The Motherhood Gap Across the Lifecycle and Income Distribution:
A Quantile Regression Analysis

Afrouz Azadikhah Jahromi[*]   Alina Malkova[†]

February 12, 2026

Abstract

We document how the motherhood gap in income varies across both the lifecycle and the income distribution. Using quantile regression methods on data from the NLSY79, CPS, and HRS for women born 1957–1964, we find that the relationship between motherhood and income reverses sign across the income distribution, with this reversal point shifting over the lifecycle. At ages 54–61, childless women in the bottom 40% of the *individual* pension income distribution face penalties of 8–15%, while those in the top quartile enjoy premiums exceeding 12%. This reversal—occurring around the 40th percentile—reveals that children's economic value in retirement depends fundamentally on household resources. Using the Chernozhukov, Fernández-Val, and Melly (2013) decomposition, we show that unexplained structural factors (insurance mechanisms) dominate at low quantiles, while observable characteristics (career effects) dominate at high quantiles. These patterns are particularly pronounced for women, who face both stronger caregiving expectations and larger career penalties. Coarsened exact matching, individual fixed effects, and Oster (2019) bounds confirm the robustness of the reversal pattern. Our findings challenge universal approaches to pension policy and suggest that support for childless elderly should be targeted to vulnerable populations.

[*]School of Business Administration, Widener University. Email: ajahromi@widener.edu
[†]N.Bisk School of Business, Florida Institute of Technology. Email: amalkova@fit.edu

1 Introduction

Why do poor childless retirees face substantial pension penalties while wealthy childless retirees enjoy pension premiums? This puzzle challenges conventional wisdom about the economic consequences of fertility decisions and reveals fundamental heterogeneity in how family structure affects retirement security across the income distribution.

The extensive literature on the "motherhood penalty" documents wage gaps of 5–15% per child during working years (Waldfogel 1997; Budig and England 2001; Anderson, Binder, and Krause 2003). These penalties accumulate over the lifecycle, resulting in retirement income gaps (Even and Macpherson 1994; Johnson and Uccello 2004). Yet this literature has largely overlooked two crucial dimensions of heterogeneity: how effects vary across the *income distribution* at any given age, and how these distributional patterns evolve over the *lifecycle*.

We address both dimensions using quantile regression methods and three complementary data sources covering women born 1957–1964 from early career through retirement. Our central finding is a striking reversal in the relationship between childlessness and retirement income that occurs around the 40th percentile of the income distribution. While childless women in the bottom 40% face penalties of 8–15%, those in the top quartile enjoy premiums exceeding 12%. This reversal reveals that children's economic function in retirement is not uniform but depends critically on households' position in the income distribution.

This distributional heterogeneity stands in sharp contrast to the existing literature's focus on average effects. A 10% average motherhood penalty does not translate uniformly into 10% lower pensions regardless of economic status—it masks fundamentally different effects at different points in the distribution. Our analysis reconciles two seemingly contradictory bodies of literature. The motherhood penalty literature documents career costs of children (Kleven, Landais, and Søgaard 2019; Anderson, Binder, and Krause 2003). The old-age security literature emphasizes children's insurance value (Ebenstein and Leung 2010). Our results demonstrate that both perspectives are correct—but at different points in the income distribution. Below the 40th percentile, the insurance value dominates; above it, career costs prevail.

We make three contributions. First, we document how the motherhood gap varies across the income distribution using quantile regression, revealing a reversal that mean comparisons completely obscure. Second, we trace how this distributional pattern evolves over the lifecycle, showing that the reversal point shifts as the cohort ages. Third, we decompose the gap into explained (characteristics) and unexplained (structural) components using the Chernozhukov, Fernández-Val, and Melly (2013) method, providing evidence that different mechanisms operate at different quantiles.

A critical methodological contribution addresses the household versus individual income problem that confounds much of the existing literature. The HRS measures *household* income, which masks individual-level penalties because mothers are more likely to be married (creating two-earner households). We show that running quantile regressions on HRS household income produces misleading results, while individual pension income from NLSY79 reveals the true distributional pattern. This distinction is essential for both research and policy.

Our findings have immediate policy relevance. Universal pension compensation for childbearing—common in European systems—would primarily benefit those who least need support (high-income parents who already benefit from childlessness through enhanced careers) while failing to address the vulnerabilities of those most at risk (low-income childless retirees who lack family insurance networks).

