s, our IV predicts a *lower* state-level labor share, again highlighting that states that would soon receive pro-union policy shocks were not historically worker-friendly. Over the treatment period, the sign of this relationship flips and then remains positive over the rest of the sample period. Again, the only period of sustained increase in the relationship between the state-year top ten and labor shares to the IV is during the treatment period.

G.5. Robustness checks

In this section, we rule out a number of potential violations of our exclusion restriction, which says that any other determinants of inequality are independent of the change in union density induced by our policy variables. Potential confounding variables include the change in manufacturing employment, omitted determinants of new unionization following the Wagner act, other policies such as taxes and minimum wages, and finally, any independent role of egalitarian norms or beliefs. In the subsections below we present evidence ruling out these alternative mechanisms.

G.5.1. Controlling for contemporaneous and pre-period difference in manufacturing We start with the role of manufacturing, which we view as the most important potential confound. As we showed in Appendix Table G.2, states that have a larger manufacturing share of employment in the pre-treatment period have larger values for our IV variable, so we have reason for concern.

The first three columns of Appendix Table G.5 show how our top-ten 2SLS results vary as we add manufacturing controls. The first column of this table reproduces the baseline result, col. 1 of Appendix Table G.3, for ease of comparison. In col. 2, including contemporaneous state manufacturing share of employment and its interaction with the two treatment windows reduces the first-stage F statistic somewhat and increases the coefficient on union density from 0.62 to 0.7. In col. 3, controlling for 1920-era manufacturing share of employment also reduces the first-stage F (to just below ten), with little effect on the second-stage point-estimate. Interestingly, while adding these controls for manufacturing employment weakens the first stage given its high correlation with the policy shock variables, contemporaneous or historical manufacturing employment does *not* appear to be an alternative mechanism for reducing top-ten-share inequality during our treatment periods. The coefficients on the interactions of both manufacturing variables with the two treatment windows are positive (significantly so for the first window), suggesting manufacturing-heavy states (all else, including the policy shock variables, equal) predicts *higher* inequality during our treatment period.

The first three columns of Appendix Table G.6 perform the parallel analysis when labor-share is the outcome. As expected, the effects on the first-stage are identical, though in the case of labor share the second-stage point-estimates are more stable, and the manufacturing controls and interactions have coefficients close to zero.

We perform similar robustness tests in Appendix Figures G.7, G.10 and G.11. These test demonstrate robustness of our estimated relationship between the pooled IV and union density, top-ten share, and labor share, respectively, to including the same controls for manufacturing employment. Echoing the results in the tables, the first stage is somewhat noisier, but follows the same general shape. While the reduced-form relationships between the inequality outcomes and the pooled IV sometimes shift *in levels*, the large changes that occur during our treatment period remain. We also control for pre-treatment agricultural share of employment, as it is a potential confound noted in Brunet (2018), with little effect on the estimates.

A final concern related to manufacturing is that the massive shift to producing the tanks, planes, and artillery needed for the war effort may have permanently transformed some states' manufacturing sectors, making it impossible to partial out any effect of the coincident rise of unions. Appendix Figure G.12 puts the manufacturing share of employment on the *left-hand side* of the analysis, exploring whether the shocks embedded in our pooled IV variable are associated with *permanent* changes in a state's manufacturing share of employment. While a positive blip can be observed for the few years of direct American combat involvement, the effect of the IV on state's manufacturing share completely disappears by 1946, whereas the effects on union density and inequality remain sticky. In fact, from 1910 to 1955 there is no sustained change in the relationship between a state's manufacturing employment and our IV variable: states with greater values for the IV are clearly more reliant on manufacturing employment, but the relationship is steady for over forty years. Beginning in the late 1950s, which is well after our treatment period, the relationship begins a slow and steady decline.

To summarize, our key findings are robust to controlling flexibly for contemporaneous manufacturing employment, as well as allowing pre-period differences in manufacturing employment to have a different effect in each year. These checks are important because of the strong positive relationship between the IV and state-level manufacturing employment. Moreover, the policy shocks we use as identification appear to have no lasting effect on states' manufacturing employment, consistent with the papers cited in Section G.3.2. States with large values for the IV are more manufacturing intensive before, during, and after our treatment period. It thus appears that manufacturing employment neither confounds nor mediates the relationship between the IV and union density or that between the IV and our inequality measures.

G.5.2. Using pre-treatment-period strikes as an alternative instrument We view the Wagner Shock (i.e., the number of union members gained in a state from 1935 to 1038 via recognition strikes and NLRB elections) as the second most serious threat to the IV analysis, considering that/given that it may be driven by local factors (e.g., friendly state governments, unobserved increases in local labor demand, or other local economic conditions) that might have their own independent effect on inequality. We do not observe coincident changes in the relationship between the pooled IV and Democrats in the governor's mansion. Appendix Figure G.13 shows that in fact there is no systematic relationships between the two variables over the course of our long sample period—it is possible that even within party, IV states during the treatment period enjoyed more worker-friendly political environments (or other local factors conducive to union organizing) in a manner difficult to observe.

We thus turn to a more comprehensive check on this possible endogeneity concern. As we showed in section G.3 and Appendix Figure G.9, states that gained the most union members immediately after the Wagner Act passed had long harbored the greatest latent demand for unions (at least as proxied by strike activity). Yet until the mid 1930s, this demand did not translate to greater density because the government consistently sided with management, with no formal protection of the right to organize.

Based on this logic, we substitute the Wagner shock in our IV with a measure of *pre-period* demand for unions: the (per capita) number of strikes in a state from 1921-1928, the years immediately before our first year of union density data in 1929. Whatever economic or political factors that might have contaminated the Wagner Act variable as an IV are unlikely to exist in this earlier period. While FDR was neutral if not friendly toward unions, Warren G. Harding's inauguration in 1921 ushered in an intense anti-union period at the federal level. Conversely, we might worry that union-friendly Democratic governors such as Michigan's Frank Murphy or Pennsylvania's George Howard Earle III played a role in the organization of industrial giants GM and U.S. Steel in the late 1930s, these states were controlled by Republicans in the 1920s. Finally, whatever local economic conditions prevailed in these states in the mid and late 1930s (specifically, the end of the Great Depression and the start of the Roosevelt Recession) are unlikely to reflect conditions during this pre-crash Roaring Twenties period. In summary, this measure reflects state-level demand for unions among workers (which we argue is long-standing and slow-moving), but is purged of any local effects specific to the mid- and late-1930s that may affect our outcomes of interest.

In Appendix Table G.7, we replicate the first-stage and 2SLS results using this measure of latent union demand instead of the Wagner shock. The war-spending shock remains unchanged. While the first-stage is less precise, the point-estimates are comparable to those in Appendix Table G.1, and the resulting 2SLS point estimates are also similar to their baseline estimates in Appendix Tables G.3.

G.5.3. Korean-War placebo tests Over 5 million U.S. military personnel served in the Korean War between 1950 and 1953, and as in World War II the government organized defense production to support the military campaign. As in World War II, the government issued wage and price controls during the conflict to address concerns that rising industrial production would spark inflation. In its geographic impact, defense production during the Korean War also mirrored that during World

War II. Appendix Figure G.14 shows that the correlation across states in per capita defense spending during the two conflicts was over 0.8, not surprising give certain states specialized in the production of ships, tanks or planes.

While industrial production during the two conflicts was similar in geographic impact and the use of price and wage controls, during the Korean War the federal government did *not* attach pro-union conditions to the receipt of defense contracts. In fact, perhaps due to the more antagonistic view of labor during this period (after the Taft-Hartley Act of 1947 and during the McCarthy era when many unions were being charged with communist sympathy), union leaders argued they were being excluded from the defense-production process during the Korean War (Stieber, 1980). In fact, in 1951, CIO representatives ended their participation in the Wage Stabilization Board with a dramatic walk-out.

For these reasons, the Korean War serves as a useful placebo test to determine whether defense production and wage stabilization alone (and not the pro-union policies that accompanied them during World War II) is sufficient to increase union density and reduce inequality. Appendix Figure G.15 shows that states that enjoyed Korean-War spending saw no increase in union density between 1954 and 1949 (the point-estimate is small and in fact "wrong"-signed). Similarly, the reduced-form relationship between our inequality measures and Korean-War related defense spending are also insignificant (Appendix Figures G.16 and G.17).

G.5.4. Other robustness checks The remaining rows of Appendix Tables G.5 and G.6 focus on robustness to other policies that might reduce inequality. Of course, these could be "bad controls" in that, say, greater union density might lead to states to increase the minimum wage or pass other worker-friendly policies. Nonetheless, robustness to these controls would help show the centrality of union density in moving our inequality measures during our treatment period. Furthermore, the 1930s and 1940s is a moment of historically active policy-making at the federal and state level, so it is important to show robustness to controlling flexibly for these policies.

Col. (4) of both tables adds as a control the share of tax units filing a federal income tax return in each state-year (and, as always, its interaction with the two treatment windows), as this share increases substantially during the war years and as such could have its own effect on the income distribution (a large public-finance literature shows that even pre-tax measures of inequality can be shaped by taxes). As we have alluded to already, local politics could be a confound, and col. (5) thus controls in the same manner for whether the state has a Democratic governor. The next two columns focus on state-level economic policy, in particular the minimum wage (which states can raise above the federal minimum) and a state-year "policy liberalism index" developed by Caughey and Warshaw (2016).

The next two columns refer to the local effects of major federal interventions. While our IV makes use of America's industrial support of the Allies, from December 1941 onward, the U.S. was also an active *military* partner, and the loss of so many working-age men to the armed forced may have had effects on labor markets during our key period. We thus control for mobilization rates by state from 1942 to 1945, and as usual its interactions with the treatment periods. In column (10), we control for per capita New Deal spending in each state in the same manner.

The final two columns adds additional state-year level covariates. Column (11) adds state top marginal tax rates on income, as described above, and Column (12) allows the state-year level measure of skill shares to have a separate effect in each treatment period, rather than a constant effect as in our main specification.

None of these controls meaningfully change the 2SLS coefficient for the labor-share outcome. The one outcome sensitive to these controls for the top-ten outcome is the IRS share, which is not surprising as the top-ten and the IRS share are drawn from the same data source and thus some mechanical correlation is likely present. Even so, it remains negative and significant. Moreover, none of these additional robustness checks reduce the first-stage F statistic below ten.

G.5.5. Did World War II create egalitarian norms? Finally, we consider a widely held view that the massive economic and military mobilization during World War II created lasting, egalitarian social norms that helped keep inequality in check for several decades.[75] If such sentiment came in part from actual war-related production, then it is a factor both correlated with our policy shock and related to inequality and thus threatens our identification.

We respond to this claim in three ways. First, we look at Gallup questions asking people how the war changed their views, in an attempt to see if aggregate changes in sentiment support the "egalitarian social norms" hypothesis. Our results are surprising (at least to us). We find no evidence that the war created the pro-labor or pro-worker sentiment that we would expect if egalitarian norms were an important constraint on inequality in the immediate post-war period. For example, in 1945, 56 percent of Gallup respondents tell pollsters that their view of labor unions is *worse* than before the war, while only 19 percent say the same of business owners and managers.[76]

75. Goldin and Margo (1992), Piketty and Saez (2003), and Goldin and Katz (2008) are among highly-cited works in economics that speculate as to the war creating egalitarian social norms.

76. In a March 1945 poll, Gallup asked: "Is your attitude toward labor unions today more

The Gallup question that is most directly related to how the war shaped respondents' views about fairness, deservingness, and income is from a June 1945 survey asking respondents both who they think has done the best financially during the war and who *should* be doing better?[77] There is an overwhelming consensus that workers have made out well, as 62 percent choose workers as the group that has done best, compared to only 19 percent that chooses white-color professionals and managers/owners of businesses. Moreover, 38 percent of Gallup subjects say that these well-off occupation groups *should* have done better during the war, compared to only nine percent saying the same about workers.

While these aggregate sentiments cast some *a priori* doubt on the egalitarian-social-norms hypothesis, our second response to the argument is to check if respondents in states hit with the two policy shocks are more likely to say that the war changed their views in a worker-friendly manner. In Table G.8 we regress a dummy variable coded as one if the respondent said they think workers and the poor should be doing better than they are against the pooled IV (col. 1), only the Wagner shock (col. 2), only the war-spending shock (col. 3), and both variables entered in the same regression (col. 4). In all cases, the coefficients of interest are very close to zero and insignificant. The remaining four columns perform the same exercise, but for the respondent saying that business owners/managers and professionals should be doing better. We again find small coefficients, with the only marginally significant results suggesting that respondents in Wagner-shock states are *more* sympathetic to business and professional interests.

Our third response considers a related "norms" argument: even if the war did not change Americans' stated views on what constitutes a fair income distribution, wartime wage structures altered worker reference points, and this process constrained post-war inequality (see, e.g., Kahneman, Knetsch, and Thaler (1986) on how respect for reference points constrains labor-market equilibria). The wages set by the NWLB and the 1942 Stabilization Act were more egalitarian than those that prevailed in

or less favorable than it was before the war?" to which 56 percent answered "less favorable," 24 percent "the same," and 20 percent "more favorable." Gallup asked in the same survey the analogous question, with "owners and managers of business concerns" in place of "labor unions." In response to this question, only 19 percent answered "less favorable," 49 percent "the same" and 32 percent "more favorable."

77. These are questions 10a and 10b from the June 1-5, 1945 survey. The wording of question 10a is "What class or group of people in this country has done best financially during the war compared to what they made before the war?" The follow-up question (10b) is: "Do you think any class or group of people in this country is NOT making as much money as it should? [capitalization in the original]."

the pre-war economy. While the government officially lifted them in 1946, workers and managers may have simply grown accustomed to this new, more compressed wage structure.

Yet, the immediate post-war years seem an unlikely moment for reference points or expectations to have much bite. First, inflation spiked briefly after the war, which should have quickly eroded any nominal wage stickiness.[78] Second, labor churn reached an all-time high after the war. U.S. military personnel shrunk from over 12 million in 1945 to only 1.5 million by 1947, meaning that over ten million Americans suddenly entered the potential labor supply.[79] Similarly, non-farm payroll contracted by two million (or by 4.9 percent) in the single month of September 1945, a record that would stand in both absolute and percentage terms until the Covid-19-related layoffs in April 2020.[80] Thus, even if workers had formed strong reference points concerning wages during the war, those workers may not have been in the same job or even still in the labor force a few years or even months later. Finally, the War Industries Board during World War I *also* imposed wage controls in war production, though without any of the pro-union policies that accompanied the World War II effort. If norms born from wage controls limit post-war inequality growth, we should have expected a similar, though muted, dampening of inequality in the years after the war, as U.S. involvement lasted only 19 months, compared to 44 in World War II. Instead, the 1920s ushered in historic growth of top-share income inequality.[81]

We thus conclude that in the immediate post-World-War-II era, unions were not particularly popular, and if anything war-era defense production had burnished the reputation of business over that of workers. Nevertheless, war-era policy made unions powerful (both in terms of millions of new members and solid revenue streams via automatic maintenance-of-membership and dues check-off), and over the next few decades they played an important role in maintaining historically low levels of inequality.

