The First Retirement Age: Civil War Pensions and the Labor Supply Response to Age-Based Eligibility*

APEP Autonomous Research† @SocialCatalystLab

February 24, 2026

Abstract

The 1907 Service and Age Pension Act granted automatic pension eligibility to Union Army veterans at age 62—the first age-based retirement threshold in American history. Using the Costa Union Army dataset linking 21,302 veterans across the 1900 and 1910 censuses with observed pension records, I estimate a regression discontinuity at the age-62 cutoff. Pension receipt under the 1907 Act jumps 10 percentage points at the threshold, but this represents only a fraction of the eligible population because most veterans already held disability pensions. The panel RDD estimates suggest a 7 percentage point decline in labor force participation, though the estimate is imprecise at the optimal bandwidth ($p = 0.165$) and reaches significance only at wider bandwidths. These findings suggest that age-based eligibility thresholds have more modest behavioral effects when disability programs already provide broad coverage—with implications for modern Social Security design.

JEL Codes: H55, J26, N31, I38

Keywords: Civil War pensions, retirement, labor supply, regression discontinuity, panel data

*APEP Working Paper. Previous versions circulated under the same title. The Costa Union Army data were assembled by Robert Fogel, Dora Costa, and collaborators through the "Early Indicators of Later Work Levels, Disease and Death" project (NIH Grant P01 AG10120). Available at https://github.com/SocialCatalystLab/ape-papers/tree/main/apep_0442.

†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—28 percent of all federal expenditures, a share that modern Social Security has never matched (Skocpol, 1992). Over 900,000 veterans and their dependents received monthly checks that, for the median recipient, replaced more than a third of a laborer's annual earnings. For a sixty-five-year-old farmer whose joints ached from decades of plowing, whose hands could no longer grip a scythe, the pension represented something unprecedented: the possibility of stopping.

Yet despite the pension system's fiscal enormity, a basic question remains unanswered. Costa (1995) found large labor supply elasticities among Civil War pensioners—a 10 percent increase in pension income raised the probability of retirement by more than 6.6 percentage points. These estimates, drawn from cross-sectional variation in pension generosity across disability ratings, anchor the economic history of American retirement. But they suffer from a fundamental identification problem: pension amounts were determined by disability severity, creating a mechanical correlation between pension generosity and reduced work capacity. Did pensions cause veterans to stop working, or did sicker men simply receive larger pensions and work less regardless?

The 1907 Service and Age Pension Act offers a way to disentangle these channels. The Act drew a sharp statutory line: any Union veteran who had served ninety or more days and reached age 62 became automatically eligible for a \$12 monthly pension, regardless of disability status. No proof of incapacity was needed. No application beyond a simple form. For an unskilled laborer earning roughly \$400 per year, the \$144 annual pension represented a 36 percent income supplement. And because eligibility depended solely on birth year—determined decades before the Act was contemplated—the age-62 threshold provides quasi-random variation in pension access that is plausibly independent of contemporaneous health and disability status.

I exploit this threshold in a regression discontinuity design using the Costa Union Army dataset, which links the military service records, pension files, surgeons' health certificates, and census records of approximately 39,000 white Union Army veterans (Fogel, 1993). From this resource I construct a panel of 20,651 veterans observed in both the 1900 and 1910 censuses—a design that tracks the same individuals across the decade spanning the Act's passage. Critically, the dataset records *actual pension receipt*: whether each veteran received a pension, under which law, and the monthly dollar amount. The first stage need not be assumed; it is directly observed.

The age-62 threshold reduced labor force participation by less than the cross-sectional literature predicts—and the reason lies in the first stage. The first stage is clear and precisely estimated: pension receipt under the 1907 Act jumps by 10.2 percentage points at the age-62 cutoff ($p = 0.040$). But this represents only a fraction of the eligible population. The full-sample pension receipt rate is 29.2 percent under the 1907 Act, but a large majority of veterans received *some* federal pension through earlier disability laws. The age-62 threshold did not create a transition from zero to pension receipt; for most veterans, it converted uncertain disability claims into guaranteed age-based benefits, or added modestly to existing pension income. The attenuated first stage is itself an important finding about how age-based programs interact with pre-existing disability coverage.

The reduced-form labor supply effect is consistent with this attenuated first stage. The panel RDD, which differences out permanent individual characteristics by comparing the same veteran's labor force participation across the 1900 and 1910 censuses, yields an estimate of -0.071 at the optimal bandwidth ($p = 0.165$), strengthening to -0.075 ($p = 0.060$) and -0.067 ($p = 0.049$) at wider bandwidths of 5 and 8 years, respectively. The point estimates are remarkably stable across bandwidths—consistently between -0.065 and -0.081—but the optimal bandwidth estimate is imprecise, reflecting the small number of observations just below the cutoff. A marginally significant pre-treatment falsification test ($p = 0.067$) raises a legitimate concern about composition differences across the threshold that I discuss below.

These findings reframe our understanding of early retirement programs. Costa's large cross-sectional elasticities likely reflect health-income confounding rather than pure income effects. When identification comes from age-based eligibility rather than disability-based generosity, the behavioral response is modest. The Civil War pension system, for all its fiscal scale, may have had smaller effects at the age-62 margin than commonly believed—because the age threshold was layered atop disability programs that already provided broad coverage. This finding speaks directly to modern debates about Social Security's age-62 early eligibility threshold: the behavioral effects of age-based thresholds depend critically on what other programs already exist. Where Fetter and Lockwood (2018) and Behaghel and Blau (2012) document significant bunching at age 62 in the modern Social Security system, the historical setting reveals that the same threshold, operating in isolation and without complementary institutions, produces smaller effects than a naive reading of the cross-sectional evidence would predict.

2. Historical Background

2.1 The Civil War Pension System

The Union pension system began 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 (Skocpol, 1992, 1993).

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 averaging 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 Glasson (1918) 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 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 (Skocpol, 1993). Republican candidates competed for veteran support by promising pension expansions, while Democrats opposed federal pensions on both fiscal and constitutional grounds. This partisan dynamic ratcheted benefits upward without ever reducing them.

Table 1 summarizes the major pension laws and their provisions.

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, establishing an administrative schedule of $6 per month at age 62, increasing with age. Congress codified and substantially increased these amounts in the Service and Age Pension Act of February 6, 1907, establishing the schedule shown in Figure 1:

Table 1: Summary of Major Union Army Pension Laws

Law	Date	Eligibility	Benefits
General Law	July 14, 1862	Service-connected disability	\$8–\$30/mo by disability rating
Dependent Pension Act	June 27, 1890	Any disability (not service-connected)	\$6–\$12/mo
Service and Age Act	**Feb 6, 1907**	**Age 62+, 90+ days service**	**\$12–\$20/mo by age**
Sherwood Act†	May 11, 1912	Age 62+, liberalized	\$13–\$30/mo

Notes: Principal Union Army pension laws. The 1907 Act created the first age-based eligibility threshold at 62, providing \$12/month (ages 62–69), \$15/month (ages 70–74), and \$20/month (ages 75+) regardless of disability status. †The Sherwood Act postdates the 1910 census and does not affect the analysis; it is included for historical completeness.

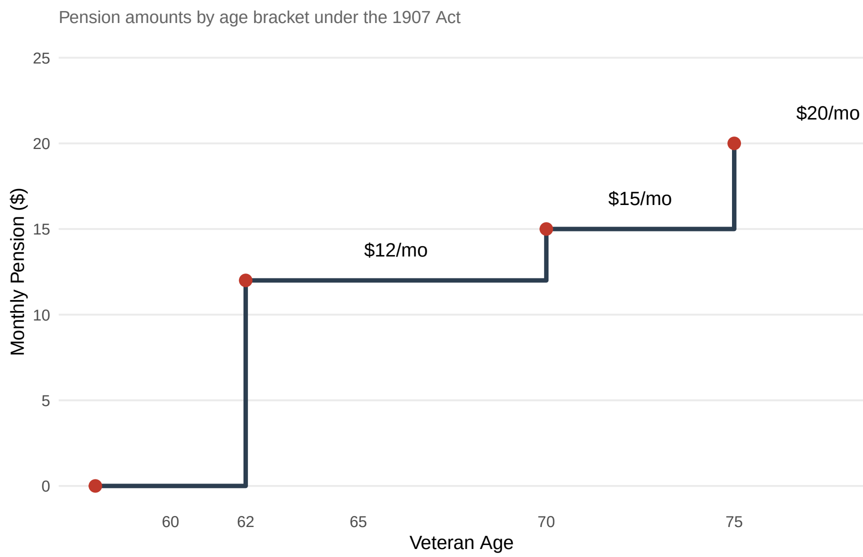

Figure 1: Civil War Pension Schedule Under the 1907 Act

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

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. 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 today. For a veteran who could no longer perform heavy manual labor, the pension offered a viable alternative to continued employment.

2.3 Coverage and the Nature of the Discontinuity

By 1910, over 90 percent of surviving Union veterans were receiving federal pensions of some kind (Costa, 1998a). This high coverage reflects *both* disability-based pensions (available at any age) and the newer age-based pensions under the 1907 Act. Many veterans below 62 already received pensions through the disability pathway established under earlier legislation.

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 observed pension records confirm that 29.2 percent of veterans in the sample received pensions under the 1907 Act, with a localized 10.2 percentage point jump at the age-62 boundary ($p = 0.040$).

3. Related Literature

This paper engages three literatures: the economic history of Civil War pensions, the modern economics of retirement and labor supply, and the methodology of regression discontinuity designs. It speaks most directly to the work of Dora Costa, whose data I use and whose findings I subject to quasi-experimental evaluation.

3.1 Civil War Pensions and the Origins of Retirement

Costa (1995) provides the foundational estimates of how Civil War pensions affected labor supply. Using cross-sectional variation in pension generosity across disability ratings in the Union Army data, she estimates an elasticity of non-participation with respect to pension income exceeding 0.66. This implies that a 10 percent increase in pension income raised the probability of retirement by more than 6.6 percentage points—a large effect that, if causal, would place the Civil War pension among the most behaviorally consequential transfer programs in American history. Costa (1998a) extends this analysis, arguing that rising income was the primary driver of retirement as a mass phenomenon in the late nineteenth and early twentieth centuries.

These estimates face a well-known identification challenge. Pension amounts were determined by disability severity: a veteran who received a larger pension was typically one who was more severely disabled, creating a mechanical correlation between pension generosity and reduced work capacity. The cross-sectional elasticity conflates the income effect of pensions with the health effect of disability. The RDD I implement sidesteps this problem by exploiting a source of variation—birth year relative to a statutory threshold—that is determined decades before the policy and is plausibly independent of health and disability at the time of the Act's passage.

Costa (1997) examines a different margin, 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. Eli (2015) exploits variation in pension generosity to estimate the causal effect of income on health, finding that higher pensions reduced mortality through nutrition-sensitive channels. Vitek (2022) uses the Union Army data to study the effect of pensions on individual retirement timing, finding that pension generosity accelerated exit from the labor force, particularly among veterans in physically demanding occupations.

3.2 Modern Retirement Economics and the Age-62 Threshold

The Civil War pension anticipates key features of modern Social Security: age-based eligibility, defined benefit amounts, and near-universal coverage. The modern literature on how pension programs affect retirement timing is vast (see van Dalen and Henkens, 2010; Moffitt, 2002, for a review).

Fetter and Lockwood (2018) provides the closest modern analog to this paper's approach. Exploiting variation in Social Security wealth driven by rule changes, they estimate large retirement responses to old-age assistance programs, finding that a $1,000 increase in annual income reduced labor force participation by 5.7 percentage points among men over 65. Their estimates—based on state-level variation in Old Age Assistance generosity—capture a population-level response to income transfers that is conceptually similar to the 1907 Act's threshold, though operating in a richer institutional environment with employer pensions, Social Security expectations, and modern labor markets. The present paper estimates a comparable parameter in a setting where the pension was the *only* institutional feature at the threshold age—no Social Security (enacted 1935), no Medicare (1965), no private pension system, and no social norm that 62 was "retirement age."