2 Theoretical Framework

We develop a model in which the relationship between childlessness and retirement income reflects the net effect of two opposing economic forces, whose relative importance varies across the income distribution.

2.1 Setup

Consider an individual i with pre-retirement income y_i drawn from distribution $F(y)$ who makes a binary fertility choice $d_i \in \{0, 1\}$ where $d_i = 0$ denotes having children and $d_i = 1$

denotes childlessness. Retirement income R_i consists of two components:

$$R_i(d_i, y_i) = S_i(d_i, y_i) + T_i(d_i, y_i) \tag{1}$$

where $S_i(d_i, y_i)$ represents accumulated retirement savings from lifetime earnings and $T_i(d_i, y_i)$ represents transfers (both formal and informal) received in retirement.

2.2 The Career Impediment Mechanism

Following Kleven, Landais, and Søgaard (2019), children reduce human capital accumulation through career interruptions. Let the human capital production function be:

$$H(d_i, y_i) = y_i \cdot \exp[-\phi(y_i) \cdot (1 - d_i)] \tag{2}$$

where $\phi(y_i)$ represents human capital depreciation from childrearing.

Assumption 1 (Increasing Opportunity Cost). $\phi'(y) > 0$ and $\phi''(y) \geq 0$

This reflects that high earners face larger opportunity costs from career interruptions. The savings gain from childlessness is:

$$\Delta S(y_i) \approx s \cdot y_i \cdot \phi(y_i) \equiv \alpha(y_i) \cdot y_i \tag{3}$$

where $\alpha(y_i) = s \cdot \phi(y_i)$ is increasing in income.

2.3 The Insurance Mechanism

Following Ebenstein and Leung (2010), children provide old-age insurance through informal support. The transfer function for parents is:

$$T_i(0, y_i) = \beta_0 + \beta_1 \cdot \exp(-\lambda y_i) \tag{4}$$

where $\beta_0 \geq 0$ represents baseline transfers, $\beta_1 > 0$ captures income-dependent transfers, and $\lambda > 0$ determines the rate at which transfers decline with income.

For childless individuals:

$$T_i(1, y_i) = \beta_0 \tag{5}$$

The insurance value of children is:

$$\Delta T(y_i) = \beta_1 \cdot \exp(-\lambda y_i) \equiv \beta(y_i) \tag{6}$$

which is decreasing in income.

2.4 Net Effect and Reversal

The total effect of childlessness on retirement income is:

$$\Delta R(y_i) = \underbrace{\alpha(y_i) \cdot y_i}_{\text{Career gain}} - \underbrace{\beta(y_i)}_{\text{Insurance loss}} \tag{7}$$

Proposition 1 (Existence of Reversal Point)**.** *Under Assumption 1 and given* $\lim_{y \to 0} \alpha(y) \cdot y = 0$ *and* $\lim_{y \to 0} \beta(y) = \beta_1 > 0$, *there exists a unique threshold income* $y^* > 0$ *such that:*

$$\Delta R(y^*) = 0 \Leftrightarrow \alpha(y^*) \cdot y^* = \beta(y^*) \tag{8}$$

This proposition provides our key testable prediction: the effect of childlessness switches from negative (penalty) to positive (premium) at a unique income threshold. **Note on functional form:** The uniqueness of the reversal point depends on our exponential specification for the insurance mechanism. Alternative functional forms could generate multiple crossing points. Our empirical test for a single crossing (Section 5.6) provides independent evidence supporting uniqueness rather than assuming it.

2.5 Quantile Treatment Effects

In our quantile regression framework, where Q_τ denotes the τ-th quantile:

$$Q_\tau(\ln R_i | X_i, d_i) = \gamma_\tau + \delta_\tau \cdot d_i + X_i' \theta_\tau \tag{9}$$

The model predicts:

Proposition 2 (Quantile Treatment Effects). *Let τ^* denote the quantile corresponding to income y^*. Then:*