78. Annual inflation during the war years averaged 5.1 percent, and was even lower at 3.3 percent between 1943-1945, whereas it averaged over 11 percent in 1946-1947. See .

79. See Acemoglu, Autor, and Lyle (2004).

80. See .

81. Goldin and Margo (1992) note that skill premia appear to briefly compress during the First World War in the US but then quickly bounce back, and they also highlight the difference with the Second World War.

References

Acemoglu, Daron, David H Autor, and David Lyle (2004). "Women, war, and wages: The effect of female labor supply on the wage structure at midcentury". *Journal of political Economy* 112.3, pp. 497–551.

Brunet, Gillian (2018). "Stimulus on the Home Front: The State-Level Effects of WWII Spending".

Caughey, Devin and Christopher Warshaw (2016). "The dynamics of state policy liberalism, 1936–2014". *American Journal of Political Science* 60.4, pp. 899–913.

Davis, Mike (1999). *Prisoners of the American dream: Politics and economy in the history of the US working class*. Verso.

Eidlin, Barry (2018). *Labor and the Class Idea in the United States and Canada*. Cambridge University Press.

Fishback, Price and Joseph A Cullen (2013). "Second World War spending and local economic activity in US counties, 1939–58". *The Economic History Review* 66.4, pp. 975–992.

Goldfield, Michael (1989). "Worker insurgency, radical organization, and New Deal labor legislation". *American Political Science Review* 83.4, pp. 1257–1282.

Goldin, Claudia and Robert Margo (Feb. 1992). "The Great Compression: The Wage Structure in the United States at Mid-Century". *The Quarterly Journal of Economics* 107, pp. 1–34.

Goldin, Claudia Dale and Lawrence F Katz (2008). *The race between education and technology*. Harvard University Press.

Jaworski, Taylor (2017). "World War II and the Industrialization of the American South". *The Journal of Economic History* 77.4, pp. 1048–1082.

Kahneman, Daniel, Jack L Knetsch, and Richard Thaler (1986). "Fairness as a constraint on profit seeking: Entitlements in the market". *The American economic review*, pp. 728–741.

Lichtenstein, Nelson (1995). *Walter Reuther: The most dangerous man in Detroit*. University of Illinois Press.

— (2003). *Labor's War at Home: The CIO in World War II*. Temple University Press.

Loomis, Erik (2018). *A History of America in Ten Strikes*. The New Press.

Naidu, Suresh and Noam Yuchtman (2018). *Labor Market Institutions in the Gilded Age of American Economic History*. Tech. rep. National Bureau of Economic Research.

Nelson, Donald (1946). *Arsenal of Democracy: The story of American war production*. Harcourt, Brace and Company.

Piketty, Thomas and Emmanuel Saez (2003). "Income inequality in the United States, 1913–1998". *The Quarterly journal of economics* 118.1, pp. 1–41.

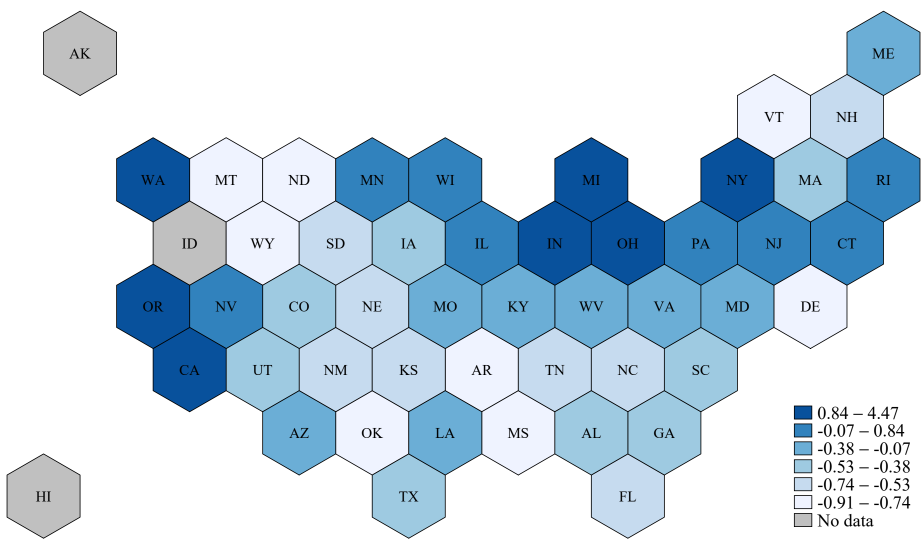
Rhode, Paul W, James M Snyder Jr, and Koleman Strumpf (2017). *The Arsenal of democracy: production and politics during WWII*. Tech. rep. National Bureau of Economic Research.

Riker, William H (1979). *Soldiers of the States*. Ayer Publishing.

Stieber, Jack (1980). "Labor's Walkout from the Korean War Wage Stabilization Board". *Labor History* 21.2, pp. 239–260.

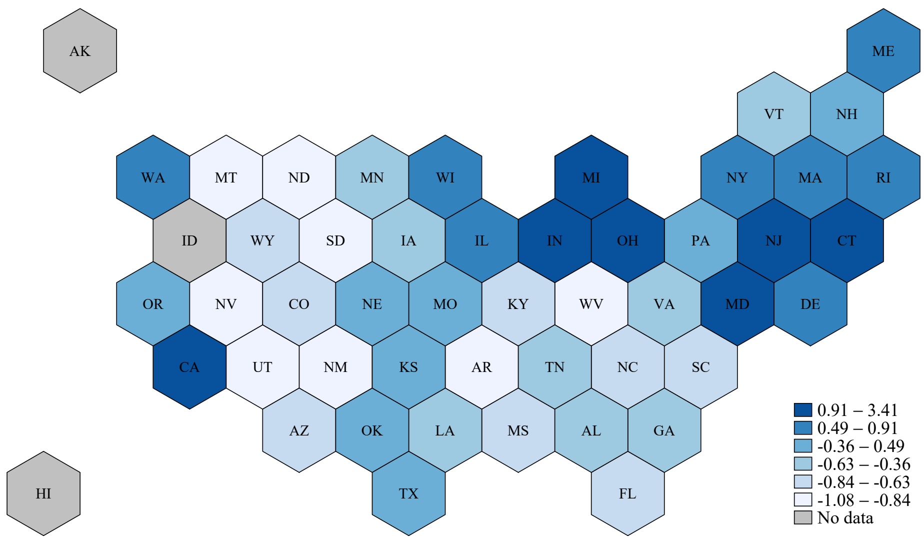
White, Ahmed (2016). *The last great strike: Little Steel, the CIO, and the struggle for labor rights in New Deal America*. Univ of California Press.

Appendix Figure G.1: Map of states by levels of the "Wagner" policy shock



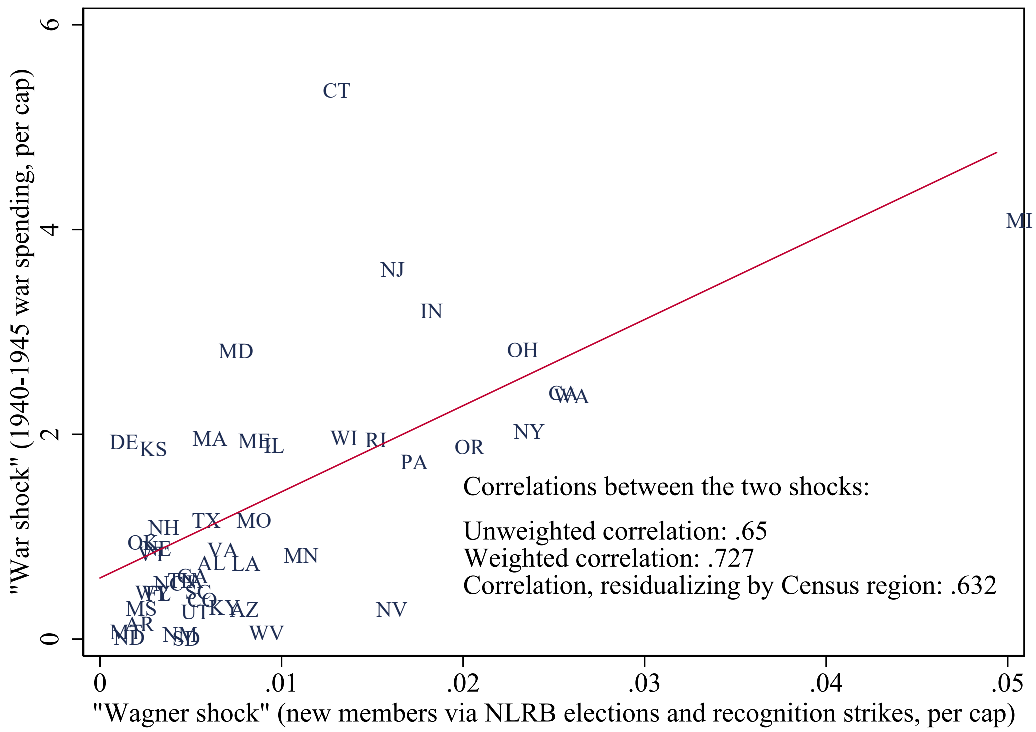
Notes: The "Wagner" policy shock is defined as the number of union members added from 1935 to 1938 via NLRB elections and successful recognition strikes, divided by 1930 state population. We then standardize this measure (subtract the mean and divide by the standard deviation).

APPENDIX FIGURE G.2: MAP OF STATES BY LEVELS OF THE "WAR-SPENDING" POLICY SHOCK



Notes: The "war-spending" policy shock is defined as the value of World-War-II defense contracts (from 1940-1945) divided by 1930 state population. We then standardize this measure (subtract the mean and divide by the standard deviation).

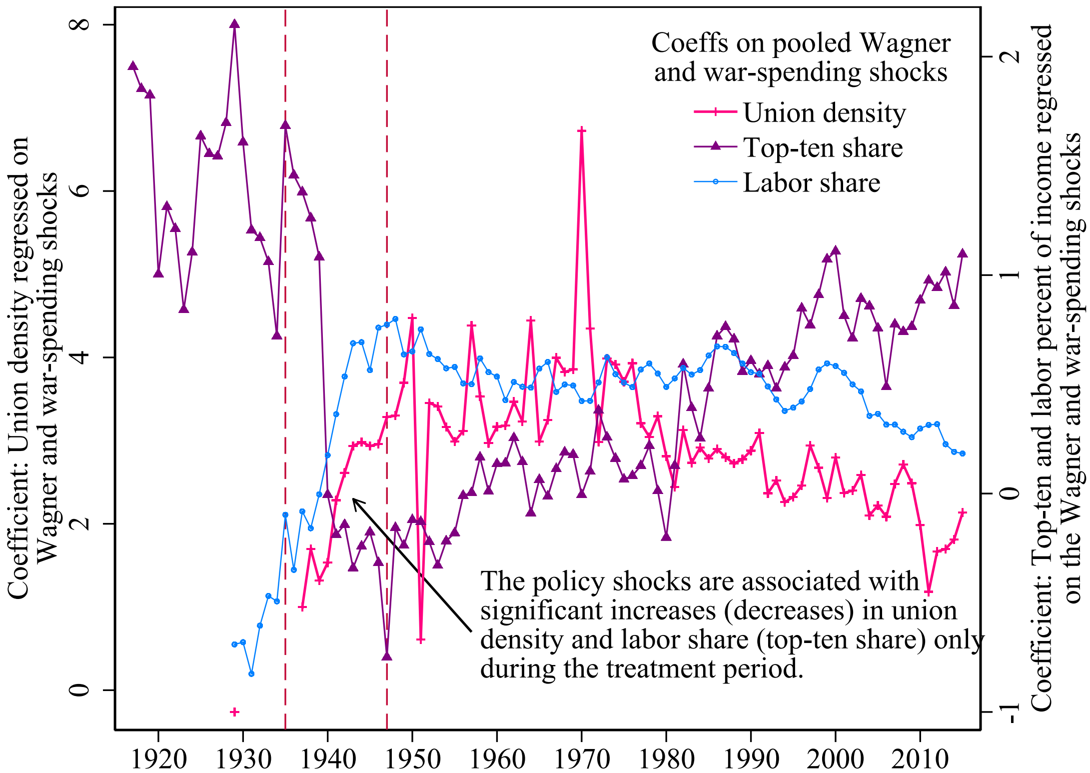
APPENDIX FIGURE G.3: CORRELATION OF THE TWO POLICY SHOCKS



Data sources: See Appendix Section G.1 for information on the construction of the two policy shock variables.

Notes: On the x-axis is the (per capita) number of new union members by state, in the five years immediately following the passage of the National Labor Relations ("Wagner") Act. On the y-axis is the total value (in 1942 dollars) of military contracts given to firms, by state, from 1940 to 1945. The raw correlation reported is merely the fitted line depicted in the graph. The weighted correlation weights observations by 1930 population, and the residualized correlation is the unweighted correlation after controlling for four Census regions.