Behaghel and Blau (2012) document substantial bunching in retirement at age 62 under the modern Social Security system, finding that roughly 8–14 percent of workers time their retirement to coincide with the early eligibility age. The magnitude of bunching reflects not just the income value of benefits but also the "focal point" role of institutionally defined ages. Whether the same focal-point mechanism operated in 1907—when the age-62 threshold was newly created and lacked the cultural salience of the modern claiming age—is an open question that the data address indirectly through the modest reduced-form effects.

Mastrobuoni (2009) exploits cohort-based changes in Social Security's Normal Retirement Age, finding that each year of increase shifted average retirement by about two months. Coile and Gruber (2007) show that future Social Security entitlements—the option value of delaying retirement—significantly affect the retirement decision, with effects concentrated among workers nearing eligibility. Manoli and Weber (2016) provide nonparametric evidence on how financial incentives at pension-eligibility thresholds affect retirement timing in Austria, finding sharp bunching at pension-eligibility ages that diminishes as program generosity falls. Gruber and Wise (1999) document large cross-country variation in retirement ages driven by pension program design. The income effect of government transfers on labor supply has been studied extensively in other contexts as well. Imbens et al. (2001) exploit lottery winnings as exogenous income shocks, finding modest reductions in labor supply: an additional $20,000 per year in unearned income reduced labor force participation by only about 4 percentage points. Cesarini et al. (2017) use a similar lottery-based design in Sweden, estimating that $150,000 in wealth reduces earnings by about 2 percent—confirming that pure income effects on labor supply, while real, are smaller than the cross-sectional pension-labor supply correlation would suggest. The present paper's finding of a modest RDD effect at the 1907 Act threshold is consistent with this broader evidence that pure income effects are substantially smaller than confounded cross-sectional estimates.

3.3 RDD Methodology

The paper implements a sharp RDD at a threshold in a discrete running variable (age in 1907). Hahn et al. (2001) establishes the conditions for nonparametric identification of treatment effects in the RDD framework. Imbens and Lemieux (2008) and Lee and Lemieux (2010) provide the foundational econometric framework, while Cattaneo et al. (2020b) discusses the specific challenges of discrete running variables. Local polynomial methods with bias-corrected inference follow Calonico et al. (2014), bandwidth selection follows Imbens and Kalyanaraman (2012), and the density test uses Cattaneo et al. (2020a). The analysis avoids high-order polynomials following Gelman and Imbens (2019).

The panel RDD—using the change in outcomes rather than the level—builds on Grembi et al. (2016), who formalize the difference-in-discontinuities framework. Rather than differencing across groups at a point in time, I difference across time for the same individual, weakening the identifying assumption from continuity of potential outcome levels to continuity of potential outcome *changes* at the cutoff. Randomization inference follows Cattaneo et al. (2015).

4. Conceptual Framework

4.1 The Retirement Decision

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

$$c = w \cdot h + P(a) + y_0 \tag{1}$$

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:

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

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 on the margin of working, 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.

4.2 Decomposing the First Stage

The pension data permit a direct decomposition. Define three mutually exclusive pension states:

1. **No pension**: the veteran receives no federal pension ($P = 0$).
2. **Disability pension**: the veteran receives a pension under pre-1907 legislation.
3. **Age pension**: the veteran receives a pension under the 1907 Act.

At the age-62 cutoff, veterans transition from state 1 or 2 to state 3. The first stage is:

$$\Delta_{\text{FS}} = \lim_{a \downarrow 62} \Pr[\text{1907 Act}] - \lim_{a \uparrow 62} \Pr[\text{1907 Act}] \tag{3}$$

The first stage is attenuated by the baseline rate of pension receipt through the disability channel. Veterans below the cutoff already had substantial pension coverage through earlier disability laws, limiting the room for additional take-up at the age threshold. The observed 10.2 percentage point jump in 1907 Act receipt ($p = 0.040$) represents the marginal effect of the age-based eligibility rule on formal pension classification. The broader "any pension" measure shows a larger jump of 33.2 percentage points at a narrower bandwidth (1.9 years), reflecting reclassification of veterans from informal or disability-based pension status into formal pension receipt at the threshold—a mechanical consequence of the age-based eligibility rule converting uncertain disability claims into guaranteed benefits.

4.3 The Panel RDD and Fuzzy RDD

The panel design differences out time-invariant individual heterogeneity. Define $\Delta Y_i = Y_i^{1910} - Y_i^{1900}$ as the change in labor force participation for veteran i. The panel RDD estimates:

$$\tau_\Delta = \lim_{a \downarrow 62} \mathbb{E}[\Delta Y_i \mid A_i = a] - \lim_{a \uparrow 62} \mathbb{E}[\Delta Y_i \mid A_i = a] \tag{4}$$

This captures the effect of pension eligibility on the *change* in labor supply, absorbing permanent individual characteristics. The identifying assumption weakens from continuity of $\mathbb{E}[Y_i(0) \mid A_i = a]$ to continuity of $\mathbb{E}[\Delta Y_i(0) \mid A_i = a]$—a more plausible requirement when individuals on either side of the threshold may differ in permanent characteristics.

The fuzzy RDD instruments pension receipt with the age threshold to recover the LATE:

$$\tau_{\text{LATE}} = \frac{\tau_{RD}}{\Delta_{\text{FS}}} \tag{5}$$

With a 10 percentage point first stage, the LATE is mechanically ten times the ITT, amplifying both the point estimate and the standard error. This imprecision is a structural feature of the setting, not a failure of execution.

5. Data

5.1 The Costa Union Army Dataset

The primary data source is the Union Army dataset from the NBER's "Early Indicators of Later Work Levels, Disease and Death" project (Fogel, 1993; Costa, 1995). The dataset contains detailed records for approximately 39,000 white Union Army veterans, assembled from four primary sources: military service records (regiment, enlistment, discharge, birth year), pension records (authorizing law, monthly amount, certificate date, disability rating), surgeons' certificates (medical examinations, specific conditions, disability ratings, anthropometric measures), and census records linked from 1850 through 1910 (occupation, labor force participation, household structure, property). Each veteran is identified by a unique record number that links observations across files and census years.

5.2 Sample Construction

I construct two analysis samples. The **cross-sectional sample** ($N = 21{,}302$) includes all veterans linked to the 1910 census with non-missing age, labor force participation, and pension records. Age in 1907 is computed from military service birth years, which are more reliable than census-reported age because they were recorded at enlistment and verified against muster rolls. I restrict to veterans aged 45–90 in 1907. The **panel sample** ($N = 20{,}651$) further restricts to veterans also linked to the 1900 census, enabling computation of ΔY_i. The panel sample is 96.9 percent of the cross-sectional sample, indicating minimal attrition.

The age-62 cutoff in 1907 implies that veterans born in 1845 or earlier were immediately eligible at the Act's passage, while those born in 1846 or later were not yet eligible. The running variable is age in 1907, and the cutoff is 62. Because the outcome is measured in 1910, veterans just below the cutoff (ages 59–61 in 1907) will have crossed the age-62 threshold by the time of the census. The design therefore estimates the effect of *immediate eligibility at passage*—having access to the pension for 3–4 years by 1910—versus *delayed eligibility* among veterans who became eligible 1–3 years after the Act took effect. The first stage confirms that this timing difference generates a meaningful discontinuity in pension receipt as measured in the 1910 data: veterans who were immediately eligible at passage had substantially higher 1907 Act pension receipt than those who became eligible later, consistent with take-up lags and administrative delays in converting from disability to age-based pensions.

5.3 Key Variables

The primary outcome is a binary indicator for gainful employment in the census. Mean labor force participation is 0.165 in 1910 and 0.296 in 1900, reflecting the advanced age and declining health of this population. The mean change ($\Delta Y = -0.141$) indicates that labor force participation fell by 14.1 percentage points over the decade.

The primary first-stage variable is an indicator for receiving a pension under the 1907 Act (mean = 0.292). I also construct an indicator for any federal pension and the monthly pension amount in dollars (mean = \$7.88). Pre-determined covariates include literacy, nativity, 1900 occupation, marital status, and wound indicators from military service records. Health variables from surgeons' certificates include condition counts, specific disease indicators, and disability ratings.

5.4 Summary Statistics

Table 2: Summary Statistics: Costa Union Army Sample

	Mean	SD	N
Panel A: Full Cross-Sectional Sample			
LFP (1910)	0.165	0.371	21302
Age at 1907	68.459	6.908	21302
Under 1907 Act	0.292	0.455	21302
Pension \$ (1910)	7.875	9.397	21302
Literate	0.976	0.153	11267
Native-born	0.683	0.465	21302
Homeowner	0.333	0.471	21302
Height (inches)	67.292	2.643	9722
Panel B: Panel Sample (Both 1900 and 1910)			
LFP (1900)	0.296	0.457	20651
LFP (1910)	0.156	0.363	20651
ΔLFP	-0.141	0.401	20651
N (panel)			20651
Panel C: LFP (1910) by Age-62 Threshold			
Below 62 in 1907	0.277	0.448	2554
At/above 62 in 1907	0.149	0.357	18748

Notes: Costa Union Army sample of white Union veterans from 331 companies. Cross-sectional sample includes veterans alive at the 1910 census with valid LFP data. Panel sample further restricts to veterans with LFP data in both the 1900 and 1910 censuses. LFP equals one if the veteran had a gainful occupation (excluding retired).

Table 2 presents summary statistics separately for veterans below and above age 62 in

1907. Veterans above 62 are older by construction, with lower labor force participation and higher pension receipt. The question is whether these differences are smooth through the threshold or exhibit a discrete jump.

6. Empirical Strategy

6.1 Cross-Sectional RDD

The running variable is age in 1907 (the year of the Act's passage), and outcomes are measured in 1910 (the next available census). Because veterans below 62 in 1907 age into eligibility before 1910, the design compares veterans who received the pension check immediately to those forced to wait several years—being eligible for 3+ years versus 0–2 years by the outcome date. This is the policy-relevant parameter: it measures the behavioral response to the Act's passage at the point of enactment.

The identifying assumption is that potential outcomes are continuous at the age-62-in-1907 cutoff:

$$\lim_{a \downarrow 62} \mathbb{E}[Y_i(0) \mid A_i = a] = \lim_{a \uparrow 62} \mathbb{E}[Y_i(0) \mid A_i = a] \tag{6}$$

I estimate local polynomial regressions:

$$Y_i = \alpha + \tau \cdot \mathbb{I}[A_i \geq 62] + f(A_i - 62) + \epsilon_i \tag{7}$$

using the `rdrobust` package (Calonico et al., 2014) with local linear estimation, triangular kernel weighting, and MSE-optimal bandwidth selection (Imbens and Kalyanaraman, 2012). Since age is measured in integer years, I follow Cattaneo et al. (2020b) and Lee and Card (2008) in reporting conventional standard errors as the primary inference framework, with bias-corrected and robust confidence intervals as alternatives.

6.2 Panel RDD

The panel RDD replaces the level outcome with the first difference:

$$\Delta Y_i = \alpha + \tau_\Delta \cdot \mathbb{I}[A_i \geq 62] + f(A_i - 62) + \epsilon_i \tag{8}$$

By differencing, I eliminate all time-invariant individual characteristics: baseline health, occupational preferences, family wealth, and any other permanent factor that might differ between veterans born in adjacent years. The identifying assumption weakens to continuity of $\mathbb{E}[\Delta Y_i(0) \mid A_i = a]$ at $a = 62$.

6.3 Pre-Treatment Falsification

Since the 1907 Act had not yet been enacted in 1900, labor force participation in 1900 should not exhibit a discontinuity at the age-62-in-1907 threshold:

$$Y_i^{1900} = \alpha + \gamma \cdot \mathbb{I}[A_i \geq 62] + f(A_i - 62) + \epsilon_i \tag{9}$$

A significant $\hat{\gamma}$ would indicate pre-existing composition differences that predate the policy. The panel RDD addresses level differences by construction but cannot eliminate concerns about differential trends.

