Unions and Inequality Over the Twentieth Century: New Evidence from Survey Data

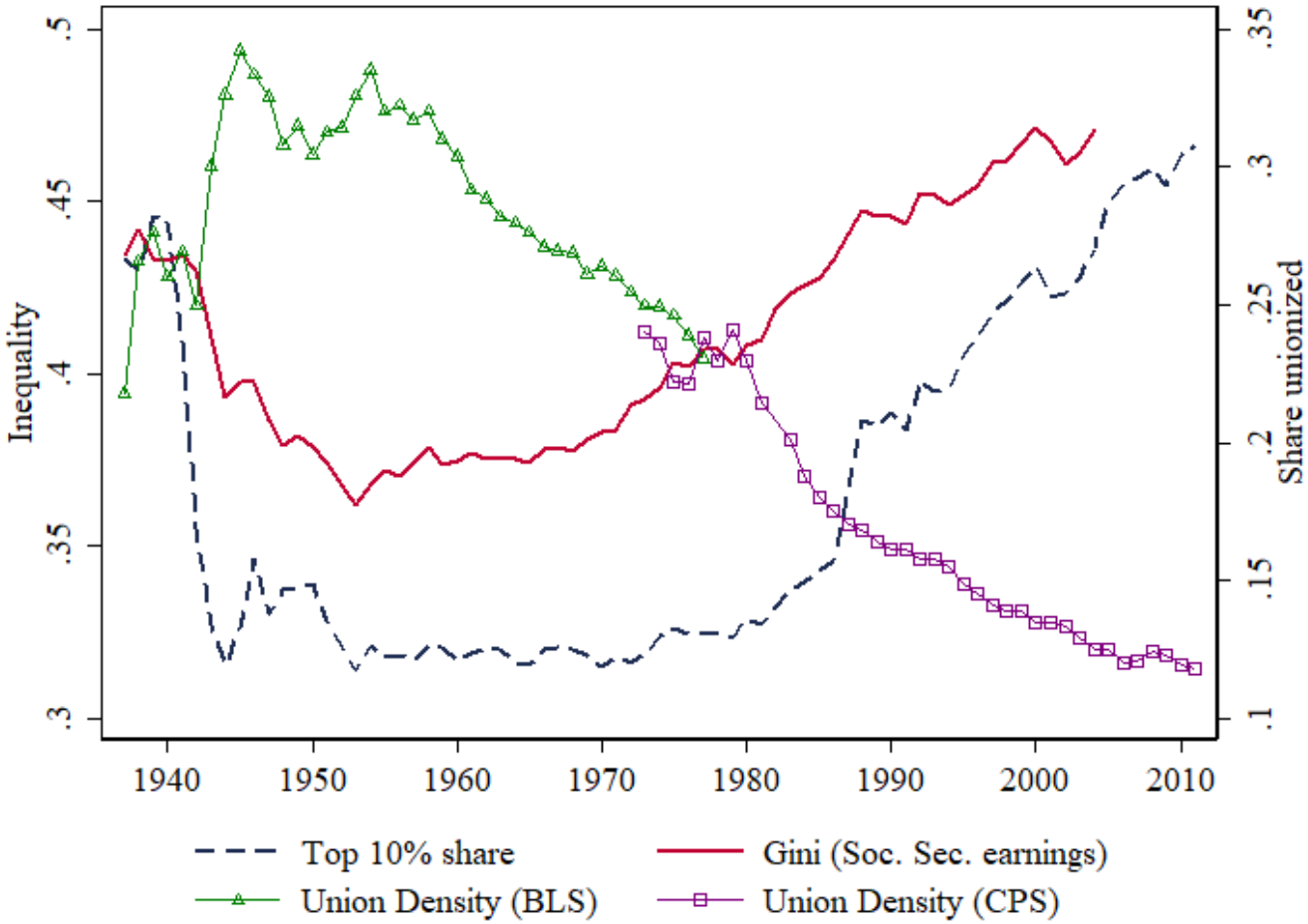
H. Farber, D. Herbst, I. Kuziemko, and S. Naidu

July 2018

Unions and inequality, past and present

- Large literature argues for causal effect of unions on inequality (Card (2001); DiNardo *et al.* (1996); Western and Rosenfeld (2011)).
  - Even larger literature emphasizes market forces as determinants of inequality (Autor *et al.*, 2008).
- Time series variation over the 20th century quite *suggestive* of inverse relationship between union density and income inequality.

Top income shares and union density



Limitations of this picture

- Prevailing (macro) view is that increasing inequality and de-unionization are jointly driven by market forces (Acemoglu *et al.* (2001); Dinlersoz and Greenwood (2012); Açıkgöz and Kaymak (2014))
- But data limitations have prevented investigation of alternative explanations
  - Census Bureau doesn't have a consistent question on union membership until 1970s CPS.
  - Earlier data come from *aggregate* union reports (with major data quality issues)
  - Pre-CPS, cannot look at union wage premium, demographic selection into unions, or state-level trends
- This project tries to address these limitations.

Our paper

- Develop new household-level dataset from over 500 Gallup and ANES opinion polls beginning in 1936
  - Include union membership, demographics, political views
  - Span period of unions' rise, heyday, and decline
- Document empirical patterns of pre-CPS union membership
  - Selection into unions by race and education
  - Union wage premium over time
- Directly estimate unions' effect on inequality
  - Distributional regressions (RIFs) separately by year (micro effect).
  - Plenty of reasons to think micro effect misses many macro effects (unemployment, tech. change, returns to capital/skilled labor).
    - Time-series regressions as in Goldin and Katz (2009)
    - State-year panel regressions
    - No clean instrument, but pattern of results consistent.

Preview of Results

- Highlights from descriptive results
  - Stable union premium over time (10-20%)
  - u-shaped selection pattern: mid-century union members were relatively less skilled than either today or pre-WW2.
  - Consistent with a causal theory of unions-inequality relationship
- Distributional regressions, time-series regressions, and state-year panel regressions all suggest increased union density lowers inequality.

Outline

1. Background and Data
   - Gallup opinion polls
   - New estimates of U.S. union density
2. Who joined unions? And why?
   - Selection into unions by education
   - Selection into unions by race
3. Estimates of the union premium over the last nine decades
   - Did unions benefit certain groups more?
4. Unions and the income distribution
   - Unconditional quantile regression analysis
   - Time-series analysis
   - State-year panel regressions

What do Unions do (to inequality)?

- Traditional mechanism: union wage premium lowers skill differentials.
  - Old view: union premium a monopoly distortion. (Rees 1963).
  - More recent view: union premium reflects product market rent-sharing. (e.g. Abowd and Lemieux 1993).
  - Even more recent: union premium counteracts monopsony power (Benmelech et al. 2018)
- Early literature (Lewis 1960) emphasized inequality between covered and uncovered sectors.
  - Widespread belief that unions *increased* inequality.
  - Freeman and Medoff 1984: within-sector compression bigger than between-sector difference.
- But unions can affect income distribution beyond effect on its members wages ("threat" effect, negative spillovers, political economy mechanisms, fairness norms).

Gallup survey data

- Over 500 surveys, $N \approx 980,000$
- Repeated cross-section at $\approx$ monthly frequency from 1936-1986
- Nationally representative*, cell-weighted by race-region Sampling
- Include union membership, gender, race, education, age, occupation, and a variety of political opinions

Typical Gallup codebook (1946)

371-T

Form T

21. It has been suggested that the Government take over most of the wheat and fats in this country to feed people in other nations during the next 60 days. This would leave very little for people in this country during that time. Would you favor or oppose such a plan?
Col. 43 — [1] Favor [2] Oppose [3] No Opinion

22. In politics, as of today, do you consider yourself a Republican, Democrat, Socialist, or Independent?
Col. 44 — [1] Rep. [2] Dem. [3] Socialist [4] Independent

23. What is your occupation? (Record *SPECIFIC* occupation, not just industry or name of organization worked for).
Write in description of occupation: See code.
Cols. 45, 46
(If housewife, widow or student, record occupation of head of family. If retired or unemployed, give former occupation).

24a. Do you remember FOR CERTAIN whether or not you voted in the 1944 presidential election?
Col. 1 — [1] Yes, voted [2] No, didn't vote [3] No, too young to vote [V] Don't remember

IF YES, VOTED, ask:
Col. 7 — b. Did you vote for Dewey, Roosevelt or Thomas?
[4] Dewey [5] Roosevelt [6] Thomas [7] Other

25. What is the last grade or class you completed in school?
Col. 48 — [1] No schooling [2] Grammar school (grades 1 through 8) [3] High school, incomplete (9th, 10th or 11th grade) [4] High school, graduated (12th grade) [5] College, incomplete [6] College, graduated
What type of college? See code