APPENDIX FIGURE G.4: REGRESSING DENSITY AND INEQUALITY OUTCOMES ON THE POOLED POLICY SHOCK VARIABLE



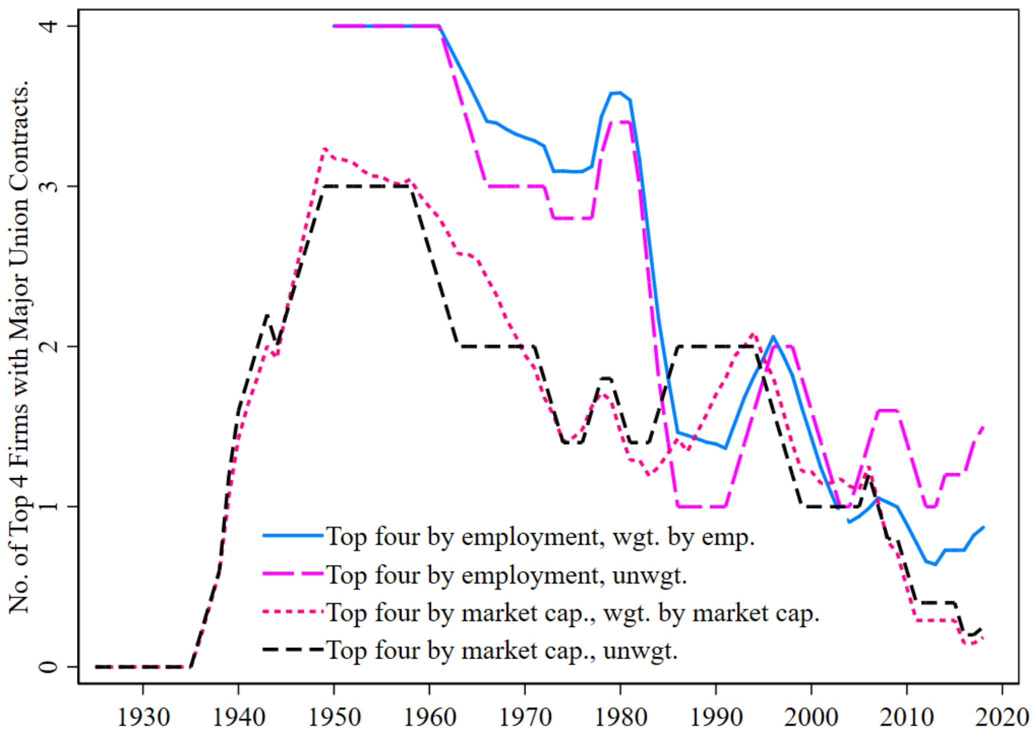
Data sources: Union density data from Gallup and CPS, except for 1929 (see Section V.C and Appendix B for construction of 1929 density, which follows Cohen, Malloy, and Nguyen (2016)). Top-ten income data are from Frank (2015). See Appendix H for construction of state-level labor share of net income.

Notes: Each point on this graph is the estimated coefficient β_y from the following regression:

$$\sum_{y\leq 2014} \beta_y IV_s \mathbb{1}_t^{t=y} + \lambda_{r(s)t} + e_{st}$$

where Y_{st} is the outcome variable (state-year union density, top-ten income share, or labor share of income); IV_s is the pooled policy shock variable (our "Wagner Act shock" and our "war-spending shock" both standardized, then summed); $\mathbb{1}_t^{t=y}$ are year fixed effects; the summation runs over all years y in the sample period (1929 and 1937-2014 for union density; 1929-2014 for labor share; 1917-2014 for top-ten income share); and $\lambda_{r(s)t}$ is a vector of Census $region \times year$ fixed effects. Note that these regressions are equivalent to regressing, separately for each year, the outcome variable on the IV and region fixed effects. We multiply union density by 100 to be on the same scale as labor share. However, in most tables (e.g., Tables II through IV) density is between zero and one to conserve table space by avoiding coefficients with multiple zeros after the decimal point.

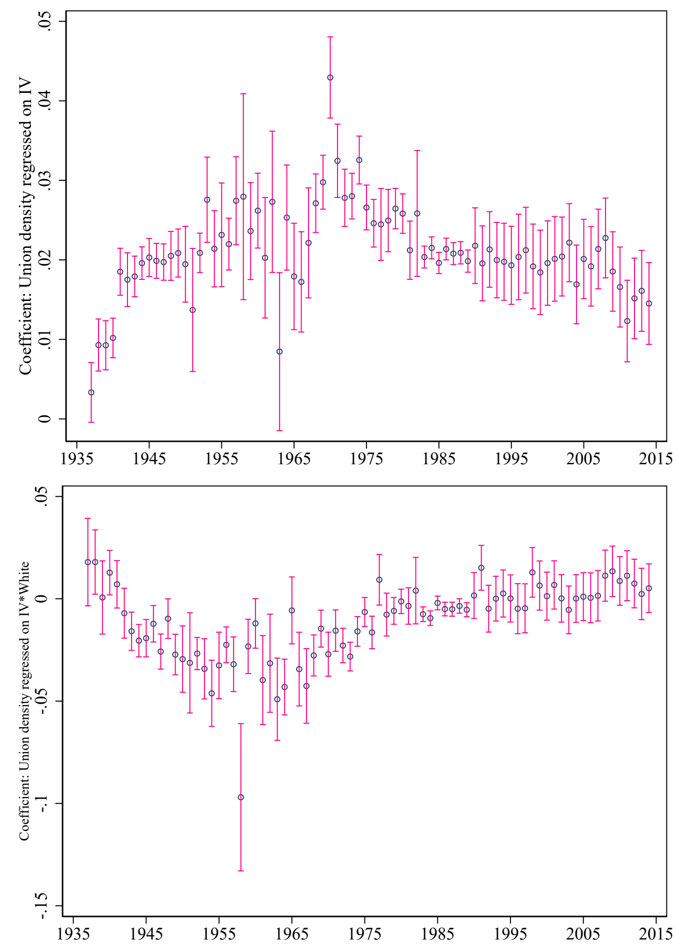
APPENDIX FIGURE G.5: SHARE OF "SUPERSTAR" FIRMS THAT ARE UNIONIZED



Data sources: Market capitalization data from CRSP; employment data from Compustat. Firms are identified by PERMNO in CRSP and GVKEY in Compustat.

Notes: This Figure shows the number of the top-four firms with major union contracts by market capitalization and employment. We identify the union contract status of each of these firms by looking up each of the top four firms on the OLMS collective bargaining agreement website at and the Catherwood library at Cornell , which together have copies of many major collective bargaining agreements filed with the Department of Labor. If they have any contract listed, we then use historical sources to identify when the first independent union contract (excluding company unions) covered the bulk of their core business. For top four firms with no entry in these two Department of Labor databases, we consult a variety of historical sources to confirm that they were never unionized. More details available on request.

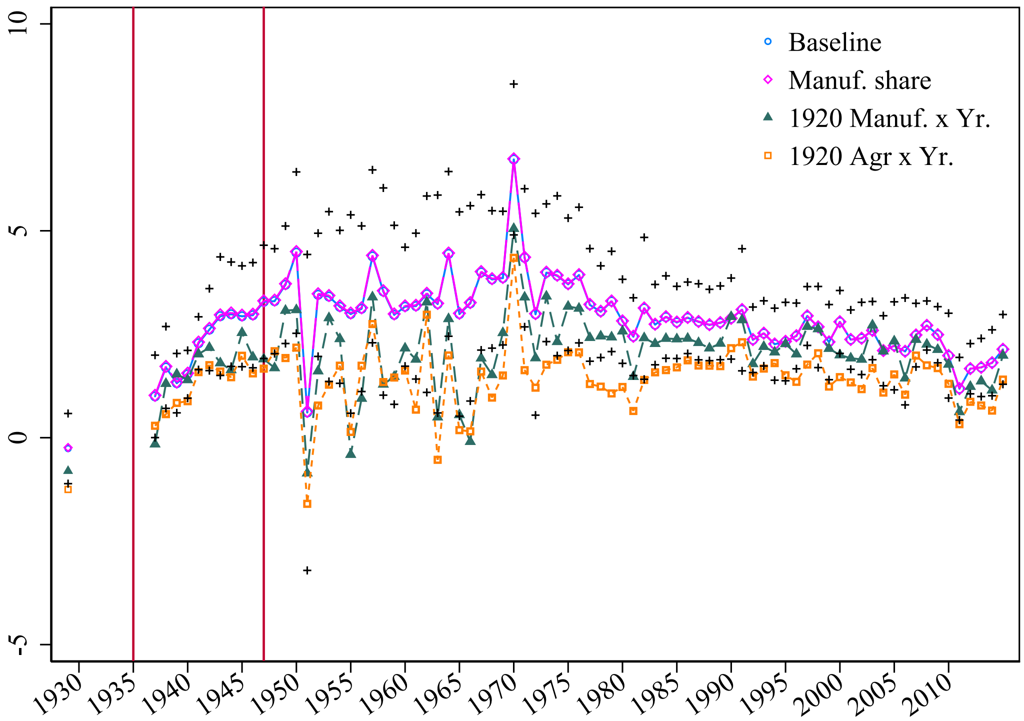
APPENDIX FIGURE G.6: IV EFFECT ON HOUSEHOLD UNIONIZATION AND SELECTION INTO UNIONS BY RACE



Data sources: Household data from Gallup and CPS, as described in Section II.C, Appendix C and B.

Notes: Panel A shows coefficients α_y from the following regression: $Union_{hst} = \sum_{y \leq 2014} \alpha_y \mathbb{1}^{t=y} IV_s + \gamma_1 Female_h^R + f(age_h^R) + \lambda_{r(s)t} + e_{hst}$ while Panel B shows coefficients β_y from the following regression: $Union_{hst} = \sum_{y \leq 2014} \alpha_y \mathbb{1}^{t=y} IV_s + \beta_y White_h^R \times IV_s \times \mathbb{1}^{t=y} + \gamma_1 Female_h^R + f(age_h^R) + \lambda_{r(s)t} + e_{hst}$ where $Union_{hst}$ is our measure of household union status, IV_s is the pooled policy shock variable; $\mathbb{1}_t^{t=y}$ are year fixed effects; the summation runs over all years y in the sample period for which we have race and state (i.e. 1937); $White_h^R$ denotes white respondents; $Female^R$ denotes femal respondents; $f(age_h^R)$ is a quadratic in respondent age; and $\lambda_{r(s)t}$ is a vector of Census $region \times year$ fixed effects. Panel A shows, by year, the effect of the IV on household union status, while Panel B shows the effect of the IV on differential selection into unions by White households.

APPENDIX FIGURE G.7: REGRESSING UNION DENSITY ON THE POOLED POLICY SHOCKS IV



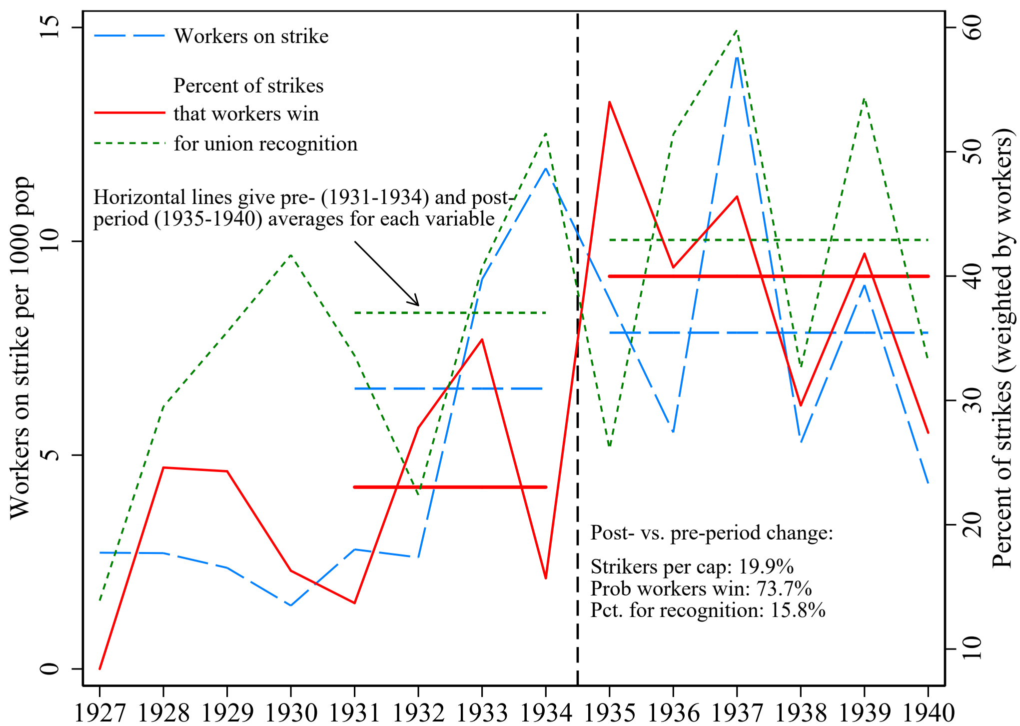
Data sources: See notes to Figure G.4. For construction of the manufacturing share, see Appendix C.

Notes: The first series reproduces the baseline estimates from Figure G.4. That is, it plots the β_t coefficients from the following regression:

$$Union_{st} = \sum_t \beta_t IV_s \mathbb{1}_t^{t=y} + \lambda_{r(s)t} + e_{st},$$

where $Union_{st}$ is state-year union density, $IV_s\mathbb{1}_t^{t=y}$ is the IV interacted with a year-t fixed effect, and $\lambda_{r(s)t}$ is a vector of $region \times year$ fixed effects. The second series adds the contemporaneous manufacturing employment share $Manuf_{st}$ to the baseline equation. The third series adds to the baseline equation the controls $\sum_{y\leq 2014} Manuf_{st}\mathbb{1}^{t=y}$ (i.e., allow contemporaneous manufacturing employment to have a different effect in each year). The fourth series to the baseline equation adds the controls $\sum_y Manuf_s^{1920}\mathbb{1}^{t=y}$,, given the evidence in Appendix Table G.2 that manufacturing share of employment is a key predictor and thus potential a confounder of our IV variable. This control allows the 1920 state-level manufacturing share to have its own effect in each year. The final series adds to the baseline equation the controls $\sum_{y\leq 2014} Agr_s^{1920}\mathbb{1}^{t=y}$, which allows the 1920 state-level agricultural share of employment to have its own effect in each year.

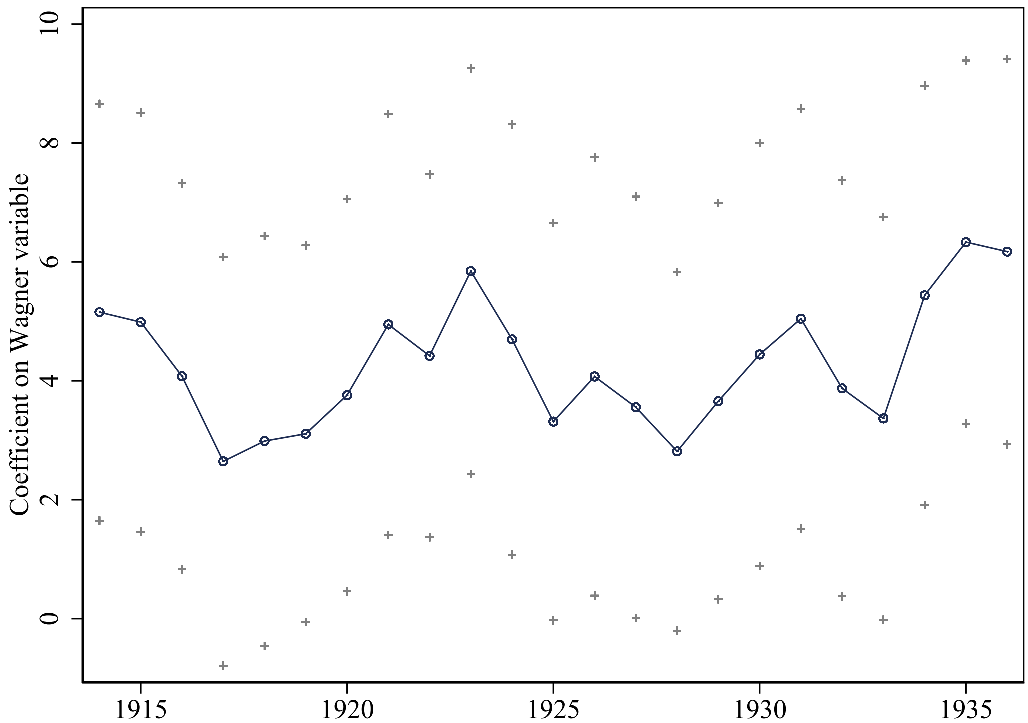
Appendix Figure G.8: Strike activity before and after the 1935 National Labor Relations Act (NLRA)



Data sources: All data are taken from BLS publications: Peterson (1937), Peterson (1938), Peterson (1939), Peterson (1940), and Division (1941).