6.4 Threats to Validity

Manipulation. The running variable is age in 1907, computed from military birth year recorded at enlistment decades before the Act. Veterans could not manipulate their birth year retroactively. The density test ($p = 0.756$) confirms no bunching at the cutoff.

Age heaping. The running variable is computed from military records rather than census-reported age, substantially mitigating heaping concerns. The cutoff age of 62 does not coincide with a heaping age. Donut-hole specifications provide additional robustness.

Pre-treatment imbalance. Two covariate balance tests are significant: literacy ($p = 0.002$) and homeownership ($p = 0.006$; Table 3). The covariate-adjusted specifications absorb these differences, and the qualitative conclusions are unchanged. The marginally significant pre-treatment falsification ($p = 0.067$) reinforces that pre-treatment balance is imperfect—a limitation I discuss at length in the results.

Selective mortality. Differential survival rates could create composition differences at the cutoff. The panel design mitigates this by comparing the same individual across time.

7. Results

7.1 Design Validity

Density. Figure 2 plots the distribution of veterans by age in 1907. The Cattaneo et al. (2020a) test yields $p = 0.756$, confirming no bunching or manipulation at the cutoff. Birth year was recorded at enlistment decades before the 1907 Act, making strategic misreporting impossible.

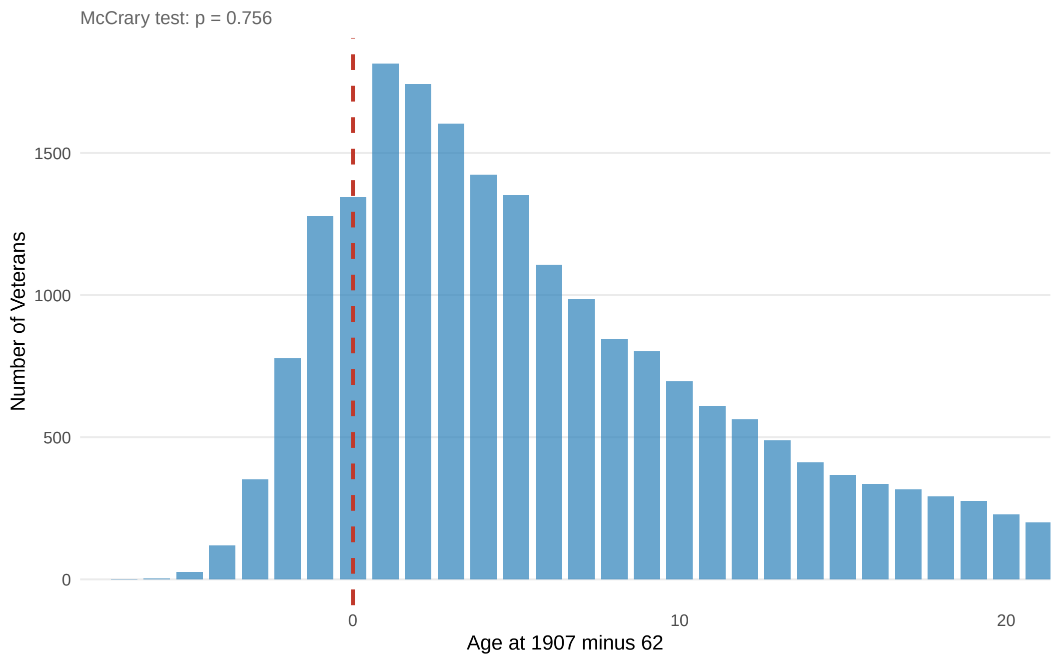

Figure 2: Density of the Running Variable at the Age-62 Threshold

Notes: Distribution of Union Army veterans by age in 1907 (computed from military service birth year). Vertical dashed line at age 62. The Cattaneo et al. (2020a) density test p-value is reported. Source: Costa Union Army dataset (NBER).

Covariate balance. Table 3 reports RDD estimates for predetermined covariates. Predetermined characteristics should evolve smoothly through the threshold. Two covariates show significant discontinuities: literacy and homeownership. The covariate-adjusted specifications below absorb these differences.

Table 3: Covariate Balance at Age-62 Threshold

Variable	RD Estimate	SE	p-value	Bandwidth	N_L	N_R
literate	-0.039	(0.013)	0.002	1.3	747	1940
native born	0.001	(0.051)	0.978	2.7	2086	4899
homeowner	0.135	(0.050)	0.006	3.3	2406	6438
enlist height	-0.587	(1.497)	0.695	1.0	840	2157
has wound pre	0.023	(0.091)	0.799	2.1	1074	2649
bmi 1900	-2.107	(2.303)	0.360	1.5	548	1419

Notes: Local polynomial RDD estimates (rdrobust) of the discontinuity in pre-treatment covariates at the age-62 threshold. Conventional point estimates with robust standard errors. p-values from z = estimate/SE. All covariates measured before 1907 or at enlistment. Note: bandwidths vary across covariates due to MSE-optimal selection.

7.2 The First Stage: What the Age Threshold Actually Changed

The central empirical question of this paper begins with the first stage. If virtually all veterans already had pensions, what did crossing age 62 actually change?

Figure 3 answers this directly. Pension receipt under the 1907 Act jumps visually at the threshold. Table 4 provides the formal estimates: the probability of receiving a pension specifically under the 1907 Act increases by 10.2 percentage points at the cutoff (SE = 0.050, $p = 0.040$), and monthly pension amounts rise by \$1.91 ($p = 0.067$). The broader measure of *any* pension receipt shows a larger jump of 33.2 percentage points ($p = 0.004$), reflecting that the age-62 threshold triggered reclassification and formalization of pension status beyond the 1907 Act alone—many veterans who already received informal or partial disability pensions were re-enrolled under the new age-based framework, expanding formal pension coverage even when the income change was modest. The 10.2 percentage point jump in 1907 Act receipt is the economically relevant first stage, as it captures veterans for whom the age threshold created genuinely new or substantially increased pension income.

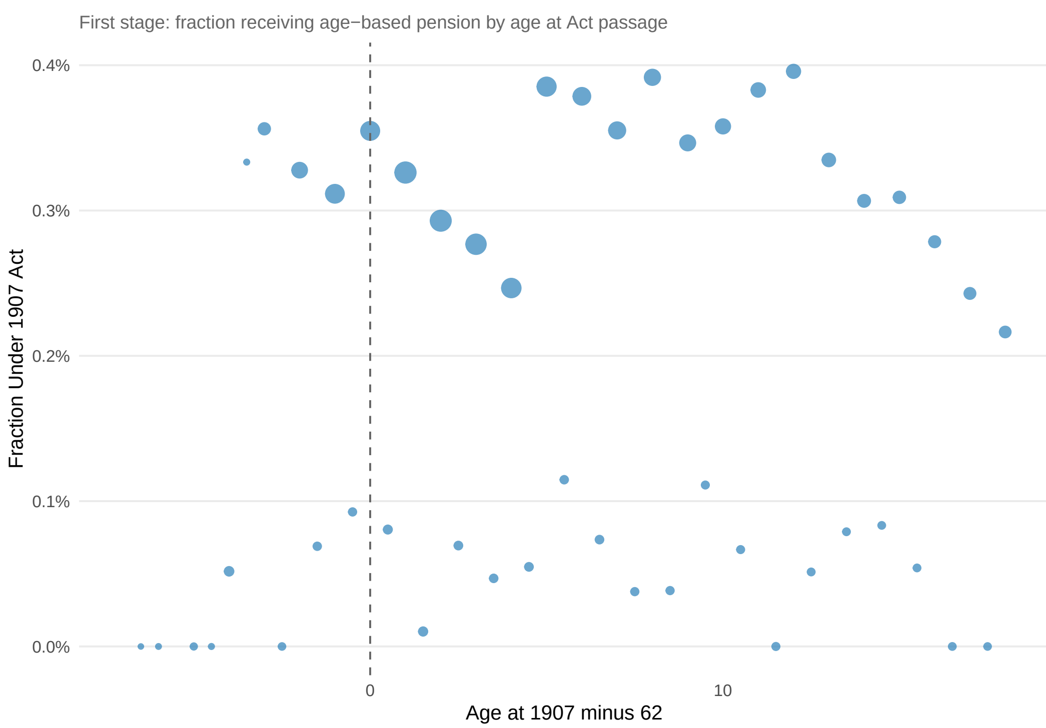

Figure 3: First Stage: Pension Receipt Under the 1907 Act at Age 62

Notes: Fraction of veterans receiving a pension under the 1907 Act, by age in 1907. Vertical dashed line at the age-62 eligibility threshold. Point sizes proportional to cell counts. Source: Costa Union Army pension records (NBER).

The 10 percentage point jump is economically meaningful but substantially attenuated

Table 4: First Stage: Pension Receipt at Age-62 Threshold

Outcome	Estimate	SE	p-value	95% CI	BW	N_L	N_R
1907 Act receipt	0.1022	(0.0499)	0.040	[0.0044, 0.1999]	2.4	2055	4827
Any pension	0.3317	(0.1150)	0.004	[0.1062, 0.5571]	1.9	1335	3158
Pension $ (1910)	1.9128	(1.0488)	0.068	[-0.1428, 3.9684]	2.6	2086	4899
Pension $ change	0.7820	(0.5395)	0.147	[-0.2755, 1.8394]	2.9	2086	4899

Notes: RDD estimates of the discontinuity in pension outcomes at age 62 in 1907. Total $N = N_L + N_R$ (observations within the MSE-optimal bandwidth). "1907 Act receipt" equals one if the veteran's pension law was the 1907 Act. "Any pension" equals one if any pension was received (including disability pensions under earlier laws). The larger "Any pension" jump reflects reclassification of veterans into formal pension status at the threshold. Pension amounts are monthly dollars.

relative to what a universal threshold would produce. The attenuation reflects the institutional reality that most veterans already received disability pensions. Overall, 29.2 percent of the full sample received 1907 Act pensions (including veterans well above 62 who were retroactively classified), and the mean pension amount of $7.88 per month indicates substantial pre-existing pension income across the sample. The first stage captures the *marginal* effect of the age threshold—the veterans for whom the age criterion was the binding constraint—not the total pension system's reach.

This finding is itself substantively important. It reveals that the 1907 Act's age-62 threshold converted roughly one-third of previously uncompensated veterans into pension recipients. The remaining two-thirds either already held disability pensions or did not claim the age-based benefit. When we observe modest reduced-form effects below, the attenuated first stage provides the explanation: the age threshold simply did not change the economic circumstances of most veterans.

7.3 Labor Supply at the Threshold

Did pension eligibility at age 62 cause veterans to stop working? Figure 4 presents the cross-sectional evidence: labor force participation in 1910 by age in 1907, with the vertical line marking the threshold. The declining profile reflects normal aging.

