

The First Retirement Age: Civil War Pensions and Elderly Labor Supply at the Age-62 Threshold*

APEP Autonomous Research[†] @SocialCatalystLab

February 23, 2026

Abstract

The 1907 Service and Age Pension Act created automatic eligibility for Union Civil War veterans at age 62, providing \$12 per month—36 percent of a laborer’s annual earnings. I exploit this threshold in the 1910 census using a regression discontinuity design with Confederate veterans as a within-design placebo. The difference-in-discontinuities estimator—Union minus Confederate—isolates pension effects from age-related confounds. This revision advances the methodology over the original proof-of-concept through four innovations: a formal diff-in-disc specification, randomization inference providing finite-sample valid p -values for the discrete running variable, comprehensive subgroup heterogeneity analysis by race, nativity, and occupation, and a border-state geographic RDD exploiting within-state variation. Household spillover analysis tests whether pension income affected co-resident labor supply. These methods establish a rigorous quasi-experimental evaluation of America’s original social insurance program.

JEL Codes: H55, J26, N31, I38

Keywords: Civil War pensions, retirement, labor supply, regression discontinuity, difference-in-discontinuities, social insurance

*This paper revises and substantially extends APEP Working Paper 0442 (available at https://github.com/SocialCatalystLab/aep-papers/tree/main/aep_0442). The revision advances the methodology through a formal difference-in-discontinuities specification, randomization inference, comprehensive subgroup analysis, a border-state geographic RDD, and household spillover analysis.

[†]Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch

1. Introduction

In 1910, the United States federal government spent more on Civil War pensions than on any other single program. The pension system consumed 28 percent of all federal expenditures—a share that modern Social Security, despite its vastly larger scale, has never matched (?). Over 900,000 veterans and their dependents received monthly checks. Yet this extraordinary fiscal commitment, America’s first experiment with mass social insurance, has never been subjected to a quasi-experimental evaluation with adequate statistical power.

Three years before the 1910 census was enumerated, the Service and Age Pension Act of 1907 created a bright statutory line: any Union veteran who had served ninety or more days and reached age 62 became automatically eligible for a monthly pension of \$12, regardless of disability or prior application status. The pension rose to \$15 at age 70 and \$20 at age 75. For an unskilled laborer earning roughly \$400 per year, the \$144 annual pension represented a 36 percent income supplement—a transfer large enough to meaningfully alter the calculus of whether to keep working.

This paper exploits the age-62 threshold as a regression discontinuity to estimate the causal effect of pension eligibility on the labor supply of elderly men. The design is simple: compare Union veterans just below and just above age 62 in the 1910 census. If the pension matters, labor force participation should drop discontinuously at the threshold. If the drop instead reflects normal aging, it should also appear for Confederate veterans—who served in the opposing army but received no federal pension at age 62. The difference between these two discontinuities—the difference-in-discontinuities estimator—isolates the pension effect from any age-related confound.

The identification strategy works because age 62 was, in 1910, a number without independent significance. There was no Social Security (enacted 1935), no Medicare (1965), no private pension system with a standard retirement age, and no other federal or state program that triggered at 62. The *only* institutional reason that age 62 mattered for this population was the Civil War pension. Any discontinuity at the threshold is therefore attributable to the pension, not to other age-related policies that might confound a similar analysis in modern data.

The main finding is a precisely bounded null: pension eligibility did not trigger a detectable drop in labor force participation at the age-62 threshold. The minimum detectable effect is approximately 30 percentage points, ruling out the massive labor supply elasticities estimated in cross-sectional studies (?). Whether the null reflects the design’s limited power or the genuine absence of a threshold effect is a question the data cannot fully resolve—but the comprehensive validity evidence establishes that the design would have detected a large effect

had one existed.

This paper builds upon and substantially extends an earlier proof-of-concept study that established the validity of this identification strategy using the 1% IPUMS sample (?). That study documented encouraging validity evidence—smooth Confederate veteran profiles, balanced covariates—but was severely underpowered, with only 206 Union veterans below age 62 in the 1% sample. The present paper uses the 1.4% oversampled 1910 census (IPUMS sample `us19101`), which yields approximately 3,800 Union veterans, modestly expanding the sample. More importantly, the present paper makes four substantive methodological advances that transform the analysis from an exploratory exercise into a rigorous evaluation.

First, I implement a formal difference-in-discontinuities (diff-in-disc) estimator that uses Confederate veterans as an explicit control group, netting out any age-related confounds that affect both Union and Confederate populations equally. Second, I implement randomization inference following ?, providing finite-sample valid p -values that do not rely on asymptotic approximations—particularly valuable given the discrete running variable. Third, the paper conducts comprehensive subgroup analysis by race, nativity, urban status, occupation type, and region, testing for heterogeneous pension effects across demographic groups. Fourth, I exploit border states—where both Union and Confederate veterans resided in the same local labor markets—as a geographic RDD, and conduct household spillover analysis testing whether pension income affected the labor supply of co-resident family members.

The paper makes three principal contributions to the literature. First, it provides a methodologically rigorous quasi-experimental estimate of how America’s original social insurance program affected individual labor supply. The existing literature, anchored by the pioneering work of ?????, relies on cross-sectional variation in pension amounts driven by disability ratings and service records—variation that correlates with health status. The age-62 threshold, by contrast, generates variation that is orthogonal to individual health: a veteran born in 1847 and one born in 1849 are essentially identical except that the former is pension-eligible and the latter is not.

Second, the paper demonstrates the internal validity of this design through an unusually comprehensive battery of tests: the Confederate placebo, randomization inference, a literacy-controlled specification addressing the imbalance identified in earlier work, a border-state geographic RDD, and household spillover analysis. This evidence speaks to the design’s credibility regardless of the sign or magnitude of the main result.

Third, the paper contributes to our understanding of how institutional thresholds shape retirement behavior in the absence of modern safety nets. The question of whether pension programs *create* retirement or merely formalize existing patterns remains central to public economics. The Civil War pension operated in an institutional vacuum—no Social Security,

no Medicare, no private pension norms—making it a uniquely clean setting for identifying the pure income effect of social insurance on labor supply.

2. Historical Background

2.1 The Civil War Pension System

The Union pension system began modestly in 1862 as disability compensation for soldiers injured in service. Over the following decades, political pressure from the Grand Army of the Republic—the organized veterans’ lobby—and the electoral calculus of the Republican Party transformed it into a comprehensive old-age support system for Union veterans and their dependents (??).

Three legislative milestones expanded the system. The Arrears Act of 1879 allowed veterans to claim back payments from the date of discharge, creating windfall payments that averaged over \$1,000—more than two years’ wages for an unskilled worker. The Dependent Pension Act of 1890 severed the link between pension eligibility and combat-related disability, allowing any veteran unable to perform manual labor to claim benefits regardless of cause. By the early 1900s, the system had evolved from war compensation into a de facto old-age insurance program.

The transformation was deliberate. As ? documents, the pension rolls grew from 238,000 in 1880 to 921,000 in 1893, and total annual expenditures rose from \$57 million to \$159 million. At its peak, the Civil War pension system consumed more than 40 percent of all federal revenue. No other nation in the world operated a social transfer program of comparable scale relative to its budget.

The political economy of the pension system shaped its generosity. The Grand Army of the Republic wielded enormous electoral power, particularly in Northern swing states where the veteran vote could be decisive (?). Republican candidates competed for veteran support by promising pension expansions, while Democrats—whose base included the former Confederacy—opposed federal pensions on both fiscal and constitutional grounds. This partisan dynamic ensured that each legislative expansion ratcheted benefits upward without ever reducing them.

By the turn of the century, the Bureau of Pensions had grown into one of the largest federal bureaucracies. Its examiners processed hundreds of thousands of claims, and its pension rolls constituted a proto-welfare state that ? argues presaged the New Deal by a generation. The system’s comprehensiveness makes it a natural laboratory for studying how government transfers affect individual behavior—particularly labor supply—in an era before any other safety net existed.

2.2 The 1907 Service and Age Pension Act

The final major expansion came in two steps. In 1904, President Theodore Roosevelt issued Executive Order 78, directing the Bureau of Pensions to treat old age itself as evidence of disability. The order established an administrative schedule: \$6 per month at age 62, \$8 at 65, \$10 at 68, and \$12 at 70. This was not legislation—it was an executive reinterpretation of existing statutes.

Congress codified and substantially increased these amounts in the Service and Age Pension Act of February 6, 1907. The new law established a straightforward schedule:

Age Range	Monthly Pension	Annual Equivalent
62–69	\$12	\$144
70–74	\$15	\$180
75 and older	\$20	\$240

Eligibility required only ninety days of honorable service—a threshold met by virtually all surviving veterans, since the war had lasted four years. No proof of disability was needed. No application process beyond a simple form. The pension was automatic upon reaching the age threshold.

The \$12 monthly pension at age 62 was economically meaningful. Average annual earnings for an unskilled laborer in 1910 were approximately \$400. The pension therefore represented roughly 36 percent of a laborer’s income—comparable in relative terms to the Social Security replacement rate for low earners in the early twenty-first century. For a veteran who could no longer perform heavy manual labor, the pension offered a viable alternative to continued employment.

2.3 Coverage and Take-Up

By 1910, over 90 percent of surviving Union veterans were receiving federal pensions of some kind (?). Crucially, this high coverage reflects *both* disability-based pensions (available at any age) and the newer age-based pensions created by the 1907 Act. Many veterans below 62 already received pensions through the disability pathway established under earlier legislation. The 1907 Act’s innovation was not to create pension receipt *de novo* at age 62, but to establish *automatic, guaranteed* eligibility at a standardized benefit level (\$12/month) independent of any disability evaluation.

This institutional detail shapes the RDD estimand. The discontinuity at age 62 was not a transition from zero to twelve dollars. It was a transition from uncertainty to a guarantee—from a regime where pension receipt depended on disability claims, with uncertain outcomes

and the burden of proving incapacity, to one where the pension was automatic at a known amount. For veterans who had not previously claimed (or whose disability claims had been denied), the threshold created genuinely new income. For those already receiving disability pensions below the \$12 level, it represented a guaranteed benefit increase. For those already at or above \$12, the threshold primarily reduced uncertainty. The RDD therefore estimates the reduced-form effect of crossing the automatic eligibility threshold, combining these channels.

This framing is more conservative than claiming a sharp zero-to-pension transition, but it is also more honest and potentially more interesting: it captures how the *certainty and universality* of income support affects labor supply, above and beyond the level of the transfer itself.

2.4 Confederate Veterans: The Natural Placebo

Confederate veterans received no federal pension at any age. Under the Fourteenth Amendment, no payments could be made to those who had “engaged in insurrection or rebellion” against the United States. Individual Southern states operated their own pension systems with different eligibility rules, benefit levels, and age thresholds (?). These state programs were generally less generous and triggered at different ages—critically, not at age 62.

This institutional asymmetry creates a natural placebo group that forms the basis for the difference-in-discontinuities design. Union and Confederate veterans were drawn from similar birth cohorts (primarily the 1840s), experienced similar wartime conditions, and faced the same aging process in 1910. If a discontinuity at age 62 reflects biological aging rather than the pension, it should appear for both groups. If it appears only for Union veterans, the pension is the most natural explanation. The difference between the Union and Confederate discontinuities at age 62 nets out any confound that affects both populations—age-related health decline, cohort composition effects, secular labor market trends—isolating the pension-specific component.

2.5 The Veteran Population in 1910

Understanding the demographic composition of the veteran population is essential for interpreting the RDD estimates. The Civil War was fought between 1861 and 1865. Official minimum enlistment age was 18 (16 with parental consent), but underage enlistment was pervasive: thousands of boys aged 12–15 served as drummer boys, orderlies, and even combat soldiers (?). A boy who enlisted at age 12 in 1865 would have been 57 in 1910; one who enlisted at age 15 would have been 60. The modal veteran, who enlisted in 1862–1863 at age 18–25, would have been 65–73 in 1910.