Notes: This figure compares strike activity before and after the passage of the National Labor Relations Act (NLRA, or Wagner Act). The year of passage is marked with the vertical dashed line in the figure.

Appendix Figure G.9: State strike activity regressed on the Wagner policy shock variable by year



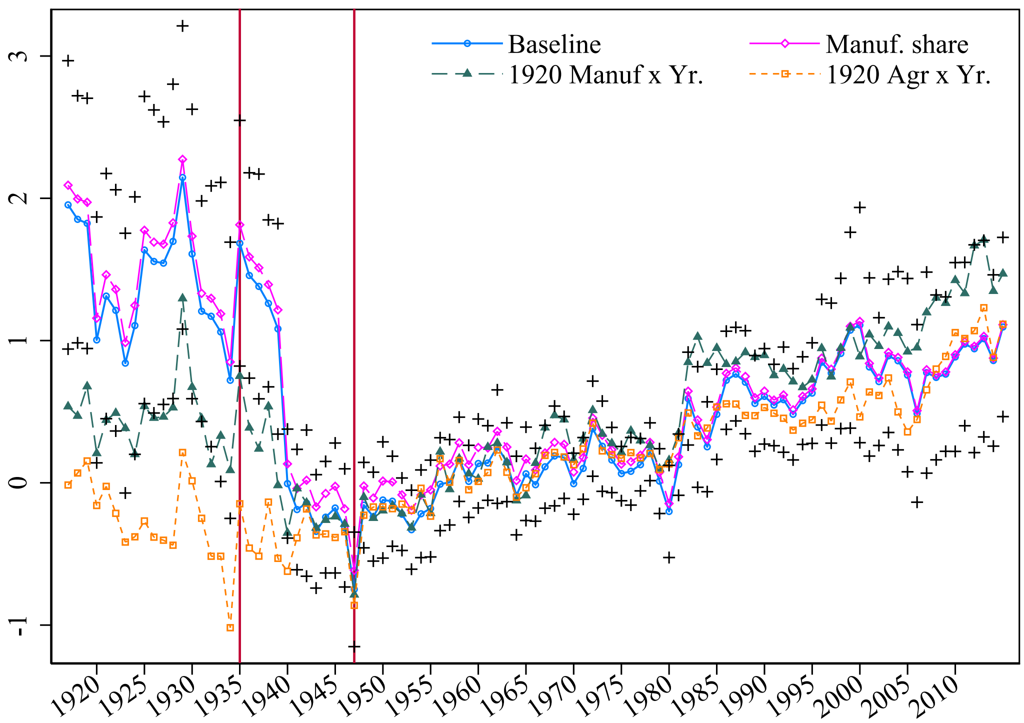
Data sources: All data are taken from BLS publication Peterson (1937).

Notes: For each year t of data, we estimate:

$$Rank\,strikes_s = \beta_t Wagner\,shock_s + \lambda_{r(s)} + e_s,$$

where *Rank strikes*$_s$ is the rank of state s in year t with respect to strikes per capita. The measure is increasing in strike activity, so the most strike-prone state in a year would have an outcome value of 47, as we have 47 states each year. $Wagner\,shock_s$ is our usual Wagner-Act policy shock variable, and $\lambda_{r(s)}$ are Census region fixed effects. We plot the estimates for β_t and whiskers mark 95-percent confidence intervals. We use rank instead of strikes per capita to more easily compare coefficients across high- and low-strike years. Note that we analyze strikes per capita, unweighted by the number of workers involved because BLS measures for workers involved are not available for the full period.

Appendix Figure G.10: Regressing top-ten-percent income share on the pooled policy shocks IV



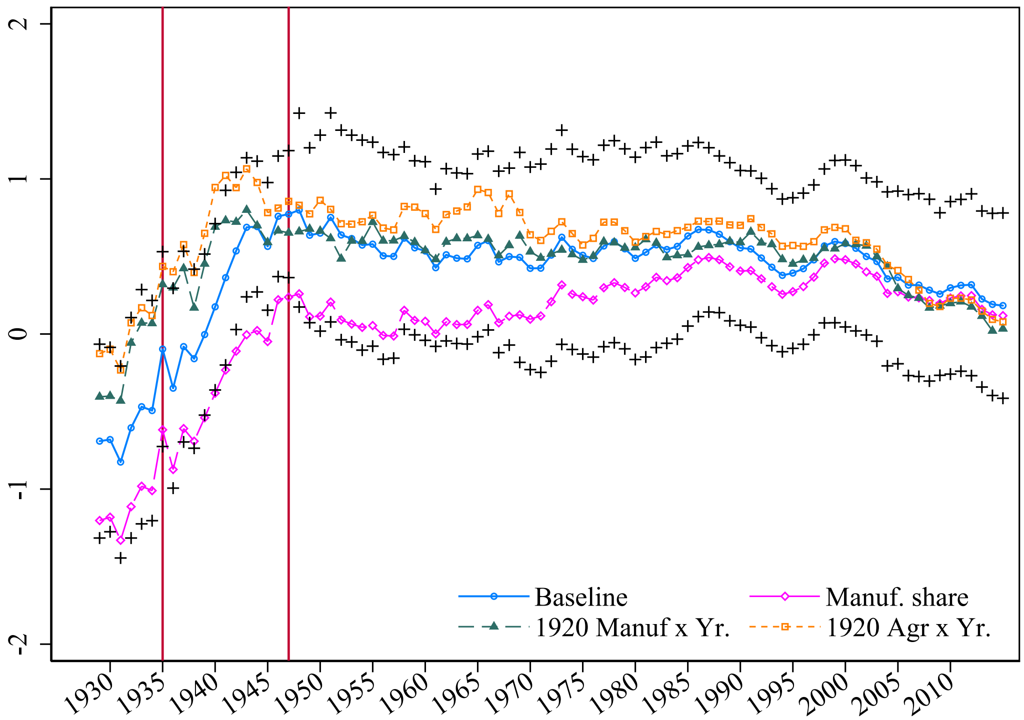
Data sources: See notes to Figure G.4. For construction of the manufacturing share, see Appendix C.

Notes: The first series reproduces the baseline estimates from Figure G.4. That is, it plots the β_t coefficients from the following regression:

$$Topten_{st} = \sum_{y \leq 2014} \beta_y IV_s \mathbb{1}_t^{t=y} + \lambda_{r(s)t} + e_{st}$$

where $Topten_{st}$ is state-year share of income accruing to the richest ten percent of tax units, $IV_s \mathbb{1}_t^{t=y}$ is the IV interacted with a year-t fixed effect, and $\lambda_{r(s)t}$ is a vector of $region \times year$ fixed effects. The second series adds the contemporaneous manufacturing employment share $Manuf_{st}$ to the baseline equation. The third series adds to the baseline equation the controls $\sum_{y \leq 2014} Manuf_{st} \mathbb{1}^{t=y}$ (i.e., allow contemporaneous manufacturing employment to have a different effect in each year). The fourth series adds to the baseline equation the controls $\sum_y Manuf_s^{1920} \mathbb{1}^{t=y}$, which allow the 1920 state-level manufacturing share to have its own effect in each year (given the evidence in Appendix Table G.2 that manufacturing share of employment is a key predictor and thus potential confounder of our IV variable). The final series adds to the baseline equation the controls $\sum_{y \leq 2014} Agr_s^{1920} \mathbb{1}^{t=y}$, which allow the 1920 state-level agricultural share of employment to have its own effect in each year.

APPENDIX FIGURE G.11: REGRESSING LABOR SHARE ON THE POOLED POLICY SHOCKS IV



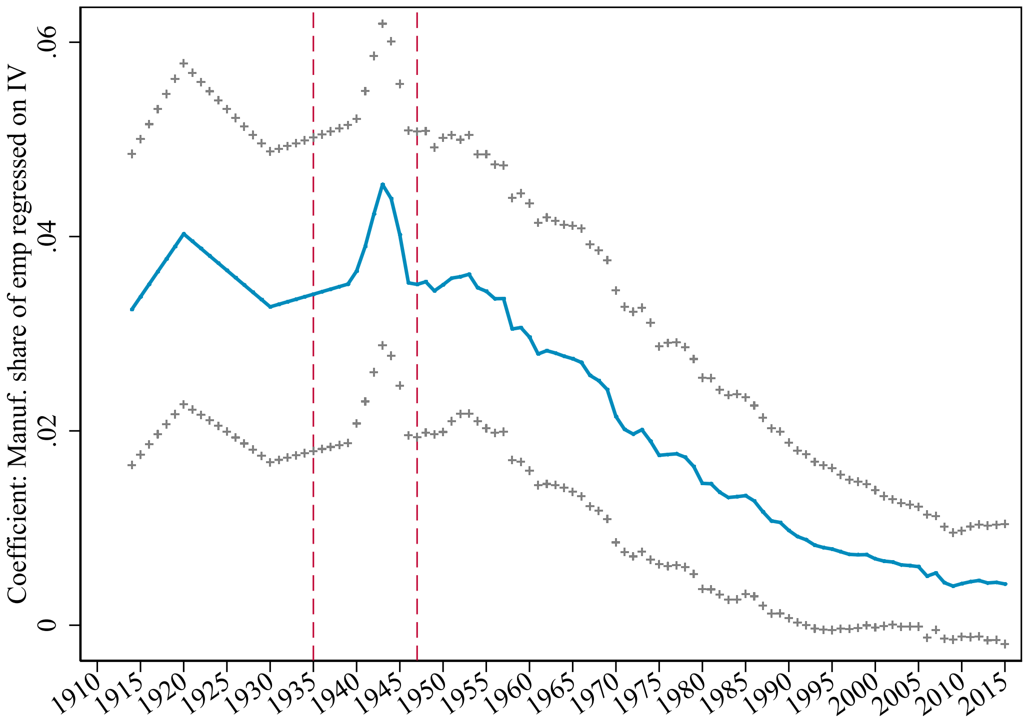
Data sources: See notes to Figure G.4. For construction of the manufacturing share, see Appendix C.

Notes: The first series reproduces the baseline estimates from Figure G.4. That is, it plots the β_t coefficients from the following regression:

$$Labor\,share_{st} = \sum_{y \leq 2014} \beta_y IV_s \mathbb{1}_t^{t=y} + \lambda_{r(s)t} + e_{st}$$

where $Labor\,share_{st}$ is state-year labor share of income, $IV_s \mathbb{1}_t^{t=y}$ is the IV interacted with a year-t fixed effect, and $\lambda_{r(s)t}$ is a vector of $region \times year$ fixed effects. The second series adds the contemporaneous manufacturing employment share $Manuf_{st}$ to the baseline equation. The third series adds to the baseline equation the controls $\sum_{y \leq 2014} Manuf_{st} \mathbb{1}^{t=y}$ (i.e., allow contemporaneous manufacturing employment to have a different effect in each year). The fourth series adds to the baseline equation the controls $\sum_y Manuf_s^{1920} \mathbb{1}^{t=y}$, which allow the 1920 state-level manufacturing share to have its own effect in each year (given the evidence in Appendix Table G.2 that manufacturing share of employment is a key predictor and thus potential confounder of our IV variable). The final series adds to the baseline equation the controls $\sum_{y \leq 2014} Agr_s^{1920} \mathbb{1}^{t=y}$, which allow the 1920 state-level agricultural share of employment to have its own effect in each year.

APPENDIX FIGURE G.12: NO SUSTAINED EFFECT OF THE IV ON STATE MANUFACTURING SHARE OF EMPLOYMENT DURING THE TREATMENT PERIOD



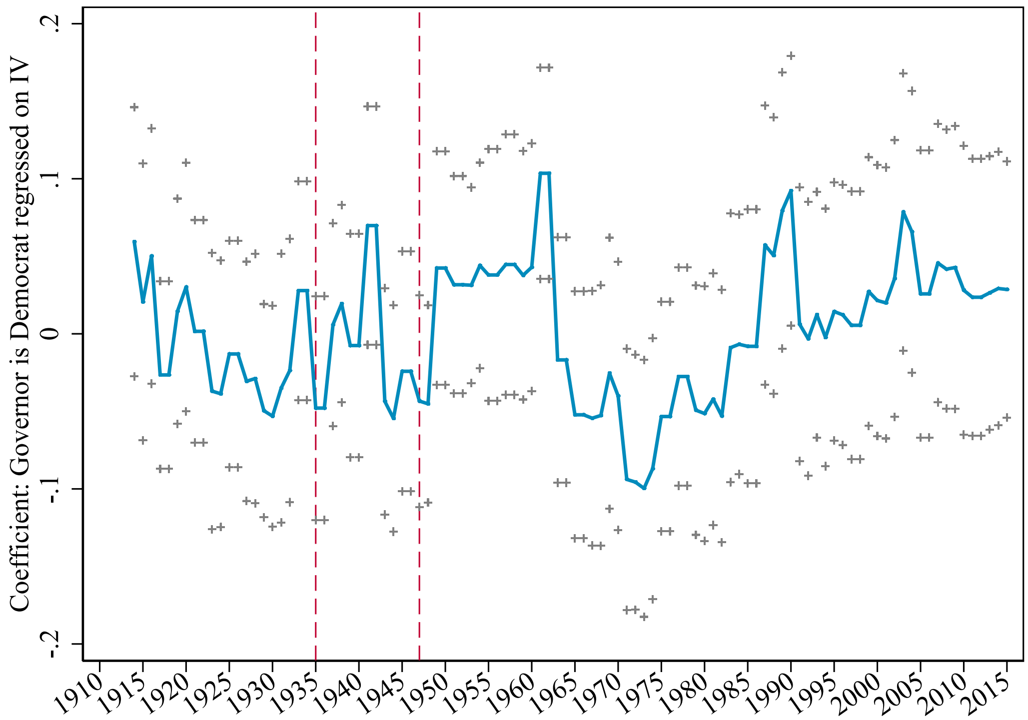
Data sources: For construction of the manufacturing share, see Appendix C.