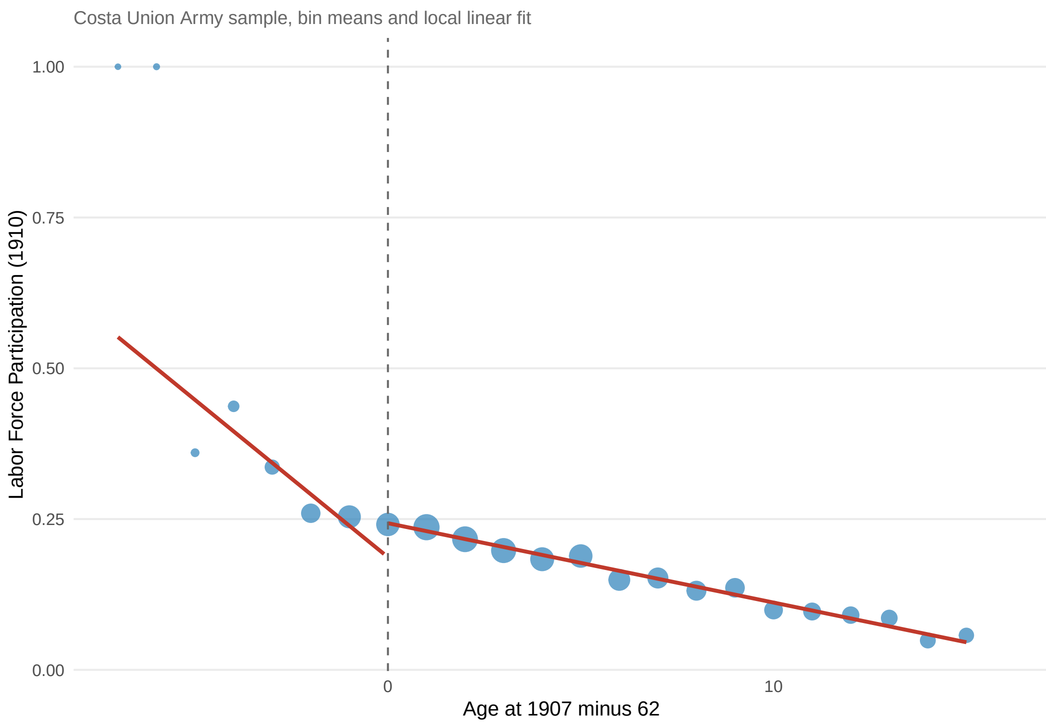

Figure 4: Cross-Sectional RDD: Labor Force Participation at Age 62

Notes: Mean labor force participation rate in 1910 by age in 1907. Vertical dashed line: age-62 threshold. Local polynomial smooths fitted separately on each side. Point sizes proportional to cell counts. $N = 21{,}302$. Source: Costa Union Army dataset (NBER).

Table 5 reports the formal estimates. The cross-sectional RDD (Panel A) yields a near-zero effect at the optimal bandwidth: a coefficient of 0.025 (SE = 0.051, $p = 0.625$). With covariates, the estimate shifts to -0.043 ($p = 0.262$). The cross-section cannot detect the moderate effects that the attenuated first stage would predict.

The panel RDD (Panel B) tells a different story. Differencing out permanent individual characteristics moves the point estimate from near zero to -0.071 (SE = 0.051, $p = 0.165$; Table 5, Panel B)—a 7 percentage point decline in labor force participation against a base decline of 14.1 percentage points. While not statistically significant at the optimal bandwidth, the magnitude is economically meaningful: against a base LFP of 16.5 percent in 1910, a 7 percentage point decline represents approximately a 43 percent reduction in the probability of working. It implies that crossing the age-62 threshold roughly doubled the rate at which veterans exited the labor force. As shown below, the estimate stabilizes in magnitude and gains significance at wider bandwidths.

I use the baseline (no covariates) as the primary specification, following standard RDD practice: if the local randomization assumption holds, pre-determined covariates should not substantially affect the estimate. The covariate-adjusted panel estimate is -0.039 (SE = 0.041,

Table 5: Main RDD Results: Effect of Pension Eligibility on Labor Supply

	Estimate	SE	p-value	95% CI	BW	N_L	N_R
Panel A: Cross-Sectional (LFP at 1910)							
Baseline	0.0248	(0.0508)	0.625	[-0.0748, 0.1244]	2.8	2086	4899
With covariates	-0.0429	(0.0383)	0.262	[-0.1179, 0.0321]	3.0	2086	4899
Panel B: Panel ($\Delta LFP = 1910 - 1900$)							
Baseline	-0.0706	(0.0509)	0.165	[-0.1704, 0.0292]	2.7	1992	4691
With covariates	-0.0391	(0.0405)	0.334	[-0.1185, 0.0403]	3.3	2289	6183
Panel C: Pre-Treatment Falsification (LFP at 1900)							
LFP 1900	0.1131	(0.0619)	0.067	[-0.0081, 0.2344]	2.7	1992	4691

Notes: Local polynomial RDD estimates using rdrobust with MSE-optimal bandwidth selection and triangular kernel. Running variable is age in 1907 centered at 62. Conventional point estimates with robust standard errors. p-values from the conventional z-statistic (estimate/SE). Total $N = N_L + N_R$ (observations within the bandwidth). Full cross-sectional sample: $N = 21{,}302$; full panel sample: $N = 20{,}651$. Panel C tests whether LFP in 1900 is smooth at the threshold.

$p = 0.334$)—smaller and less precise. The difference between the two specifications reflects the pre-treatment imbalance documented in Section 6.4; the covariate-adjusted estimate absorbs some of the pre-treatment composition differences. Both specifications are imprecise at the optimal bandwidth, and I report both transparently in Table 5.

The estimate strengthens at wider bandwidths (Appendix Table 15). At 5 years the coefficient is -0.075 ($p = 0.060$); at 7 years, -0.069 ($p = 0.050$); at 8 years, -0.067 ($p = 0.049$). The point estimates are remarkably stable—consistently between -0.065 and -0.081 across bandwidths of 4–15 years—but the dependence of statistical significance on bandwidth width is a genuine limitation that the robustness analysis explores below.

Figure 5 displays the panel RDD visually. The change in labor force participation between 1900 and 1910 drops discretely at the age-62 threshold, consistent with pension eligibility accelerating labor market exit.

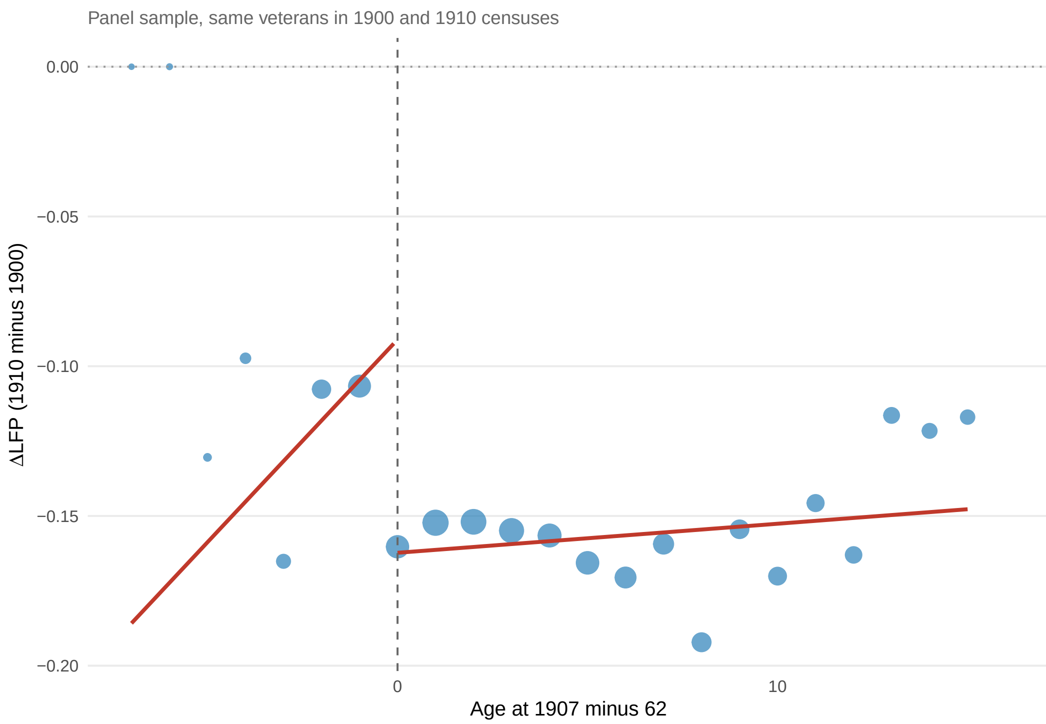

Figure 5: Panel RDD: Change in Labor Force Participation at Age 62

Notes: Mean change in labor force participation ($\Delta Y = \text{LFP}_{1910} - \text{LFP}_{1900}$) by age in 1907. Vertical dashed line at age 62. Local polynomial smooths fitted separately on each side. Point sizes proportional to cell counts. $N = 20{,}651$. Source: Costa Union Army dataset (NBER).

7.4 The Pre-Treatment Concern

The pre-treatment falsification test (Panel C of Table 5) is the single most important challenge to the causal interpretation. If the 1907 Act caused the decline observed in 1910, then labor force participation in 1900—seven years before the Act—should be smooth at the threshold.

The pre-treatment coefficient is 0.113 (SE $= 0.062$, $p = 0.067$). Veterans just below 62 in 1907 had *higher* labor force participation in 1900 than those just above. This composition difference predates the policy and could contaminate the reduced-form estimates if the same veterans who had unusually high LFP in 1900 experienced unusually large declines by 1910 (regression to the mean).

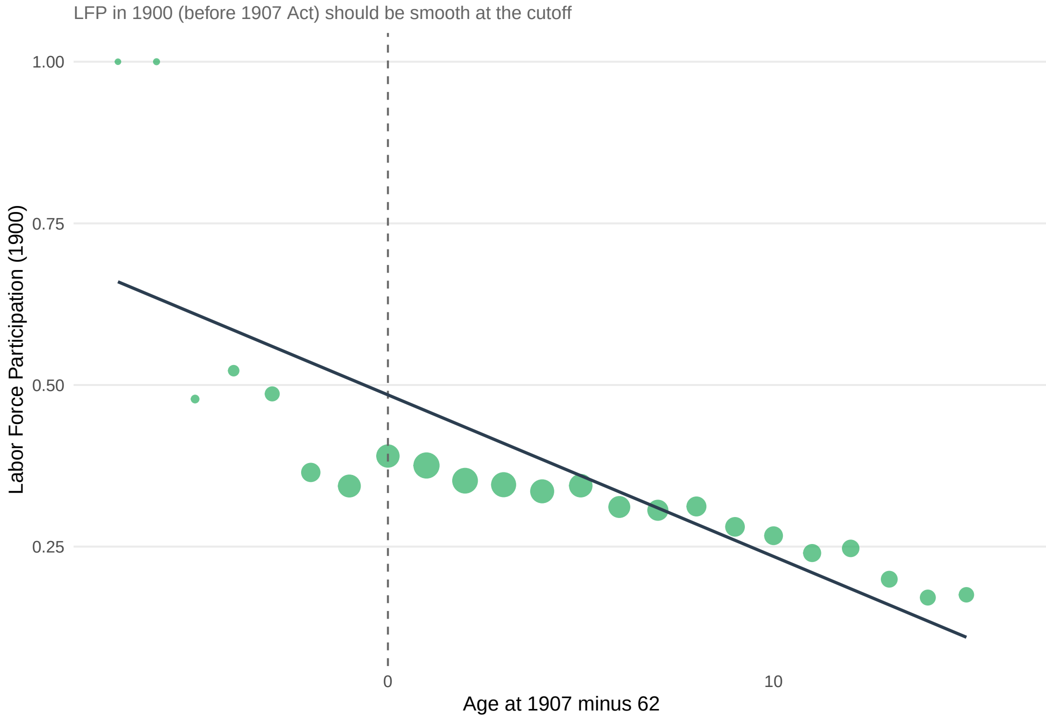

Figure 6: Pre-Treatment Falsification: LFP in 1900 at the Age-62 Threshold

Notes: Mean labor force participation rate in 1900 by age in 1907. Vertical dashed line at the age-62 threshold. Under the null, the regression should be smooth through the cutoff. Source: Costa Union Army dataset (NBER).

Three considerations temper this concern. First, the coefficient is marginally significant at the 10 percent level but not the 5 percent level. Second, the panel RDD differences out the *level* of the pre-treatment imbalance; the concern operates only through differential *trends*. Third, the covariate-adjusted panel estimate of -0.039 partially absorbs the pre-treatment difference, producing a smaller but still negative point estimate. Nevertheless, I report this result prominently because the credibility of any empirical analysis depends on honest engagement with its weaknesses.