26a. Are you (or is your husband), a member of a labor union?
[1] Yes [2] No

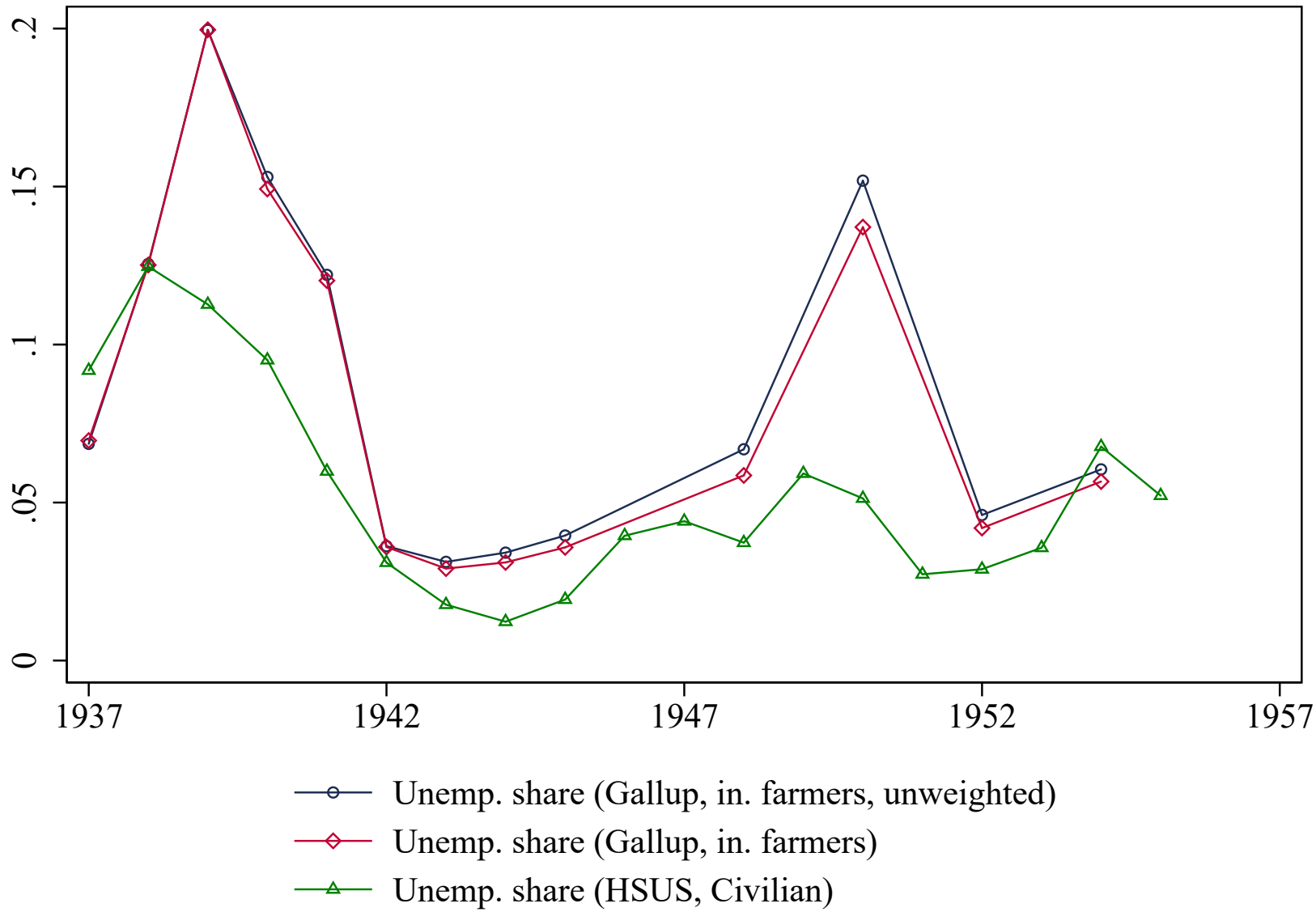
If YES, ask:
Col. 49 — b. Which one?
[3] C.I.O. [4] A.F. of L. [5] Other.....

27a. Is there a telephone in your home? See code
[1] Yes [2] No

IF YES, ask:
b. Is the telephone listed either under your name or the name of a member of your immediate family?
Col. 50 — [1] Yes [2] No

ASK ONLY OF MEN:
28. Did you serve in any branch of the armed forces in World War II?
Col. 51 — [1] Yes [2] No

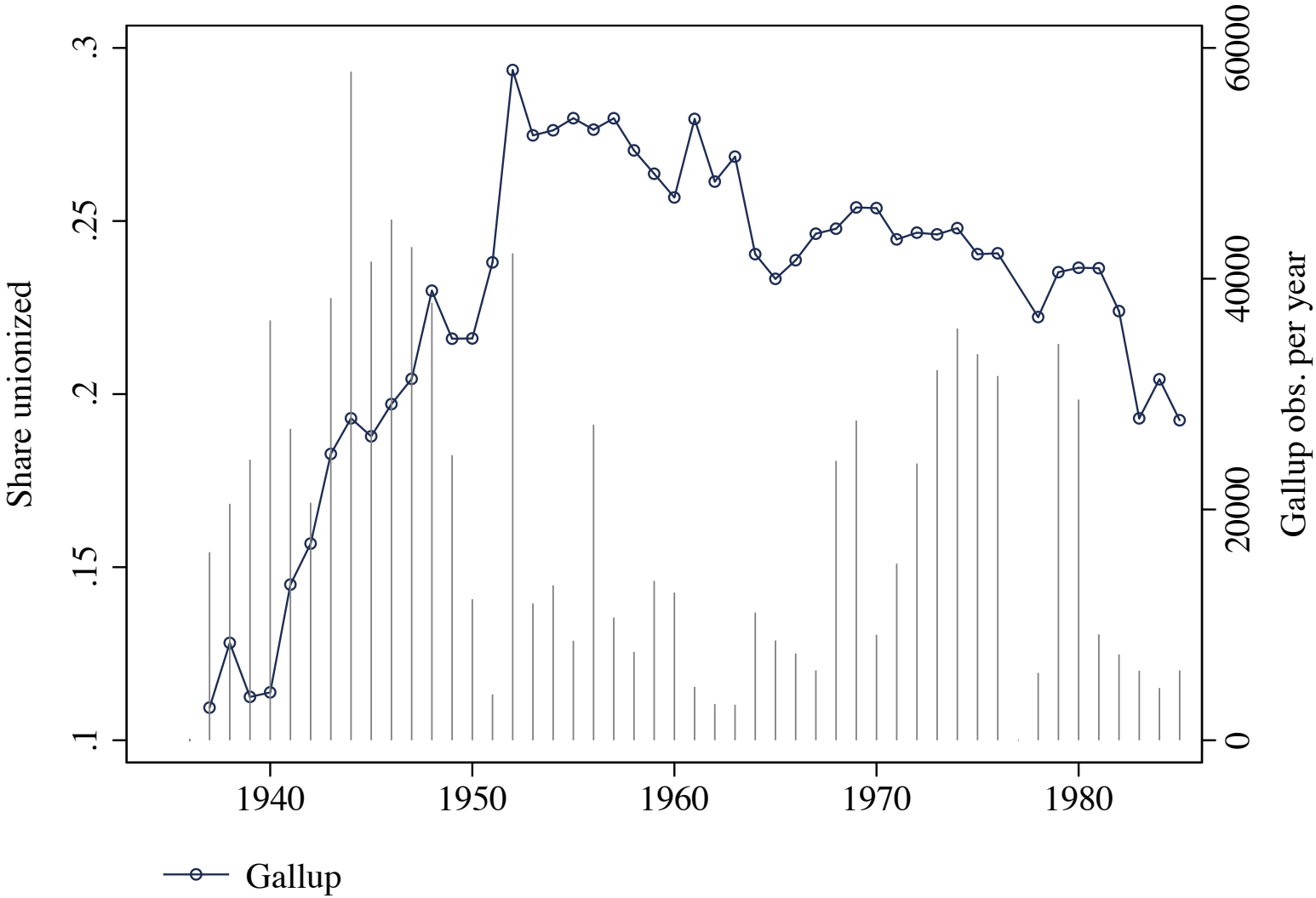
Check: Unemployment Reasonably Close to HSUS



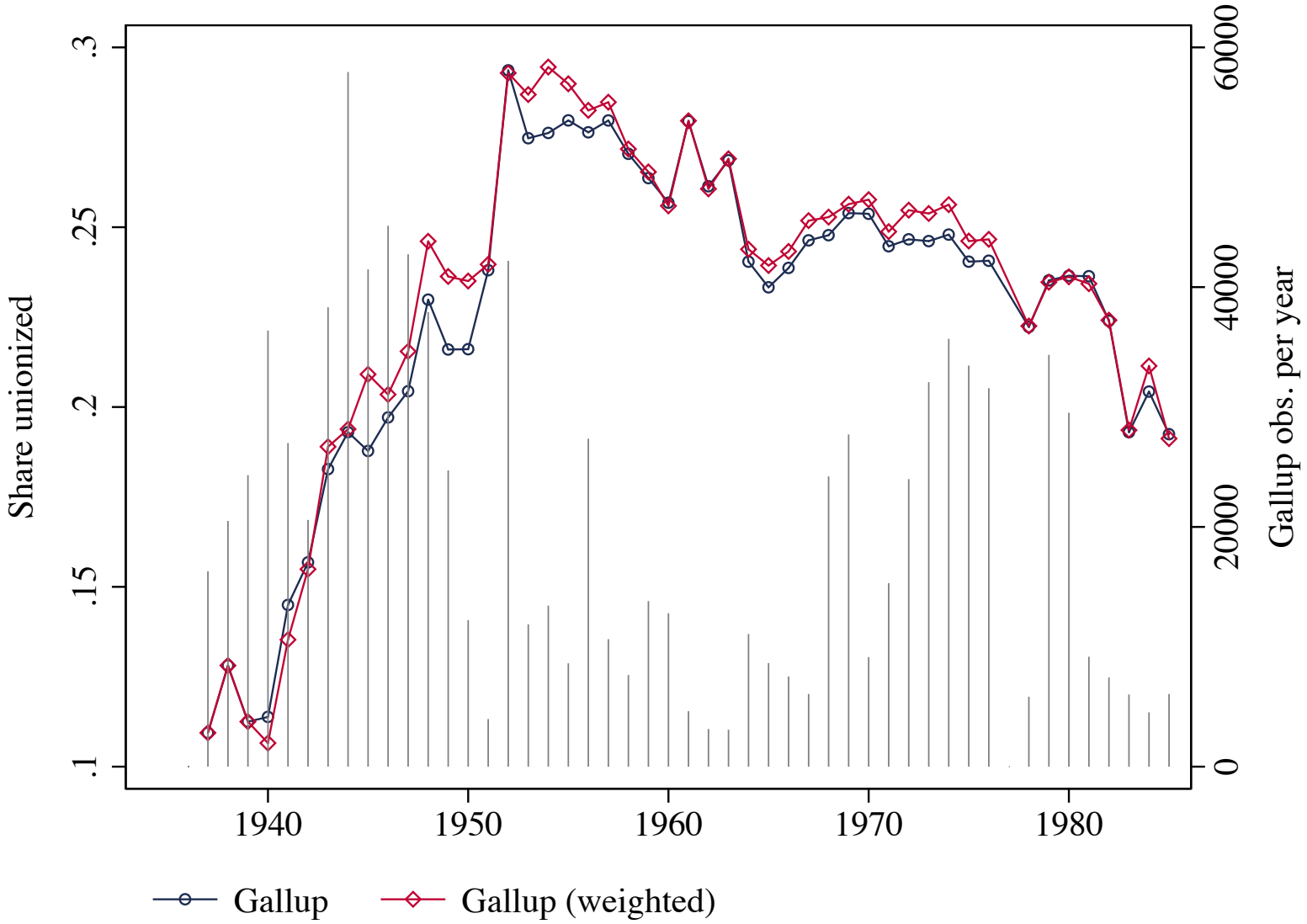
Typical Gallup union question

- "Are you (or is your husband) a member of a labor union?"
  - Most (but not all) years choices are: "neither," "yes, I am," "yes, he is," "yes, both are."
  - We harmonize so that union household coded as one if *either* is a member, zero otherwise.
- Implied unit is an individual or couple though we use it to proxy household union status.
- Limitation: Gallup does not ask industry, so we cannot break into public vs. private-sector union.
- Use other surveys, especially ANES, with union household question: "Does anyone in the household belong to a labor union?" Other Data