Notes: In this analysis, we follow our baseline specification, but instead consider state manufacturing share of employment as the outcome. That is, the figure plots the β_t coefficients (and their 95-percent confidence intervals) from the following regression:

$$\textit{Manufacturing employment share}_{st} = \sum_{y \leq 2014} \beta_y IV_s \mathbb{1}_t^{t=y} + \lambda_{r(s)t} + e_{st}$$

where all notation is as in the baseline specification. Note that until 1939, we do not have annual data on manufacturing shares and rely on interpolation between Census years. See Appendix C for more details on the construction of the manufacturing employment share data.

APPENDIX FIGURE G.13: NO SYSTEMATIC RELATIONSHIP BETWEEN THE IV AND DEMOCRATIC GOVERNORSHIPS



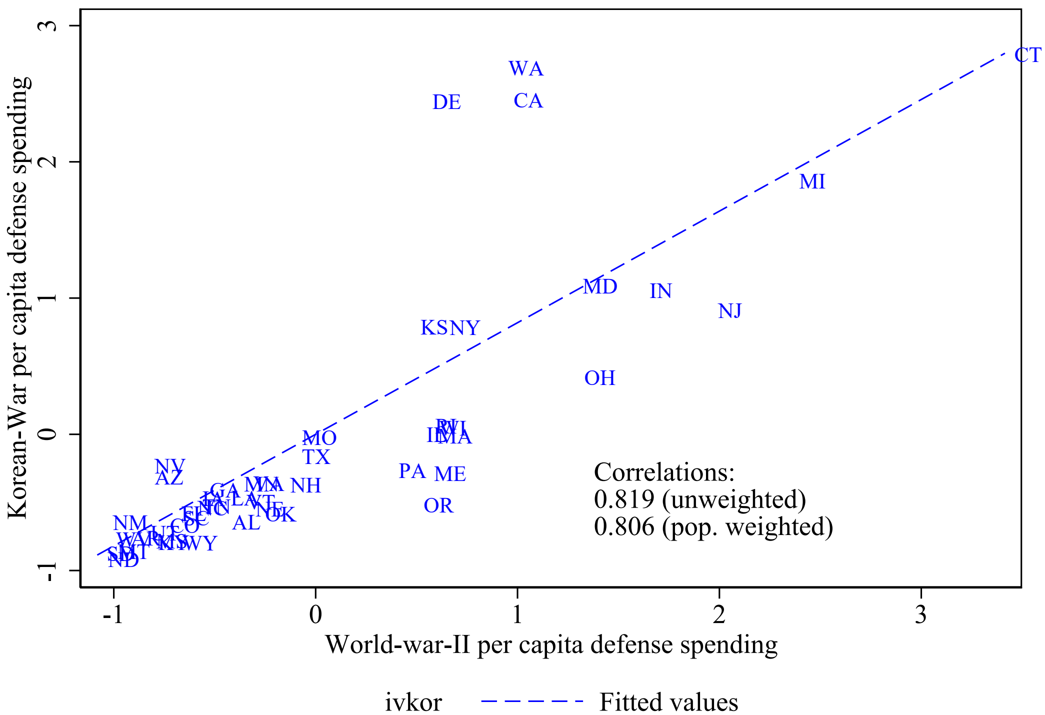
Data sources: Democratic governorships data from Besley, Persson, and Sturm (2010).

Notes: In this analysis, we follow our baseline specification in Figure G.4, but consider a binary variable coded as one if a Democrat is governor in state s in year t as the outcome. That is, the figure plots the β_t coefficients (and their 95-percent confidence intervals) from the following regression:

$$Democratic\ governor_{st} = \sum_{y \leq 2014} \beta_y IV_s \mathbb{1}_t^{t=y} + \lambda_{r(s)t} + e_{st}$$

where all notation is as in the baseline specification.

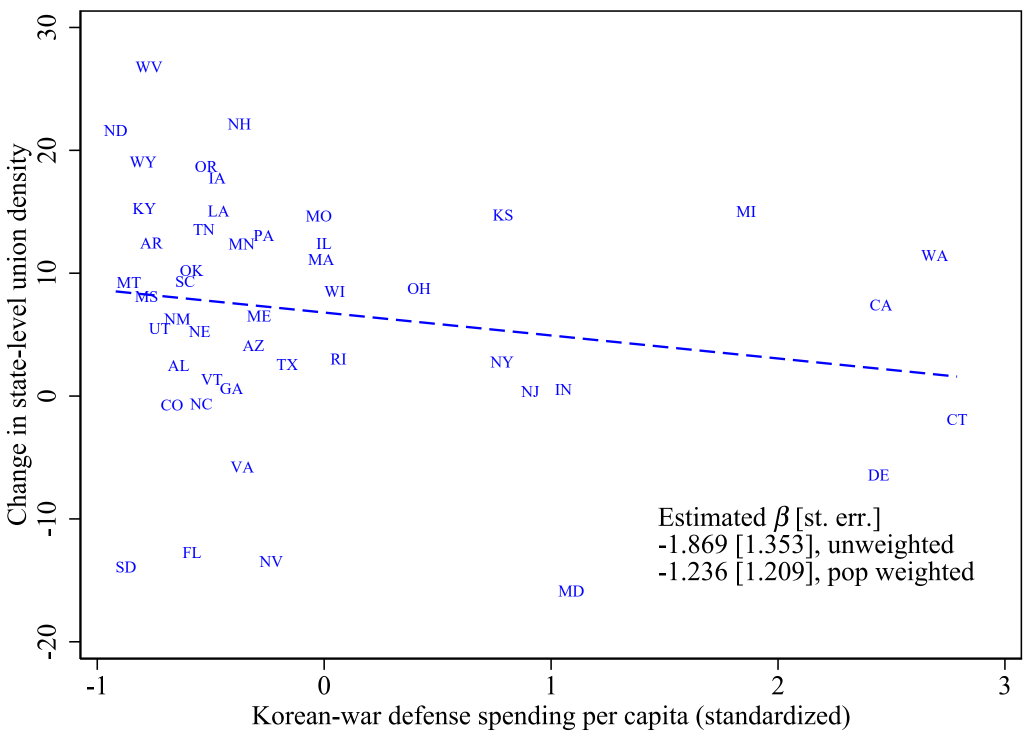
Appendix Figure G.14: Strong correlation across states in World-War-II and Korean-War defense contracts



Data sources: Data for Korean-War defense spending are from Secretary of Defense (1962), Table III.

Notes: Defense contracts during World War II are from 1940 to 1945 and during the Korean War from 1950-1953. The "pop weighted" correlation weights states by their 1930 population.

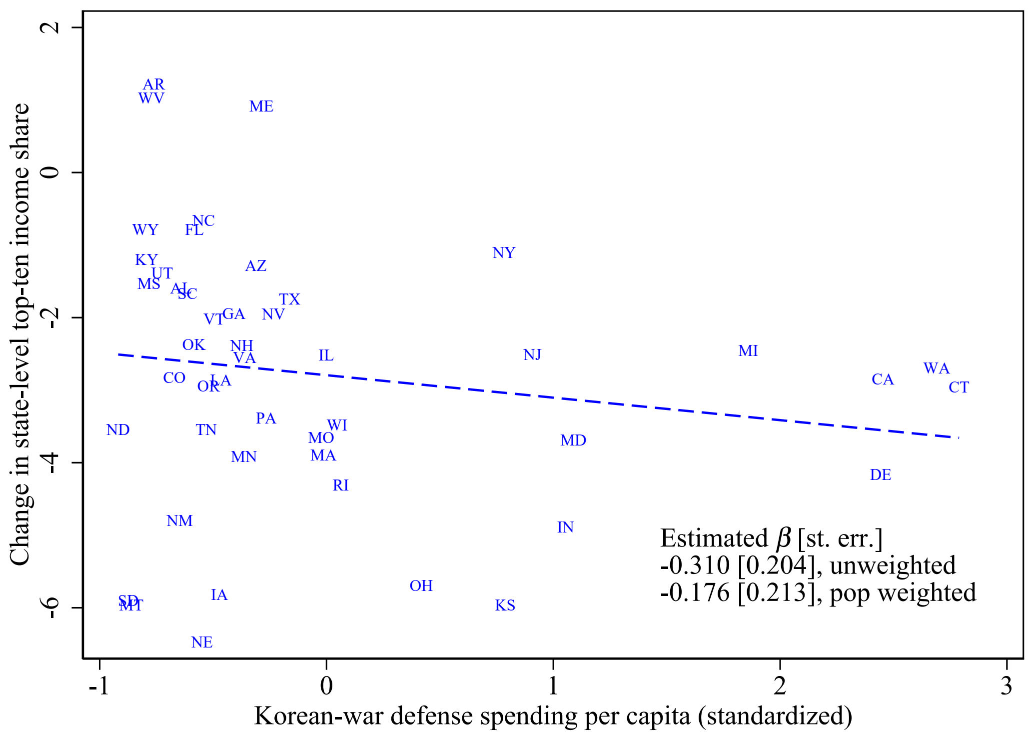
Appendix Figure G.15: No significant relationship between 1954-1949 changes in state-level union density and Korean-War contracts



Data sources: Data for Korean-War defense spending are from Secretary of Defense (1962), Table III.

Notes: In this analysis we regress the change in state-level union density in 1954-1949 as a function of state-level per capita defense spending during the Korean War (1950-1953). The "pop weighted" estimate weights states by their 1930 population.

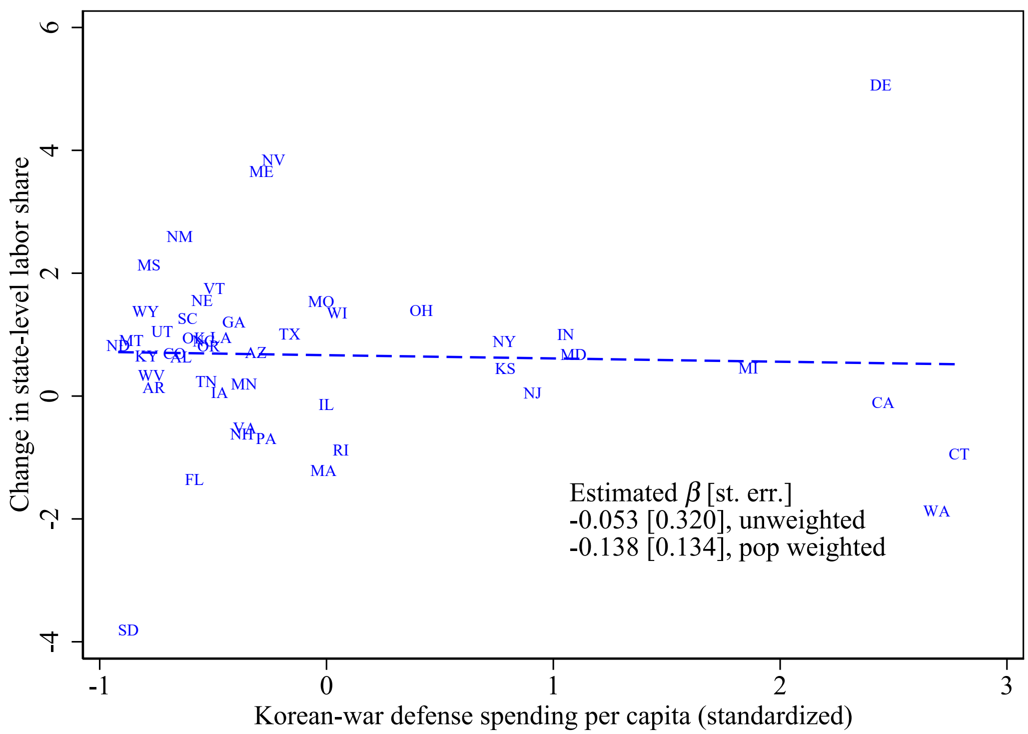
APPENDIX FIGURE G.16: NO SIGNIFICANT RELATIONSHIP BETWEEN 1954-1949 CHANGES IN STATE-LEVEL TOP-TEN SHARES AND KOREAN-WAR CONTRACTS



Data sources: Data for Korean-War defense spending are from Secretary of Defense (1962), Table III.

Notes: In this analysis we regress the change in state-level union density in 1954-1949 as a function of state-level per capita defense spending during the Korean War (1950-1953). The "pop weighted" estimate weights states by their 1930 population.

APPENDIX FIGURE G.17: NO SIGNIFICANT RELATIONSHIP BETWEEN 1954-1949 CHANGES IN STATE-LEVEL LABOR SHARES AND KOREAN-WAR CONTRACTS



Data sources: Data for Korean-War defense spending are from Secretary of Defense (1962), Table III.

Notes: In this analysis we regress the change in state-level union density in 1954-1949 as a function of state-level per capita defense spending during the Korean War (1950-1953). The "pop weighted" estimate weights states by their 1930 population.

APPENDIX TABLE G.1: FIRST-STAGE RELATIONSHIP OF THE POLICY SHOCKS AND UNION DENSITY

	Dept. variable: Change in state-level union density					
	(1)	(2)	(3)	(4)	(5)	(6)
Wagner shock x (1929-1938)	0.0465*** [0.0127]	0.0468*** [0.0135]	0.0358*** [0.0103]	0.0672*** [0.0136]	0.0567*** [0.0171]	
War shock x (1938-1947)	0.0378*** [0.0130]	0.0376*** [0.0138]	0.0350*** [0.00879]	0.0347** [0.0143]	0.0338** [0.0165]	
Wagner shock	-0.00143 [0.00196]		0.000475 [0.00146]	-0.00151 [0.00346]	-0.00311 [0.00259]	
War shock	-0.00346 [0.00323]		-0.00648** [0.00256]	-0.00356 [0.00346]	-0.00110 [0.00410]	
Change in est. state skill share	-0.0325 [0.0276]	-0.0358 [0.0303]	-0.000852 [0.0322]	-0.0318 [0.0282]	-0.0331 [0.0277]	-0.0336 [0.0274]
War shock x (1929-1938)					-0.0160 [0.0146]	
Wagner shock x (1938-1947)					0.00400 [0.0153]	
Pooled shock x (1929-1938)						0.0217*** [0.00556]
Pooled shock x (1938-1947)						0.0184*** [0.00650]
Pooled Shock						-0.00215 [0.00129]
Dept. var. mean	-0.000763	-0.000763	0.00304	-0.00109	-0.000763	-0.000763
F-stat	17.21	15.44	30.28	26.17	9.664	16.41
Weighted?	No	No	Yes	No	No	No
State FE?	No	Yes	No	No	No	No
Excl. Mich?	No	No	No	Yes	No	No
Observations	409	409	409	400	409	409

Data sources: See notes to Table IV.