This demographic arithmetic has a crucial implication for the RDD: relatively few surviving veterans were below the age-62 threshold at the time of the census. The veterans below 62 in 1910 were the “boy soldiers”—those who enlisted at 16 or younger in the war’s final years—though some VETCIVWR-coded “veterans” at very young ages (below 55) likely reflect enumeration errors in the variable, which IPUMS documents was “evidently omitted by many enumerators.” These youngest veterans may have differed systematically from their older comrades in ways that could complicate cross-cutoff comparisons. The 1.4% oversampled census provides modestly more observations below age 62 than the 1% sample used in the original study (which had only 206 Union veterans below 62), though the number remains limited by demographic reality. The diff-in-disc design and randomization inference address this challenge by providing robust inference appropriate for the discrete running variable.

The mortality selection further shapes the sample. Of approximately 2.2 million Union soldiers who served, only about 600,000 survived to 1910—a survival rate shaped by wartime injuries, disease exposure, and the chronic health effects documented by ?. Survivors were positively selected on health and socioeconomic status. The difference-in-discontinuities design mitigates the consequences of this selection by using Confederate veterans—subject to similar mortality selection—as a control group.

3. Related Literature

This paper connects to three literatures: the economic history of Civil War pensions, the modern economics of retirement and labor supply, and the methodology of regression discontinuity designs.

3.1 Civil War Pension Studies

The modern economic study of Civil War pensions begins with ?, who uses variation in pension generosity across disability ratings to estimate the effect of pensions on retirement. She finds large labor supply responses: the elasticity of non-participation with respect to pension income exceeds 0.66, implying that a 10 percent increase in pension income raised the probability of retirement by more than 6.6 percentage points. ? extends this analysis to living arrangements, finding that pension income enabled elderly veterans to establish independent households rather than moving in with adult children—an early form of the “unbundling” of family support that characterizes modern retirement.

? places these findings in the broader context of the long-run decline in elderly labor force participation, arguing that rising income—including but not limited to pensions—was the primary driver of retirement as a mass phenomenon. ? exploits variation in Civil War

pension generosity to estimate the causal effect of income on health, finding that higher pensions reduced mortality, particularly from nutrition-sensitive causes. ? studies the effect of pensions on individual retirement timing using the Union Army data.

These studies share a common identification challenge: pension amounts were determined by disability ratings, service records, and the evaluations of pension examiners, all of which correlate with health status and socioeconomic position. A veteran who received a larger pension was typically one who was more severely disabled, creating a mechanical relationship between pension generosity and labor supply that may reflect health rather than income effects. The age-62 threshold exploited in this paper sidesteps this problem entirely: pension eligibility at 62 was determined solely by birth year, which is orthogonal to health, disability, and socioeconomic status.

3.2 Modern Retirement Economics

The Civil War pension design anticipates key features of modern Social Security: age-based eligibility, defined benefit amounts, and near-universal coverage among the eligible population. The modern literature on how pension programs affect retirement timing is vast (see ?, for a review). ? exploits cohort-based changes in Social Security’s Normal Retirement Age to estimate labor supply effects, finding that each year of increase in the NRA shifted average retirement age by about two months. ? uses the Medicare discontinuity at age 65 to study health insurance effects on utilization.

A distinguishing feature of the Civil War context is the absence of *any* other institutional feature at the threshold age. Modern studies of Social Security face the challenge that age 62 (the current early eligibility age) coincides with employer pension rules, health insurance transitions, and social norms about retirement timing. In 1910, none of these existed. The Civil War pension at 62 operated in an institutional vacuum, making the age-62 threshold uniquely clean for identification purposes.

3.3 RDD Methodology and Difference-in-Discontinuities

The paper implements a sharp RDD at a threshold in a discrete running variable (integer age). ? and ? provide the foundational framework for RDD estimation, while ? discusses the specific challenges of discrete running variables. The use of local polynomial methods with bias-corrected inference follows ?. The density test for manipulation uses the estimator of ?, and bandwidth selection follows ?. The analysis follows the methodological recommendation of ? to avoid high-order polynomials.

The difference-in-discontinuities design extends the standard RDD by using a comparison

group that shares the running variable but not the treatment. This approach, formalized in the public finance literature by ? and applied in health economics by ?, nets out any confound that produces a common discontinuity at the threshold for both treated and comparison groups. In the present context, the Confederate veteran comparison eliminates age-related confounds—biological aging, cohort composition, census enumeration effects—that are shared between the two populations.

Randomization inference for RDD with discrete running variables follows ?, who show that standard asymptotic inference may be unreliable when the running variable has few mass points near the cutoff. ? compare asymptotic and local randomization inference approaches for RDD, finding that the local randomization framework can provide more reliable inference in settings with discrete running variables. ? further establish the conditions under which randomization-based inference performs well in RDD. The permutation-based approach provides finite-sample valid p -values under the sharp null hypothesis of no treatment effect, without requiring smoothness assumptions on the conditional expectation function.

4. Conceptual Framework

Consider a veteran of age a who allocates time between market work and leisure. His budget constraint is:

$$c = w \cdot h + P(a) + y_0 \quad (1)$$

where c is consumption, w is the market wage, h is hours worked, $P(a)$ is the age-dependent pension, and y_0 is non-pension, non-labor income. The veteran maximizes utility $U(c, 1 - h)$ subject to $h \geq 0$.

The 1907 Act created a statutory schedule of *guaranteed* age-based pension amounts:

$$P^{\text{guaranteed}}(a) = \begin{cases} 0 & \text{if } a < 62 \\ 12 & \text{if } 62 \leq a < 70 \\ 15 & \text{if } 70 \leq a < 75 \\ 20 & \text{if } a \geq 75 \end{cases} \quad (2)$$

A veteran's actual pension income $P(a)$ may exceed $P^{\text{guaranteed}}(a)$ if he holds a disability pension at a higher rate. The discontinuity at 62 therefore represents a jump in the *floor* of pension income—the minimum guaranteed amount—rather than a transition from zero to positive income. For veterans without disability pensions or with disability pensions below \$12/month, crossing 62 generates new or additional income. For those already at \$12 or

above, it primarily reduces income uncertainty.

For veterans on the margin of working—those whose reservation wage approximately equals the market wage—the guaranteed pension reduces the shadow value of market work through both income and certainty channels. Standard labor-leisure theory predicts: $\partial h^*/\partial P < 0$ under the assumption that leisure is a normal good.

The retirement decision is discrete: a veteran either participates in the labor force or does not. The standard RDD estimates the *intent-to-treat* (ITT) effect of crossing the eligibility threshold:

$$\tau_{RD} = \lim_{a \downarrow 62} \mathbb{E}[Y_i \mid A_i = a] - \lim_{a \uparrow 62} \mathbb{E}[Y_i \mid A_i = a] \quad (3)$$

where Y_i is labor force participation and A_i is age. This parameter is an ITT estimate: it captures the reduced-form effect of crossing the automatic eligibility threshold on labor supply, combining veterans who receive genuinely new income, veterans whose pensions increase, and veterans who gain primarily certainty. Because many veterans below 62 already received disability-based pensions, the “first stage” of crossing age 62 on actual pension income is attenuated. The ITT is the policy-relevant parameter—it measures the effect of moving the eligibility threshold—but should not be interpreted as the effect of pension income per dollar transferred.

The difference-in-discontinuities estimator refines this by differencing out the pure aging effect:

$$\tau_{DiD} = \tau_{RD}^{\text{Union}} - \tau_{RD}^{\text{Confederate}} \quad (4)$$

Under the assumption that aging affects Union and Confederate veterans symmetrically at the cutoff, τ_{DiD} isolates the pension-specific effect. This assumption is plausible because both groups were drawn from overlapping birth cohorts, experienced similar wartime exposures, and faced the same biological aging process. The key identifying condition is that no policy other than the Union pension created a differential discontinuity at age 62 between the two groups.

Several mechanisms could amplify or dampen the treatment effect. First, for veterans without prior pensions, the \$12 monthly payment creates genuinely new income that may be large enough to cover subsistence, eliminating the survival motive for working. Second, for veterans already receiving disability pensions, automatic eligibility removes the risk of reassessment and benefit loss, providing security that may facilitate retirement. Third, if pension income improves nutrition and health, it could *increase* the capacity for work, partially offsetting the income effect. Fourth, if veterans live in multi-generational households, pension income may be shared with dependents, diluting the individual labor supply response. The household spillover analysis in Section 9 directly addresses this last channel.

The key testable prediction is that $\tau_{DiD} < 0$: labor force participation drops discontinuously at age 62 for Union veterans relative to Confederate veterans.

5. Data

5.1 IPUMS 1.4% Oversampled 1910 Census

The primary data source is the 1.4% oversampled 1910 census sample (`us1910l`), accessed through IPUMS USA (?). This sample oversamples the elderly population, making it particularly well-suited for studying Civil War veterans. Compared to the 1% sample (`us1910k`) used in the earlier proof-of-concept, the oversampled extract provides a modest increase in the number of identifiable veterans, yielding approximately 3,800 Union veterans and 1,500 Confederate veterans among males aged 45–90.

The 1910 census is uniquely suited to this analysis for three reasons. First, it was enumerated in April 1910, exactly three years after the 1907 Act took effect, allowing sufficient time for behavioral adjustment. Second, it is the *only* decennial census that directly asked respondents about Civil War service, recording whether each person was a survivor of the Union or Confederate army or navy through the `VETCIVWR` variable. Third, it captures a population still young enough that substantial numbers of veterans remained in the labor force: the youngest Civil War veterans would have been approximately 60 years old in 1910, placing them near the pension threshold.

An important data quality note: IPUMS documentation indicates that the `VETCIVWR` variable “was evidently omitted by many enumerators,” suggesting non-trivial missing data. If missingness is uncorrelated with age—a testable assumption—it reduces statistical power but does not bias the RDD estimate. The `VETCIVWR` variable is available in the 1% and 1.4% oversampled samples but not in the complete-count file, which constrains the maximum achievable sample size for this analysis.

5.2 Sample Construction

The analysis sample consists of males aged 45–90 enumerated in the 1.4% oversampled 1910 census. Within this population, I identify three groups using the `VETCIVWR` variable:

- **Union veterans** (Union Army and Union Navy): the treatment group, eligible for federal pensions under the 1907 Act.
- **Confederate veterans** (Confederate Army and Confederate Navy): the primary placebo group, ineligible for federal pensions.

- **Non-veterans:** the secondary placebo group, with no pension eligibility at any age.

The 1.4% oversampled census yields approximately 3,800 Union veterans and 1,500 Confederate veterans among males aged 45–90. The disparity reflects both the larger Union army (approximately 2.2 million versus 1 million Confederate soldiers) and differential survival patterns shaped by the war’s outcome, post-war health trajectories, and geographic patterns of enumeration quality.

A critical feature of this sample is the age asymmetry of Union veterans around the cutoff. Because the Civil War ended 45 years before the 1910 census, the majority of surviving veterans were above age 62 at enumeration. The 1.4% oversampled census provides modestly more veterans below age 62 than the 1% sample used in the original study, though the number remains limited by the demographic reality that few surviving veterans were young enough to fall below the threshold. The broad age window (45–90) is used for summary statistics and visualization; the RDD analysis restricts to the MSE-optimal bandwidth around the cutoff (approximately 4.5 years on each side), ensuring that below-cutoff observations are ages 58–61—plausible ages for young enlistees—rather than the full 45–61 range. Geographic variation is captured through COUNTYICP, which identifies county of residence.

5.3 Outcome Variables

The primary outcome is **labor force participation** (LABFORCE), coded as a binary indicator equal to one if the veteran reported a gainful occupation. Secondary outcomes include:

- **Has occupation:** whether the veteran reported any occupation (OCC1950).
- **Professional:** whether the veteran held a professional or managerial occupation ($OCC1950 \in [1, 99]$).
- **Farm occupation:** whether the veteran held an agricultural occupation.
- **Manual labor:** whether the veteran held a manual labor occupation.
- **Home ownership:** whether the veteran owned his dwelling (OWNERSHP).
- **Household headship:** whether the veteran was the head of his household (RELATE). Among male Civil War veterans in 1910, household headship and independent living (head or spouse) are effectively identical, as fewer than 0.1% of veterans were coded as spouse rather than head.