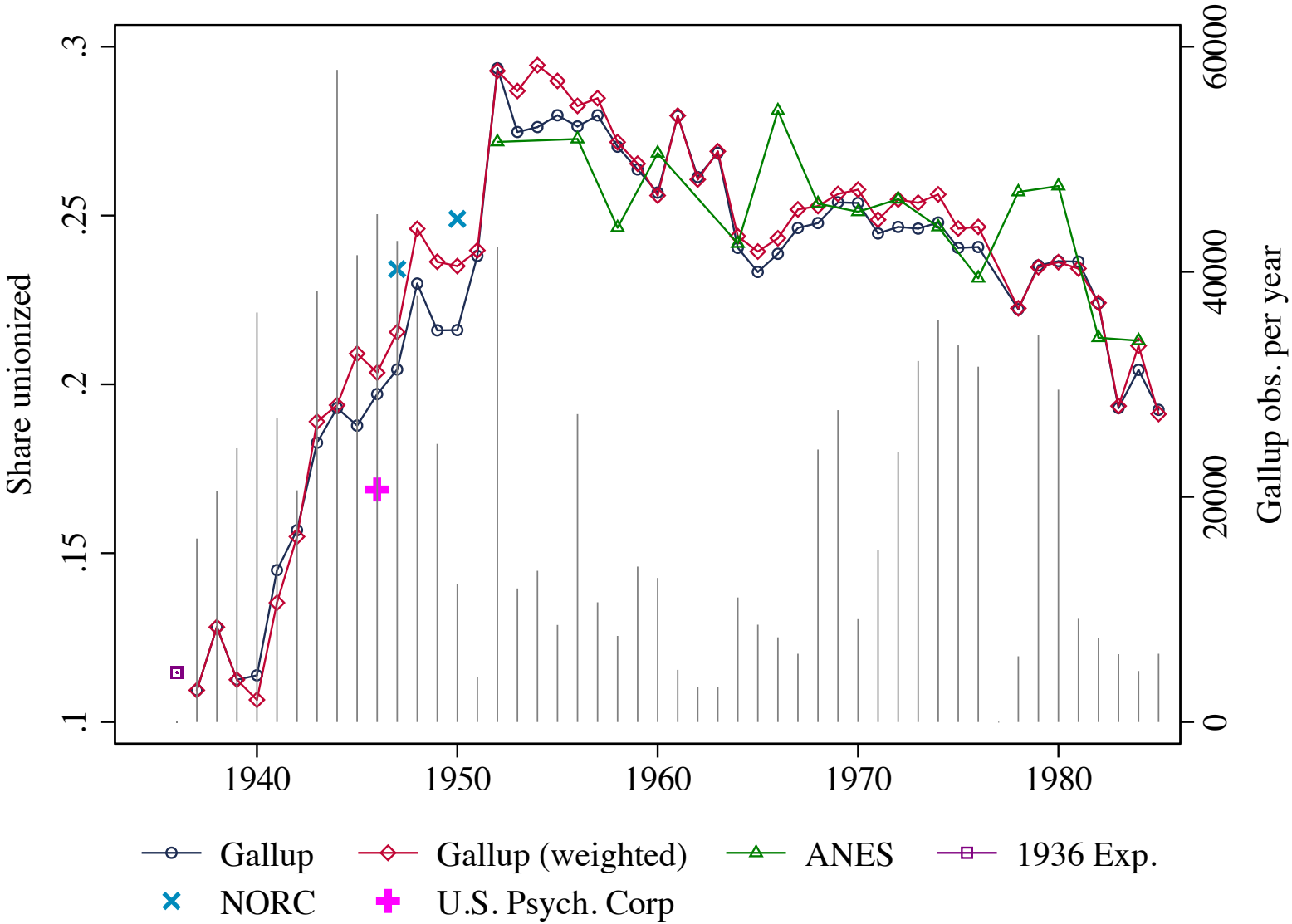
Comparing our new series to existing series



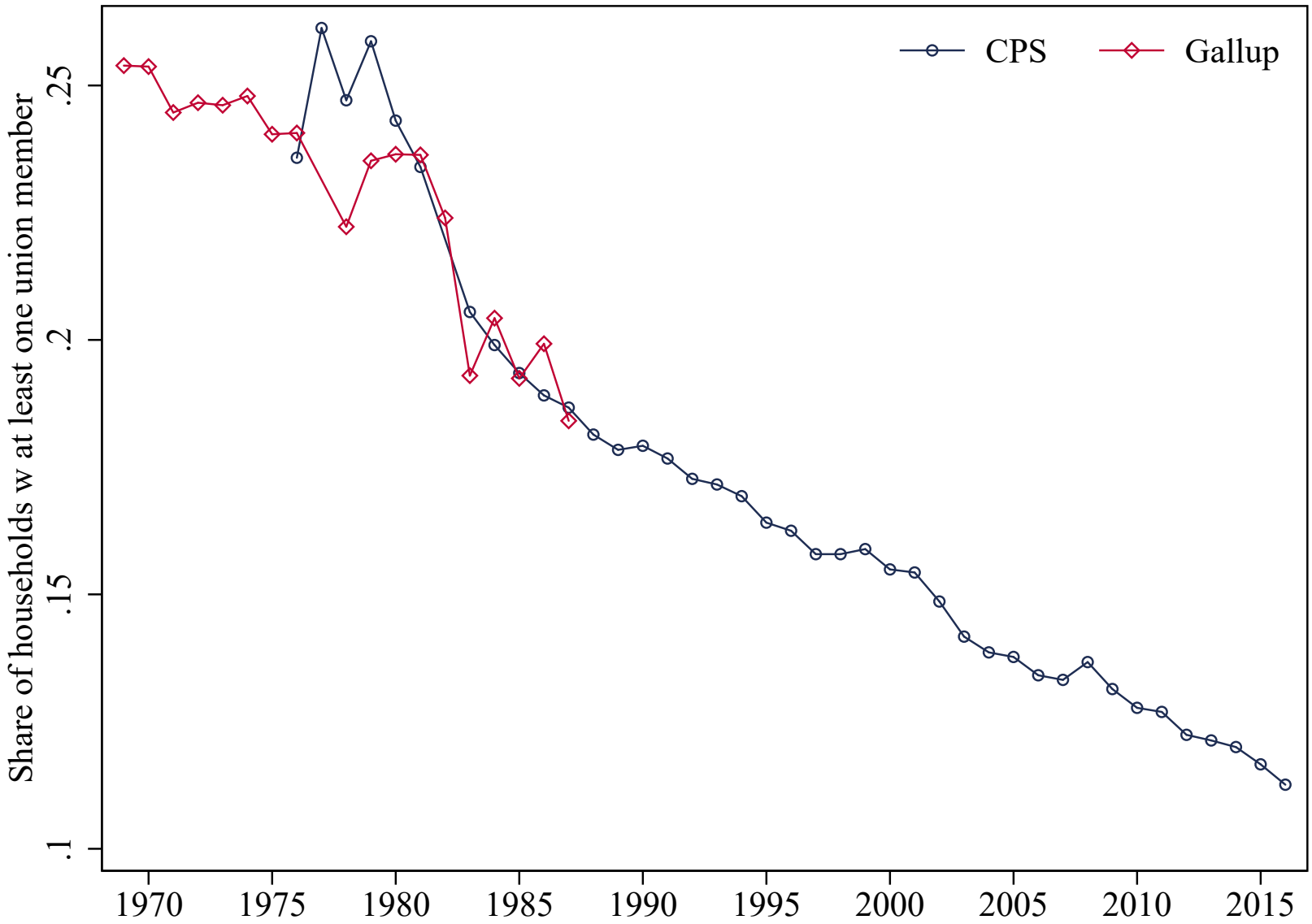
Comparing our new series to existing series



Comparing our new series to existing series



Comparison with CPS over more recent years



Outline

1. Background and Data
   - Gallup opinion polls
   - New estimates of U.S. union density
2. Who joined unions? And why?
   - Selection into unions by education
   - Selection into unions by race
3. Estimates of the union premium over the last nine decades
   - Did unions benefit certain groups more?
4. Unions and the income distribution
   - Unconditional quantile regression analysis
   - Time-series analysis
   - State-year panel regressions