Notes: Each observation is a state x nine-year interval. We standardize (subtract the mean and divide by the standard deviation) each policy shock variable so their coefficients are easier to compare. The hypothesized treatment period for the Wagner Act shock is the years immediately after its 1935 passage, but due to missing union-density data from 1930-1936, we denote 1929-1938 as its treatment period. The war-spending variable sums state-level war spending from 1940-1945, and so we denote 1938-1947 as its treatment period to use another nine-year interval and to avoid any year of the war when other war-related policies could have direct effects on inequality. Non-treatment intervals are all other non-overlapping nine-year intervals (i.e., 1947-1958, 1958-1967, etc.). In col. (3), weights refer to 1930 state population. In the final column, the pooled IV sums the two (already standardized) policy shock variables. Standard errors are clustered by state. $^{*}p = 0.1$, $^{**}p = 0.05$, $^{***}p = 0.01$.

APPENDIX TABLE G.2: CORRELATIONS OF THE POOLED IV VARIABLE WITH 1920 STATE CHARACTERISTICS

	Outcome: Pooled Wagner and war shocks IV			
	(1)	(2)	(3)	(4)
Manuf. share of employment	8.263^{***}	3.078	9.484^{***}	6.785
	[1.463]	[2.452]	[3.101]	[4.339]
Agr. share of employment		1.820	0.969	-1.206
		[3.540]	[3.364]	[3.707]
Urban share of pop.		5.572	2.376	0.558
		[3.380]	[3.222]	[3.361]
Black share of pop.		-1.100	-0.391	0.935
		[1.925]	[1.973]	[2.413]
Foreign-born share of pop.		0.772	2.647	3.006
		[3.644]	[3.847]	[3.712]
Log of 1920 state pop		0.313	0.526^{*}	0.0901
		[0.255]	[0.266]	[0.292]
Geographic FE	None	None	Region	Division
Observations	47	47	47	47

Sources: We create state-level averages using 1920 Census microdata from IPUMS (using person weights).

Notes: By construction, the mean of the dependent variable is zero in all columns, as it is the standardized sum of the two policy shock variables. $^{*}p = 0.1,^{**}p = 0.05,^{***}p = 0.01$.

APPENDIX TABLE G.3: EFFECT OF UNION DENSITY ON STATE-LEVEL INEQUALITY MEASURES, IV RESULTS

	Dept var: Top-ten income share						Dept var: Labor-share of state income					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Change in union density	-62.32***	-61.55***	-75.14***	-56.38***	-36.98**	-81.03***	33.88***	33.30***	39.94***	25.20***	26.36***	38.91***
	[10.96]	[10.89]	[17.39]	[12.74]	[14.73]	[16.67]	[6.034]	[6.037]	[5.813]	[5.302]	[6.720]	[7.733]
Wagner shock	0.345	0.0897	0.279*	0.405	0.515**		0.0712	-0.00180	0.127	0.142	0.0186	
	[0.232]	[0.151]	[0.161]	[0.310]	[0.225]		[0.0931]	[0.0964]	[0.108]	[0.146]	[0.102]	
War shock	-0.311	1.602***	-0.275	-0.337	-0.213		-0.0261	-0.192	-0.160	-0.0262	-0.0454	
	[0.346]	[0.286]	[0.268]	[0.355]	[0.361]		[0.138]	[0.159]	[0.143]	[0.135]	[0.146]	
Change in est. state skill share	5.536***	-5.497***	-3.245	-5.172***	-4.576***	-6.222***	1.394	1.279	0.0402	1.060	1.102	1.567
	[1.853]	[1.924]	[2.273]	[1.831]	[1.518]	[2.166]	[1.096]	[1.159]	[1.299]	[0.928]	[0.958]	[1.167]
Wagner shock x (1938-1947)					-2.369***						0.718	
					[0.703]						[0.669]	
War shock x (1929-1938)					-1.072						0.236	
					[0.723]						[0.311]	
Pooled Shock						0.0750						0.0126
						[0.100]						[0.0499]
Dept. var. mean	0.643	0.643	0.425	0.643	0.643	0.643	0.0320	0.0320	0.367	0.0206	0.0320	0.0320
F-stat	17.21	15.44	30.28	26.17	9.664	16.41	17.21	15.44	30.28	26.17	9.664	16.41
AR-Pvalue	0.00179	0.00194	0.00618	0.000570	0.0243	0.00469	0.00433	0.00512	0.00157	0.00344	0.0111	0.00721
State FE?	No	Yes	No	No	No	No	No	Yes	No	No	No	No
Weighted?	No	No	Yes	No	No	No	No	No	Yes	No	No	No
Excl. Mich?	No	No	No	Yes	No	No	No	No	No	Yes	No	No
Observations	409	409	409	400	409	409	409	409	409	400	409	409

Data sources: See notes to Table IV.

Notes: Each observation is a state x nine-year interval. In cols. 1-5 and 6-11 the excluded instruments are the (standardized) Wagner Act shock interacted with its treatment period (1929-1937) and the (standardized) war-spending shock interacted with its treatment period (1938-1947). In cols. 6 and 12 we sum the two policy shocks and interact this pooled IV with the two treatment periods. Non-treatment intervals are all other non-overlapping nine-year intervals (i.e., 1947-1956, 1956-1965, etc.). In cols. (3) and (9), weights refer to 1930 state population. Standard errors are clustered by state. $^{*}p = 0.1$, $^{**}p = 0.05$, $^{***}p = 0.01$.

Appendix Table G.4: Effect of union density on state-level inequality measures, reduced form results

	Dept var: Top-ten income share						Dept var: Labor-share of state income					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Wagner shock x (1929-1938)	-1.648**	-1.647**	-1.490	-2.472**	-0.747		1.082***	1.082***	1.143***	0.876**	0.978***	
	[0.747]	[0.790]	[0.934]	[1.122]	[1.175]		[0.222]	[0.235]	[0.224]	[0.384]	[0.343]	
War shock x (1938-1947)	-4.155***	-4.155***	-4.835***	-3.832***	-3.579***		1.980***	1.981***	1.921***	2.069***	1.840***	
	[0.819]	[0.868]	[0.871]	[0.831]	[1.148]		[0.454]	[0.481]	[0.397]	[0.503]	[0.566]	
Wagner shock	0.316	-0.440***	0.110	0.360	0.324**		-0.0123	-0.0383	0.171	0.0681	-0.0252	
	[0.235]	[0.109]	[0.126]	[0.296]	[0.153]		[0.179]	[0.0394]	[0.146]	[0.248]	[0.175]	
War shock	0.0664	0.531***	0.452*	0.0397	0.177		-0.110	0.265***	-0.464**	-0.131	-0.116	
	[0.297]	[0.118]	[0.230]	[0.301]	[0.223]		[0.215]	[0.0556]	[0.185]	[0.225]	[0.215]	
Change in est. state skill share	-3.160**	-3.055**	-3.031**	-3.142**	-3.155**	-3.122**	0.127	0.0882	-0.0522	0.143	0.123	0.0977
	[1.243]	[1.393]	[1.199]	[1.262]	[1.265]	[1.290]	[0.536]	[0.599]	[0.517]	[0.542]	[0.536]	[0.544]
Wagner shock x (1938-1947)					-0.978						0.220	
					[0.852]						[0.431]	
War shock x (1929-1938)					-1.568						0.194	
					[1.022]						[0.398]	
Pooled shock x (1929-1938)						-1.142***						0.600***
						[0.338]						[0.147]
Pooled shock x (1938-1947)						-2.230***						1.000***
						[0.272]						[0.274]
Pooled Shock						0.253***						-0.0686
						[0.0467]						[0.0412]
Dept. var. mean	0.637	0.637	0.426	0.637	0.637	0.637	0.0570	0.0570	0.367	0.0465	0.0570	0.0570
State FE?	No	Yes	No	No	No	No	No	Yes	No	No	No	No
Weighted?	No	No	Yes	No	No	No	No	No	Yes	No	No	No
Excl. Mich?	No	No	No	Yes	No	No	No	No	No	Yes	No	No
Observations	423	423	423	414	423	423	423	423	423	414	423	423

Data sources: See notes to Table IV.

Notes: Each observation is a state x nine-year interval. See notes to Table G.3. In cols. 6 and 12, we sum the two

APPENDIX TABLE G.5: EFFECT OF UNION DENSITY ON TOP-TEN SHARE, ROBUSTNESS CHECKS

	Dept. variable: Change in state top-ten income share											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Change in union density	-61.28***	-70.07***	-62.55***	-29.29**	-24.88***	-62.59***	-60.52***	-59.10***	-60.22***	-71.51***	-55.81***	-64.50***
	[10.71]	[13.79]	[16.54]	[14.18]	[7.983]	[11.95]	[10.50]	[12.62]	[11.09]	[11.36]	[10.66]	[11.12]
Wagner shock	0.319	0.362	0.347	0.311**	0.273	0.348	0.373	0.374	0.321	0.411*	0.328	0.412*
	[0.239]	[0.257]	[0.239]	[0.150]	[0.169]	[0.231]	[0.230]	[0.230]	[0.210]	[0.240]	[0.217]	[0.243]
War shock	-0.360	-0.386	-0.332	-0.00138	0.0591	-0.335	-0.363	-0.267	-0.234	-0.414	-0.258	-0.442
	[0.366]	[0.364]	[0.357]	[0.240]	[0.262]	[0.346]	[0.326]	[0.347]	[0.300]	[0.369]	[0.362]	[0.370]
Control variable		-1.279	-94.07	0.412	2.697	0.179	1.526	0.0725	-11.66	0.00296	0.00891	1.700**
		[3.371]	[209.5]	[1.196]	[2.691]	[0.583]	[1.046]	[0.230]	[7.725]	[0.00352]	[0.0552]	[0.672]
Control var x (1929-1938)		13.16*	846.5	9.474**	-64.63***	1.113	42.07	-0.863	4.426	-0.0278**	0.0115	-5.139**
		[7.903]	[886.1]	[4.675]	[12.36]	[2.213]	[29.75]	[1.289]	[34.74]	[0.0114]	[0.146]	[2.285]
Control var x (1938-1947)		4.635	189.9	25.78***	-30.55***	-2.068	-36.62***	-2.085	-12.03	-0.0300**	0.424	-2.118
		[9.239]	[1004.4]	[5.706]	[4.280]	[2.708]	[9.205]	[1.622]	[32.27]	[0.0129]	[0.296]	[3.412]
Dep. v. mean	0.643	0.643	0.643	0.643	0.643	0.639	0.643	0.643	0.643	0.643	0.643	0.643
F-stat	17.40	11.72	8.876	9.200	11.40	17.68	16.95	10.01	17.73	15.38	17.66	14.91
AR-Pvalue	0.00203	0.000973	0.00113	0.0611	0.0164	0.00244	0.00232	0.00167	0.00320	0.000779	0.00303	0.00128
Control var.	N/A	Manuf. emp. share	1920 Man share	1920 Ag. share	IRS share	Dem. gov.	Min wage	Liberal index	WWII mob. rate	New Deal aid	State top MTR	State skill share
Observations	409	409	409	409	409	406	409	409	409	409	409	409

Notes: Each observation is a state x nine-year interval. All specifications include the change in skill shares as a control. The first column reproduces col. (1) of Appendix Table G.3. All subsequent columns add explanatory variables $Control_{st}$, $Control_{st} \times$(1929-1938), and $Control_{st} \times$(1938-1947), where $Control_{st}$ is a state-year varying control (listed in the bottom rows of the table), and we include its interactions with the two treatment intervals. The controls are the manufacturing share of employment, the share of tax units that pay federal income tax, whether the governor is a Democrat, the state minimum wage (if it is less than the federal, it is coded as the federal), the state's policy liberalism index (Caughey and Warshaw, 2016), the total New Deal expenditure received by the state in 1933-1939 (taken from Fishback, Kantor, and Wallis, 2003), and the share of the state's young men that were drafted in the Second World War (taken from Acemoglu, Autor, and Lyle, 2004). Standard errors, in brackets, are clustered by state. $^{*}p = 0.1,^{**}p = 0.05,^{***}p = 0.01$.

APPENDIX TABLE G.6: EFFECT OF UNION DENSITY ON LABOR SHARE, ROBUSTNESS CHECKS

	Dept. variable: Change in state labor share of income											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Change in union	33.58***	30.35***	27.70***	26.17***	27.95***	37.15***	32.08***	37.28***	33.78***	34.31***	31.02***	36.26***
density	[5.989]	[7.245]	[7.701]	[8.411]	[6.422]	[5.593]	[5.570]	[7.230]	[6.103]	[6.377]	[5.898]	[6.418]
Wagner shock	0.0778	0.0846	0.0984	0.0828	-0.0219	0.0618	0.0534	0.0388	0.0754	0.0642	0.0811	0.116
	[0.0948]	[0.109]	[0.109]	[0.0989]	[0.110]	[0.106]	[0.0888]	[0.0839]	[0.0936]	[0.0979]	[0.101]	[0.0970]
War shock	-0.0137	-0.0462	-0.0279	-0.0878	0.0476	0.0175	-0.0293	-0.0815	-0.0496	-0.0559	-0.0564	-0.0986
	[0.142]	[0.123]	[0.121]	[0.134]	[0.157]	[0.160]	[0.130]	[0.134]	[0.126]	[0.123]	[0.158]	[0.152]
Control variable		-0.465	-76.11	0.0146	-5.967**	0.242	-0.265	0.168	5.725	-0.00238	-0.00569	0.651
		[2.154]	[132.6]	[0.970]	[2.928]	[0.358]	[0.498]	[0.145]	[4.476]	[0.00218]	[0.0367]	[0.425]
Control var x		-1.247	-9.964	0.100	8.359	-2.394***	15.98	-0.768	-13.00	0.00285	-0.0419	1.175
(1929-1938)		[3.406]	[318.3]	[2.659]	[8.917]	[0.911]	[13.57]	[0.682]	[14.03]	[0.00512]	[0.0708]	[1.397]
Control var x		6.324	701.7	-8.149*	11.73**	2.784	30.52***	0.637	-1.771	0.00177	-0.170	-1.618
(1938-1947)		[5.072]	[511.2]	[4.633]	[4.646]	[1.754]	[6.146]	[1.037]	[23.34]	[0.0115]	[0.155]	[1.706]
Dep. v. mean	0.0320	0.0320	0.0320	0.0320	0.0320	-0.00161	0.0320	0.0320	0.0320	0.0320	0.0320	0.0320
F-stat	17.40	11.72	8.876	9.200	11.40	17.68	16.95	10.01	17.73	15.38	17.66	14.91
AR-Pvalue	0.00437	0.00483	0.0174	0.0119	0.00696	0.00347	0.00599	0.00182	0.00327	0.00201	0.00524	0.00354
Control var.	N/A	Manuf. emp. share	1920 Man share	1920 Ag. share	IRS share	Dem. gov.	Min wage	Liberal index	WWII mob. rate	New Deal aid	State top MTR	State skill share
Observations	409	409	409	409	409	406	409	409	409	409	409	409

Notes: Each observation is a state x nine-year interval. The first column reproduces col. (7) of Appendix Table G.3. All subsequent columns add explanatory variables $Control_{st}$, $Control_{st} \times$(1929-1938), and $Control_{st} \times$(1938-1947), where $Control_{st}$ is a state-year varying control (listed in the bottom rows of the table), and we include its interactions with the two treatment intervals. The controls are the manufacturing share of employment, the share of tax units that pay federal income tax, whether the governor is a Democrat, the state minimum wage (if it is less than the federal, it is coded as the federal), the state's policy liberalism index (Caughey and Warshaw, 2016), the total New Deal expenditure received by the state in 1933-1939 (taken from Fishback, Kantor, and Wallis, 2003), and the share of the state's young men that were drafted in the Second World War (taken from Acemoglu, Autor, and Lyle, 2004). Standard errors, in brackets, are clustered by state. $^{*}p = 0.1, ^{**}p = 0.05, ^{***}p = 0.01$.