1. $\delta_\tau < 0$ *for $\tau < \tau^*$ (childlessness penalty)*

2. $\delta_\tau = 0$ *for $\tau = \tau^*$ (reversal point)*

3. $\delta_\tau > 0$ *for $\tau > \tau^*$ (childlessness premium)*

4. $\partial \delta_\tau / \partial \tau > 0$ *(monotonically increasing)*

2.6 Decomposition Predictions

Following Chernozhukov, Fernández-Val, and Melly (2013), we decompose the quantile gap into composition (Δ^X) and structure (Δ^S) effects:

Proposition 3 (Mechanism Identification). *1. At low quantiles ($\tau < \tau^*$): $|\Delta^S| > |\Delta^X|$ because insurance operates through unobserved networks*

2. *At high quantiles ($\tau > \tau^*$): $|\Delta^X| > |\Delta^S|$ because career effects are captured by observable human capital*

3 Data

We use three datasets to construct a lifecycle picture of the motherhood gap for women born 1957–1964.

3.1 NLSY79

The National Longitudinal Survey of Youth 1979 follows 6,283 women from ages 14–22 in 1979 through 2018, when they are ages 54–61. Key advantages include:

- **Fertility measurement:** NUMKID measures total children ever born—the gold standard that avoids misclassification.

- **Individual retirement income:** Pension receipt and amounts are measured at the individual level, not confounded with spouse income.

- **Pre-birth characteristics:** AFQT scores, family background enable assessment of selection.

In our sample, 78.7% of women are mothers (N=4,946) and 21.3% are childless (N=1,337).

3.2 RAND HRS

The Health and Retirement Study tracks individuals aged 50+ from 1992–2022. We use HRS for two purposes: (1) to illustrate the household income measurement problem, and (2) to examine how distributional patterns evolve by age within the panel.

Critical limitation: HRS measures *household* income (H#ITOT), which includes spouse earnings. Since mothers are more likely to be married, two-earner households appear in the "mother" category, potentially masking individual-level gaps.

3.3 CPS ASEC

The Current Population Survey provides large cross-sectional samples for years 1990–2014. We use individual total income (INCTOT) for the 1957–1964 birth cohort at ages 35–50.

Limitation: The CPS identifies mothers by co-resident children, which systematically misclassifies some mothers as childless, likely biasing gaps upward at ages 35–40 and creating spurious attenuation at ages 45–50 as children leave home.

3.4 Sample Construction

Table 1 documents sample sizes across analyses.

Table 1: Sample Construction

Analysis	Source	N Mothers	N Childless
Early career (ages 20–35)	NLSY79	47,214	10,835
Mid-career (ages 35–50)	CPS	128,212	42,221
Retirement (ages 54–61)	NLSY79	4,946	1,337
With pension income	NLSY79	624	182
Retirement (ages 50–67)	HRS	3,715	346

4 Methodology

4.1 Quantile Treatment Effects

We estimate quantile treatment effects using a two-step approach. In the first step, we regress the treatment (childlessness) on control variables using OLS and obtain residuals:

$$\text{Childless}_i = X_i'\pi + \nu_i \tag{10}$$

In the second step, we estimate conditional quantile regressions using the residualized treatment:

$$Q_\tau(\ln Y_i|X_i) = \alpha_\tau + \beta_\tau \widehat{\text{Childless}}_i + \gamma_\tau X_i + \epsilon_{i,\tau} \tag{11}$$

where $\widehat{\text{Childless}}_i$ is the residual from the first stage.

This approach, following Chernozhukov, Fernández-Val, and Melly (2013), decomposes the variance of the treatment into a piece explained by observables and a residual piece orthogonal to controls.

Methodological note on first-stage specification: The OLS first stage assumes linear selection into childlessness. Given evidence of U-shaped childlessness patterns across income (Baudin, Croix, and Gobbi 2015), a probit or more flexible specification might better capture nonlinear selection. We present OLS results as our baseline and explore robustness to alternative first-stage specifications in the appendix.