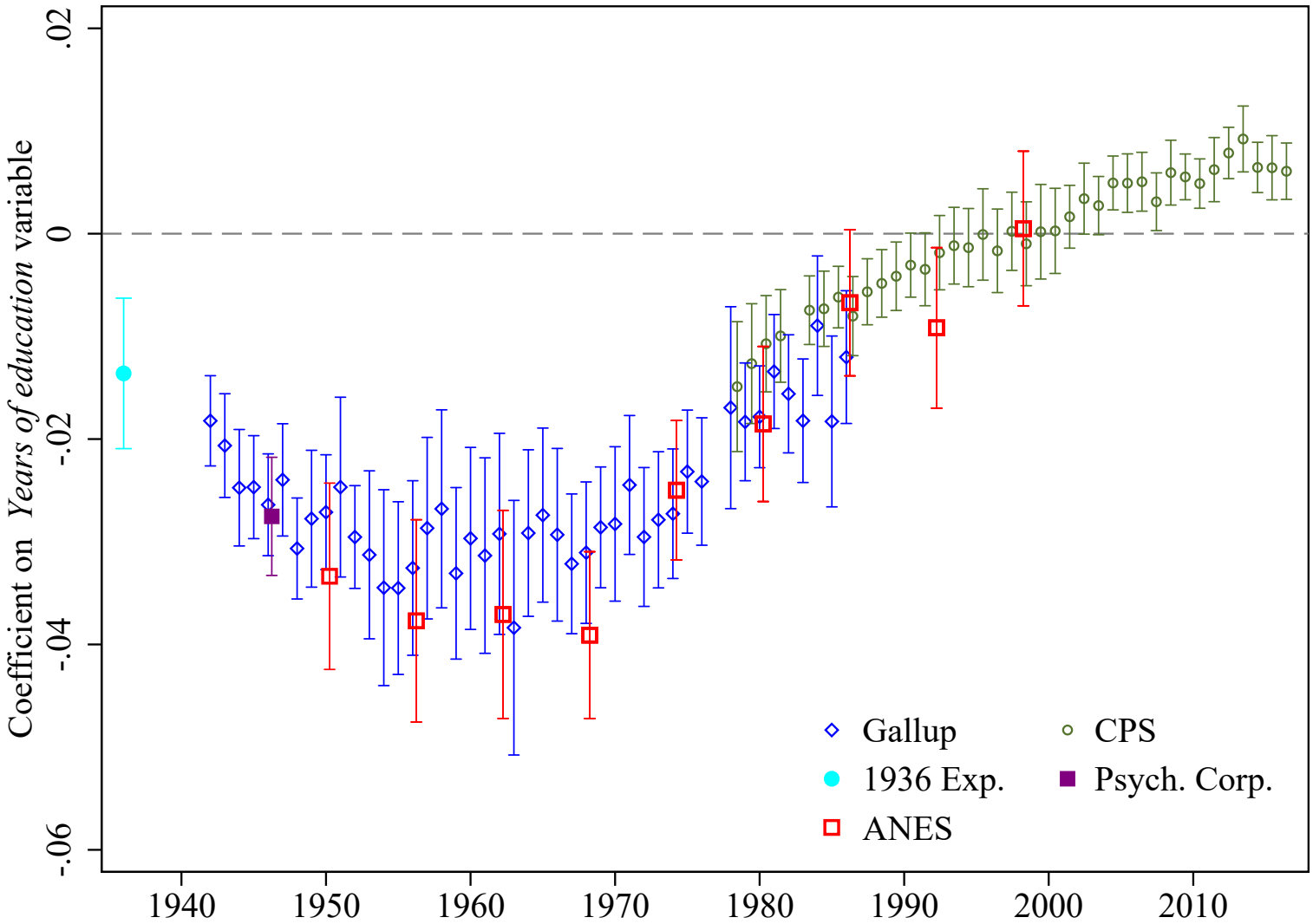
Estimating selection over time

- Covariate X of interest: education and race
- We estimate, separately by survey source and year y:

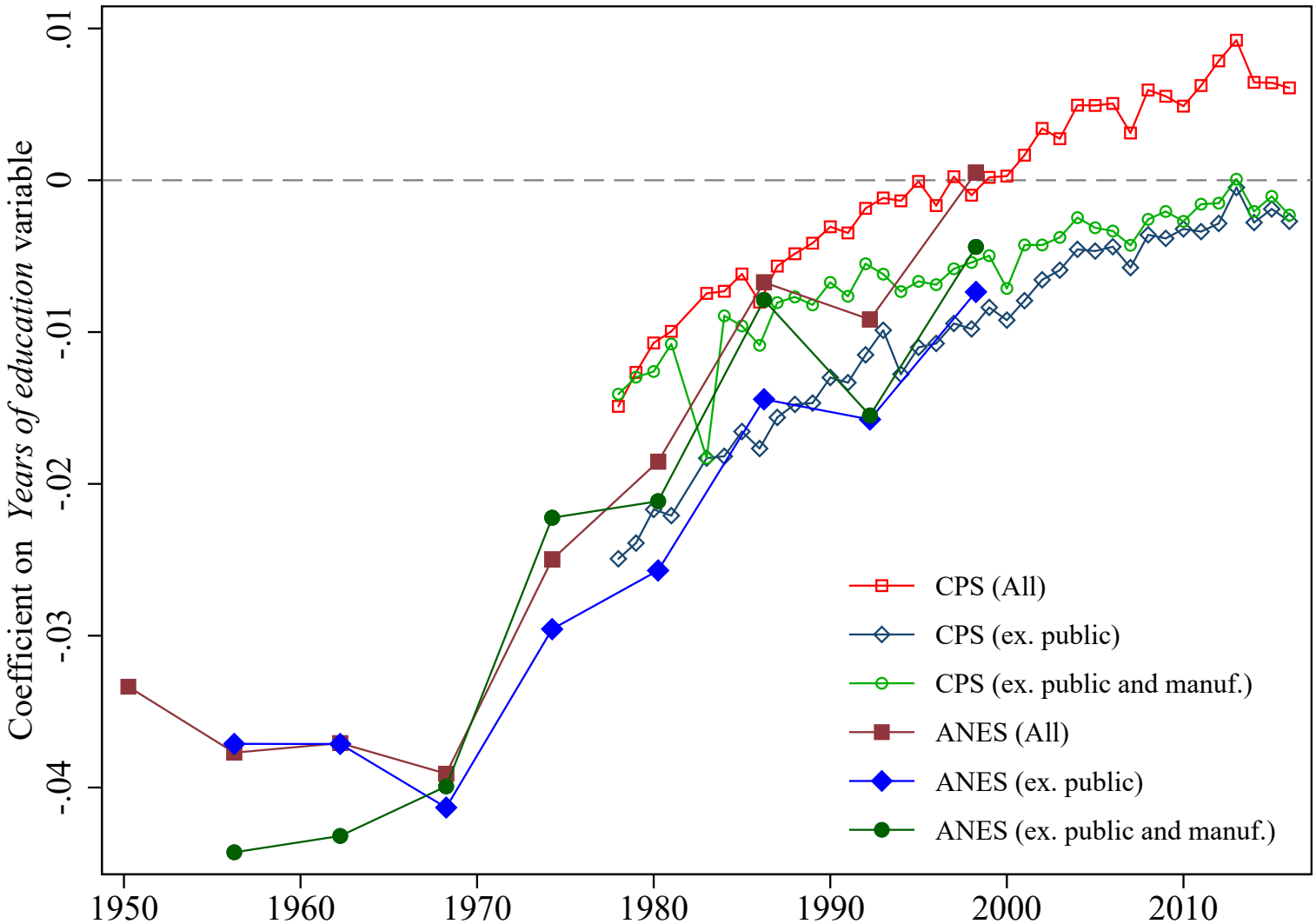
$$union_{hst} = \beta_y X_{hst} + \gamma D_{hst} + \mu_s + \nu_t + e_{hst},$$

where D_{hst} are basic demographics (age & its square, gender), X_{hst} is the covariate of interest, μ_s and ν_t are state and survey-date fixed effects.

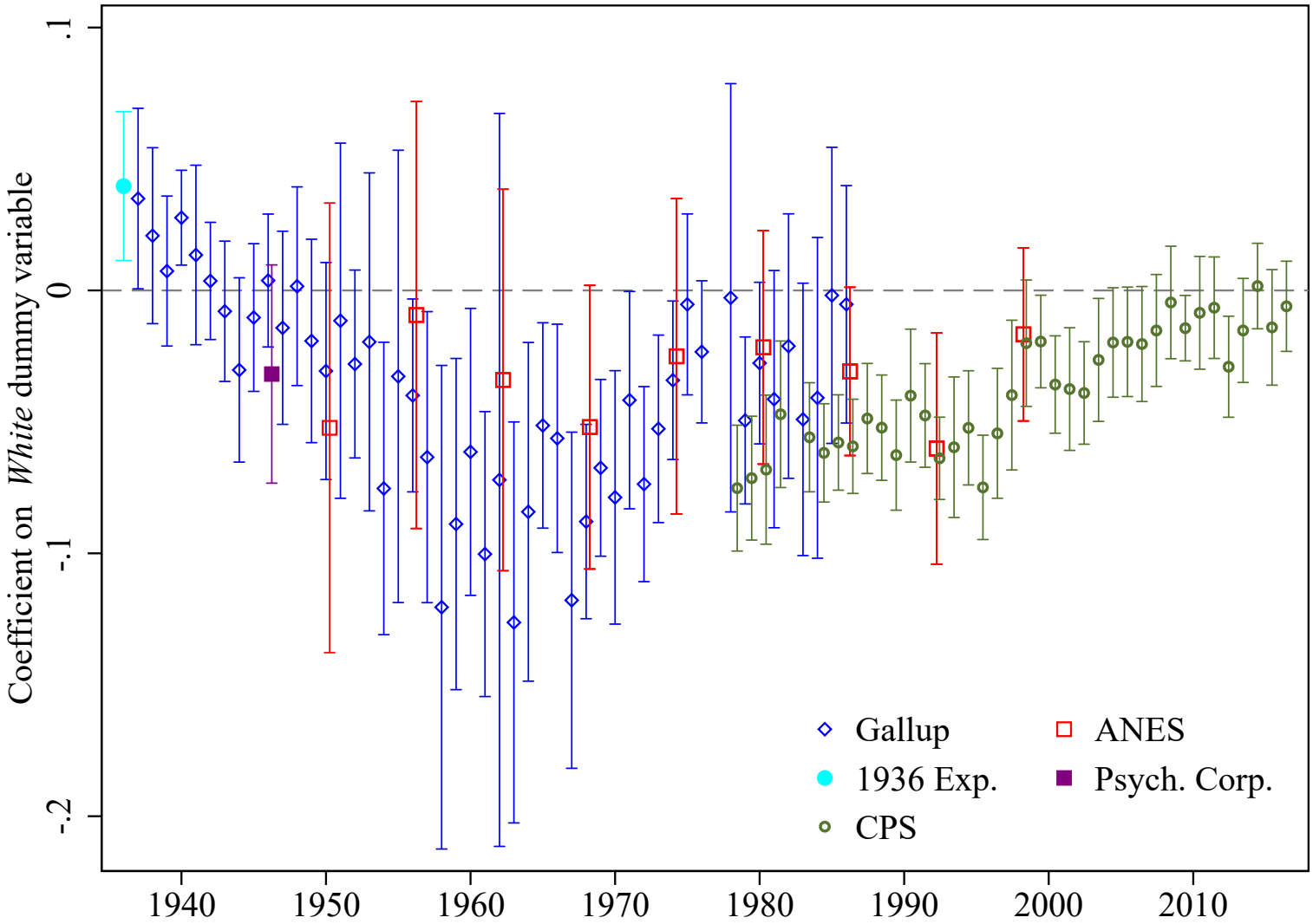
Selection into unions by education



Not driven by rise of public sector nor fall of manufacturing alone



Selection into unions by race



Driven by Great Migration? Recall we include state FE, and in Appendix, we show result looks the same if we simply drop the South.