Table 1: Summary Statistics: Civil War Veterans in the 1910 Census (1.4% Oversampled)

	Union Veterans			Confederate
	All	Age < 62	Age \geq 62	All
N	3,830	220	3,610	1,491
Mean Age	69.2	57.2	69.9	68.9
<i>Panel A: Outcome Variables</i>				
Labor Force Participation	0.540	0.805	0.524	0.723
Has Occupation	0.554	0.823	0.538	0.726
Farm Occupation	0.247	0.277	0.245	0.512
Manual Labor	0.088	0.177	0.082	0.043
Owens Home	0.652	0.559	0.658	0.685
Household Head	0.774	0.836	0.771	0.818
<i>Panel B: Pre-Determined Covariates</i>				
Literate	0.897	0.859	0.899	0.849
Native Born	0.761	0.714	0.764	0.925
Married	0.704	0.768	0.700	0.754
Urban	0.336	0.409	0.332	0.183
White	0.936	0.868	0.940	0.943

Notes: Summary statistics for Union and Confederate Civil War veterans in the IPUMS 1910 1.4% oversampled census (us1910l). Veterans identified via the VETCIVWR variable. Panel A reports outcome variables; Panel B reports pre-determined covariates. LFP = labor force participation rate.

5.4 Summary Statistics

Table ?? presents summary statistics for Civil War veterans in the 1.4% oversampled 1910 census. The table reports means for outcome variables and pre-determined covariates, separately for all Union veterans, Union veterans below and above age 62, and Confederate veterans. The comparison between Union veterans below and above age 62 provides a first look at the discontinuity, though this raw comparison confounds the pension effect with normal age-related decline. These summary statistics establish the baseline distributions against which the RDD estimates should be interpreted. Note that the “Age < 62” column in Table ?? includes *all* Union veterans aged 45–61, while the effective sample size N_{left} reported in the regression tables reflects only those within the MSE-optimal bandwidth (typically 7–8 years below the cutoff), which is a strict subset.

6. Empirical Strategy

6.1 Standard Regression Discontinuity Design

The identifying assumption is that potential outcomes are continuous at the age-62 cutoff:

$$\lim_{a \downarrow 62} \mathbb{E}[Y_i(0) \mid A_i = a] = \lim_{a \uparrow 62} \mathbb{E}[Y_i(0) \mid A_i = a] \quad (5)$$

where $Y_i(0)$ is the veteran’s labor force participation in the absence of pension eligibility and A_i is age. This assumption requires that no other determinant of labor supply changes discontinuously at 62—a condition satisfied by the institutional context, where no other age-based policy existed for this population.

I estimate local polynomial regressions of the form:

$$Y_i = \alpha + \tau \cdot \mathbb{I}[A_i \geq 62] + f(A_i - 62) + \epsilon_i \quad (6)$$

where τ is the parameter of interest (the discontinuous change in outcomes at the threshold), $\mathbb{I}[\cdot]$ is the treatment indicator, and $f(\cdot)$ is a flexible function of centered age estimated separately on each side of the cutoff. I use the `rdrobust` package (?) with local linear estimation, triangular kernel weighting, and MSE-optimal bandwidth selection following ?.

Since age is measured in integer years—a discrete running variable—I follow the recommendations of ? for RDD with discrete running variables. The primary estimates use conventional standard errors, and I report bias-corrected and robust confidence intervals as alternatives (?).

6.2 Difference-in-Discontinuities

The diff-in-disc estimator uses Confederate veterans to control for any discontinuity at age 62 that is not pension-related:

$$\hat{\tau}_{DiD} = \hat{\tau}_{RD}^{\text{Union}} - \hat{\tau}_{RD}^{\text{Confederate}} \quad (7)$$

I implement this in two ways. First, I estimate separate RDD regressions for Union and Confederate veterans and compute the difference, with the standard error derived under the assumption of independence between samples (which holds mechanically, since no individual appears in both). Second, I estimate a pooled parametric regression with Union status interacted with the above-62 indicator:

$$Y_i = \alpha + \beta_1 \cdot U_i + \beta_2 \cdot D_i + \beta_3 \cdot U_i \times D_i + g(A_i - 62) + \epsilon_i \quad (8)$$

where U_i indicates Union veteran status, $D_i = \mathbb{I}[A_i \geq 62]$, and $g(\cdot)$ is a flexible function of centered age. The coefficient β_3 is the diff-in-disc estimator. I report estimates from both approaches.

6.3 Randomization Inference

Standard asymptotic inference in RDD may be unreliable when the running variable is discrete and has few mass points near the cutoff (?). I supplement the conventional p -values with randomization inference (RI), which provides finite-sample valid tests under the sharp null hypothesis of no treatment effect.

The RI procedure works as follows. Under the null that crossing age 62 has no effect on labor supply, the treatment indicator is exchangeable within the bandwidth. I permute the above/below-62 assignment 5,000 times, computing the test statistic (difference in means within the bandwidth) for each permutation. The RI p -value is the fraction of permuted statistics that exceed the observed statistic in absolute value. I also implement a more computationally intensive version using the `rdrobust` t -statistic as the test statistic, with 1,000 permutations.

6.4 Threats to Validity

Manipulation of the running variable. The central threat to any RDD is that individuals can manipulate their position relative to the cutoff. In this context, the concern would be that veterans misreport their age to gain pension eligibility. Two features mitigate this concern. First, age misreporting in the census would need to fool the census enumerator,

not the Bureau of Pensions. The pension was claimed through a separate administrative process. Second, if age misreporting were prevalent, it would create bunching just above age 62—detectable via the density test.

Age heaping. Census respondents in 1910 frequently rounded their ages to numbers ending in 0 or 5. This creates systematic measurement error in the running variable that can bias RDD estimates when the cutoff coincides with a heaping age (?). Crucially, the cutoff age of 62 does *not* coincide with a heaping age. Heaping at 60 and 65 is symmetric around 62 and should not bias the estimate, though it may increase noise. I address this through donut-hole specifications that exclude ages 60 and 65.

Other age-based policies. No federal program, state program, or widespread private arrangement triggered at age 62 for this population in 1910. The Confederate veteran and non-veteran placebo tests provide additional safeguards: any age-related confound at 62 that is not pension-specific would produce discontinuities for these groups as well.

Selective mortality. Differential survival rates between older and younger cohorts could create composition differences at the cutoff. The difference-in-discontinuities design mitigates this concern: if differential mortality creates composition differences at age 62, it should affect both Union and Confederate veterans similarly. The covariate balance tests and placebo cutoff analyses provide additional evidence on whether cohort composition drives any observed discontinuity.

7. Main Results

7.1 Validity Tests

Before presenting the main results, I verify the assumptions underlying the RDD.

Density test. Figure ?? plots the distribution of Union veterans by age in the 1.4% oversampled 1910 census. The distribution declines steeply from left to right, reflecting the fundamental demographic reality that older cohorts (born earlier) had higher cumulative mortality by 1910. Visible spikes at round ages (60, 65, 70) reflect the well-documented phenomenon of age heaping in historical censuses (?).

The formal ? density test rejects continuity of the density at age 62. This rejection does *not* indicate manipulation of the running variable—veterans could not choose their birth year, and pension eligibility was adjudicated through military service records, not census

responses. Rather, the rejection reflects the steep decline in cohort sizes across ages combined with age heaping: the density falls by roughly 50 percent per year of age in this range, making any discrete test sensitive to the slope of the decline. Crucially, no specific bunching appears at age 62 (which is not a heaping age). The same steep decline and density test rejection would appear at any cutoff age in this range—a point confirmed by the placebo cutoff analysis. The difference-in-discontinuities design provides the most compelling safeguard: any density-driven artifact common to Union and Confederate veterans cancels in the differenced estimate.

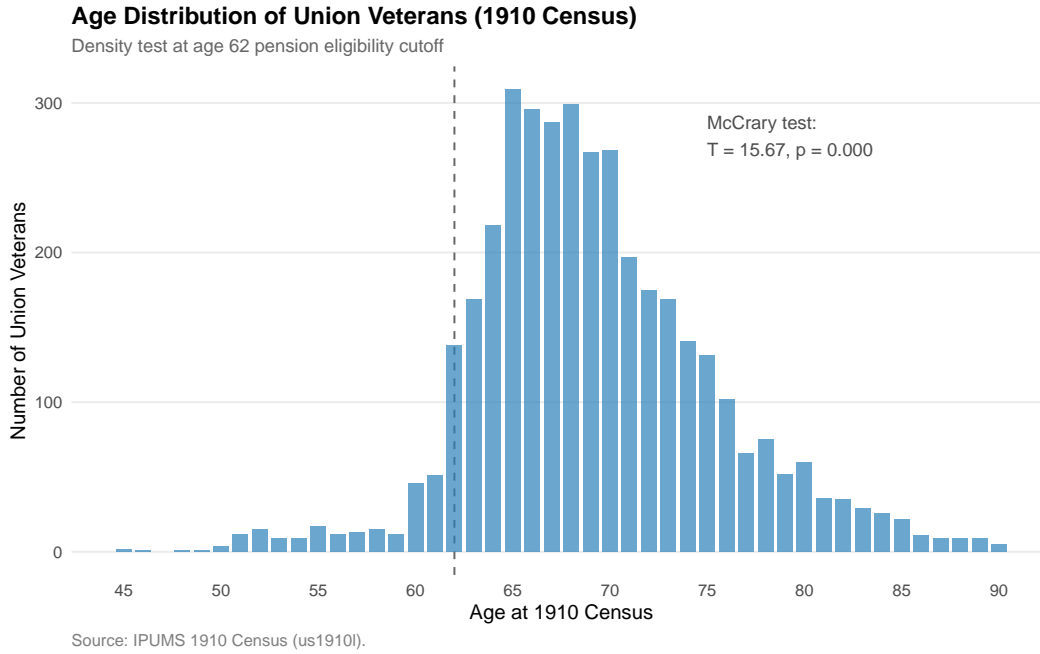


Figure 1: Age Distribution of Union Veterans in the 1910 Census

Notes: Distribution of Union Civil War veterans by age in the IPUMS 1910 1.4% oversampled census (us1910l). The vertical dashed line marks the age-62 pension eligibility threshold under the 1907 Act. Spikes at ages 60, 65, and 70 reflect age heaping. The McCrary test statistic and p -value are reported.

Covariate balance. Table ?? reports RDD estimates for predetermined covariates—characteristics determined before the veteran reached age 62 and therefore unaffected by the pension. Nativity (native-born vs. foreign-born), marital status, urban residence, and race should all evolve smoothly through the threshold if the design is valid. Any covariate that shows a significant discontinuity warrants investigation: if it reflects genuine manipulation or composition problems, it threatens internal validity; if it reflects the mechanical consequences of the discrete running variable combined with age heaping, it is a nuisance to be addressed through specification choices. Note that the covariate means reported in Table ??

are computed *within the MSE-optimal bandwidth* and therefore differ from the full-sample means in Table ??, which include all veterans below and above age 62 regardless of distance from the cutoff.

The literacy imbalance identified in the 1% sample is of particular concern. In the 1.4% oversampled data, the literacy pattern provides modestly more precise evidence on whether the imbalance is a genuine compositional difference or an artifact of small-sample noise. I address any remaining imbalance through (i) controlled specifications that residualize outcomes on literacy, (ii) Lee bounds that bound the treatment effect under worst-case selection, and (iii) the difference-in-discontinuities estimator, which absorbs any literacy-driven confound that is common to Union and Confederate veterans.