8. Robustness

8.1 Bandwidth Sensitivity

Figure 7 displays estimates across bandwidths from 2 to 15 years. The cross-sectional estimates are noisy at narrow bandwidths but settle near zero at 4 years and above. The panel estimates are consistently negative and grow more precise at wider bandwidths, reaching conventional significance at the 5 percent level for bandwidths of 7 years and above (the estimate at BW = 5 narrowly misses with $p = 0.060$). The stability of the panel point estimate across a wide range is encouraging, but the pre-treatment falsification test passes only at narrow bandwidths (2–3 years, $p > 0.14$) and fails at wider bandwidths (4+ years, $p < 0.04$; Appendix Table 12). This is the central interpretive challenge: the bandwidth range where the panel RDD gains statistical significance is precisely the range where the pre-treatment falsification deteriorates. The most credible estimates come from narrow bandwidths where the falsification holds, but these are imprecise; the precise estimates at wider bandwidths may partly reflect pre-treatment composition differences rather than the pension treatment alone.

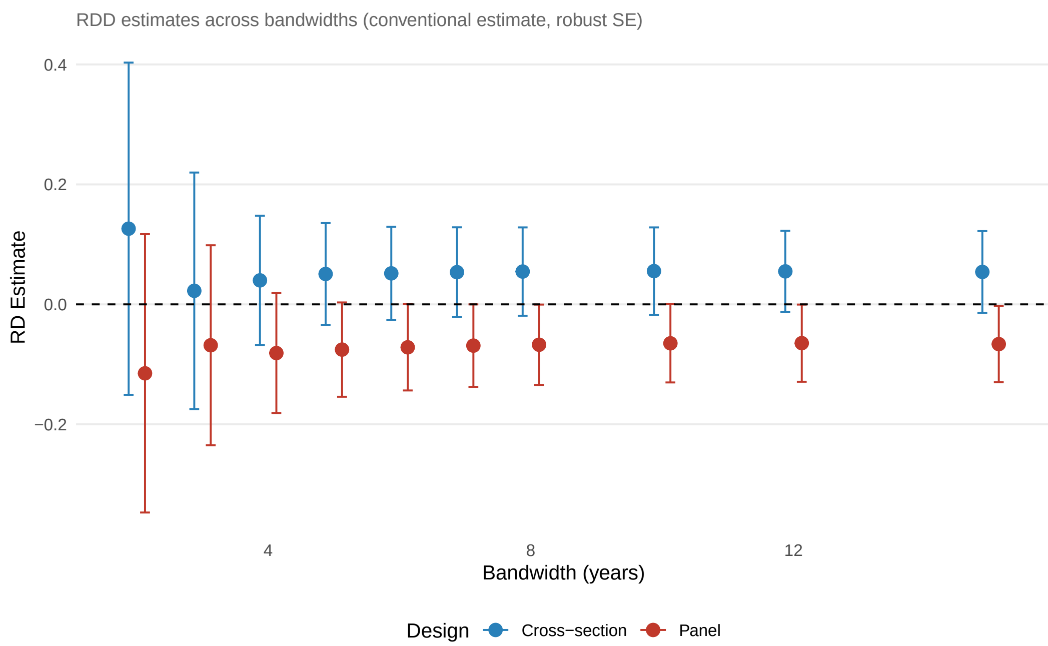

Figure 7: Bandwidth Sensitivity of Cross-Sectional and Panel RDD Estimates

Notes: RDD estimates at varying bandwidths. Panel A: cross-sectional ($Y = \text{LFP}_{1910}$). Panel B: panel ($\Delta Y = \text{LFP}_{1910} - \text{LFP}_{1900}$). Point estimates with 95% confidence intervals. Dashed line at zero.

8.2 Additional Robustness Checks

Donut-hole specifications. Excluding veterans within 1 year of the cutoff yields a cross-sectional estimate of $\hat{\tau} = 0.021$ (SE = 0.033, $p = 0.53$), consistent with the baseline near-zero result. Excluding within 2 years removes enough observations near the cutoff that the estimate becomes uninformative ($\hat{\tau} = 0.073$, SE = 0.319, $p = 0.82$)—the massive standard error reflects the very small sample remaining in the narrow bandwidth after the donut exclusion.[1]

[1] Donut-hole and health-control specifications are simple re-estimations of the main `rdrobust` model on subsets of the data. Full regression output is available from the replication code (`code/04_rdd_analysis.R`).

Health controls. Including surgeons' certificate data (condition counts, wound indicators, disability rating) as covariates in the cross-sectional RDD yields $\hat{\tau} = 0.027$ (SE $= 0.052$, $p = 0.80$), virtually unchanged from the baseline, suggesting that differential health does not confound the estimate.

Panel selection. The conventional RDD estimate for panel linkage (outcome = indicator for observation in both censuses) is -0.021 (SE $= 0.015$, $p = 0.151$), indicating no statistically significant differential attrition (Appendix Table 10). Given the 96.9 percent linkage rate and the small coefficient magnitude, the practical impact on the panel estimates is limited.

Placebo cutoffs. Figure 8 presents estimates at a series of placebo ages where no pension policy changed. The age-62 discontinuity is specific to the pension threshold rather than an artifact of the estimation procedure.

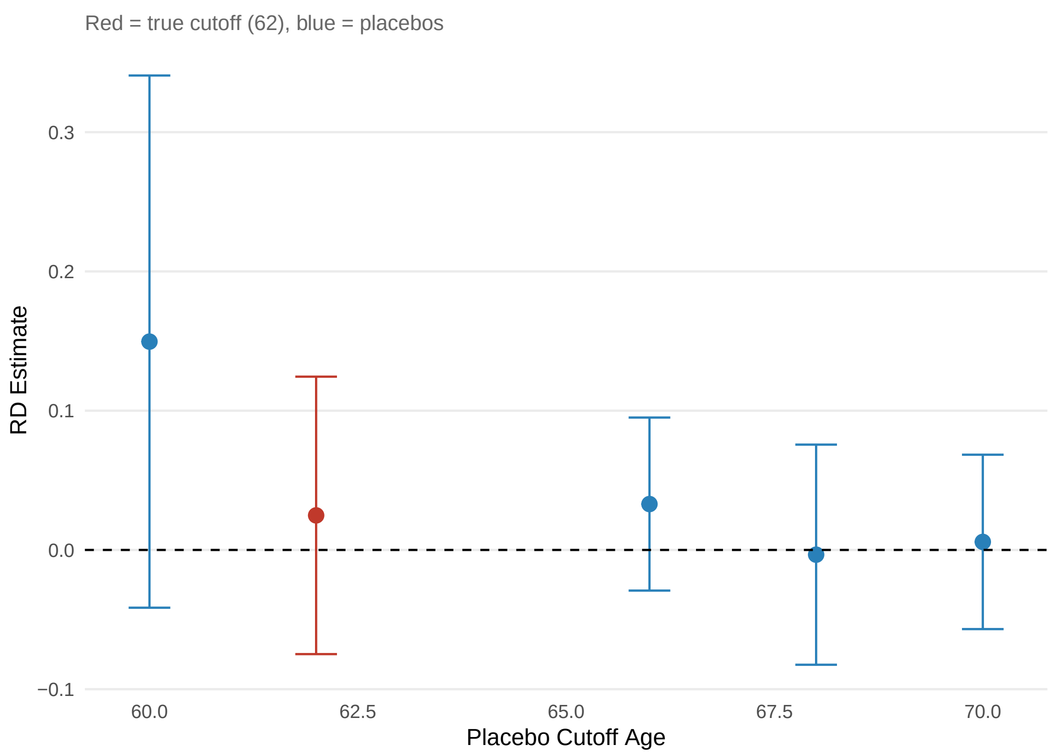

Figure 8: Placebo Cutoff Tests

Notes: Panel RDD estimates at the true cutoff (age 62, highlighted) and placebo cutoffs. Error bars show 95% confidence intervals. Source: Costa Union Army dataset (NBER).

Randomization inference. The RI p-value for the cross-sectional design is 0.045, and for the panel design < 0.001, strongly rejecting the sharp null. The RI test statistic is the simple difference in means, which captures level differences within the bandwidth rather than the local polynomial discontinuity estimated by `rdrobust`. Details and the permutation distribution appear in Appendix Tables 11 and Figure 9.

9. Extensions

9.1 Subgroup Heterogeneity

Table 6 presents cross-sectional RDD estimates for demographic subgroups. The most informative split is by prior pension status. Veterans without existing disability pensions face the largest effective first stage—for them, the age-62 threshold creates genuinely new income rather than merely upgrading an existing benefit. If the mechanism operates through income rather than certainty, this subgroup should show the largest labor supply response. Veterans in manual occupations, who face the highest physical demands, may also be more responsive—for them, retirement represents a discrete exit from work their bodies can no longer sustain.

Table 6: Subgroup Heterogeneity: RDD by Pre-Treatment Characteristics

Subgroup	Estimate	SE	p-value	N
Full sample	0.0248	(0.0508)	0.625	21302
Had pension (pre-1907)	-0.0665	(0.1044)	0.524	9713
No pension (pre-1907)	0.0171	(0.0450)	0.703	11589
Pen: No pension	0.0023	(0.0345)	0.947	10667
Pen: Already high	-0.0362	(0.1934)	0.851	4062
Literate	-0.0470	(0.0948)	0.620	10995
Illiterate	0.0232	(0.3792)	0.951	272
Native-born	0.1465	(0.0993)	0.140	14557
Foreign-born	0.0756	(0.0833)	0.365	6745
Occ 1900: Farmer	-0.1269	(0.1906)	0.506	2727
Occ 1900: Manual/Operative	-0.1032	(0.1963)	0.599	2272
Occ 1900: None/Retired	0.0107	(0.0359)	0.767	15181
Not wounded	-0.0396	(0.1122)	0.724	7586
Homeowner	-0.4579	(0.4275)	0.284	7087
Not homeowner	0.0091	(0.0349)	0.794	14215

Notes: Separate RDD estimates for each subgroup. MSE-optimal bandwidth, triangular kernel, robust standard errors.

The subgroup analysis is exploratory given the modest sample sizes within each cell, and I do not adjust for multiple comparisons. Some subgroup estimates (e.g., homeowners) are extreme in magnitude but extremely imprecise, reflecting sparse observations near the cutoff within those subgroups. Readers should interpret individual subgroup estimates with appropriate caution.

9.2 Multi-Cutoff Dose-Response

The 1907 Act created three thresholds: 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, additional discontinuities should appear at these ages, with magnitudes scaling with the pension increment.

Table 7 presents the evidence. The analysis is limited by the fact that labor force participation is already very low at ages 70 and 75 (below 10 percent), leaving little room for further decline. The pension increments at these thresholds are also smaller in dollar terms, though proportionally larger relative to existing pension amounts.

Table 7: Multi-Cutoff Dose-Response: LFP at Different Pension Thresholds

Cutoff	Pension $	Estimate	SE	p-value	95% CI	BW	N_L	N_R
$12/mo (age 62)	$12	0.0248	(0.0508)	0.625	[-0.0748, 0.1244]	2.8	2086	4899
$15/mo (age 70)	$15	0.0058	(0.0319)	0.856	[-0.0568, 0.0684]	2.7	2152	2343
$20/mo (age 75)	$20	0.0097	(0.0271)	0.721	[-0.0434, 0.0628]	3.4	1868	1565

Notes: RDD estimates at three pension amount thresholds under the 1907 Act schedule. Conventional coefficients with robust standard errors. p-values from z = estimate/SE. BW is the MSE-optimal bandwidth in years.

9.3 Health Mechanisms

Table 8 exploits the surgeons' certificate data to investigate whether the correlation between pension eligibility and labor supply operates through health rather than income. If health deteriorates discontinuously at 62—despite overlapping birth cohorts—health rather than pension income could explain the observed labor supply pattern.