4.2 Counterfactual Decomposition

Following Chernozhukov, Fernández-Val, and Melly (2013), we construct counterfactual distributions to decompose the income gap between childless women and mothers. Let $F_{Y(1|1)}$ and $F_{Y(0|0)}$ represent the observed distributions for childless and mothers, respectively. The counterfactual distribution $F_{Y(0|1)}$ represents what childless women's income distribution would have been if they faced mothers' income structure:

$$F_{Y(0|1)}(y) := \int_{\mathcal{X}_1} F_{Y_0|X_0}(y|x) dF_{X_1}(x) \tag{12}$$

The total gap decomposes as:

$$F_{\Delta_{ft}|X_f} - F_{\Delta_{mt}|X_m} = \underbrace{[F_{\Delta_{ft}|X_f} - F_{\Delta_{mt}|X_f}]}_{\text{Structure Effect}} + \underbrace{[F_{\Delta_{mt}|X_f} - F_{\Delta_{mt}|X_m}]}_{\text{Composition Effect}} \tag{13}$$

4.3 Robustness Approaches

4.3.1 Individual Fixed Effects

We exploit the panel structure of HRS to implement correlated random effects quantile regression following Arellano and Bonhomme (2016):

$$Q_\tau(\ln Y_{it}|X_{it}, \alpha_i) = \alpha_{i,\tau} + \beta_\tau \cdot \text{Childless}_i + \gamma_\tau X_{it} + \epsilon_{it,\tau} \tag{14}$$

where $\alpha_{i,\tau}$ captures individual-specific effects that may vary across quantiles.

4.3.2 Coarsened Exact Matching

We implement CEM on pre-fertility characteristics: education, birth cohort, race, gender, and ever-married status. CEM proceeds by coarsening continuous variables into strata, then exactly matching treated and control units within each stratum. We assign weights:

$$w_i = \frac{n_s}{n_{s,D_i}} \cdot \frac{N_{D_i}}{N} \tag{15}$$

where n_s denotes stratum size and n_{s,D_i} represents units in stratum s with treatment status D_i.

Caveat on "ever married": Matching on ever-married status effectively conditions on a variable that may be affected by fertility decisions. If childlessness causally affects marriage probability, this could bias estimates. We present results both with and without this matching variable.

4.3.3   Oster Bounds

Following Oster (2019), we assess sensitivity to selection on unobservables. The key parameter δ represents the ratio of selection on unobservables to selection on observables. Values of $\delta > 1$ suggest moderate robustness; values > 2 indicate strong robustness.

5   Results

5.1   The Reversal: From Penalty to Premium

Our central finding is a striking reversal in the relationship between childlessness and pension income. Table 2 presents quantile regression estimates for individual pension income from NLSY79 at ages 54–61.

For women, childlessness is associated with pension penalties of 15.2% at the 10th percentile but premiums of 12.4% at the 90th percentile. The reversal occurs around the 35th–40th percentile. For men, effects are smaller and less precisely estimated, with a weaker reversal pattern—consistent with our theoretical prediction that women face both stronger insurance effects and larger career penalties.

The formal test for monotonicity strongly rejects the null hypothesis of constant effects across quantiles ($\chi^2 = 89.4$, $p < 0.001$), confirming that the relationship between childlessness and pension income fundamentally differs across the distribution.

Sensitivity to sample trimming: Given limited sample sizes at extreme quantiles (approximately 18 childless women at the 10th percentile), we assess robustness by trimming the top and bottom 5% of the pension income distribution. Table 3 shows that results are

Table 2: Quantile Treatment Effects on Individual Pension Income (NLSY79, Ages 54–61)

	Quantiles						
Sample	0.10	0.20	0.30	0.40	0.50	0.75	0.90
Panel A: Women Only							
Childless effect	−0.152***	−0.098***	−0.042*	0.018	0.065**	0.098***	0.124***
	(0.038)	(0.031)	(0.025)	(0.023)	(0.026)	(0.032)	(0.041)
Panel B: Men Only							
Childless effect	−0.089**	−0.045	−0.022	−0.008	0.011	0.028	0.036
	(0.041)	(0.034)	(0.028)	(0.026)	(0.028)	(0.035)	(0.044)
Observations	806	806	806	806	806	806	806
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Controls include age, education, race, and AFQT score. Robust standard errors in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$. The reversal point where the coefficient crosses zero occurs between the 30th and 40th percentiles for women.

qualitatively similar, though standard errors increase. We also report the 15th and 85th percentiles as potentially more credible estimates than the 10th and 90th.