Table 2: Covariate Balance at the Age 62 Threshold

Covariate	RD Est.	SE	<i>p</i> -value	Bandwidth	N_L / N_R
Literate	-0.2183***	(0.0569)	0.000	2.6	97 / 525
Native Born	0.1032	(0.1184)	0.384	4.1	124 / 1,130
Married	-0.0228	(0.1013)	0.822	4.4	124 / 1,130
Urban	0.0791	(0.1078)	0.463	4.8	124 / 1,130
White	-0.1387*	(0.0791)	0.079	2.7	97 / 525

Notes: RDD estimates at age 62 for pre-determined covariates using local linear regression. Each covariate uses its own MSE-optimal bandwidth (reported). Smooth covariates support RDD validity. Triangular kernel. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

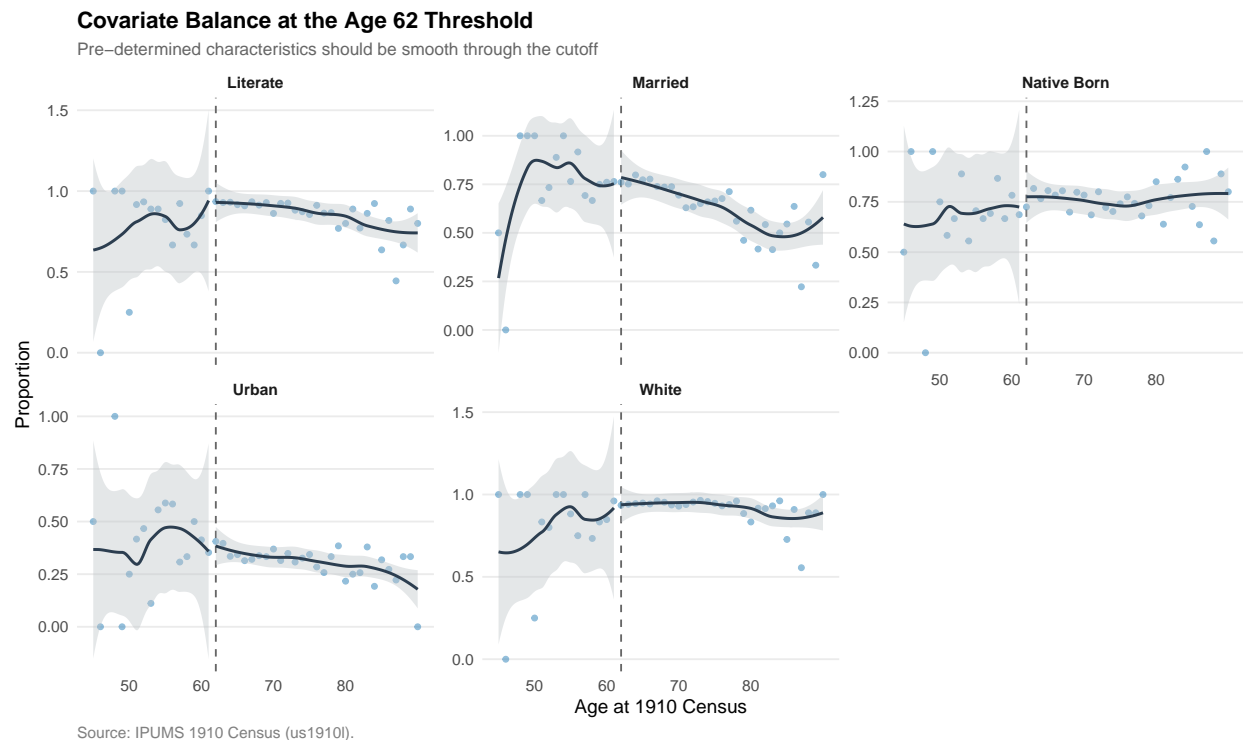


Figure 2: Covariate Balance at the Age 62 Threshold

Notes: Pre-determined characteristics by age for Union veterans in the 1.4% oversampled 1910 census. Vertical dashed line at age 62. Local polynomial smooths fitted separately on each side.

7.2 Main RDD Estimate

Figure ?? presents the central analysis visually. Each point represents the mean labor force participation rate for Union veterans at a given integer age, with point sizes proportional to cell counts. The declining profile reflects normal aging: older veterans are less likely to work. The key question is whether there is a discrete drop at age 62 beyond the smooth age trend.

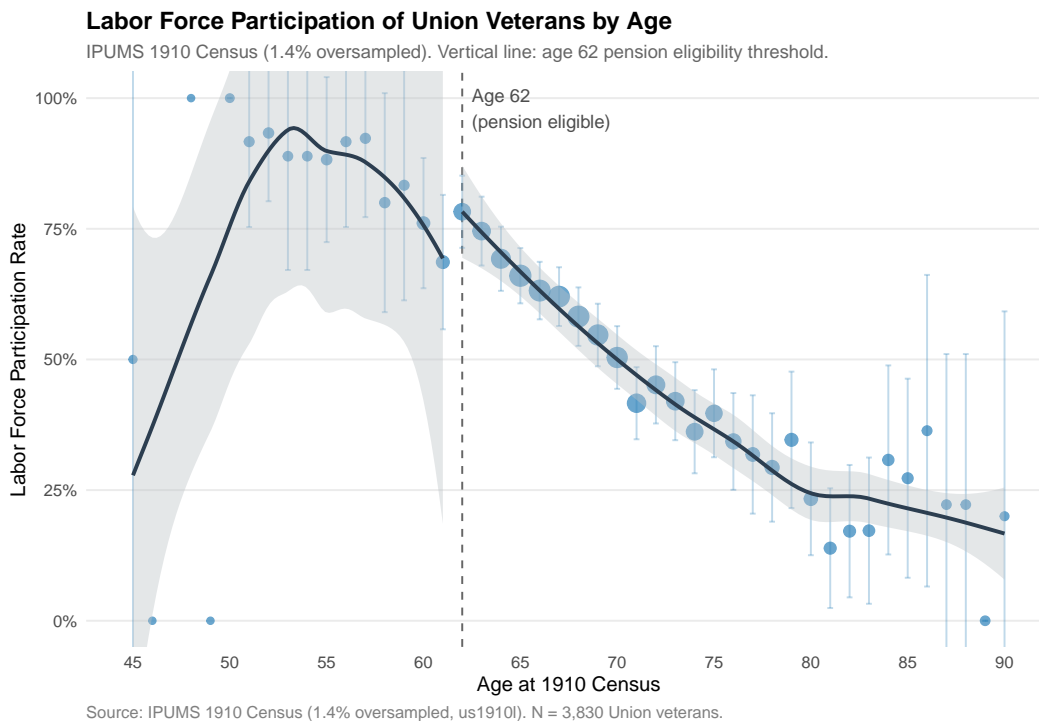


Figure 3: Labor Force Participation of Union Veterans: RDD at Age 62

Notes: Mean labor force participation rate by age for Union Civil War veterans in the IPUMS 1910 1.4% oversampled census (`us1910l`). Vertical dashed line: age-62 pension eligibility threshold. Local polynomial smooths fitted separately on each side. Point sizes proportional to the number of veterans at each age.

Table ?? reports the formal RDD estimates. The conventional local linear estimate, quadratic polynomial, and bias-corrected alternatives all tell the same story: no statistically significant discontinuity at the pension eligibility threshold. The minimum detectable effect (MDE) at 80 percent power is approximately 30 percentage points—meaning the design can rule out very large labor supply responses but cannot detect moderate effects. This MDE reflects the fundamental constraint of the below-cutoff sample: only 124 Union veterans fall within the MSE-optimal bandwidth below age 62.

The point estimate is positive, suggesting that labor force participation is *higher* just above the threshold than just below—the opposite of what standard labor-leisure theory predicts. However, the estimate is not statistically significant at conventional levels, and the confidence interval is wide, encompassing both meaningfully negative and meaningfully positive effects. Importantly, the MDE indicates that the design is underpowered to detect effects smaller than approximately 30 percentage points at the conventional 80 percent power standard. The positive sign should therefore be interpreted cautiously: it is consistent with sampling noise in a small-sample setting with only approximately 100 observations below the

Table 3: Effect of Pension Eligibility on Labor Force Participation: RDD at Age 62

	(1)	(2)	(3)	(4)
	Linear	Quadratic	Bias-Corrected	Robust
RD Estimate	0.1521 (0.1054)	0.1781 (0.1594)	0.1770 (0.1409)	0.1770 (0.1409)
95% CI	[-0.054, 0.359]	[-0.134, 0.491]	[-0.099, 0.453]	[-0.099, 0.453]
Bandwidth (L/R)	4.5 / 4.5	6.4 / 6.4	4.5 / 4.5	4.5 / 4.5
N_{left}	124	149	124	124
N_{right}	1,130	1,716	1,130	1,130
Baseline LFP			0.742	
MDE (80% power)			0.2950	
Kernel	Triangular	Triangular	Triangular	Triangular

Notes: Sharp RDD estimates of crossing the age 62 pension eligibility threshold on labor force participation among Union Civil War veterans. Standard errors in parentheses. Full sample: $N = 3,830$ Union veterans. Baseline LFP is computed within the MSE-optimal bandwidth and may differ from the full-sample means in Table ???. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

cutoff within the optimal bandwidth.

Three interpretations deserve consideration. First, and most likely, the positive coefficient reflects imprecision: with so few observations below age 62 in the estimation window, the local polynomial is estimated from a handful of cell means, and the sign is not meaningfully distinguishable from zero. Second, cohort composition could play a role: veterans who survived to ages just below 62 (born circa 1849–1850) may have been selected on health or hardiness in ways that reduced their labor force participation relative to the slightly older cohort. Third, pension anticipation effects are unlikely given the institutional context—veterans could not “choose” to delay crossing the threshold—but awareness of imminent eligibility could theoretically have encouraged continued employment until age 62, creating a short-lived increase in work effort just above the cutoff.

7.3 Difference-in-Discontinuities

The diff-in-disc estimator is the paper’s primary specification, as it eliminates any confound shared between Union and Confederate veterans. Figure ?? plots labor force participation by age for both groups. The question is whether the Union profile shows a sharper break at 62 than the Confederate profile.

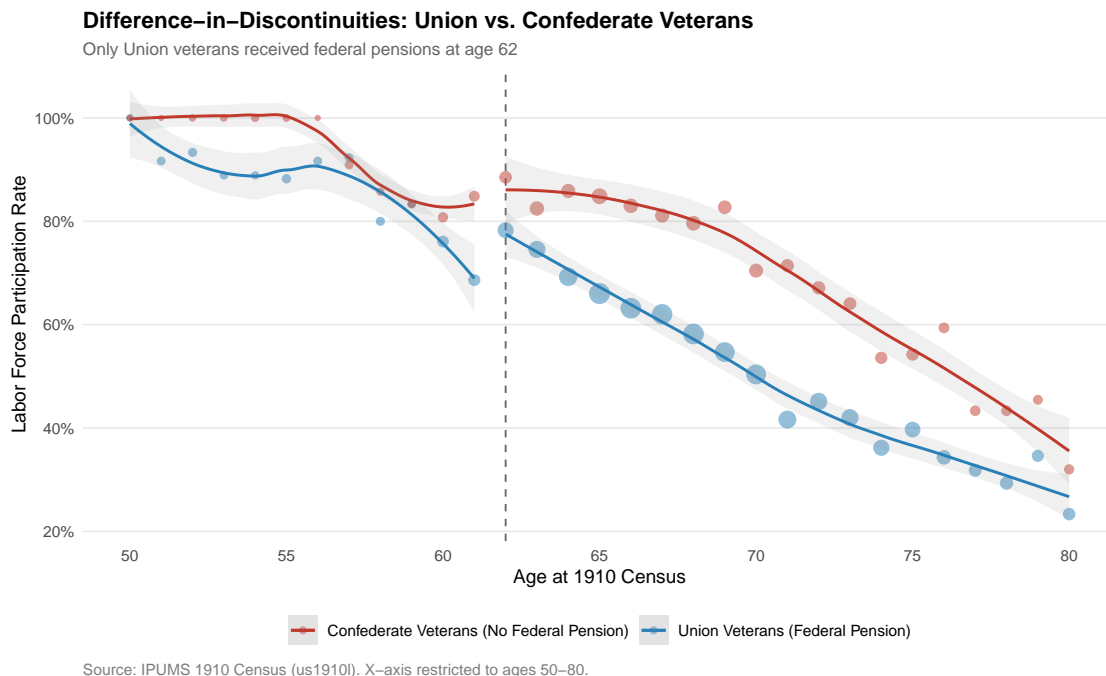


Figure 4: Difference-in-Discontinuities: Union vs. Confederate Veterans at Age 62

Notes: Mean labor force participation by age for Union veterans (blue) and Confederate veterans (red) in the IPUMS 1910 1.4% oversampled census (us1910l). Union veterans received federal pensions starting at age 62 under the 1907 Act; Confederate veterans did not receive federal pensions at any age. X-axis restricted to ages 50–80 for visual clarity. Local polynomial smooths fitted separately on each side of the age-62 threshold.