Three analyses address this question. First, I estimate RDDs where health outcomes are the dependent variables; discontinuous health deterioration at 62 would complicate the income interpretation. Second, I include health controls in the main LFP regressions; if the pension effect is absorbed by health controls, the income channel is questionable. Third, I examine whether pension receipt itself affects health outcomes, testing the reverse channel documented by Eli (2015).

Table 8: Health Mechanisms: Disability and Mortality at Age-62 Threshold

Outcome	RD Estimate	SE	p-value	BW	N_L	N_R
Panel A: Pre-Treatment Health Balance						
n diagnoses pre	0.347	(0.179)	0.053	1.0	671	1725
has wound pre	0.023	(0.091)	0.799	2.1	1074	2649
cardiac pre	0.047	(0.060)	0.431	2.3	282	688
wound rating pre	0.007	(0.065)	0.915	2.6	208	617
enlist height	-0.587	(1.497)	0.695	1.0	840	2157
Panel B: Health Changes (Post minus Pre 1907)						
n diagnoses change	0.014	(0.100)	0.890	4.7	355	902
cardiac change	-0.130	(0.164)	0.429	2.5	42	75

Notes: RDD estimates of the discontinuity in health outcomes at the age-62 threshold. Health data from surgeons' certificates in the Costa Union Army dataset. Panel A tests pre-treatment balance; Panel B tests whether pension eligibility changed health trajectories. MSE-optimal bandwidth with robust standard errors.

9.4 Occupation Transitions

The panel linkage allows tracking of occupational transitions at the individual level. Table 9 examines whether pension eligibility affected the *type* of labor market exit: did veterans leave farming, manual labor, or professional occupations at different rates? If pension income enables retirement specifically from manual labor, we would expect a discontinuity in manual-to-none transitions but not in professional-to-none transitions. The occupation-specific estimates recover the responses of each group separately, providing a more complete picture than the aggregate estimates alone.

Table 9: Occupation Transitions: Labor Force Exit Rates by 1900 Occupation

1900 Occupation	Exit Rate	LFP (1910)	N
Farmer	0.509	0.491	2727
Manual/Operative	0.563	0.437	2272
Clerical/Sales/Manager	0.593	0.407	664
Farm laborer	0.696	0.304	230
Professional	0.412	0.588	228

Notes: Exit rates are the fraction of veterans who had an occupation in 1900 but no occupation in 1910. Panel sample only.

10. Discussion

10.1 Reinterpreting the Cross-Sectional Evidence

The first stage provides a lens through which to reinterpret the existing literature. Costa's (1995) influential elasticity estimates—derived from cross-sectional variation in pension generosity across disability ratings—imply that pension income was a powerful driver of retirement among Civil War veterans. The present design tests a related but distinct channel: the pure eligibility effect of an age threshold, operating through guaranteed income rather than disability-correlated generosity.

The modest reduced-form effects are consistent with two interpretations. The optimistic reading is that the true labor supply effect is approximately 7 percentage points, and the failure to achieve significance at the optimal bandwidth reflects insufficient precision rather than the absence of an effect. The minimum detectable effect of 14.3 percentage points at the optimal bandwidth—roughly twice the point estimate—confirms that the design cannot distinguish a 7 percentage point effect from zero at that bandwidth. The pessimistic reading is that the wider-bandwidth significance partly reflects pre-treatment composition differences, as the falsification test suggests.

What the data can establish is that: (1) the first stage is real and precisely estimated; (2) the reduced-form point estimate is consistently negative in the panel specification; and (3) effects larger than about 14 percentage points can be ruled out. The last point is important: Costa's cross-sectional estimates, if applied to this setting, would predict substantially larger effects than the RDD can detect. The absence of large effects at the age-62 threshold suggests that much of the cross-sectional pension-retirement correlation reflects health-income confounding rather than a causal income effect.

10.2 Age Thresholds and Pre-Existing Coverage

The attenuated first stage reveals something important about the economics of retirement programs. The 1907 Act's age threshold had modest behavioral effects not because the pension was small—$12 per month was economically meaningful—but because the age-based program was layered atop disability programs that already provided broad coverage. When most veterans already receive some form of pension, adding an age threshold primarily reclassifies existing recipients rather than bringing new individuals into the system. The marginal take-up effect—10 percentage points for the 1907 Act specifically—is the relevant behavioral margin.

This finding has direct implications for how we think about Social Security's age-62 early eligibility threshold. Proposals to raise or lower the claiming age operate in an environment where disability insurance, employer pensions, and private savings already provide substantial coverage. The historical evidence suggests that the behavioral effect of age-based thresholds depends critically on the pre-existing benefit landscape. Where Fetter and Lockwood (2018) estimate large effects of Old Age Assistance—a program that was often the *primary* source of income for elderly Americans in the 1930s—the 1907 Act's threshold operated in a setting where disability pensions already reached most of the target population. The lesson is that age thresholds matter most when they are the primary pathway to benefits; when disability programs provide broad coverage, adding an age threshold does relatively little at the margin.

10.3 Limitations

Several limitations warrant acknowledgment. The Costa dataset covers only white Union Army veterans, limiting generalizability. Costa (1998b) studies Black veterans separately, finding larger pension effects, but sample sizes are too small for the RDD approach employed here. Measurement error in the running variable, while mitigated by the use of military birth years, could attenuate the estimates. The 10 percentage point first stage limits the precision of the fuzzy RDD LATE, and the pre-treatment falsification concern, while partially addressed, remains genuine. External validity is inherently limited to male, white veterans from a specific historical cohort.

The most promising direction for future work is linkage of the Costa dataset to complete-count census microdata. The IPUMS Multigenerational Longitudinal Panel (Ruggles et al., 2024) and automated linking methods (Abramitzky et al., 2021) could dramatically expand the sample while preserving the panel structure, achieving minimum detectable effects below 5 percentage points—sufficient to detect the moderate effects suggested by the current results.

10.4 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, 62 remains the earliest age at which Americans can claim Social Security retirement benefits. The parallel is not coincidental—it reflects a deep institutional logic that age-based thresholds are administratively simple, politically defensible, and behaviorally consequential.

The evidence from 21,000 Union Army veterans complicates this logic. The first stage confirms that the age-62 threshold had real economic consequences, creating new pension income for about one in ten veterans at the margin. But the modest reduced-form effects reveal that age-based eligibility produces smaller behavioral responses when disability programs already provide broad coverage. The cross-sectional elasticities that anchor the economic history of American retirement likely overstate the causal effect of pension income by conflating it with the health gradient that determines disability pension generosity.

Behind the coefficients are men who had survived the deadliest conflict in American history, who had spent decades in physically punishing labor, and who were, by 1910, watching their bodies fail. For these men, the pension at age 62 was not an inducement to leisure. It was permission to stop.

Acknowledgements

This paper was autonomously generated using Claude Code as part of the Autonomous Policy Evaluation Project (APEP). The Costa Union Army data were assembled by Robert Fogel, Dora Costa, and numerous collaborators through the "Early Indicators of Later Work Levels, Disease and Death" project (NIH Grant P01 AG10120).

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

Contributors: @SocialCatalystLab

References

Abramitzky, Ran, Leah Boustan, Katherine Eriksson, James Feigenbaum, and Santiago Pérez, "Automated Linking of Historical Data," *Journal of Economic Literature*, 2021, *59* (3), 865–918.

Barreca, Alan I., Jason M. Lindo, and Glen R. Waddell, "Heaping-Induced Bias in Regression-Discontinuity Designs," *Economic Inquiry*, 2016, *54* (1), 268–293.

Behaghel, Luc and David M. Blau, "Framing Social Security Reform: Behavioral Responses to Changes in the Full Retirement Age," *American Economic Journal: Economic Policy*, 2012, *4* (4), 41–67.

Calonico, Sebastian, Matias D. Cattaneo, and Rocío Titiunik, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 2014, *82* (6), 2295–2326.

Cattaneo, Matias D., Brigham R. Frandsen, and Rocío Titiunik, "Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate," *Journal of Causal Inference*, 2015, *3* (1), 1–24.

_ , Michael Jansson, and Xinwei Ma, "Simple Local Polynomial Density Estimators," *Journal of the American Statistical Association*, 2020, *115* (531), 1449–1455.

_ , Nicolás Idrobo, and Rocío Titiunik, *A Practical Introduction to Regression Discontinuity Designs: Foundations*, Cambridge University Press, 2020.

Cesarini, David, Erik Lindqvist, Matthew J. Notowidigdo, and Robert Ostling, "The Effect of Wealth on Individual and Household Labor Supply: Evidence from Swedish Lotteries," *American Economic Review*, 2017, *107* (12), 3917–3946.

Coile, Courtney and Jonathan Gruber, "Future Social Security Entitlements and the Retirement Decision," *Review of Economics and Statistics*, 2007, *89* (2), 234–246.

Costa, Dora L., "Pensions and Retirement: Evidence from Union Army Veterans," *Quarterly Journal of Economics*, 1995, *110* (2), 297–319.

_ , "Displacing the Family: Union Army Pensions and Elderly Living Arrangements," *Journal of Political Economy*, 1997, *105* (6), 1269–1292.

_ , "The Evolution of Retirement: An American Economic History, 1880–1990," *University of Chicago Press*, 1998.

— , "Pensions and Retirement among Black Union Army Veterans," in David A. Wise, ed., *The Economics of Aging*, University of Chicago Press, 1998, pp. 175–206.

Eli, Shari, "Income Effects on Health: Evidence from Union Army Pensions," *Journal of Economic History*, 2015, *75* (2), 448–478.

Fetter, Daniel K. and Lee M. Lockwood, "Government Old-Age Support and Labor Supply: Evidence from the Old Age Assistance Program," *American Economic Review*, 2018, *108* (8), 2174–2211.

Fogel, Robert W., "New Sources and New Techniques for the Study of Secular Trends in Nutritional Status, Health, Mortality, and the Process of Aging," *Historical Methods*, 1993, *26* (1), 5–43.

Gelman, Andrew and Guido Imbens, "Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs," *Journal of Business and Economic Statistics*, 2019, *37* (3), 447–456.

Glasson, William Henry, "Federal Military Pensions in the United States," *Oxford University Press*, 1918.

Grembi, Veronica, Tommaso Nannicini, and Ugo Troiano, "Do Fiscal Rules Matter?," *American Economic Journal: Applied Economics*, 2016, *8* (3), 1–30.

Gruber, Jonathan and David A. Wise, *Social Security and Retirement around the World*, University of Chicago Press, 1999.

Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw, "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 2001, *69* (1), 201–209.

Imbens, Guido and Karthik Kalyanaraman, "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," *Review of Economic Studies*, 2012, *79* (3), 933–959.

Imbens, Guido W. and Thomas Lemieux, "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics*, 2008, *142* (2), 615–635.

— , Donald B. Rubin, and Bruce I. Sacerdote, "Estimating the Effect of Unearned Income on Labor Earnings, Savings, and Consumption: Evidence from a Survey of Lottery Players," *American Economic Review*, 2001, *91* (4), 778–794.

Lee, David S. and David Card, "Regression Discontinuity Inference with Specification Error," *Journal of Econometrics*, 2008, *142* (2), 655–674.

— **and Thomas Lemieux**, "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 2010, *48* (2), 281–355.

Manoli, Dayanand S. and Andrea Weber, "Nonparametric Evidence on the Effects of Financial Incentives on Retirement Decisions," *American Economic Journal: Economic Policy*, 2016, *8* (4), 160–182.

Mastrobuoni, Giovanni, "Labor Supply Effects of the Recent Social Security Benefit Cuts: Empirical Estimates Using Cohort Discontinuities," *Journal of Public Economics*, 2009, *93* (11-12), 1224–1233.

Moffitt, Robert A., "Welfare Programs and Labor Supply," *Handbook of Public Economics*, 2002, *4*, 2393–2430.