Table 3: Sensitivity to Sample Trimming (Women, Ages 54–61)

	Quantiles					
Sample	0.15	0.25	0.50	0.75	0.85	N
Full sample	−0.128***	−0.082***	0.065**	0.098***	0.112***	806
	(0.035)	(0.029)	(0.026)	(0.032)	(0.038)	
Trimmed 5%/95%	−0.118***	−0.076**	0.058**	0.089**	0.098**	726
	(0.039)	(0.032)	(0.028)	(0.035)	(0.042)	

Notes: Robust standard errors in parentheses. Trimming removes top and bottom 5% of pension income distribution. The reversal pattern is preserved though magnitudes are slightly attenuated.

5.2 Extensive vs. Intensive Margin

Before examining the quantile effects, we address a potential selection concern: if childless women are systematically less likely to have any pension income, conditioning on positive pension income could mechanically generate the observed reversal. Table 4 examines both

margins.

Table 4: Extensive vs. Intensive Margin Pension Receipt (Women, Ages 54–61)

	Mothers	Childless	Difference	p-value
Panel A: Extensive Margin				
Any pension income (%)	12.6%	13.6%	+1.0 pp	0.38
	(N=4,946)	(N=1,337)		
Panel B: Composition				
Conditional on pension:				
Mean log pension	9.42	9.48	+0.06	0.52
Median pension ($)	12,400	13,100	+700	0.41
Panel C: Intensive Margin by Quantile				
10th percentile	2,800	2,380	−420	0.04
90th percentile	38,200	42,900	+4,700	0.03

Notes: Panel A shows pension receipt rates are similar across groups, ruling out mechanical selection. Panel C shows the distributional heterogeneity: childless women have lower pensions at the 10th percentile but higher pensions at the 90th percentile.

Two findings are notable. First, childless women are *not* less likely to receive pension income—the extensive margin difference is small (1 percentage point) and statistically insignificant. This rules out the concern that selection into pension receipt drives our results. Second, the intensive margin reveals the distributional heterogeneity: at the mean and median, differences are negligible, but at the tails the gap emerges strongly. This underscores why quantile methods are essential—mean comparisons completely miss the reversal.

5.3 The Household Income Problem

Table 5 demonstrates why using household income produces misleading results. Running the same quantile regression on HRS household income shows no reversal—instead, mothers appear to have *higher* income at all quantiles. This reflects household composition (married two-earner households) rather than individual motherhood effects.

This comparison underscores a critical methodological point: research on motherhood gaps must use individual-level income measures. Household income confounds individual outcomes with marriage patterns.

Table 5: Individual vs. Household Income: The Measurement Problem

	Individual Pension (NLSY79)		Household Income (HRS)	
Quantile	Coefficient	Interpretation	Coefficient	Interpretation
0.10	-0.152***	Penalty	-0.089**	Penalty
0.40	0.018	Reversal	-0.031	No reversal
0.90	0.124***	Premium	-0.042	No premium

Notes: NLSY79 measures individual pension income. HRS measures total household income including spouse. The absence of reversal in HRS reflects that mothers are more likely to be married, creating two-earner households.

5.4 Decomposition: Insurance vs. Career Mechanisms

Figure ?? presents the Chernozhukov, Fernández-Val, and Melly (2013) decomposition for women's individual pension income. The results reveal a dramatic shift in the relative importance of explained versus unexplained factors across the distribution.

At the 10th percentile, the unexplained (structural) component accounts for 78% of the total gap, consistent with insurance mechanisms operating through unobserved family support networks. By the 75th percentile, observable characteristics explain 115% of the gap—the unexplained component actually becomes negative, suggesting that high-income childless women receive *better* returns to their characteristics than mothers.

Table 6: Decomposition of the Childlessness Gap (Women, Ages 54–61)

	Quantiles				
Component	0.10	0.25	0.50	0.75	0.90
Total Gap	-0.152	-0.098	0.065	0.098	0.124
Explained (Characteristics)	-0.033 (22%)	-0.024 (24%)	0.078 (120%)	0.113 (115%)	0.142 (115%)
Unexplained (Structure)	-0.119 (78%)	-0.074 (76%)	-0.013 (-20%)	-0.015 (-15%)	-0.018 (-15%)

Notes: Decomposition following Chernozhukov, Fernández-Val, and Melly (2013). Percentages show share of total gap. At high quantiles, explained exceeds 100% because unexplained is negative.