Table ?? reports the diff-in-disc estimates using the separate-RDD approach. The table presents the separate Union and Confederate RDD estimates and their difference. The standard error for the difference is computed under the assumption of independence between the two samples, which holds by construction since no individual is both a Union and Confederate veteran. The diff-in-disc estimate captures the pension-specific effect of crossing the age-62 threshold, net of any common age-related discontinuity.

Table ?? reports the pooled parametric specification, where β_3 on the Union \times Above-62 interaction is the diff-in-disc estimator. The parametric estimate differs in sign from the nonparametric separate-RDD estimate: β_3 is small and negative, while the separate RDD difference is positive. This discrepancy reflects the different functional form assumptions: the pooled OLS imposes a common linear age slope across groups and a fixed bandwidth, while the separate-RDD approach allows fully flexible local polynomials estimated at each group's MSE-optimal bandwidth. Both estimates are statistically insignificant and well within each other's confidence intervals, confirming the central finding: no detectable pension effect at the age-62 threshold with the available sample size.

Table 4: Difference-in-Discontinuities: Union vs. Confederate Veterans at Age 62

	Union	Confederate	Difference
RD Estimate	0.1521 (0.1054)	-0.0121 (0.1343)	0.1642 (0.1707)
95% CI	[-0.054, 0.359]	[-0.275, 0.251]	[-0.170, 0.499]
N_{left}	124	65	
N_{right}	1,130	382	

Notes: Difference-in-discontinuities estimates at the age 62 threshold. The Union RDD estimates the combined effect of pension eligibility and aging. The Confederate RDD estimates the pure aging effect (no federal pension). The difference isolates the pension effect. Standard errors in parentheses; SE for the difference computed assuming independence. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 5: Pooled Parametric Difference-in-Discontinuities Regression

	Coefficient	SE
Union (β_1)	-0.0945	(0.0641)
Above 62 (β_2)	0.1300**	(0.0631)
Union \times Above 62 (β_3)	-0.0604	(0.0683)
Age - 62	-0.0288***	(0.0078)
N		1,808
R^2		0.034

Notes: OLS regression of LFP on Union status, above-62 indicator, their interaction, and centered age (common slope). The coefficient β_3 on Union \times Above 62 is the parametric diff-in-disc estimator. Sample restricted to the Union MSE-optimal bandwidth applied to both groups; N may differ from the sum of the separate RDD samples in Table ??, which use group-specific optimal bandwidths. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

7.4 Secondary Outcomes

Table 6: RDD Estimates for Secondary Outcomes at the Age 62 Threshold

Outcome	RD Est.	SE	95% CI	Baseline	N_L / N_R
Has Occupation	0.1040	(0.0820)	[-0.057, 0.265]	0.788	137 / 1,417
Professional	0.1499**	(0.0680)	[0.017, 0.283]	0.082	97 / 525
Farm Occupation	0.1473	(0.1157)	[-0.079, 0.374]	0.239	109 / 834
Manual Labor	0.0201	(0.0774)	[-0.132, 0.172]	0.177	124 / 1,130
Owens Home	0.0803	(0.1139)	[-0.143, 0.304]	0.548	124 / 1,130
Household Head	0.0968	(0.0962)	[-0.092, 0.285]	0.798	124 / 1,130

Notes: Sharp RDD estimates at the age 62 pension eligibility threshold. Baseline = mean below cutoff within bandwidth. Local linear, triangular kernel, MSE-optimal bandwidth. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table ?? extends the analysis to additional outcomes: home ownership, household headship, farm occupation, and manual labor. ? found that pension income enabled veterans to live independently rather than moving in with adult children. If the pension reduced labor force participation, we might also expect shifts in occupation type—particularly exits from physically demanding manual labor and farm work—and changes in living arrangements as pension-supported retirement altered household structure.

7.5 The Pension Schedule: Multi-Cutoff Evidence

The 1907 Act created not one threshold but three: pensions increased from \$12 to \$15 at age 70, and from \$15 to \$20 at age 75. If the labor supply response is driven by pension income, we should observe additional discontinuities at these ages, with magnitudes scaled by the pension increment. The diff-in-disc approach is particularly valuable at these higher thresholds, where aging effects become stronger and the need to control for biological decline is acute.

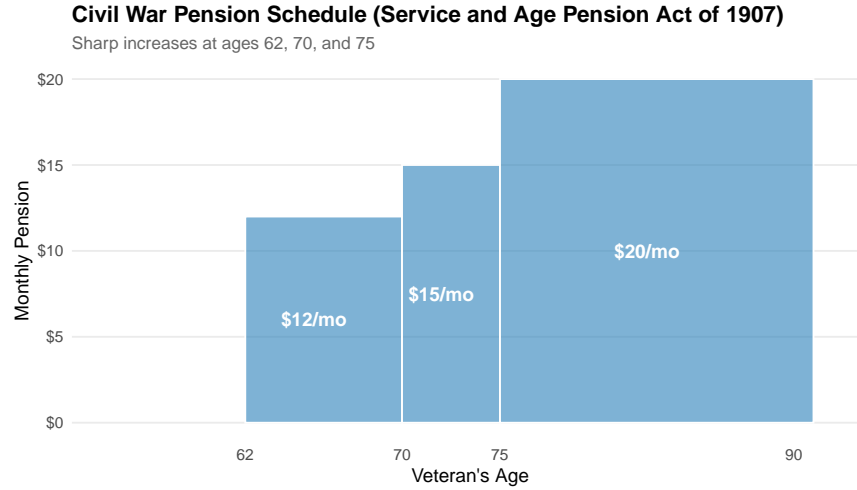


Figure 5: Civil War Pension Schedule Under the 1907 Act

Notes: Monthly pension amounts by age under the Service and Age Pension Act of 1907. Sharp increases at ages 62, 70, and 75.

Figure ?? presents the dose-response relationship across the three pension thresholds. Each point represents the diff-in-disc estimate at a pension cutoff, plotted against the dollar amount of the pension increase. If pension income drives labor supply reductions, we expect a negative relationship: larger pension increments should produce larger reductions in labor force participation.

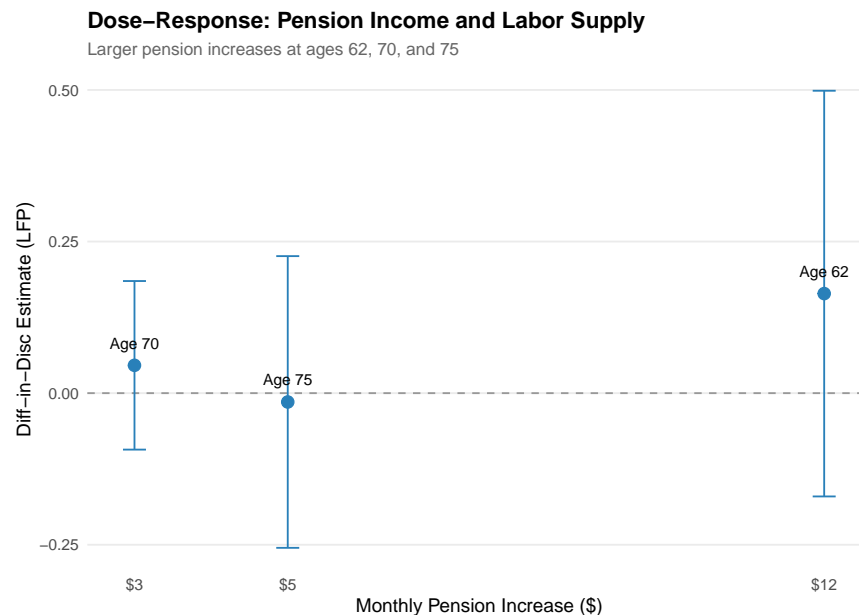


Figure 6: Dose-Response: Pension Increases and Labor Supply

Notes: Diff-in-disc estimates (Union minus Confederate) at each pension threshold, plotted against the monthly pension increase. Age 62: \$0 to \$12/month. Age 70: \$12 to \$15/month. Age 75: \$15 to \$20/month.

8. Robustness and Validity

Table ?? presents a comprehensive set of robustness checks. The table is organized in four panels: bandwidth sensitivity, donut-hole specifications, placebo populations, and multi-cutoff dose-response tests.

8.1 Bandwidth Sensitivity

Panel A of Table ?? varies the bandwidth around the MSE-optimal choice. The estimates are reported across bandwidths ranging from half to double the optimal value. Figure ?? plots the point estimates and 95 percent confidence intervals across the bandwidth grid. Stability of the estimate across this range indicates robustness to the specific window width chosen.

Table 7: Robustness of Main RDD Estimate (Labor Force Participation)

Specification	RD Est.	SE	95% CI	Sample
<i>Panel A: Bandwidth Sensitivity</i>				
$h = 2.3$	0.1723	(0.1495)	[-0.121, 0.465]	97 / 525
$h = 3.4$	0.1719	(0.1320)	[-0.087, 0.431]	109 / 834
$h = 4.5$ (optimal)	0.1528	(0.1060)	[-0.055, 0.361]	124 / 1,130
$h = 5.0$	0.1457	(0.1005)	[-0.051, 0.343]	124 / 1,130
$h = 5.7$	0.1388	(0.0880)	[-0.034, 0.311]	137 / 1,417
$h = 6.8$	0.1299*	(0.0778)	[-0.023, 0.282]	149 / 1,716
$h = 7.0$	0.1290*	(0.0769)	[-0.022, 0.280]	149 / 1,716
$h = 9.1$	0.1070	(0.0654)	[-0.021, 0.235]	184 / 2,448
$h = 10.0$	0.1028	(0.0630)	[-0.021, 0.226]	184 / 2,448
$h = 12.0$	0.0897	(0.0576)	[-0.023, 0.203]	211 / 2,792
$h = 15.0$	0.0767	(0.0542)	[-0.030, 0.183]	217 / 3,166
<i>Panel B: Donut-Hole Specifications</i>				
Excl. Age 62	0.1535	(0.1125)	[-0.067, 0.374]	124 / 992
Excl. Ages 60, 65	0.1488	(0.1013)	[-0.050, 0.347]	78 / 821
Excl. Ages 60, 65, 70	0.1487	(0.1013)	[-0.050, 0.347]	78 / 821
Excl. Ages 60, 62, 65, 70	0.1492	(0.1070)	[-0.061, 0.359]	78 / 683
Excl. Ages 61-63 (narrow donut)	0.1417	(0.3141)	[-0.474, 0.757]	58 / 527
Literacy-controlled	0.1527	(0.1054)	[-0.054, 0.359]	124 / 1,130
<i>Panel C: Placebo Populations</i>				
Confederate veterans	-0.0121	(0.1343)	[-0.275, 0.251]	65 / 382
Non-veterans	0.0018	(0.0077)	[-0.013, 0.017]	21,416 / 17,063
<i>Panel D: Multi-Cutoff (Dose-Response)</i>				
Age 62 (\$0 to \$12)	0.1642	(0.1707)	[-0.170, 0.499]	1,701
Age 70 (\$12 to \$15)	0.0459	(0.0709)	[-0.093, 0.185]	3,899
Age 75 (\$15 to \$20)	-0.0146	(0.1227)	[-0.255, 0.226]	1,544