Unpacking Selection

- Aggregate u shape in selection driven by inverse-u in aggregate density, not just time-quadratic.
- Same relationship between density and selection exists at state-year level, conditional on state + year FE.
- When union density is high union members are much more low-skilled.

Selection and Drivers of Inequality

As union density increases, the marginal member is increasingly negatively selected

- Consistent with causal effect of unions on inequality.
  - Inequality increases as low-skilled no longer receive union premia
- Inconsistent with some SBTC models (e.g. Acemoglu *et al.* (2001))
  - Predict unions should become *less* skilled over time, as high skill workers opt-out of union

Outline

1. Background and Data
   - Gallup opinion polls
   - New estimates of U.S. union density
2. Who joined unions? And why?
   - Selection into unions by education
   - Selection into unions by race
3. Estimates of the union premium over the last nine decades
   - Did unions benefit certain groups more?
4. Unions and the income distribution
   - Unconditional quantile regression analysis
   - Time-series analysis
   - State-year panel regressions

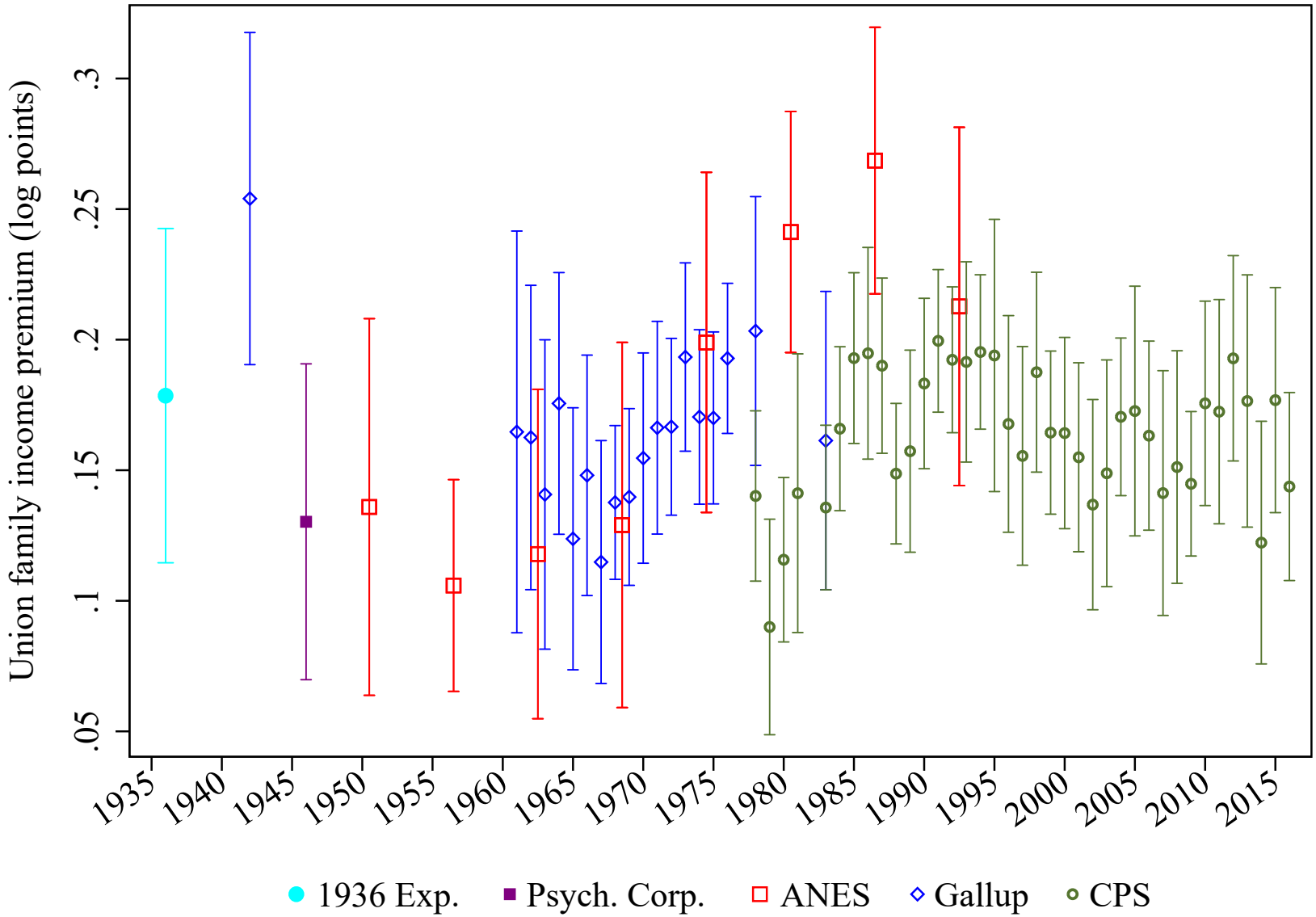
Household income premium

- Polling data typically ask about *household* income, whereas union wage usually estimated on *individual earnings*.
- We estimate the following household income function, *separately by year and data source*:

$$\log(y_{hst}^{HH}) = \beta Union_h^{HH} + \gamma_1 Black_h^R + \gamma_2 Female_h^R + f(age^R) + g(education_h^R) + k(employment_h) + \lambda_s + \mu_t + e_{hst},$$

- where h denotes household, R respondent, s state, and t survey number.
- In Appendix we show results controlling for *occupation of household head*, but categories vary across surveys.
- Note an implicit assumption is assortative matching across households. All covars have expected signs (see Appendix).

Union Family Income Premium, 1936–2016.



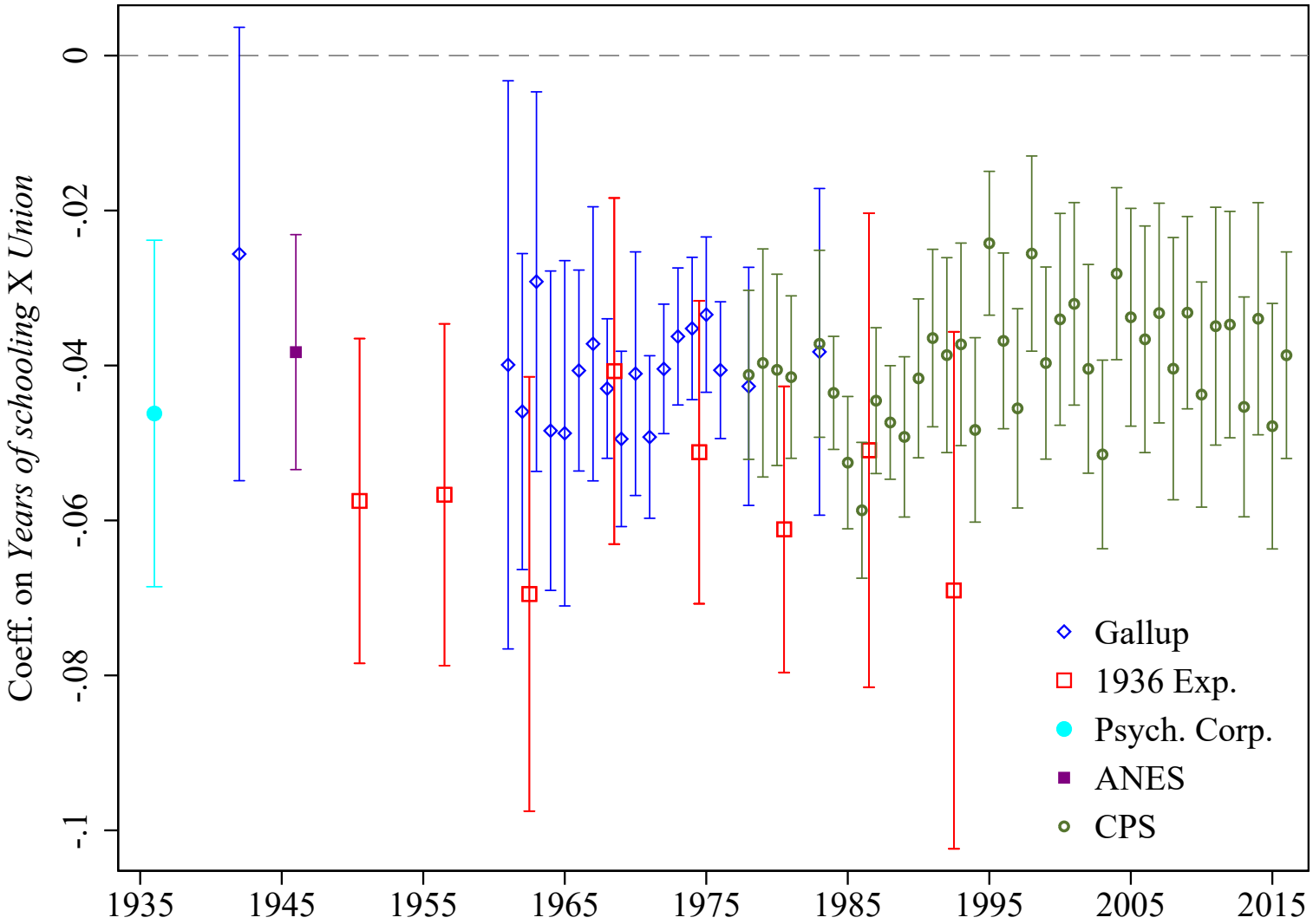
Summary of union premium results

- Evidence suggests that union households were far better off than their education or demographics would suggest. Premium reasonably steady (given wide std. errors) over the 1936 to 2016 period.
- How to compare family premium with traditional wage premium?
  - Family premium *not* driven by more workers per household.
  - Union households have slightly more workers, mechanically, but ANES data suggest little or no bias.
- Also show significant non-wage benefits: union household more likely to have paid vacations back in 1940s (see Appendix), especially low-status households.
- Rents: Union households more likely to say hard to find as good a job as this one.