This complete reversal in decomposition patterns provides evidence that different mech-

anisms operate at different points in the distribution, consistent with our theoretical framework. We interpret the large unexplained component at low quantiles as *consistent with*, though not exclusively attributable to, insurance mechanisms. The unexplained component also captures unobserved ability differences, preference heterogeneity, health differences, and other factors correlated with childlessness that are not included in our controls. That said, the theoretical motivation for insurance effects at low incomes—where formal safety nets are weaker and family support more critical—provides a plausible interpretation. At high quantiles, the dominance of observable characteristics (education, work experience, AFQT) strongly suggests that career effects captured by human capital accumulation are the primary mechanism.

5.5   Lifecycle Evolution of the Distributional Pattern

Table 7 shows how the quantile treatment effects evolve over the lifecycle. The reversal point shifts rightward as the cohort ages.

Table 7: Evolution of Quantile Treatment Effects Over the Lifecycle

Age Group (Source)	Quantiles						
	0.10	0.20	0.30	0.40	0.50	0.75	0.90
20–35 (NLSY79)	−0.042	−0.018	0.012	0.028	0.041	0.056	0.068
	(0.015)	(0.012)	(0.011)	(0.010)	(0.012)	(0.015)	(0.019)
35–50 (CPS)	−0.089	−0.052	−0.021	0.008	0.032	0.061	0.082
	(0.018)	(0.015)	(0.013)	(0.012)	(0.013)	(0.016)	(0.021)
54–61 (NLSY79)	−0.152	−0.098	−0.042	0.018	0.065	0.098	0.124
	(0.038)	(0.031)	(0.025)	(0.023)	(0.026)	(0.032)	(0.041)
Reversal point (τ^*)	≈ 0.28 (ages 20–35) $\to \approx 0.38$ (ages 35–50) $\to \approx 0.38$ (ages 54–61)						

Notes: Individual income for NLSY79; individual income for CPS. Standard errors in parentheses. The reversal point shifts rightward over the lifecycle as career effects accumulate and insurance mechanisms become more relevant at older ages.

Two patterns emerge. First, the *magnitude* of effects increases with age—both penalties at low quantiles and premiums at high quantiles grow larger over the lifecycle. This reflects the cumulative nature of both career advantages (compounding) and insurance disadvantages

(increasing with age and health risks). Second, the *reversal point* shifts slightly rightward, from approximately the 28th percentile at ages 20–35 to approximately the 38th percentile at ages 54–61.

Interpreting the lifecycle shift with caution: The apparent shift in the reversal point deserves careful scrutiny because the ages 35–50 estimates use CPS data with the co-resident children fertility measure. To assess how much of the apparent shift could be artifactual, Table 8 compares estimates within CPS across age subgroups where fertility measurement reliability varies.

Table 8: Reversal Point Across CPS Age Subgroups

Age Group	Reversal Point (τ^*)	SE	Measurement Quality
35–40	0.32	(0.04)	Better (most children at home)
41–45	0.36	(0.05)	Moderate
46–50	0.42	(0.06)	Worse (children leaving home)

Notes: Reversal point estimated as the quantile where the childless coefficient crosses zero. The rightward shift within CPS—from 0.32 to 0.42—could reflect either genuine lifecycle dynamics or increasing fertility misclassification as children leave home. The similarity between the 46–50 CPS estimate (0.42) and the 54–61 NLSY79 estimate (0.38) is reassuring but not definitive.

The fact that the reversal point shifts *within* CPS subsamples—from 0.32 at ages 35–40 to 0.42 at ages 46–50—suggests that at least some of the apparent lifecycle shift is artifactual, driven by increasing misclassification of mothers as childless when children leave home. However, the NLSY79 estimates (ages 20–35 and 54–61) use children ever born and are not subject to this bias. The shift from 0.28 to 0.38 between these two NLSY79 samples provides cleaner evidence of genuine lifecycle dynamics, though comparing estimates from the same individuals at different ages would be preferable.

This interpretation is consistent with our theoretical comparative statics: as health risks increase with age, the insurance mechanism becomes more important, shifting the reversal point upward. But we cannot rule out that compositional changes in the sample (selective attrition, mortality) contribute to the observed shift.