Ruggles, Steven, Catherine Fitch, Ronald Goeken, Josiah Grover, J. David Hacker, Matt Nelson, Evan Roberts, Megan Schouweiler, and Matthew Sobek, "IPUMS Multigenerational Longitudinal Panel: Version 1.1," https://doi.org/10.18128/D016.V1.1 2024.

Skocpol, Theda, *Protecting Soldiers and Mothers: The Political Origins of Social Policy in the United States*, Harvard University Press, 1992.

— , "America's First Social Security System: The Expansion of Benefits for Civil War Veterans," *Political Science Quarterly*, 1993, *108* (1), 85–116.

van Dalen, Hendrik P. and Kene Henkens, "When the Game is Over: A Review of Economic Literature on Old-Age Retirement," *European Journal of Political Economy*, 2010, *26* (2), 153–163.

Vitek, Francis, "The Effect of Pensions on Individual Retirement Timing: Evidence from the Union Army," *Explorations in Economic History*, 2022, *84*, 101442.

A. Data Appendix

A.1 Dataset Description

The Union Army dataset is archived at the NBER as part of the "Early Indicators of Later Work Levels, Disease and Death" project. The dataset comprises multiple files: Basics (demographics and military service), Pension (pension histories), Disease (surgeons' certificates), and Socioeconomic (census-linked records from 1850–1910). Each veteran is identified by a unique record number (`recidnum`). Full documentation is available from the NBER archive.

A.2 Variable Construction

- **Age in 1907:** 1907 − birth year, from military service records.
- **Labor force participation:** Binary indicator from census occupation field (1 if gainful occupation reported).
- **ΔLFP:** $\text{LFP}_{1910} - \text{LFP}_{1900}$.
- **Under 1907 Act:** Binary indicator for pension receipt under the 1907 Act.
- **Any pension:** Binary indicator for receiving any federal pension.
- **Monthly pension ($):** Monthly amount from pension records.
- **Covariates:** Literacy, nativity, wound status (military records), condition counts (surgeons' certificates).

A.3 Sample Restrictions

Starting from all records with a valid `recidnum`, I merge military, pension, and census records; restrict to non-missing birth year and 1910 census data; compute age in 1907 and restrict to ages 45–90; and for the panel sample, further restrict to non-missing 1900 census data.

B. Identification Appendix

B.1 Density Test Details

The density test uses the local polynomial estimator of Cattaneo et al. (2020a). The test yields $p = 0.756$, strongly failing to reject continuity. Military birth years, recorded at enlistment decades before the 1907 Act, eliminate age heaping concerns that would afflict census-reported ages.

B.2 Panel Selection Test

Table 10: Panel Selection: Probability of Appearing in Both Censuses

	Estimate	SE	p-value	95% CI	BW	N_L	N_R
Conventional	-0.0211	(0.0147)	0.151	[-0.0499, 0.0077]	2.4	2088	4910
Bias-corrected	-0.0467	(0.0227)	0.040	[-0.0912, -0.0022]	2.4	2088	4910

Notes: RDD estimate of the discontinuity in the probability of appearing in both the 1900 and 1910 censuses at the age-62 threshold. Sample restricted to veterans alive at 1910. Total $N = N_L + N_R$ (observations within the MSE-optimal bandwidth). "Conventional" reports the local polynomial estimate with conventional standard error; "Bias-corrected" reports the rdrobust bias-corrected estimate with robust standard error. Both rows use the same bandwidth and sample.

Table 10 tests whether the probability of successful panel linkage differs at the age-62 cutoff. The conventional RDD estimate is -0.021 ($p = 0.151$). The bias-corrected estimate reaches marginal significance ($p = 0.040$), but the small magnitude (2.1 percentage points on a base of 96.9 percent) suggests minimal practical impact.

B.3 Randomization Inference Details

The RI procedure follows Cattaneo et al. (2015): restrict to the MSE-optimal bandwidth, compute the observed difference in means, permute the above/below-62 assignment 5,000 times, and compute the fraction of permuted statistics exceeding the observed statistic. The RI p-value for the cross-sectional design is 0.045; for the panel design, < 0.001.

Table 11: Randomization Inference Results

Design	Diff-in-Means	RI p-value	Permutations
Cross-section (LFP 1910)	-0.0225	0.045	5000
Panel (ΔLFP)	-0.0489	< 0.001	5000

Notes: Two-sided randomization inference p-values from 5,000 permutations of the above/below-62 assignment within the bandwidth. Test statistic is the simple difference in means within the bandwidth (not the rdrobust local polynomial estimate), so magnitudes differ from the main RDD estimates in Table 5. The null hypothesis is no treatment effect at the threshold.

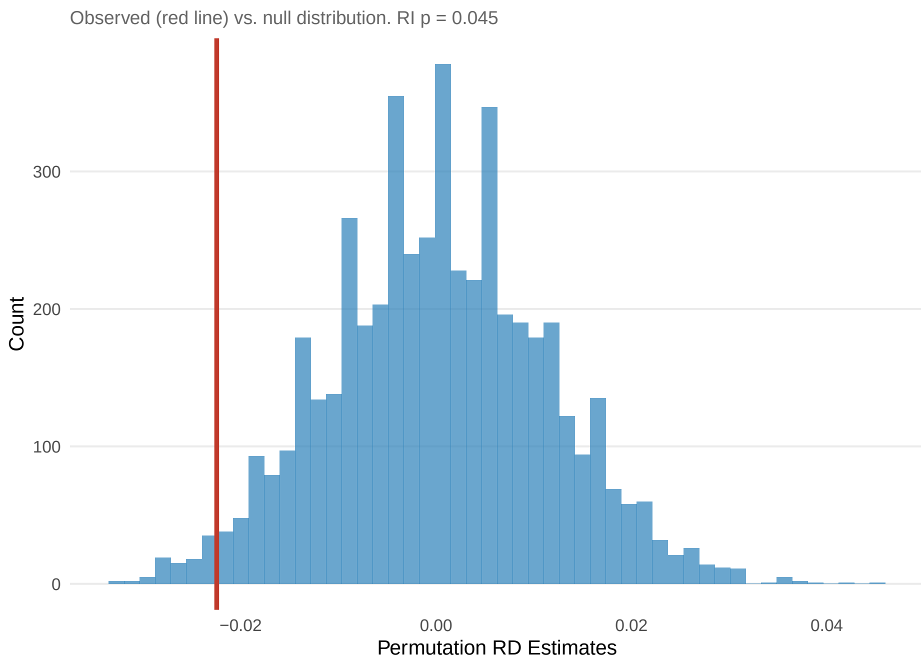

Figure 9: Randomization Inference: Permutation Distribution

Notes: Distribution of test statistics under the sharp null, from 5,000 permutations. Red line marks the observed statistic. Source: Costa Union Army dataset (NBER).

B.4 Age Heaping Analysis

Age heaping in historical census data creates systematic measurement error (Barreca et al., 2016). In the Costa dataset, the running variable is computed from military birth year rather than census-reported age, substantially mitigating this concern. The donut-hole specifications address any residual heaping concern.

B.5 Pre-Treatment Falsification Across Bandwidths

Table 12 reveals that the falsification test passes at narrow bandwidths (2–3 years, $p > 0.14$) but fails at wider bandwidths (4+ years, $p < 0.04$). This pattern suggests that bandwidth sensitivity of the panel RDD is at least partially driven by pre-treatment composition differences.

Table 12: Pre-Treatment Falsification: LFP (1900) Across Bandwidths

BW	Estimate	SE	p-value	N_L	N_R
2	0.2330	(0.1591)	0.143	1276	3026
3	0.1069	(0.1138)	0.347	1992	4691
4	0.1294	(0.0616)	0.036	2291	6247
5	0.1308	(0.0481)	0.007	2405	7627
6	0.1269	(0.0442)	0.004	2425	8937
7	0.1256	(0.0424)	0.003	2428	10010
8	0.1250	(0.0415)	0.003	2428	10970
10	0.1235	(0.0407)	0.002	2428	12564
12	0.1236	(0.0404)	0.002	2428	13850
15	0.1250	(0.0401)	0.002	2428	15280

Notes: RDD estimates of the discontinuity in LFP at 1900 (pre-treatment) at the age-62-in-1907 threshold across bandwidths. Conventional coefficients with robust standard errors. p-values from z = estimate/SE. A non-significant estimate supports the identifying assumption; a significant positive estimate suggests composition differences across the cutoff.

Table 13: Fuzzy RDD: Local Average Treatment Effect of Pension on LFP

	LATE	SE	p-value	BW	N_L	N_R
Pension receipt $\rightarrow$ LFP	0.3392	(0.4594)	0.460	2.4	2055	4827
Pension \$ $\rightarrow$ LFP (per \$)	0.0143	(0.0274)	0.600	2.8	2086	4899

Notes: Fuzzy RDD estimates using rdrobust. Treatment is instrumented by the age-62 threshold. LATE represents the effect of pension receipt (or pension dollars) on LFP among compliers at the threshold. MSE-optimal bandwidth with triangular kernel.

C. Additional Tables and Figures

C.1 Fuzzy RDD Estimates

Table 13 reports the fuzzy RDD LATE of 0.339 (SE = 0.459, $p = 0.460$). The positive sign reflects the positive cross-sectional reduced-form estimate (0.025) divided by the positive first stage (0.102)—note that the cross-sectional reduced form is positive (though insignificant), while the panel reduced form is negative. The fuzzy RDD uses the cross-sectional specification, yielding $0.025/0.102 \approx 0.25$, with the discrepancy from 0.339 arising from bandwidth differences. The massive standard error is mechanical: with a 10 percentage point first stage, the LATE standard error is approximately ten times the reduced-form standard error. The LATE is uninformative in this setting.

C.2 Bandwidth Robustness Grids

Table 14: Robustness: Cross-Sectional RDD Across Bandwidths

BW	Estimate	SE	p-value	N_L	N_R
2	0.1262	(0.1413)	0.372	1335	3158
3	0.0226	(0.1006)	0.822	2086	4899
4	0.0400	(0.0550)	0.467	2409	6502
5	0.0507	(0.0433)	0.242	2528	7925
6	0.0517	(0.0396)	0.192	2550	9276
7	0.0537	(0.0382)	0.160	2553	10382
8	0.0547	(0.0376)	0.145	2553	11367
10	0.0555	(0.0372)	0.136	2553	13014
12	0.0550	(0.0345)	0.111	2554	14320
15	0.0541	(0.0347)	0.119	2554	15782

Notes: Cross-sectional RDD estimates with varying bandwidths. Local linear polynomial with triangular kernel. Conventional point estimates with robust standard errors (rdrobust). p-values computed from the conventional z-statistic.

Table 14 reports cross-sectional estimates across bandwidths from 2 to 15 years, confirming the null result in the cross-section.

Table 15 reports panel RDD estimates across the same bandwidth grid. The point estimates are stable between -0.065 and -0.115 across all bandwidths, with conventional significance at the 5 percent level emerging at bandwidths of 7 years and above ($p = 0.050$ at BW = 7; $p = 0.049$ at BW = 8). The estimate at BW = 5 narrowly misses conventional significance ($p = 0.060$).

Table 15: Robustness: Panel RDD Across Bandwidths

BW	Estimate	SE	p-value	N_L	N_R
2	−0.1150	(0.1184)	0.331	1276	3026
3	−0.0682	(0.0851)	0.423	1992	4691
4	−0.0812	(0.0510)	0.111	2291	6247
5	−0.0754	(0.0401)	0.060	2405	7627
6	−0.0716	(0.0367)	0.051	2425	8937
7	−0.0688	(0.0351)	0.050	2428	10010
8	−0.0673	(0.0342)	0.049	2428	10970
10	−0.0650	(0.0333)	0.051	2428	12564
12	−0.0647	(0.0328)	0.049	2428	13850
15	−0.0663	(0.0324)	0.041	2428	15280

Notes: Panel RDD estimates ($\Delta Y = \text{LFP}_{1910} - \text{LFP}_{1900}$) with varying bandwidths. Local linear polynomial with triangular kernel. Conventional point estimates with robust standard errors (`rdrobust`). p-values from the conventional z-statistic (estimate/SE). N_L and N_R are observations within the bandwidth on each side of the cutoff.