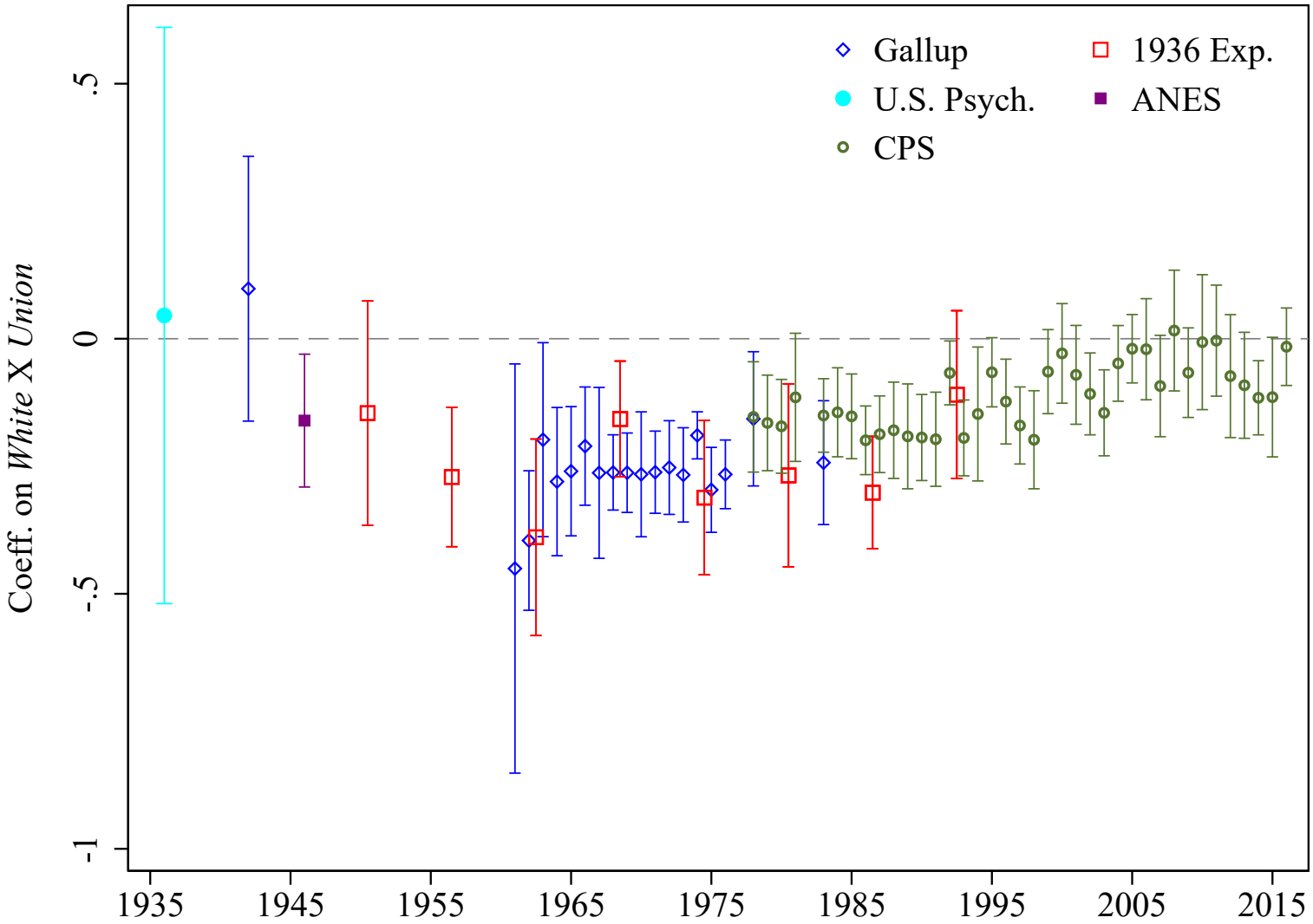
Did unions benefit certain groups more?

- We interact the $Union^{HH}$ var with, respectively, years of education and then a *white* dummy.

Differential premium by years of education



Differential premium by race



Outline

1. Background and Data
   - Gallup opinion polls
   - New estimates of U.S. union density
2. Who joined unions? And why?
   - Selection into unions by education
   - Selection into unions by race
3. Estimates of the union premium over the last nine decades
   - Did unions benefit certain groups more?
4. Unions and the income distribution
   - Unconditional quantile regression analysis
   - Time-series analysis
   - State-year panel regressions

Unions and inequality

- Have shown in the micro data that mechanisms by which unions might reduce inequality active in the historical period.
  - High union density associated with lower-skilled union membership and unions increase relative wages of low-skilled workers.
- Now look directly at income distribution
  - Distributional regression exercise, as in Firpo et al. 2009.
  - Aggregate time series analysis.
  - State-year panel data analysis.

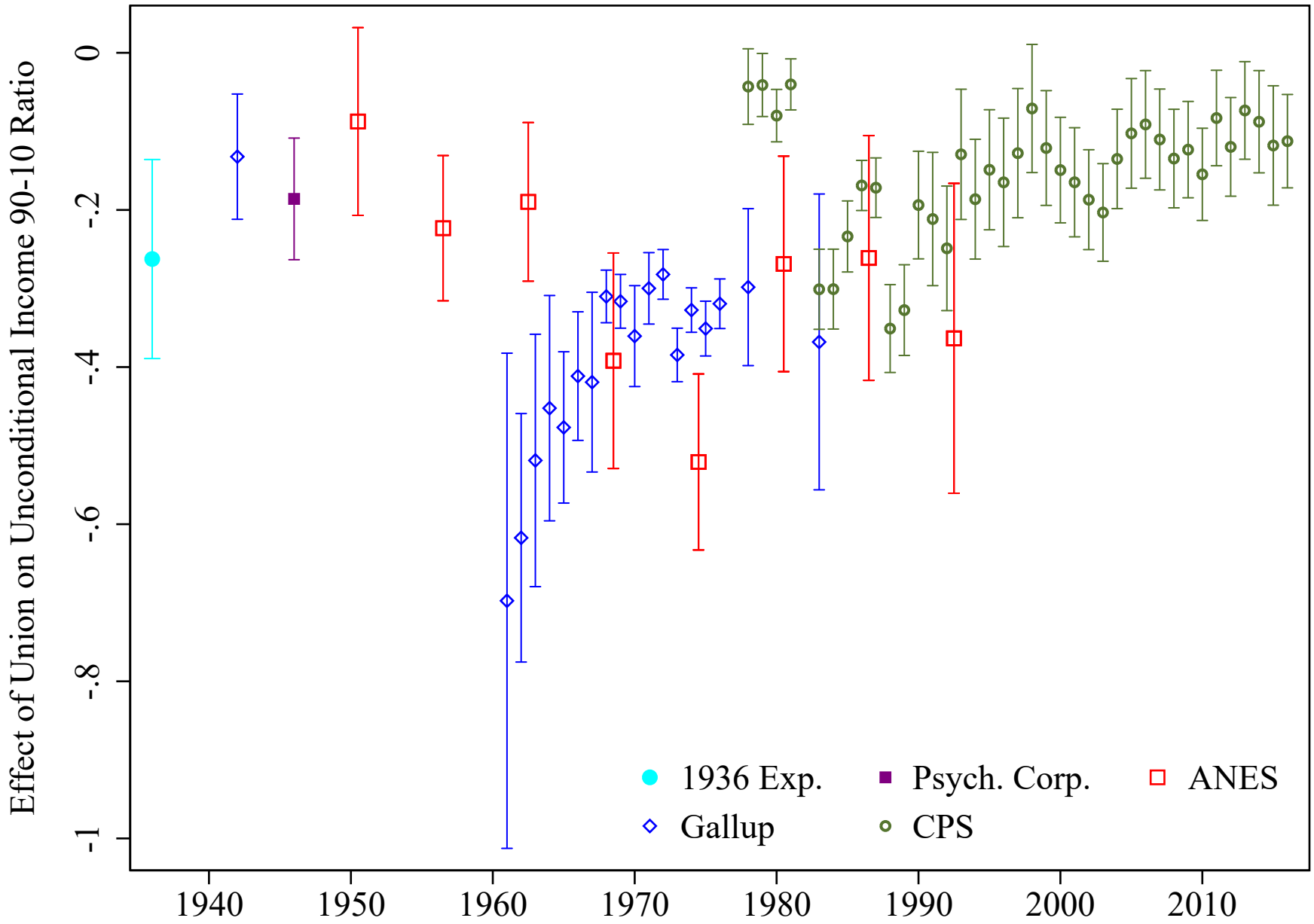
Unconditional quantile regression analysis

- Adapt methodology from Firpo et al. (2009).
- Let $v(F)$ be some distributional statistic (e.g. 90-10, Gini); u a dummy for union status; y family income; X covariates.
- If so, can decompose distribution $F(y, X, u)$ as