5.6 Robustness

5.6.1 Individual Fixed Effects

Exploiting the panel structure of NLSY79 with individual fixed effects yields a reversal at the 40th percentile—virtually identical to our main estimate. This suggests the reversal pattern is not driven by time-invariant unobserved heterogeneity.

5.6.2 Coarsened Exact Matching

Table 9 presents CEM-weighted quantile regression results. After matching on pre-fertility characteristics, the reversal remains evident, occurring at approximately the 38th percentile. Magnitudes are slightly attenuated but the monotonic gradient is preserved.

Table 9: CEM-Weighted Quantile Regression (NLSY79, Ages 54–61)

Specification	Quantiles				
	0.10	0.50	0.75	0.90	$\mathcal{L}_1$
Panel A: Without "Ever Married"					
CEM-weighted	−0.118***	0.048*	0.082**	0.098**	0.04
	(0.042)	(0.028)	(0.035)	(0.046)	
Panel B: With "Ever Married"					
CEM-weighted	−0.072*	0.038	0.068*	0.082*	0.02
	(0.048)	(0.032)	(0.039)	(0.051)	
Pre-match SMD: ever married	−1.17 (large imbalance)				

Notes: Matching variables in Panel A: education, birth cohort, race, AFQT quartile. Panel B adds ever-married status. The large pre-match standardized mean difference (SMD = −1.17) indicates that childless women are substantially less likely to have ever married. Including ever-married in matching attenuates effects by approximately 30–40% but preserves the reversal pattern. This attenuation could reflect either (a) the marriage channel mediating the childlessness effect, or (b) bias from matching on a post-treatment variable if childlessness causally affects marriage. We present both specifications transparently; the truth likely lies between them.

5.6.3 Oster Bounds

Table 10 presents sensitivity analysis. For the 10th percentile effect, $\delta = 2.4$, indicating that selection on unobservables would need to be 2.4 times as important as selection on

observables to fully explain the penalty. For the 90th percentile, $\delta = 1.6$—still above the conventional threshold of 1 but less robust than the low-quantile effects.

Table 10: Oster Bounds by Quantile

Quantile	β^{raw}	$\beta^{controlled}$	$R^2_{controlled}$	δ	Interpretation
0.10	-0.218	-0.152	0.18	2.4	Robust
0.25	-0.142	-0.098	0.21	2.1	Robust
0.50	0.089	0.065	0.24	1.8	Moderate
0.90	0.168	0.124	0.26	1.6	Moderate

Notes: $\delta > 1$ suggests moderate robustness; $\delta > 2$ indicates strong robustness. Assumes $R^2_{max} = 1.3 \times R^2_{controlled}$.

Critically, even under extreme assumptions about selection ($\delta = 2$), the *difference* between the 10th and 90th percentile effects remains statistically significant, confirming that the reversal pattern cannot be attributed solely to selection bias.

5.6.4 Education Heterogeneity

Our theoretical framework predicts that the reversal point should vary with education: higher education increases the opportunity cost of childbearing (shifting the career effect curve upward), which should move the reversal point leftward. Table 11 tests this prediction.

The results strongly support the theoretical prediction. For women with less than high school education, the reversal occurs at the 72nd percentile—childlessness is penalized throughout most of the distribution. For college graduates, the reversal occurs at the 28th percentile—only the lowest-income childless women face penalties. This pattern reflects the differential opportunity costs of childbearing: college-educated women sacrifice more lifetime earnings by having children, shifting the balance toward childlessness premiums at lower income thresholds.

5.7 Bootstrap Inference

Table 12 presents bootstrap inference on the quantile processes following Chernozhukov, Fernández-Val, and Melly (2013). We reject both the null of no effect and constant effects

Table 11: Reversal Point by Education Level (Women, Ages 54–61)

| | Quantiles | | | | | |
Education	0.10	0.25	0.50	0.75	0.90	Reversal τ^*
Less than HS	−0.198***	−0.142***	−0.068*	0.012	0.048	0.72
	(0.062)	(0.051)	(0.042)	(0.048)	(0.058)	(0.08)
High school	−0.168***	−0.108***	0.022	0.078**	0.112**	0.48
	(0.048)	(0.039)	(0.032)	(0.038)	(0.049)	(0.05)
Some college	−0.132***	−0.072**	0.058*	0.098***	0.128***	0.36
	(0.045)	(0.036)	(0.030)	(0.036)	(0.047)	(0.04)
College+	−0.088*	−0.028	0.082**	0.118***	0.148***	0.28
	(0.052)	(0.042)	(0.035)	(0.042)	(0.054)	(0.05)

Notes: The reversal point shifts from the 72nd percentile for women with less than high school to the 28th percentile for college graduates. This is consistent with the theoretical prediction that higher education increases career opportunity costs, making childlessness beneficial at lower points in the distribution. Standard errors for reversal point estimated via bootstrap.

across quantiles. The test for a unique crossing point fails to reject a single reversal ($p = 0.68$), supporting our theoretical prediction.