APPENDIX TABLE G.7: USING 1920S STRIKES INSTEAD OF THE WAGNER ACT VARIABLE AS AN IV

	Dept var: Nine-year changes in....		
	(1)	(2)	(3)
	Union density	Top-ten share	Labor share
Change in union density		-114.5***	39.48***
		[32.44]	[14.96]
Strikes 1921-1928 x (1929-1938)	0.0375**		
	[0.0163]		
War shock x (1938-1947)	0.0351***		
	[0.0130]		
Strikes 1921-1928	-0.00215	-0.125	-0.0310
	[0.00228]	[0.284]	[0.111]
War shock	-0.00110	0.164	0.0147
	[0.00260]	[0.275]	[0.122]
Dept. var mean	-0.000763	0.643	0.0320
F-stat		6.635	6.635
AR-Pvalue		0.000211	0.00554
Observations	409	409	409

Notes: The regressions in this table are identical to, respectively, col. (1) of Appendix Table G.1 and cols. (1) and (7) of Appendix Table G.3, except that strikes per capita from 1921 to 1928 is used instead of the Wagner shock variable. Standard errors clustered by state. $^{*}p = 0.1,^{**}p = 0.05,^{***}p = 0.01$.

APPENDIX TABLE G.8: ARE RESPONDENTS IN STATES HIT WITH POLICY SHOCKS MORE LIKELY TO EXPRESS PRO-WORKER VIEWS?

	Dept. var: Subject says these groups should be doing better (x100)							
	Workers, laborers, poor				Biz owners, managers, profs.			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pooled IV	-0.443				1.082			
	[0.335]				[1.081]			
Wagner shock		-0.600		-0.201		0.457		-2.928
		[0.501]		[0.801]		[1.420]		[1.942]
War shock			-0.949	-0.774			4.022*	6.572**
			[0.842]	[1.293]			[2.186]	[3.207]
Mean, dept. var.	8.966	8.966	8.966	8.966	37.58	37.58	37.58	37.58
Observations	2911	2911	2911	2911	2911	2911	2911	2911

Sources: Data taken from Gallup survey fielded June 1-5, 1945.

Notes: Each observation is a respondent. The outcome variable is based on the second part of a two-part question (items 10a and 10b) . The wording of 10a is “What class or group of people in this country has done best financially during the war compared to what they made before the war?”. The follow-up question 10b reads: “Do you think any class or group of people in this country is NOT making as much money as it should? [capitalization in the original].” $^{*}p = 0.1,^{**}p = 0.05,^{***}p = 0.01$.

H. Construction of historical state-year labor share of net income

The standard state-year measure of labor share from the Bureau of Economic Analysis (BEA) extends back only to 1963. In this Appendix, we introduce a modified state-year measure of labor share (that is, a state-year measure of the labor share of *net national income*) that we construct beginning in 1929. We describe our methodology (in particular how it builds off of past work) and compare our measure to other measures at the aggregate and state-year level.

The BEA has constructed the state-year labor share of GDP, but only since 1963, and labor shares of personal income since 1929. The former doesn't span the entire time period of our sample; the latter ignores corporate income and so is not consistent with the macroeconomic accounts. In this section we describe how we construct labor shares of *net* income using the information in the components of personal income and allocations of national corporate income available from the BEA, following Piketty-Saez-Zucman 2019 (PSZ). At the end, we show how our measured labor share of net national income compares with the BEA GDP measures during the years they overlap. We use the components of personal income available since 1929, together with allocations of national pre-tax corporate retained earnings, to construct a measure of Net National Income at the state-year level from 1929 onwards. We remind readers of the differences below:

- Gross Domestic Product (GDP) = labor income paid by firms +taxes on production (indirect taxes) + capital income paid by firms. This production-side measure is available at the aggregate level back to 1929 and at the state-year level back to 1963.

- Gross National Income (GNI) = labor income earned by residents + taxes on production (indirect taxes) + capital income earned by residents + out-of-state income transferred to residents. This income-side measure is based on flows to residents of a jurisdiction (state or country). In a closed economy this is identical to GDP, but in an open economy there can be differences.

- Net National Income (NNI) = GNI - depreciation. The definition used by PSZ (p. 561) includes indirect taxes paid to government as income. PSZ then apportion indirect tax income to individuals based on labor and capital incomes minus savings. Other authors exclude indirect taxes from net national income (Rognlie 2015). We will exclude indirect tax income due to data limitations, as we do not observe disaggregated savings in our historical period and so cannot apportion it.

- Personal Income (PI) = NNI - indirect taxes - contributions to government social insurance + transfers from government and business (e.g. insurance payouts) - corporate profits. Personal income is also an income side measure, but deducts income that cannot be spent (indirect taxes, contributions to government insurance, and corporate retained earnings) This measure is available at the state-year level from the BEA back to 1929. See for a more detailed description.

Recall that in a closed economy, income received by individuals is equal to payments to factor owners, so GNI = GDP. Net National Income, however, is theoretically closer to a welfare measure (Weitzman 1976). NNI deducts depreciation from GNI, which is the loss of value to capital holders owing to wear and tear and obsolescence of capital goods. Nobody can consume or save depreciation, so it is deducted from GNI to get measures of income received by capital owners. The capital share of GDP (gross operating surplus in the BEA GDP accounts) reflects capital's importance in production, but net capital income reflects the income accruing to owners of capital. The labor share of net income is the total wage income received by residents plus employer contributions plus labor's share of self-employment income, all divided by total income received by residents and all measured prior to any taxation.

In an open economy, besides depreciation, GDP and NNI can also differ due to differences between the location of production and the location of individuals. GDP records the payments made to workers and owners, regardless of where they live from firms in a jurisdiction. NNI records the payments received by workers and owners living in a jurisdiction, regardless of where the firm paying them is located. The inclusion of income earned from out-of-state is particularly important for measures of inequality because out-of-state income (especially capital income) is potentially large for the rich. While the macroeconomics literature has focused on labor's share of GDP, the literature on distributional accounts and inequality has focused on NNI, and we follow the latter literature here.

Personal income, on the other hand, summarizes all the disposable income received by residents in a state. The labor share of personal income captures labor income as a share of all incomes paid to residents of a state, including transfer payments (but excluding government insurance payments). It is not clear how to handle transfer income: one could either include it in labor income or remove it from total income. We choose the latter in order to focus on a pre-tax measure. Further, personal income is inconsistent with the national accounts, as it includes capital income paid to owners (i.e. interest, rent, and dividends) but excludes retained corporate earnings. As a result the labor share of personal income could be significantly higher than labor share based on the national accounts (indeed the labor share of personal income approaches one during World War II). These difficulties in interpreting the labor share of personal income make it a less-than-ideal measure for estimating the effects of unionization on the distribution of factor income. One benefit of using this measure, however, is that personal income was a focus of economic measurement prior to World War II, so its components are available at the state-year level over a much longer time period. In contrast, state-level GDP components are only available beginning in 1963.

We therefore focus on labor share of NNI, which has three advantages over either GDP or Personal Income, as it is a measure that is a) constructible back to 1929, b) consistent with the national accounts, and c) comparable to the recent literature on inequality.

H.1. Data Availability and Construction of Measures

At the state-year level, the divergence between production and incomes can be considerable, due to much factor income being paid to out-of-state agents and much income being derived from out-of-state asset holdings and transfers. One could imagine unionization having different impacts on NNI vs GDP. While GDP reflects how the organization of production compensates suppliers of capital and labor whereever they are, NNI reflects how residents receive capital and labor income. However, there is no measure of state-year GDP prior to 1963. We can, however, construct a NNI-based measure from the BLS/BEA estimates of **personal income** and its components, which exist back to 1929.

From the definitions above, we can see that NNI = personal income plus corporate net retained earnings plus contributions for government social insurance minus asset income minus transfers. We do not observed any state-level allocations of corporate net retained earnings, which are components of the capital share of GDP and NNI. At the national level, the income from assets held elsewhere + transfers roughly equals the corporate retained earnings plus social insurance contributions, so GDP is nearly equal to NNI + depreciation as shown in: . Yet, this may not be true at the state level, so the differences between GDP and NNI may be quite large.

Total labor compensation includes wage and salary payments plus employer contributions. The BEA data historically tracked income labelled as "Supplements to wages and salaries" which combines two accounts: "Contribution for government social insurance" and "Employer contributions for employee pension and insurance funds". In measuring Personal Income, the "Contribution for government social insurance" is deducted to get the labor compensation component of personal income because it is not realized as personal income. We have these components separately, so we can add back "contributions for government social insurance" to worker compensation. As discussed above, we deduct government transfers.

The remaining issue for calculating NNI is allocating corporate net retained earnings (before taxes). We assume that national corporate net retained earnings are allocated proportionally to interest/dividend/rental income. PSZ assign corporate net retained earnings across the income distribution. They allocate this income in proportion to corporate equity holdings imputed from dividends and capital gains reported on tax returns and provide auxiliary evidence that this assumption is warranted. Unfortunately, we do not see state-level capital gains income at all nor do we see dividends income by itself in the BEA accounts. Instead we observe the total payments of interest, dividends, and rental income, including imputed rent on housing.

We can assess the plausibility of our allocation rule using the IRS state-level SOI data from 2001-2017, which report capital gains, dividends, interest income, as well as real estate taxes paid. We follow Saez and Zucman (2016) and convert real estate taxes paid into a value of housing stock by assuming a uniform effective tax rate of 1%, and then we convert the implied stock of housing wealth into a rental income flow by multiplying by the PSZ rate of return on gross housing in that year. The average is 7% over this sample period. We can then compare a state's share of dividends plus interest plus rental income with a state's share of equity plus dividends. Figure H.1 shows the year-by-year regression coefficients. In every year between 2001 and 2017, the coefficient from a bivariate regression of state share of dividends + interest + rental income on a state's share of dividends plus capital gain income is greater than 0.9 and generally statistically indistinguishable from 1. This result suggests that the assumption that net retained earnings is allocated across states in proportion to personal capital income (excluding capital gains) is quite reasonable.

Under the assumption that a state's share of corporate retained earnings is proportional to a state's share of dividend, interest, and rental income, our calculation of NNI is given by:

(H.1)

$$Y_{st}^{NNI} = Y_{st}^{PI} + SocIns_{st} - Transfers_{st} + \frac{Y_{st}^{PI,K}}{\sum_{s' \in S} Y_{s't}^{PI,K}} NetCorpRetainedEarn_t$$

Note that if all corporate profits (minus depreciation) are paid out as dividends, they will be included in the personal income paid as capital income $Y_{st}^{PI,K}$, and net corporate retained earnings is 0. Then NNI will be identical to personal income plus social insurance minus transfers. We calculate pre-tax net corporate retained earnings as the sum of net private saving by domestic business plus corporate taxes paid, both of which are available in the national accounts back to 1929. We also make use of PSZ's estimate of aggregate capital income, $Y_t^{NNI,K}$ as a check and construct an alternative measure of pre-tax net corporate retained earnings as the difference between national capital income and the sum of state-level personal capital income, so $NetCorpRetainedEarn_t = Y_t^{NNI,K} - \sum_{s' \in S} Y_{s't}^{PI,K}$. We find that these two measures of net corporate retained earnings are extremely close, and the differences are likely due to the adjustment for sales taxes in PSZ.

H.2. Construction of the aggregate series

* Data sources:

- Annual estimates (1929-2018) of the compensation of employees is obtained from the FRED.
- Annual estimates (1929-2018) of the GDP are also obtained from the FRED.
- Annual estimates (1929-2018) of NNI were obtained from PSZ, for comparison.

* Calculation

We calculate the aggregate labor share by dividing total compensation of employees by GDP.

H.3. Construction of the state-year series

* Data sources

- Data on state-level personal income (1929-2018), as well as wages and salaries, supplements to wages and salaries, and proprietor's income is obtained from the BEA—Annual State Personal Income and Employment, Personal Income and Employment by Major Component (SAINC4).
- Data on the national Personal Income, GDP, net private saving by domestic business, federal taxes on corporate income, and state and local taxes on corporate income in each year was obtained from the FRED.

* Calculation

  We define the labor share of NNI in state s in year t as:

$$Labor\, share_{st} = \frac{wages_{st} + supplements_{st} + 0.66 proprietors\, income_{st}}{Y_{st}^{NNI}},$$

where Y^{NNI} is calculated as in Equation (H.1). Imputing 2/3 of proprietor's income to labor income is standard for advanced countries (see Krueger, 1999, Gollin, 2002 or Johnson, 1954). We discussed this measure with BEA staff, who confirmed that there is no exact way to get corporate retained earnings at the state-year level, mostly because corporate income is not reported (and hasn't ever been reported) at the state level to the BEA, except for a few highly regulated sectors. We confirm that our measure is highly correlated with the standard BEA measure when they overlap. We further consulted the BEA to see if there was a way to construct a more comprehensive denominator (i.e. including corporate retained earnings) and were told that this would not be feasible with existing data.

H.4. Results

Figure H.2 plots different measures of aggregate labor share over our sample period. The numerator is identical in all measures, but the denominator corresponds to different definitions of income. The figure shows the labor share of Personal Income, Personal Income Without Transfers (which approaches one during World War II), our measure of Net National Income, the measure of labor share of Net National Income from Piketty, Saez, and Zucman (2018) for comparison, and labor share of GDP. Our measure tracks the PSZ measure very closely, with the difference being the deduction of indirect taxes from our NNI measure due to lack of data on how the sales tax incidence (i.e., consumption) is divided between capital and labor income. The labor share of GDP measure is generally the lowest, consistent with depreciation being added to the GDP denominator but not affecting the numerator.