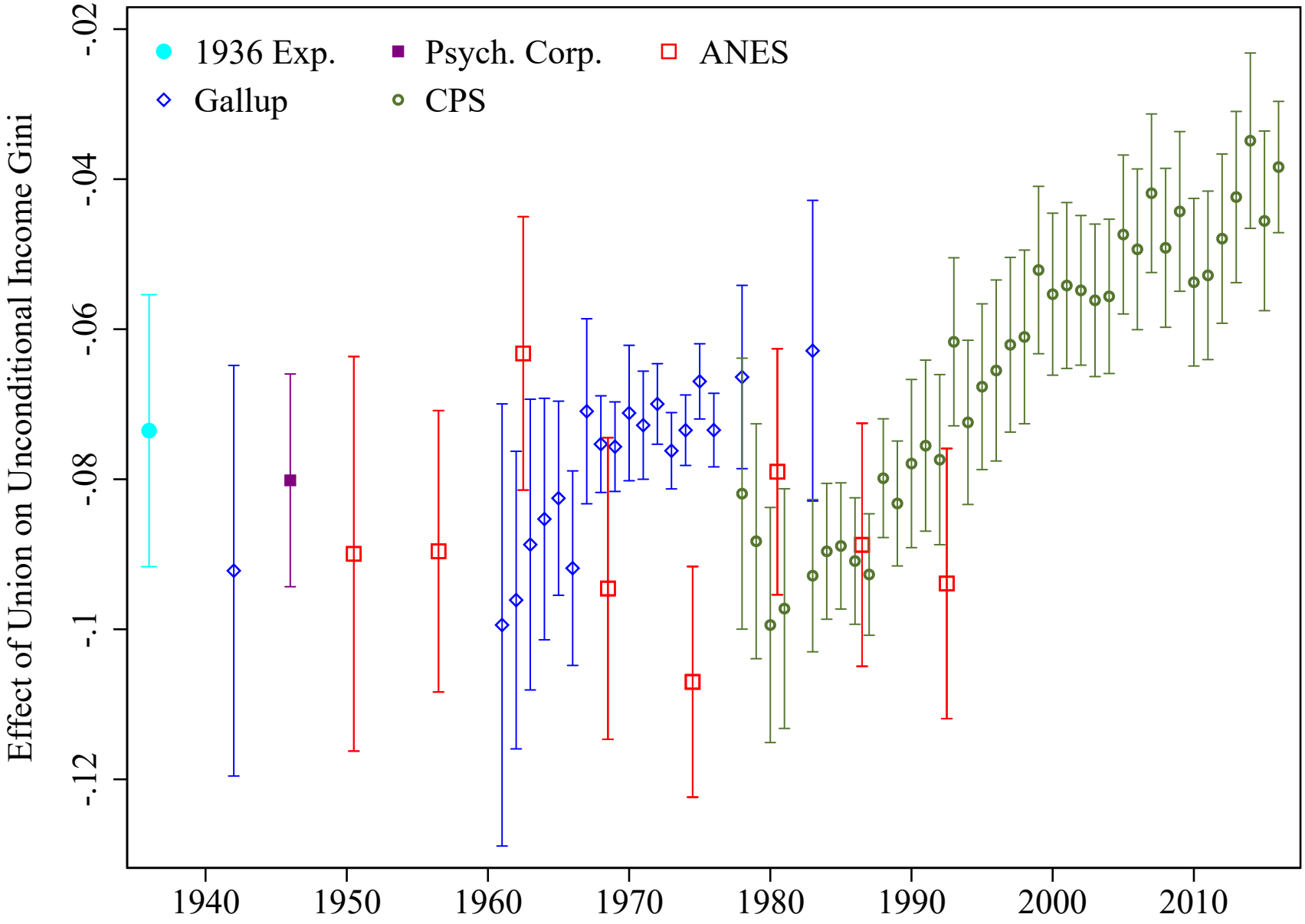
$$Pr(u = 1)F(y, X|u = 1) + \big[1 - Pr(u = 0)\big]F(y, X|u = 0).$$

- OLS coefficient from regression of $RIF(v, F, y_i)$ on union u_i gives estimate of $\frac{dv(F)}{dPr(u=1)}$
- **Interpretation:** How much would inequality fall if you increased the share of union members, *holding the joint distribution of wages and covariates constant.*

Results for 90-10 HH income ratio



Results for Gini coefficient



Time-series analysis

- We use national time-series at either the annual or decadal level.
  - See Katz Murphy (1992), Autor et al. (1998), Goldin and Katz (2009) and Autor et al. (2008) in analysis of skill shares on skill premia (SBTC and polarization literature).
- We essentially adopt their specifications, but add union density (averaged between BLS and Gallup measures) as an additional explanatory variable and explore other outcome variables.
  - College premium (decadal until CPS in 1964).
  - Log 90/10 ratio (decadal until CPS in 1964).
  - Gini coefficient from Social Security earnings (annual, 1937-2004).
  - Top10 share (annual since 1937).

Unions and inequality: Annual time-series

	Dependent variable:			
	Coll. premium		90/10 ratio	
	(1)	(2)	(3)	(4)
Union Density	-1.360**	-1.536**	-1.449***	-1.456**
	[0.559]	[0.741]	[0.451]	[0.678]
Mean, dept. var	0.512	0.512	1.376	1.376
Gallup edu. control?	No	No	No	No
Addit. controls?	No	Yes	No	Yes
Cubic polynomial?	Yes	Yes	Yes	Yes
Observations	49	49	49	49

All specifications include skill shares. Additional Controls are: Federal minimum wage, national unemployment rate, and top marginal tax rate. Newey-West Standard Errors.

Unions and inequality: Annual time-series

	Dependent variable:			
	Gini coeff.		Top 10 share	
	(1)	(2)	(3)	(4)
Union Density	-0.141**	-0.116**	-35.45**	-16.83
	[0.0553]	[0.0552]	[14.83]	[12.13]
Mean, dept. var	0.410	0.410	35.848	35.848
Gallup edu. control?	Yes	Yes	Yes	Yes
Addit. controls?	No	Yes	No	Yes
Cubic polynomial?	Yes	Yes	Yes	Yes
Observations	65	65	70	70

All specifications include skill shares. Additional Controls are: Federal minimum wage, national unemployment rate, and top marginal tax rate. Newey-West Standard Errors.

State-year panel regressions

- In general, the inequality literature hasn't taken a state-year panel approach.
  - In the SBTC literature, concern is the college-educated will migrate to state-years with high college premia, leading to reverse causality.
  - Similarly, unions target places with high profit margins, low-skill workers might migrate to places with unions.
- Nonetheless, we try to exploit variation within states across time:
  - Absorbing state and year FE, and controlling for policy environment, business cycle, industry mix, and skill shares.

Regressing state-year inequality measures on state-year union density

	Coll. premium		90/10 ratio	
	(1)	(2)	(3)	(4)
Household union share	-0.442***	-0.434***	-0.296**	-0.244**
	[0.118]	[0.104]	[0.118]	[0.0953]
Mean, dept. var.	0.490	0.497	1.386	1.398
Industry shares	No	Yes	No	Yes
State-spec. quad.	No	Yes	No	Yes
Income covars.	No	Yes	No	Yes
Policy covars.	No	Yes	No	Yes
Observations	1640	1505	1640	1505

All regs have state and South X year fixed effects. Industry shares are 1-digit employment shares. Income covars are Log GDP/capita and share of hh filing taxes. Policy covars are minimum wage and "policy liberalism" index. SEs clustered by state. $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$

Regressing state-year inequality measures on state-year union density

	Top 10 Share		Gini coeff.	
	(1)	(2)	(3)	(4)
Household union share	-0.0635***	-0.0738***	-5.342***	-3.147**
	[0.0235]	[0.0235]	[2.065]	[1.399]
Mean, dept. var.	0.376	0.378	36.61	36.96
Industry shares	No	Yes	No	Yes
State-spec. quad.	No	Yes	No	Yes
Income covars.	No	Yes	No	Yes
Policy covars.	No	Yes	No	Yes
Observations	1640	1505	3107	2723

All regs have state and South X year fixed effects. Industry shares are 1-digit employment shares. Income covars are Log GDP/capita and share of hh filing taxes. Policy covars are minimum wage and "policy liberalism" index. SEs clustered by state. $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$

How much of the change in inequality can unions explain?

Share explained by Δ union density, using β values from...

	Total Δ	RIF (high)	RIF (low)
—1940-1960			
Union density	0.113		
Male 90/10 ratio	-0.326	0.128	0.0377
Gini coefficient	-0.0585	0.162	0.0892
—1970-2004			
Union density	-0.123		
Male 90/10 ratio	0.465	0.0980	0.0288
Gini coefficient	0.0875	0.118	0.0651

How much of the change in inequality can unions explain?

Share explained by Δ union density, using β values from...

	Total Δ	Time-series	State panel
—1940-1960			
Union density	0.113		
Skill premium	-0.0926	1.894	0.529
Male 90/10 ratio	-0.326	0.504	0.0844
Gini coefficient	-0.0585	0.204	0.142
Top ten percent	-12.77	0.137	0.0278
—1970-2004			
Union density	-0.123		
Skill premium	0.212	0.904	0.252
Male 90/10 ratio	0.465	0.385	0.0645
Gini coefficient	0.0875	0.149	0.104
Top ten percent	12.13	0.158	0.0319

Concluding thoughts

- New data from political polls allows examination of unionized labor markets during heyday of union power
- We find descriptive results consistent with a causal effect of unions on inequality
  - union premium remains relatively stable over time
  - selection increasingly *less* skilled as union density increases
- Direct estimates of unions' effect on inequality negative and significant across a variety of identification strategies
- With care, historical political polling data of considerable value for economic history

THE END

Magnitudes

- Implied effect on Gini roughly similar over all three methods. From 1970–2004, decline in union density could "explain" 9-20% of rise in Gini.
- Implied effect on 90/10 varies across methods: smaller for RIF and state-year (10–20% explained), but large for annual time-series (35%).
- Implied effect for college premium: 13% from state-year, 30% for annual time-series
- Implied state-year effect for top 10 percent income share explains between 2.5 and 5 of increase, implied annual effect explains between 30 and 60.
- 12% decline between 1970-2004 roughly symmetric explanatory power as 11% increase between 1940-1960.
- One point to remember is that state-year reduction in inequality doesn't "aggregate up" to national reduction in inequality.