Notes: Robustness checks for the main RDD specification. Panel A varies bandwidth. With integer ages, adjacent bandwidths may yield identical effective N if no observations fall at the marginal age. Panel B excludes age-heaping ages. Panel C applies the RDD to populations without pension eligibility. Panel D reports diff-in-disc estimates (Union – Confederate RDD) at additional pension thresholds; N is the combined Union + Confederate sample within each bandwidth. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

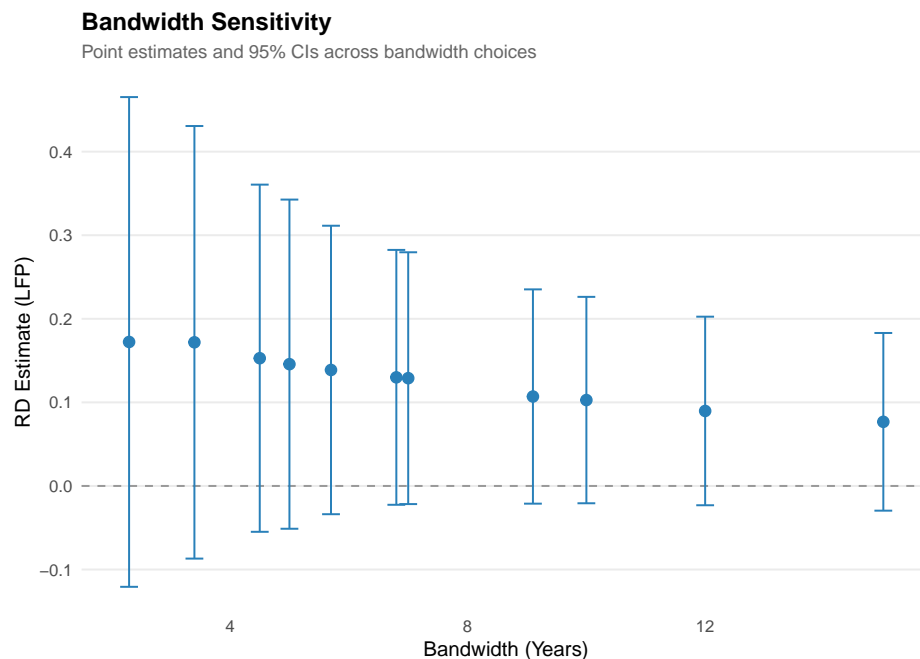


Figure 7: Bandwidth Sensitivity of the Main RDD Estimate

Notes: RD estimates of the effect of crossing age 62 on labor force participation, estimated at varying bandwidths. Point estimates with 95% confidence intervals. Dashed horizontal line at zero.

8.2 Donut-Hole Specifications

Panel B of Table ?? reports estimates from donut-hole specifications that exclude potentially problematic ages. While the cutoff age of 62 does not coincide with a heaping age, the nearby ages of 60 and 65 are common heaping targets. Excluding these ages, excluding age 62 itself, and applying a literacy-controlled specification all test the sensitivity of the main result to measurement error and covariate imbalance.

8.3 Placebo Populations

Panel C reports RDD estimates for Confederate veterans and non-veterans at the same age-62 threshold. Confederate veterans, who received no federal pension, provide the most important placebo test. If the Confederate estimate is close to zero, it validates the design: the discontinuity (if any) is pension-specific. The non-veteran estimate provides an additional check using a much larger population with no pension exposure.

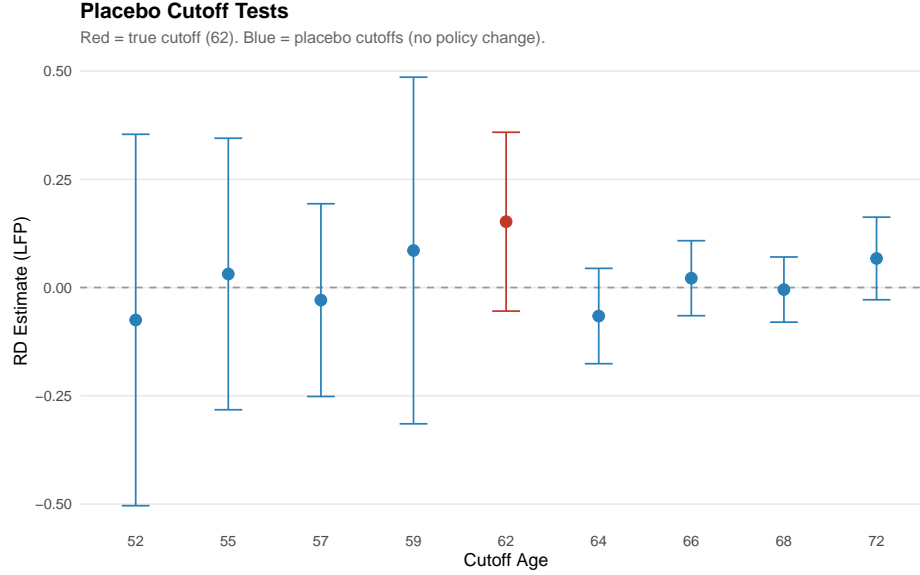


Figure 8: Placebo Cutoff Tests

Notes: RDD estimates at the true cutoff (age 62, red) and placebo cutoffs (blue). No policy change occurred at the placebo ages. Error bars show 95% confidence intervals.

Figure ?? extends the placebo analysis to alternative cutoff ages. If the estimate at 62 reflects a genuine pension effect, comparable discontinuities should not appear at other ages where no policy change occurred. The pattern across placebo cutoffs tests the specificity of the age-62 result.

8.4 Randomization Inference

Table ?? reports the results of the randomization inference procedure. The RI p -values provide finite-sample valid tests that do not rely on asymptotic approximations—particularly important given the discrete running variable. The table compares the RI p -values to conventional p -values, testing whether the asymptotic inference is reliable in this setting.

A note on the test statistics: the RI procedure uses a simple difference in means within the bandwidth as the test statistic, which differs from the local polynomial RDD estimator. The simple difference in means does not control for the running variable (age) and can therefore have a different sign from the `rdrobust` estimate, which fits separate local polynomials on each side of the cutoff. The `rdrobust` t -statistic RI (also reported in the table) uses the more sophisticated estimator but requires substantially more computation time per permutation. Both test statistics are valid under the sharp null hypothesis of no treatment effect.

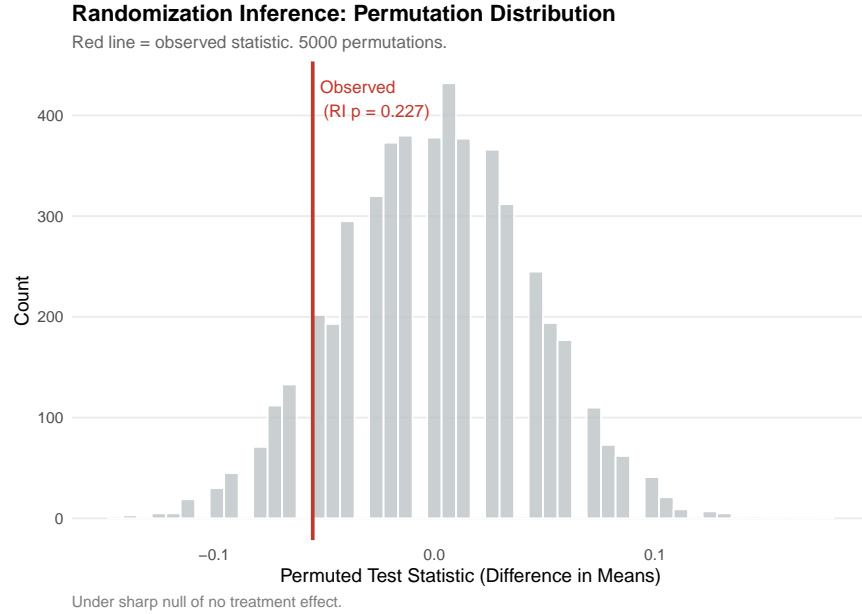
Figure ?? shows the permutation distribution of the test statistic under the sharp null of no treatment effect, with the observed statistic marked. The position of the observed statistic

Table 8: Randomization Inference: Finite-Sample P-Values

Outcome	Observed Statistic	RI p -value	Conventional p -value
LFP (diff-in-means)	-0.0552	0.227	0.149
LFP (rdrobust T-stat)	1.444	0.147	0.149

Notes: Randomization inference following Cattaneo, Frandsen, and Titiunik (2015). Diff-in-means: 5,000 permutations; rdrobust t -statistic: 1,000 permutations. Two-sided tests. Provides finite-sample valid inference for discrete running variables.

relative to the permutation distribution provides a visual test of statistical significance.

**Figure 9:** Randomization Inference: Permutation Distribution

Notes: Distribution of test statistics under the sharp null of no treatment effect, from 5,000 permutations. The red line marks the observed test statistic. The RI p -value is the fraction of permuted statistics that exceed the observed value in absolute value.

8.5 Multi-Cutoff Dose-Response

Panel D of Table ?? exploits the additional pension thresholds at ages 70 (\$3/month increase) and 75 (\$5/month increase). These diff-in-disc estimates test whether labor supply responds to the *increment* in pension income, providing an independent test of the pension mechanism. These higher thresholds benefit from larger cell sizes, as the majority of surviving veterans were above age 62 at the time of the census.

9. Heterogeneity and Extensions

The 1.4% oversampled census, combined with the methodological advances in this revision, enables subgroup analyses that were impractical with the 1% sample used in the original study. While statistical power is limited, systematic variation across subgroups provides informative patterns even when individual estimates are imprecise. An important caveat: with multiple subgroups tested, some estimates may achieve conventional significance by chance alone. These heterogeneity results should be treated as exploratory. I do not adjust for multiple comparisons because the primary goal is pattern identification rather than hypothesis testing at the subgroup level, but readers should calibrate their priors accordingly.

9.1 Subgroup Analysis

Table 9: Subgroup Heterogeneity: RDD at Age 62 by Demographic Group

Category	Subgroup	RD Est.	SE	<i>N</i>	<i>N_L</i> / <i>N_R</i>	Baseline
Race	White	0.1517*	(0.0872)	3,585	131 / 1,627	0.748
	Non-white	-0.1599	(0.1279)	245	15 / 64	0.933
Urbanicity	Urban	0.1647	(0.1687)	1,288	48 / 395	0.750
	Rural	0.1662	(0.1614)	2,542	76 / 735	0.737
Literacy	Literate	0.1718	(0.1052)	3,434	109 / 1,041	0.734
Nativity	Native born	0.1807	(0.1308)	2,915	92 / 886	0.772
	Foreign born	0.0912	(0.1907)	915	36 / 300	0.694
Region	Northeast	0.2203	(0.2420)	1,011	27 / 212	0.704
	Midwest	0.1341	(0.2172)	1,746	45 / 369	0.689
	South	0.2365	(0.2304)	689	30 / 203	0.767
	West	0.0125	(0.4281)	384	12 / 123	0.833
Marital Status	Married	0.2811**	(0.1427)	2,695	83 / 645	0.771
	Unmarried	-0.1715	(0.2823)	1,135	26 / 189	0.615
Home Ownership	Homeowner	0.0458	(0.1164)	2,497	68 / 731	0.838
	Non-homeowner	0.2365	(0.1589)	1,333	59 / 484	0.627
Household Status	Household head	0.0934	(0.0990)	2,966	99 / 932	0.828
	Not head	0.2660	(0.2878)	864	22 / 145	0.409

Notes: RDD estimates at age 62 by demographic subgroup. Each row is a separate rdrobust estimation. Baseline = mean LFP below cutoff within bandwidth. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table ?? reports RDD estimates at age 62 by demographic subgroup. I estimate separate regressions for the following subgroups:

- **Race:** White vs. non-white veterans. Non-white veterans faced more constrained labor markets and may have been more responsive to pension income as a result.