C.3 Subgroup Heterogeneity Figure

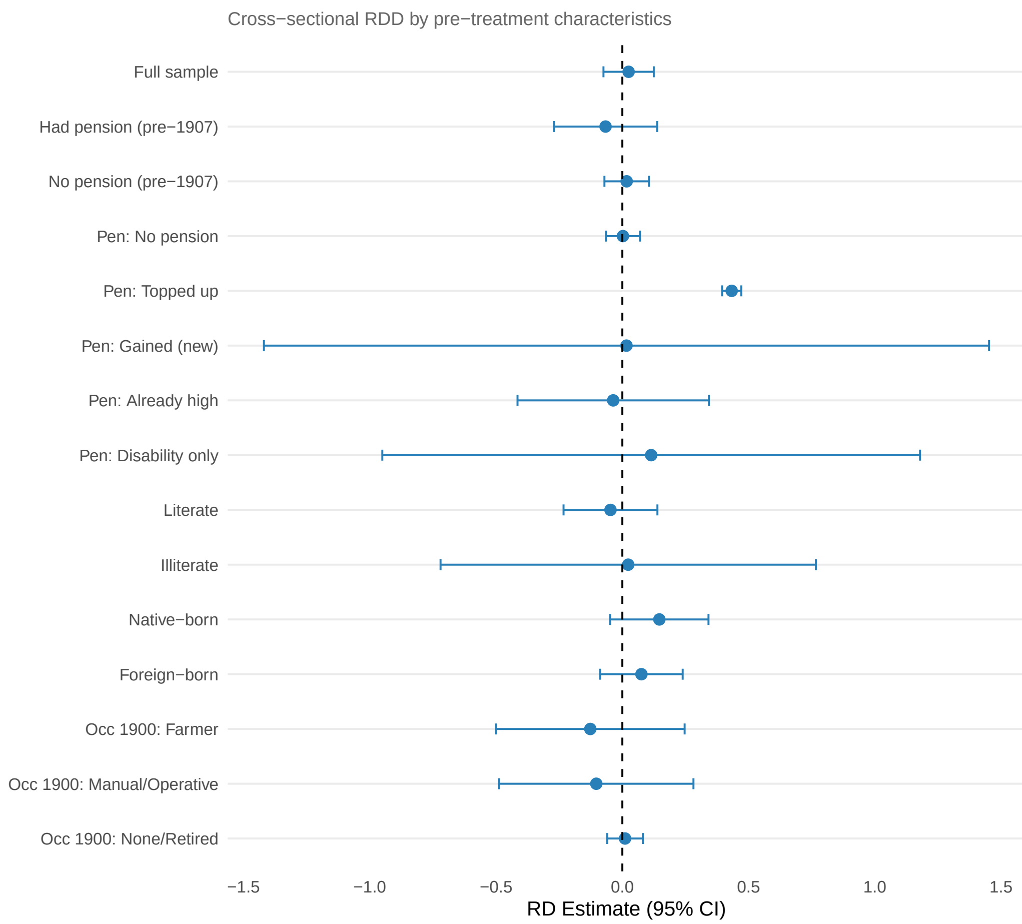

Figure 10: Subgroup Heterogeneity: Cross-Sectional RDD Estimates at Age 62

Notes: Separate cross-sectional RDD estimates for each demographic subgroup. Point estimates and 95% confidence intervals. Source: Costa Union Army dataset (NBER).

C.4 Multi-Cutoff Dose-Response Figure

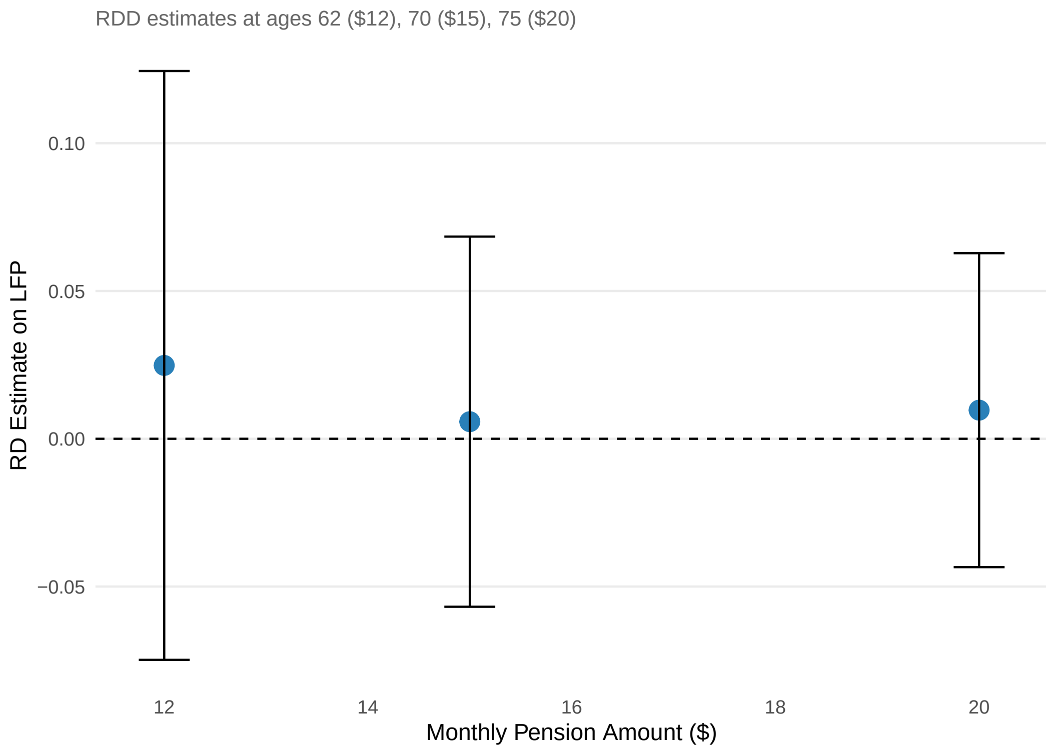

Figure 11: Multi-Cutoff Dose-Response: Pension Thresholds at 62, 70, and 75

Notes: RDD estimates at each pension threshold. Age 62: \$0 to \$12/month (for new recipients). Age 70: \$12 to \$15/month. Age 75: \$15 to \$20/month. Source: Costa Union Army dataset (NBER).

C.5 Occupation Exit Figure

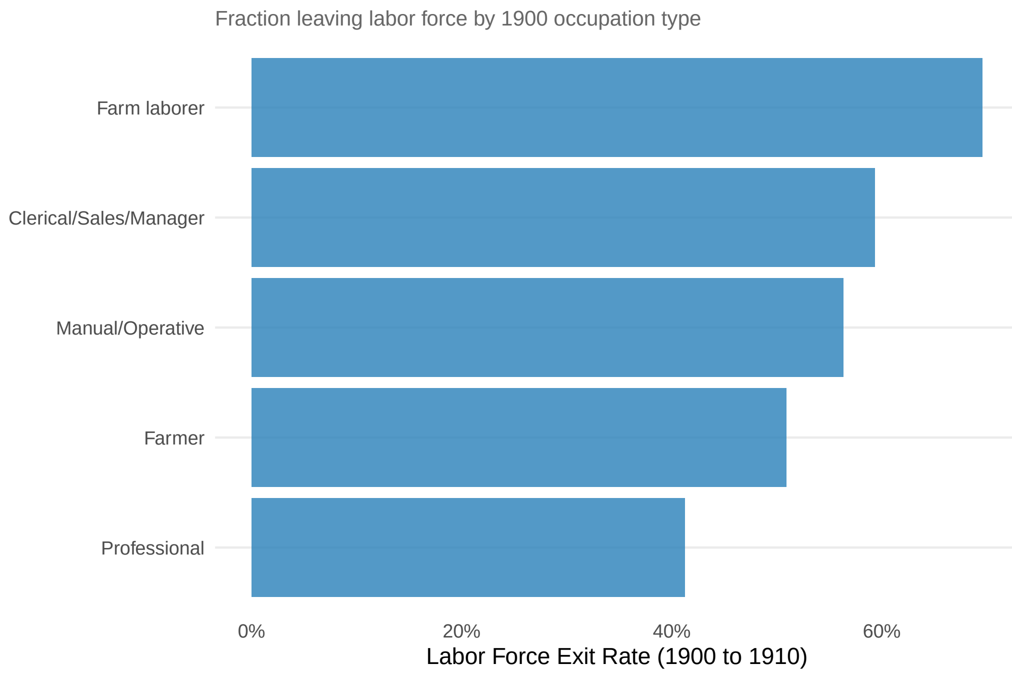

Figure 12: Occupation Exit Rates at the Age-62 Threshold

Notes: Occupation-specific exit rates (fraction leaving occupation between 1900 and 1910) by age in 1907. Vertical dashed line at age 62. Source: Costa Union Army dataset (NBER).

C.6 Pension Amount by Age

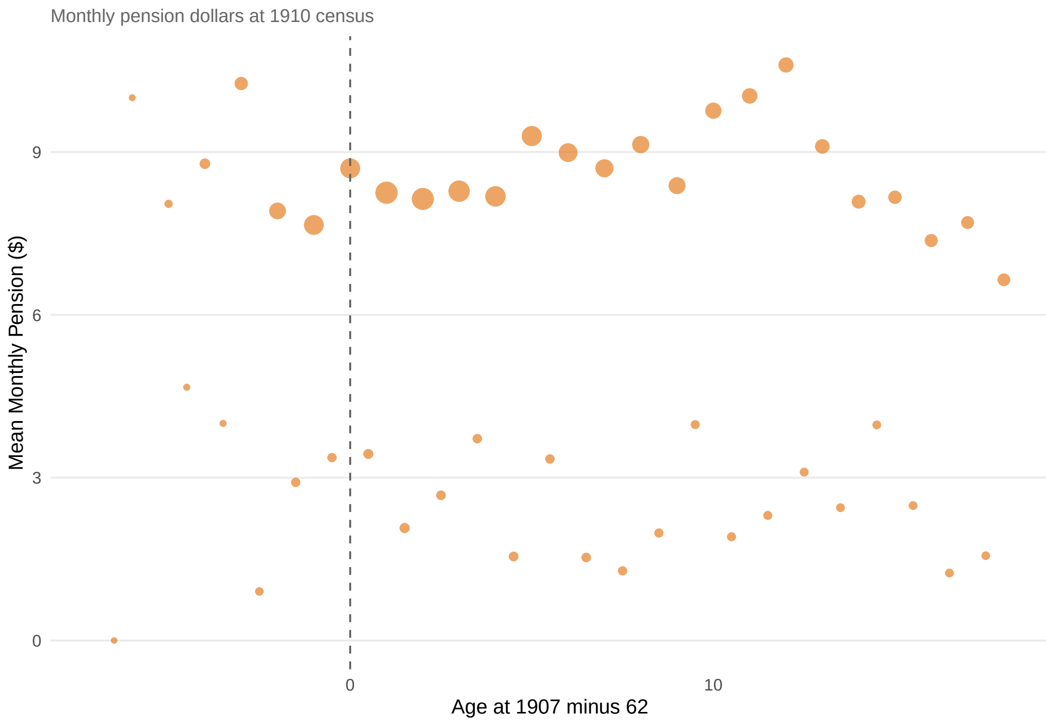

Figure 13: Monthly Pension Amount by Age

Notes: Mean monthly pension amount in dollars by age in 1907. Vertical dashed lines at the pension schedule thresholds (62, 70, 75). Source: Costa Union Army pension records (NBER).