Comparing Gallup and Census in 1940

	Gallup	Census	Census	Gallup	Census
Black	0.0290	0.0895	0.0906	0.0325	0.0357
Female	0.338	0.505	0.344	0.341	0.343
Age	40.45	39.61	40.06	40.40	40.55
HS Graduate	0.493	0.278	0.266	0.494	0.290
Northeast	0.0835	0.0660	0.0629	0.0946	0.0854
Mid Atlantic	0.262	0.253	0.241	0.297	0.327
East Central	0.207	0.187	0.186	0.235	0.252
West Central	0.176	0.127	0.129	0.200	0.175
South	0.118	0.258	0.263	0	0
Rocky Mountain	0.0751	0.0284	0.0308	0.0851	0.0418
Pacific Coast	0.0784	0.0754	0.0818	0.0888	0.111
College Graduate		0.0472	0.0499	0.0709	0.0543
Gender/HH adj?	No	No	Yes	No	No
Ex. S/SW?	No	No	No	Yes	Yes
Observations	148290	736832	736832	130400	544375

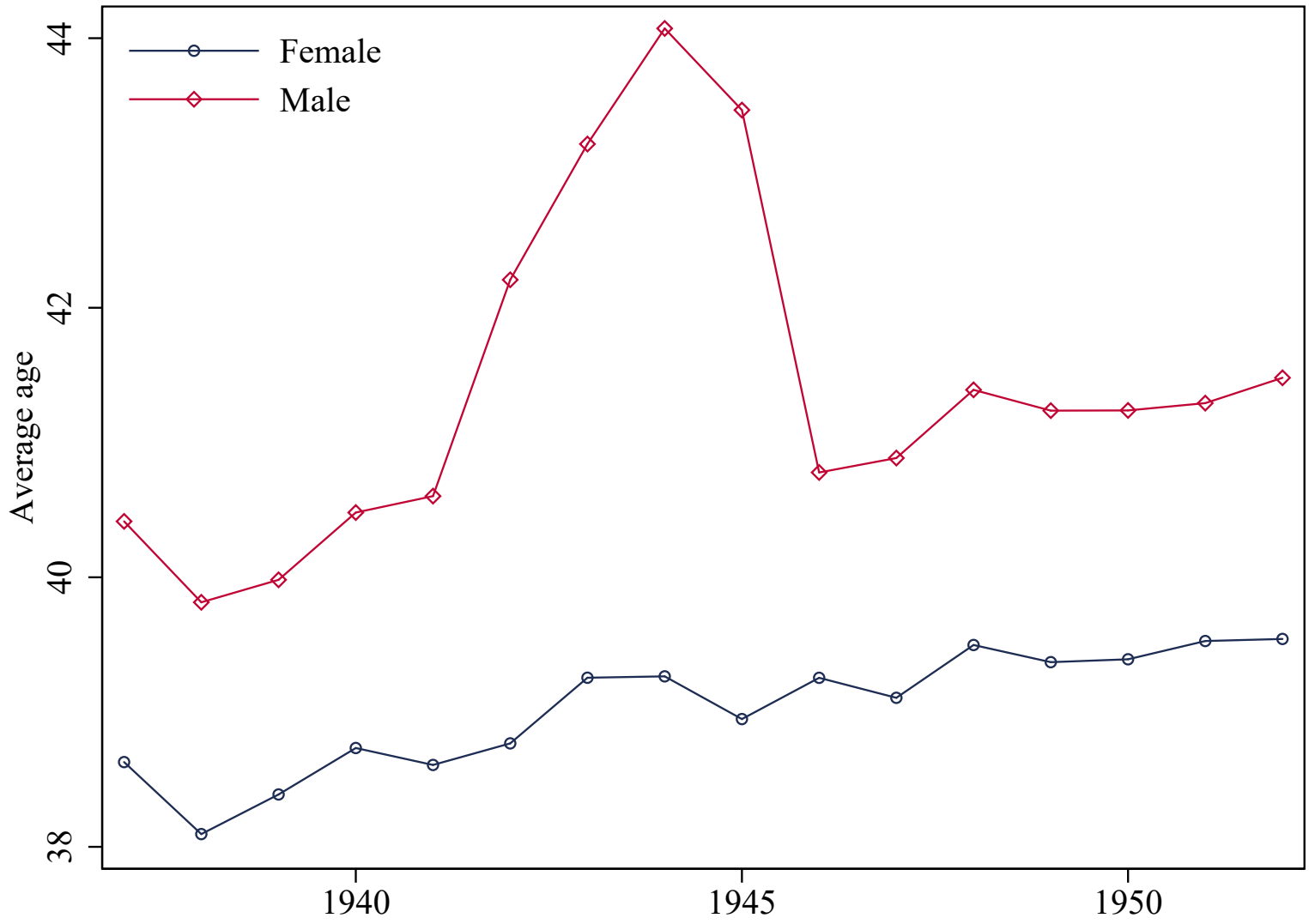
	Gallup	Census	Census	Gallup	Census
Professional	0.0780	0.113	0.122	0.0793	0.129
Farmer	0.209	0.156	0.159	0.185	0.109
Propietors, managers, officials	0.0104	0.0928	0.0875	0.0106	0.0933
Clerks (white collar)	0.294	0.0535	0.0539	0.301	0.0609
Skilled workmen and foremen	0.0906	- -	- -	0.0953	- -
Sales workers	- -	0.0462	0.0457	- -	0.0499
Craftsmen	- -	0.142	0.139	- -	0.153
Operatives	- -	0.146	0.147	- -	0.159
Unskilled or semi-skilled labor	0.190	- -	- -	0.200	- -
Laborers	- -	0.0932	0.0973	- -	0.0944
Service workers (priv. HH)	- -	0.0103	0.0105	- -	0.00626
Other service workers	- -	0.0477	0.0468	- -	0.0508
No answer, N/A, etc.	0.0826	0.0999	0.0920	0.0836	0.0949
Survey wgts?	No	Yes	Yes	No	Yes
Gender wgts?	No	No	Yes	No	Yes
One obs. per HH?	N/A	No	Yes	N/A	Yes
Ex. S/SW?	No	No	No	Yes	Yes
Observations	148290	736832	736832	130400	544375

Comparing Gallup and Census, 1950–1980

	1950		1960		1970		1980	
	Census	Gallup	Census	Gallup	Census	Gallup	Census	Gallup
South Share	0.242	0.117	0.259	0.138	0.271	0.247	0.296	0.256
—*South*								
Female	0.516	0.505	0.521	0.518	0.529	0.507	0.529	0.503
Age	44.61	44.31	45.07	47.64	45.94	46.35	45.20	46.13
Black	0.200	0.0849	0.182	0.147	0.160	0.129	0.159	0.160
HS grad.	0.294	0.373	0.366	0.372	0.473	0.529	0.619	0.635
—*Non-South*								
Female	0.515	0.504	0.517	0.512	0.528	0.506	0.528	0.503
Age	46.67	43.75	45.96	45.87	46.27	45.38	45.28	44.10
Black	0.0530	0.0454	0.0611	0.0586	0.0709	0.0614	0.0782	0.0874
HS grad.	0.385	0.473	0.450	0.531	0.579	0.659	0.710	0.755
Observ.	296223	182171	5388972	95064	2444218	138098	7475162	128507

Back

Gallup data can pick up high-frequency changes in demographics



Acemoglu, D., Aghion, P. and Violante, G. L. (2001). Deunionization, technical change and inequality. In *Carnegie-Rochester conference series on public policy*, Elsevier, vol. 55, pp. 229–264.

Açikgöz, Ö. T. and Kaymak, B. (2014). The rising skill premium and deunionization. *Journal of Monetary Economics*, **63**, 37–50.

Autor, D. H., Katz, L. F. and Kearney, M. S. (2008). Trends in us wage inequality: Revising the revisionists. *The Review of economics and statistics*, **90** (2), 300–323.

Card, D. (2001). The effect of unions on wage inequality in the us labor market. *ILR Review*, **54** (2), 296–315.

DiNardo, J., Fortin, N. M. and Lemieux, T. (1996). Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. *Econometrica*, **64** (5), 1001–1044.

Dinlersoz, E. M. and Greenwood, J. (2012). *The Rise and Fall of Unions in the US*. Tech. rep., National Bureau of Economic Research.

Goldin, C. D. and Katz, L. F. (2009). *The race between education and technology*. Harvard University Press.

Other surveys

- Gallup and ANES allow us to look across large spans of years.
- But we uncovered a few more sources of union micro data, which we use, mostly as a check on Gallup and ANES.
  - A 1936-1937 BLS Consumption survey asks if you have spent anyone in the household has spent money on union dues, which we use to generate a household union status.
  - The U.S. Psychological Corporation conducted a 1946 survey that asks union status, family income and standard covariates.
  - NORC occasionally has surveys with all these covariates as well.

Back

Gallup sampling before 1950

- Berinsky (2006) provides great detail: quota-based sampling of *voters*. Details
- From 1950 onward, more effort to reach representative sample of Americans and to provide weights to correct. We construct our own weights for pre-1950 data.
- Before 1942, we can only adjust by region and race: $WhitexSouth$ (4 cells). From 1942, we adjust by $White \times Educ \times South$ (16 cells). Census Comparison
- Match existing unemployment series (e.g. Roosevelt recession) quite well, and pick up WW2 deployment in age distribution. Gallup Unemp.

Does union sector have lower residual income variance?

- A classic result (Card (2001)) from individual-earnings data in CPS is, relative to union sector:
  1. greater *total* earnings variance in non-union sector
  2. greater *explained* earnings variance in union-sector (covariates do more work)
  3. but, nonetheless, greater *residual* earnings variance in non-union sector.
- As differential union effects by race and education suggest, we replicate (2) in our household data. Union sector has lower total income variance (not shown).
- Over our ninety-year period, ratio of union residual variance to non-union is significantly below one.

Ratio of residual income variance: union v. non-union