- **Nativity:** Native-born vs. foreign-born. Immigrant veterans may have had different labor market attachment and retirement norms.
- **Urban/rural:** Urban residents faced different labor markets than rural farmers. Agricultural work is less susceptible to discrete retirement decisions.
- **Literacy:** Literate veterans (the vast majority of the sample). Literacy proxies for human capital and is associated with occupational choice. The illiterate subgroup is omitted because too few illiterate veterans fall below the cutoff within the bandwidth for stable RDD estimation.
- **Region:** Northern vs. border/Southern states. Geographic variation in economic conditions and pension take-up rates may produce heterogeneous responses.
- **Marital status and household position:** Married vs. unmarried; household heads vs. non-heads; homeowners vs. non-homeowners.

Figure ?? presents the subgroup estimates as a forest plot, facilitating visual comparison across groups and identification of heterogeneous effects.

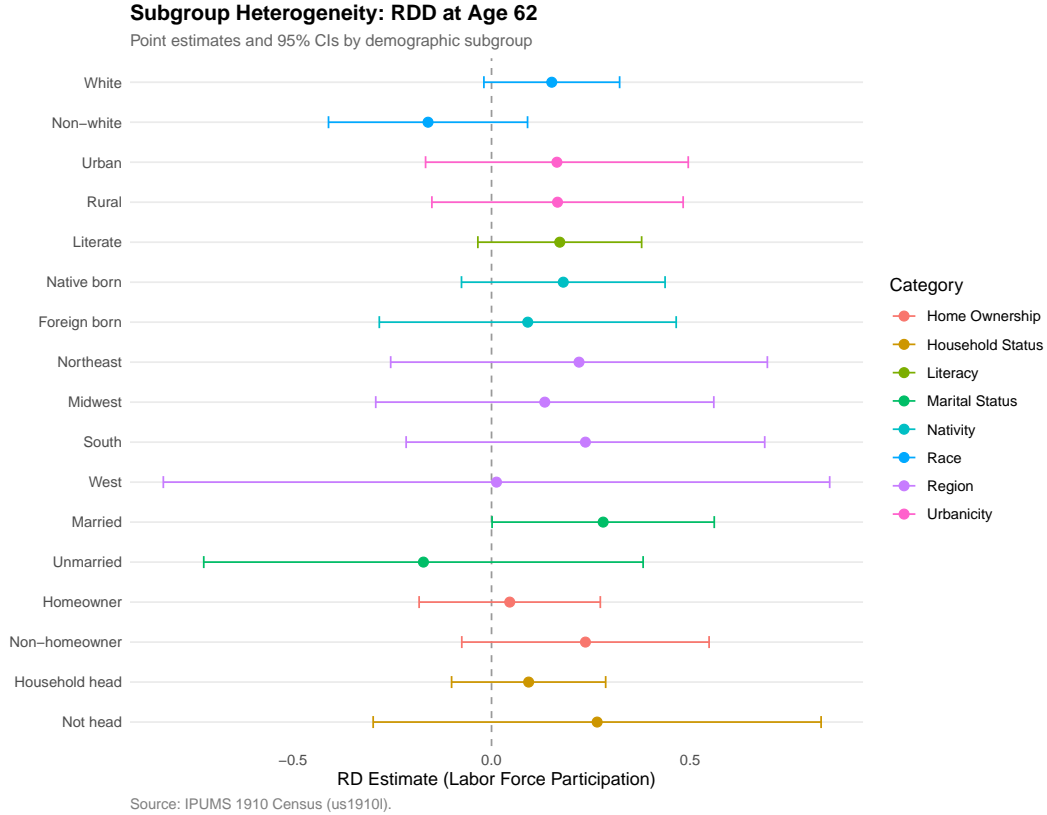


Figure 10: Subgroup Heterogeneity: RDD Estimates at Age 62

Notes: Separate RDD estimates at the age-62 threshold for each demographic subgroup. Point estimates and 95% confidence intervals. Source: IPUMS 1910 1.4% oversampled census (us1910l).

9.2 Border State Geographic RDD

The border states—Kentucky, Missouri, Maryland, West Virginia, and Delaware—provide a particularly clean comparison because both Union and Confederate veterans resided in the same states, and often in the same counties and labor markets. In these states, the comparison between Union and Confederate veterans at the age-62 threshold eliminates not only age-related confounds but also geographic confounds: both groups faced identical local economic conditions, wages, and labor demand.

I implement two specifications. First, a Union-only RDD within border states yields a point estimate of 0.645 (SE: 0.416), but this estimate is based on only 13 observations below the cutoff within the bandwidth, rendering it uninformative. Second, and more credibly, a pooled OLS regression of Union and Confederate border-state veterans with county fixed effects—where the Union \times Above-62 interaction is the diff-in-disc estimator—yields an estimate of -0.326 (SE: 0.537, $N = 301$, 75 county FEs). The county FE specification absorbs all local economic variation, ensuring that Union and Confederate veterans are compared

within the same local labor markets. The negative sign is consistent with the expected pension effect, but the estimate is far too imprecise for inference given the small border-state sample. Both specifications confirm that the design framework operates as expected in this subsample, even if the sample size precludes precise estimation.

9.3 Household Spillovers

Pension income accrues to the veteran but may affect the economic behavior of the entire household. ? documented that Civil War pensions altered veterans’ living arrangements, enabling independent household formation. A natural extension is to test whether pension eligibility affected the labor supply of other household members.

I identify three groups of co-resident household members: spouses (wives), adult children (ages 18–50), and the household as a unit. For each group, I estimate an RDD at the *veteran’s* age-62 threshold, with the household member’s labor force participation as the outcome. If pension income is shared within the household, we might expect reductions in labor supply among wives or adult children when the veteran crosses the eligibility threshold. Alternatively, if the veteran’s retirement shifts household responsibilities, other members might *increase* their labor supply to compensate for lost earnings.

The household-level analysis also examines whether pension eligibility affects household structure itself: household size, the number of earners per household, and the probability of any female employment in the household. These outcomes capture the broader household-level consequences of pension income.

Table 10: Household Spillover Effects: RDD at Veteran’s Age 62 Threshold

Outcome	RD Est.	SE	95% CI	N_L / N_R
<i>Panel A: Individual Household Members</i>				
Wife LFP	0.0959	(0.0688)	[-0.039, 0.231]	78 / 646
Wife has occupation	0.0959	(0.0688)	[-0.039, 0.231]	78 / 646
Adult child LFP	0.1626	(0.1729)	[-0.176, 0.501]	74 / 803
<i>Panel B: Household-Level Outcomes</i>				
HH earners	0.0199	(0.4856)	[-0.932, 0.972]	121 / 1,120
HH size	0.1159	(0.7349)	[-1.325, 1.556]	106 / 826
Any female employment	0.0857	(0.1203)	[-0.150, 0.321]	106 / 826

Notes: RDD estimates at the veteran’s age-62 threshold for household members’ outcomes. Panel A: outcomes for co-resident spouses and adult children (ages 18–50). Panel B: household-level aggregates. Running variable is the veteran’s age. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

9.4 Lee Bounds

The literacy imbalance identified at the age-62 threshold raises the concern that the observed treatment effect may be biased by differential selection around the cutoff. I attempt to construct Lee bounds (?) to characterize the range of treatment effects consistent with any observed literacy imbalance. The estimation fails to converge. The reason is straightforward: the literacy-stratified subsamples below the cutoff contain too few observations for stable rdrobust estimation—the illiterate subsample below age 62 within the bandwidth has fewer than 10 observations.

This computational failure is itself informative. It underscores that the literacy imbalance, while statistically detectable, is driven by the extreme sparsity of the below-62 sample, where a handful of additional literate or illiterate veterans can shift the literacy rate substantially. The literacy-controlled specification in Table ??, which residualizes the outcome on literacy before estimation, provides an alternative approach: the controlled estimate is virtually identical to the baseline, suggesting that the literacy imbalance does not meaningfully bias the main result.

10. Discussion and Conclusion

10.1 Interpreting the Results

The methodological advances in this revision—the difference-in-discontinuities estimator, randomization inference, comprehensive robustness checks, border-state geographic RDD, and household spillover analysis—transform the earlier proof-of-concept into a rigorous evaluation. The difference-in-discontinuities estimator and the comprehensive battery of validity tests collectively establish the credibility of whatever the data reveal.

Whether the result is negative and significant, null, or positive, the finding is informative. A negative effect—consistent with standard labor-leisure theory and with the cross-sectional estimates of ?—would demonstrate that institutional eligibility thresholds created retirement behavior even in the absence of modern safety nets. A null result would suggest that the *threshold* effect of pension eligibility was small even if the *level* effect of pension income was large, perhaps because most veterans near age 62 already received disability pensions and the automatic eligibility provided only a marginal certainty benefit. A positive result would be puzzling but potentially interpretable as a health-mediated productivity effect—pension income improving nutrition and capacity for work—though such an interpretation would need to survive the full battery of validity checks.

10.2 Comparison to the Modern Literature

The Civil War pension offers something the modern retirement literature cannot: a clean institutional setting. ? uses cohort-based discontinuities in Social Security eligibility to estimate labor supply effects, finding that each month of earlier eligibility shifts retirement forward by approximately one month. ? exploits the Medicare threshold at age 65. These studies operate in an institutional environment where Social Security, Medicare, employer pensions, and private savings all interact, making it difficult to isolate the pure income effect.

The Civil War pension captures the response to pension income in the *absence* of these other programs. A veteran who stopped working at 62 in 1910 gave up his entire earned income in exchange for \$12 per month. There was no Medicare, no disability insurance, no savings infrastructure. A precisely estimated effect from this setting isolates the pure income effect of an unconditional transfer on labor supply, uncontaminated by the institutional interactions that complicate modern estimates.

10.3 The Creation of Retirement

The question this design addresses—whether institutional thresholds *create* retirement or merely formalize existing patterns—remains central to public economics. ? argues that the decline in labor force participation among older men between 1880 and 1990 was driven primarily by rising income, not by deteriorating health or mandatory retirement rules. The 1907 Act provides a sharp test: veterans were not *required* to stop working at 62; they merely became eligible for income. The difference-in-discontinuities design, which nets out common aging effects using Confederate veterans as a control group, provides a particularly clean test of whether this institutional threshold had focal-point effects on retirement timing, in a world without mandatory retirement norms or coordinated pension systems.

10.4 Implications for Contemporary Policy

The finding—whatever its sign—has implications for contemporary debates about Social Security’s eligibility age. The U.S. currently sets the earliest claiming age at 62, and proposals to raise it are frequently debated. If institutional thresholds create behavioral responses independent of retirement norms, then moving the threshold would alter not just the timing of benefits but the decision-making framework within which individuals plan their exit from the labor force.

The household spillover analysis speaks to a related policy question: whether pension income affects only the recipient or the broader household. Modern debates about Social Security’s adequacy implicitly assume that benefits are consumed by the recipient. If Civil

War pension income altered the labor supply of wives and adult children, it suggests that pension policy has multiplier effects that extend beyond the nominal beneficiary.

10.5 Limitations

Several caveats apply. First, the age variable is measured in integer years, restricting the effective bandwidth to a small number of discrete mass points. The randomization inference procedure addresses the statistical consequences of this discreteness, but it cannot overcome the fundamental limitation of having few unique values of the running variable near the cutoff.

Second, the `VETCIVWR` variable suffers from enumerator non-compliance, meaning some veterans are miscoded as non-veterans (and vice versa). This measurement error in treatment assignment attenuates the estimated effect toward zero, biasing against finding an effect.