Figure H.3 shows the series for Personal Income, NNI, and GDP (for the post 1963 years) separately for two high union density states and two low union density states. In all cases, while labor share of personal income is greater than labor share of NNI and labor share of NNI is greater than labor share of GDP, all series roughly track each other.

Panel B of Table A.10 shows the same repertoire of state-year panel specifications shown in the paper for other outcomes with the labor share of Net National Income as the outcome. As discussed in the main text, coefficients are uniformly positive and significant, although the inclusion of state-specific quadratic trends lowers the sign and significance somewhat. Table shows the same specifications, with the labor share of Personal Income as the outcome. Coefficients are again uniformly positive and significant, although somewhat less stable and more sensitive to the state-specific time quadratic than the NNI based labor share regressions.

We next compare our results with GDP-based labor share to those with NNI-based labor share, for the years for which we have comparable data. Table H.2 shows two specifications, one without any controls save state and year (and south X year) fixed effects, the other with all the controls in column 6 of Panel B of Table A.10. We show the full sample with our measure of labor share, and then restrict attention to the post-1963 sample where the BEA's labor share of GDP measure is available. While the effect of union density of labor share of GDP is positive and of comparable magnitude (albeit not significant) to the full-sample in the specification without controls, the inclusion of all the controls (particularly the state-specific quadratics) makes the effect close to 0 and insignificant. Our labor share of NNI measure, while smaller in both specifications compared to the full sample, remains statistically significant in both. Much more of the production side of income statistics (as opposed to the income side) at the state-level is interpolated in the 5 years in between Economic Censuses, and this smoothing may be one reason for the difference in precision as well as differential robustness to controls in the two measures.

References

Gollin, Douglas (2002). "Getting income shares right". *Journal of political Economy* 110.2, pp. 458–474.

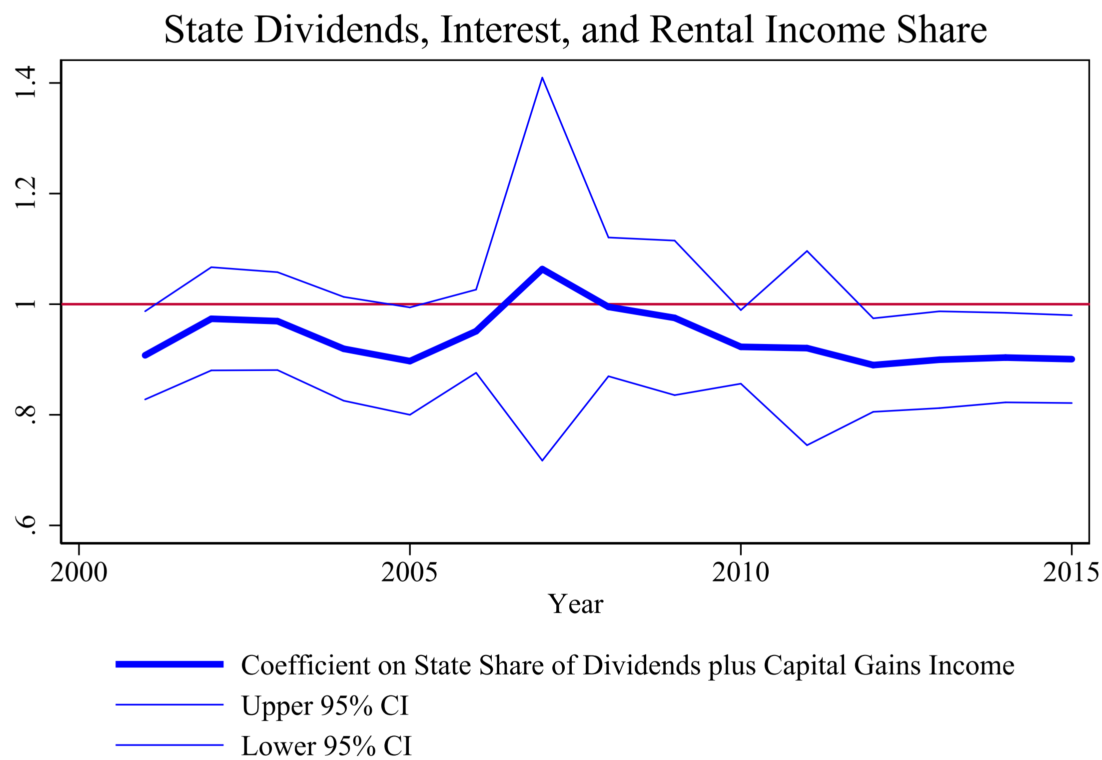
Johnson, D Gale (1954). "The functional distribution of income in the United States, 1850-1952". *The Review of Economics and Statistics*, pp. 175–182.

Krueger, Alan B (1999). "Measuring labor's share". *American Economic Review* 89.2, pp. 45–51.

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman (2018). "Distributional national accounts: methods and estimates for the United States". *The Quarterly Journal of Economics* 133.2, pp. 553–609.

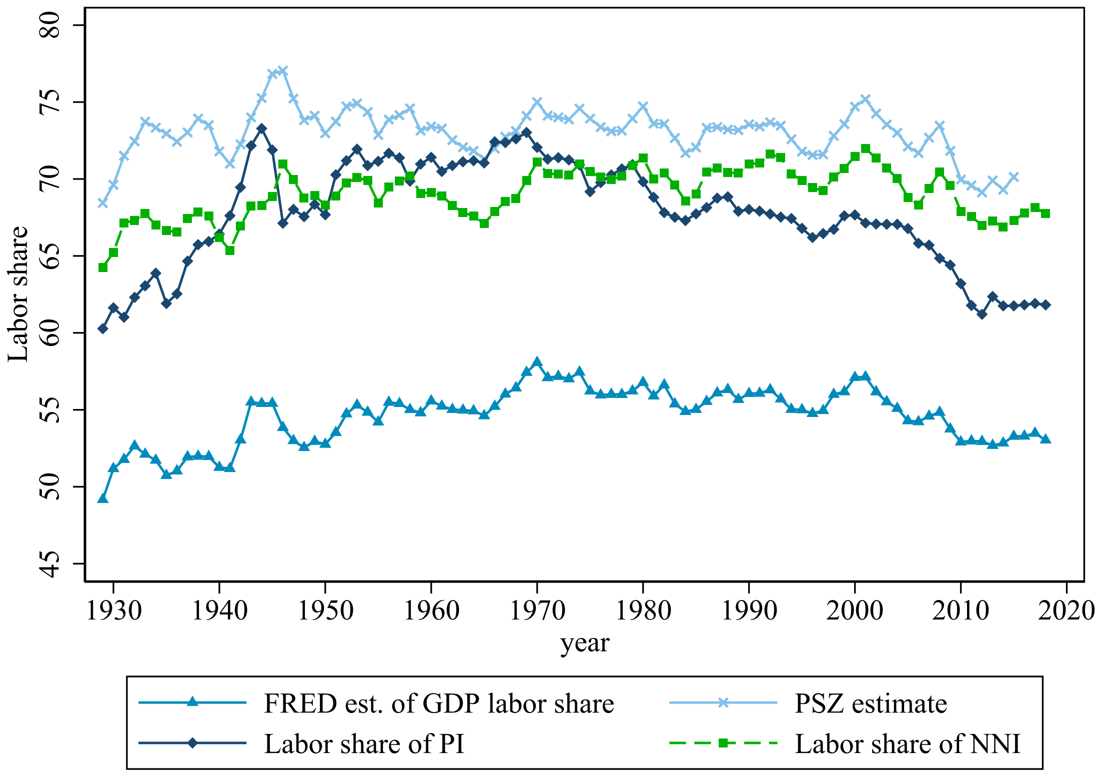
Saez, Emmanuel and Gabriel Zucman (2016). "Wealth inequality in the United States since 1913: Evidence from capitalized income tax data". *The Quarterly Journal of Economics* 131.2, pp. 519–578.

APPENDIX FIGURE H.1: SIMILARITY OF SHARES OF CAPITAL GAINS PLUS DIVIDENDS AND SHARES OF DIVIDENDS, INTEREST, AND RENTAL INCOME.



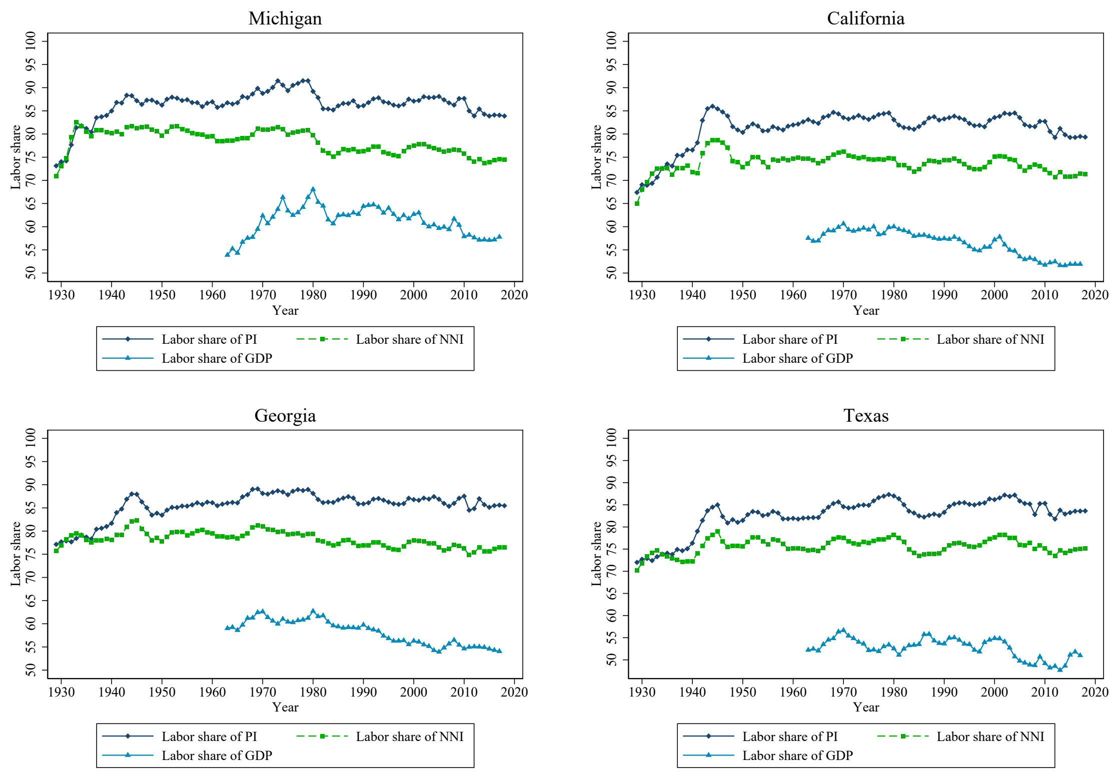
Notes: This graph shows coefficients β^t and confidence intervals from separate regressions of the form $X_{st}^{y_{kgains}+y_{div}} = \alpha_t + \beta^t X_{st}^{y_{div}+y_{int}+y_{rental}} + \epsilon_s$ where X_{st}^{y} denotes the share of taxable income y accruing to residents of state s in year t. Data from IRS Statistics on Income, with rental income y_{rental} calculated from real estate taxes paid as described in the text.

APPENDIX FIGURE H.2: TIME SERIES OF AGGREGATE LABOR SHARE MEASURES



Notes: This graph plots the time-series estimate of our constructed labor share of net income measure (which excludes indirect taxes) against the Piketty, Saez, and Zucman (2018) measure of labor share (which includes indirect taxes) and the GDP and Personal Income based measures of labor share from the BEA.

APPENDIX FIGURE H.3: TIME SERIES OF LABOR SHARE: HIGH AND LOW UNION DENSITY STATES



Note: In this figure we plot our state-year estimate of net income based labor share and compare it to the personal income based and GDP based labor income shares for two high union density states (Michigan and California) and two low union density states (Georgia and Texas).

APPENDIX TABLE H.1: STATE-YEAR LABOR SHARE OF PERSONAL INCOME AS A FUNCTION OF UNION DENSITY (ALL YEARS)

	Dep't var: Labor share of personal income						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Household union share	4.063***	6.176***	6.690***	4.478**	4.478**	3.821*	1.374
	[1.338]	[2.026]	[2.162]	[2.057]	[1.933]	[1.982]	[0.916]
Observations	3,537	3,537	3,537	3,537	3,537	3,537	3,537
Min Year	1937	1937	1937	1937	1937	1937	1937
Max. Year	2014	2014	2014	2014	2014	2014	2014
SouthXyear FE	Yes	Yes	Yes	Yes	Yes	No	Yes
Split-Sample IV	No	Yes	Yes	Yes	Yes	Yes	Yes
Income covars.	No	No	Yes	Yes	Yes	Yes	Yes
Industry Shares	No	No	No	Yes	Yes	Yes	Yes
Policy covars.	No	No	No	No	Yes	Yes	Yes
RegionXyear FE	No	No	No	No	No	Yes	No
State-spec. quad.	No	No	No	No	No	No	Yes

Data sources: Labor share of Personal Income constructed from BEA accounts as described in Appendix H text. For specification descriptions and other variables see notes to Table III.

APPENDIX TABLE H.2: STATE-YEAR LABOR SHARE AS A FUNCTION OF UNION DENSITY (FOR 1963+, WHEN WE HAVE GDP LABOR SHARE)

	Dependent Variable:					
	NNI		GDP (63+)		NNI (63+)	
	(1)	(2)	(3)	(4)	(5)	(6)
Household union share	5.567***	1.090	2.517	-0.521	3.822	1.876
	[1.870]	[1.029]	[3.875]	[2.049]	[2.868]	[1.304]
Observations	3,537	3,537	2,395	2,395	2,395	2,395
Min Year	1937	1937	1963	1963	1963	1963
Max. Year	2014	2014	2014	2014	2014	2014
SouthXyear FE	Yes	Yes	Yes	Yes	Yes	Yes
Split-Sample IV	Yes	Yes	Yes	Yes	Yes	Yes
Industry Shares	No	Yes	No	Yes	No	Yes
Policy covars.	No	Yes	No	Yes	No	Yes
RegionXyear FE	No	Yes	No	Yes	No	Yes
State-spec. quad.	No	Yes	No	Yes	No	Yes

Data sources: Labor share of GDP from BEA. For specification descriptions and other variables see notes to Table III.