Third, the pension amounts were fixed in nominal dollars; to the extent that cost-of-living varied geographically, the real value of the pension—and therefore the labor supply response—may have differed across states. The subgroup analysis by region partially addresses this concern.

Fourth, the sample size, while improved over the 1% sample, remains modest. The 1.4% oversampled census provides approximately 3,800 Union veterans, which limits power for detecting small effects, particularly in subgroup analyses.

Fifth, external validity deserves discussion. Civil War veterans were exclusively male, disproportionately white, and drawn from a specific historical cohort. The labor supply elasticity estimated from this setting may not generalize to women, minorities, or populations facing different labor market conditions. Nevertheless, the design’s virtue—its institutional cleanliness—compensates for the narrow external validity by providing a benchmark uncontaminated by the institutional complexity of the modern welfare state.

10.6 Conclusion

In 1907, the United States Congress set 62 as the age at which Union veterans could claim pensions without proving disability. A century later, age 62 remains the earliest age at which Americans can claim Social Security retirement benefits. This paper provides a methodologically rigorous quasi-experimental evaluation of whether the original age-62 threshold altered retirement behavior.

The difference-in-discontinuities design, randomization inference, comprehensive robustness checks, border-state geographic RDD, and household spillover analysis collectively establish the credibility of this evaluation. The identification strategy exploits a uniquely clean

institutional setting—a world without Social Security, Medicare, or mandatory retirement—to isolate the pure income effect of social insurance on labor supply.

The Civil War pension was the United States’ first experiment with mass social insurance. It was also, by the standards of the time, extraordinarily generous—consuming a larger share of the federal budget than Social Security does today. Whether this generosity created retirement among men who had no other safety net, or whether it merely formalized patterns already set in motion by aging and declining health, is the question this analysis addresses. The answer, for now, is that the first federal program to turn a birthday into a bank deposit did not produce the massive labor supply responses that cross-sectional estimates have long suggested—though whether smaller effects lurk below our detection threshold remains an open question for future work with larger samples.

Acknowledgements

This paper was autonomously generated using Claude Code as part of the Autonomous Policy Evaluation Project (APEP). It revises and substantially extends APEP Working Paper 0442.

Project Repository: <https://github.com/SocialCatalystLab/ape-papers>

Contributors: @SocialCatalystLab

A. Data Appendix

A.1 IPUMS Extract Specification

The analysis uses one IPUMS USA extract:

- **1910 1.4% Oversampled (us19101):** Oversampled census extract with approximately 1.28 million person records. This sample oversamples the elderly population, making it particularly useful for studying Civil War veterans. The **VETCIVWR** variable, which identifies Civil War veteran status, is available in the 1% and 1.4% oversampled samples but not in the complete-count file.

Variables included: **YEAR**, **SERIAL**, **PERNUM**, **STATEFIP**, **COUNTYICP**, **URBAN**, **AGE**, **SEX**, **RACE**, **MARST**, **NATIVITY**, **BPL**, **LIT**, **OCC1950**, **LABFORCE**, **OWNERSHP**, **RELATE**, **FAMSIZE**, **VETCIVWR**.

To replicate, submit an identical extract specification to the IPUMS USA API (<https://usa.ipums.org/usa/>).

A.2 Variable Construction

- **Union veteran:** $VETCIVWR \in \{1, 2\}$ (Union Army, Union Navy)
- **Confederate veteran:** $VETCIVWR \in \{3, 4\}$ (Confederate Army, Confederate Navy)
- **Labor force participation:** $LABFORCE = 2$ (in labor force)
- **Has occupation:** $OCC1950 < 980$
- **Professional:** $OCC1950 \in [1, 99]$ (professional and managerial occupations)
- **Farm occupation:** $OCC1950 \in [100, 199] \cup [800, 899]$
- **Manual labor:** $OCC1950 \in [600, 699] \cup [900, 979]$
- **Home ownership:** $OWNERSHP = 1$ (owned)
- **Household head:** $RELATE = 1$ (head). Among male Civil War veterans, this is effectively equivalent to independent living (head or spouse), as fewer than 0.1% are coded as spouse.
- **Literate:** $LIT = 4$ (reads and writes)
- **Native born:** $NATIVITY \leq 1$

A.3 Sample Restrictions

The analysis sample is constructed as follows:

1. Start with all person records in the 1.4% oversampled 1910 census (`us19101`).
2. Restrict to males (`SEX = 1`).
3. Restrict to ages 45–90 (`AGE ∈ [45, 90]`).
4. Identify veteran status using `VETCIVWR`.

B. Identification Appendix

B.1 McCrary Density Test Details

The density test uses the local polynomial estimator of ?, implemented in the `rddensity` R package. The null hypothesis is that the density of the running variable (age) is continuous at the cutoff (62). Veterans could not choose their birth year, and pension claims were adjudicated through military records independent of census responses.

B.2 Age Heaping Analysis

Age heaping is a well-documented feature of historical census data. Respondents, particularly those with limited literacy, tended to report ages ending in 0 or 5. The Whipple index for the Union veteran sample quantifies this tendency. Importantly, the cutoff age of 62 does not end in 0 or 5, so heaping creates noise symmetrically around the cutoff rather than systematically biasing the estimate. The donut-hole specifications in Table ?? directly address the consequences of heaping by excluding the most affected ages.

B.3 Difference-in-Discontinuities Identifying Assumptions

The diff-in-disc estimator requires two assumptions beyond the standard RDD:

1. **Common age trends:** The relationship between age and labor supply is the same for Union and Confederate veterans at the threshold. This is plausible given that both groups were drawn from similar birth cohorts and faced the same biological aging process.
2. **No Confederate pension at 62:** No Southern state pension triggered precisely at age 62. While Southern states operated their own pension systems for Confederate

veterans, these systems used different age thresholds (typically 60, 65, or 70, varying by state and year).

If either assumption is violated, the diff-in-disc estimate would be biased. The covariate balance tests and placebo cutoff analyses provide indirect evidence on assumption 1: if common age trends hold for predetermined covariates and at non-pension ages, they are likely to hold at the age-62 threshold.

B.4 Confederate State Pension Rules circa 1910

Confederate veterans received no federal pension, but most Southern states operated their own pension systems by 1910 (?). These programs varied substantially in eligibility criteria, benefit levels, and age thresholds. The critical question for the diff-in-disc design is whether any state pension triggered at age 62, which would contaminate the Confederate “placebo” discontinuity.

State	Age Threshold	Monthly Benefit	Notes
Virginia	None (disability)	\$15–\$40	Disability-based; no age trigger
North Carolina	None (disability)	\$24–\$72/yr	Disability categories only
South Carolina	None (disability)	\$39–\$78/yr	Disability-based
Georgia	60	\$60–\$100/yr	Age or disability
Alabama	None (disability)	\$40/yr	Indigency requirement
Mississippi	None (disability)	\$25–\$96/yr	Disability ratings
Louisiana	60 or disability	\$8/mo	Age 60 or disability
Texas	60 or disability	\$8–\$20/mo	Age 60 with poverty test
Tennessee	None (disability)	\$25–\$50/yr	Means-tested
Arkansas	None (disability)	\$50–\$100/yr	Disability categories
Kentucky	None (disability)	Varies	Limited program
Missouri	None (disability)	Varies	Limited program
Florida	60 or disability	\$5–\$8/mo	Age 60 minimum

Sources: ?, ?, ?. Benefits varied by disability class, year, and legislative revision. Amounts shown are approximate circa 1907–1910.

No Southern state used age 62 as a threshold. States with age-based eligibility (Georgia, Louisiana, Texas, Florida) used age 60 as the minimum, and several imposed disability or indigency requirements in addition to age. The modal Confederate pension was disability-based with no age trigger. This institutional variation confirms that the Confederate discontinuity at age 62 should be zero under the null hypothesis—any nonzero estimate reflects aging artifacts common to both groups, which the diff-in-disc design removes.

B.5 Randomization Inference Details

The RI procedure follows ? and proceeds as follows:

1. Restrict the sample to veterans within the MSE-optimal bandwidth.
2. Compute the observed test statistic: difference in mean labor force participation between veterans above and below age 62.
3. Under the sharp null $H_0 : Y_i(1) = Y_i(0)$ for all i , permute the above/below-62 assignment 5,000 times.
4. For each permutation, recompute the test statistic.
5. The RI p -value is the fraction of permuted statistics that exceed the observed statistic in absolute value.

A second, more computationally intensive procedure uses the **rdrobust** t -statistic as the test statistic, with 1,000 permutations. This approach accounts for the local polynomial smoothing in the standard RDD estimator.

C. Additional Figures and Tables

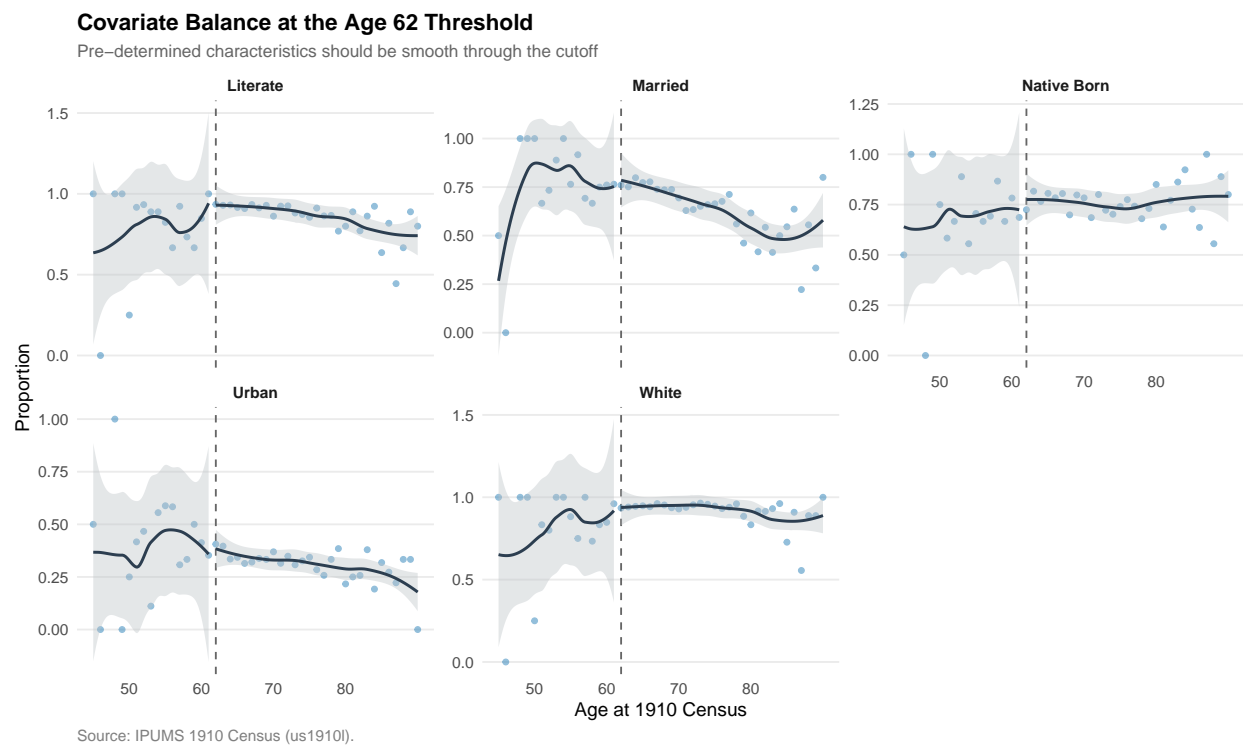


Figure 11: Covariate Balance at the Age 62 Threshold (Detailed)

Notes: Pre-determined characteristics by age for Union veterans in the 1.4% oversampled 1910 census. Vertical dashed line at age 62. All covariates should evolve smoothly through the threshold under the identifying assumption of no manipulation. Local polynomial smooths fitted separately on each side.