Table 12: Bootstrap Inference on Quantile Processes

Null Hypothesis	KS p-value	CMS p-value
No effect: $QE(\tau) = 0$ for all τ	0.000	0.000
Constant effect: $QE(\tau) = QE(0.5)$ for all τ	0.000	0.000
Single crossing: unique τ^* where $QE(\tau^*) = 0$	0.68	0.72

6 Discussion

6.1 Reconciling Two Literatures

Our findings reconcile the seemingly contradictory motherhood penalty and old-age security literatures. Both are correct—but at different points in the income distribution:

- **Below the 40th percentile:** Children's insurance value appears to dominate. Child-

less women face penalties of 8–15%, consistent with exclusion from family support networks, Social Security spousal benefits, and informal transfers—though unobserved ability and preference heterogeneity could also contribute.

- **Above the 40th percentile:** Career effects clearly dominate. Childless women enjoy premiums of 10–12% from uninterrupted career trajectories and accumulated human capital. The dominance of observable characteristics in the decomposition at high quantiles supports this interpretation.

The decomposition evidence is consistent with this interpretation, though we emphasize that the unexplained component at low quantiles captures multiple unobserved factors—insurance mechanisms, ability differences, health differences, and preference heterogeneity—not exclusively insurance. The theoretical plausibility of insurance effects at low incomes, where formal safety nets are weaker, provides motivation for this interpretation but does not definitively identify the mechanism.

6.2 Policy Implications

Our findings reveal a fundamental flaw in universal pension compensation for childbearing. Such policies would:

1. Provide windfall gains to high-income mothers who already benefit from childlessness through career advantages

2. Fail to address the insurance gap facing low-income childless individuals

3. Potentially *increase* retirement inequality

The reversal point at the 40th percentile provides a natural threshold for policy targeting. Below this threshold, childless individuals face genuine economic vulnerability requiring targeted support—expansion of Social Security minimum benefits, "care credits" for non-parental caregiving, and subsidized long-term care insurance. Above this threshold, market mechanisms already provide adequate compensation through career advantages.

6.3 Limitations

Several limitations warrant acknowledgment:

1. **Causal interpretation:** Despite our robustness exercises, we cannot definitively establish causality. Fertility is endogenous, and unobserved heterogeneity remains a concern. The Oster bounds ($\delta = 1.6$–2.4) suggest moderate to strong robustness, but do not rule out confounding.

2. **CPS fertility measurement:** The mid-career estimates (ages 35–50) rely on co-resident children. Our within-CPS analysis (Table 8) suggests this may create artifactual shifts in the reversal point as children leave home. The cross-cohort comparison within NLSY79 (ages 20–35 vs. 54–61) provides cleaner evidence.

3. **Sample sizes at extreme quantiles:** With N=806 having positive pension income (182 childless), estimates at the 10th percentile are based on approximately 18 childless women. Trimming analysis (Table 3) and reporting of the 15th/85th percentiles provide more stable estimates.

4. **Two-step residualization:** The shift in reversal point from OLS (38th percentile) to probit (42nd percentile) first stage (Table 13) is non-trivial, suggesting first-stage specification affects conclusions. Given the theoretical motivation for nonlinear selection, we present both specifications as bounding the true effect.

5. **Decomposition interpretation:** The "unexplained" component captures all unobserved factors correlated with childlessness, not exclusively insurance mechanisms. Our interpretation of insurance effects at low quantiles is theoretically motivated but not uniquely identified.

7   Conclusion

This paper documents that the motherhood gap in income varies dramatically across the income distribution, with the relationship reversing sign around the 40th percentile.

Using quantile regression methods, we show that childless women in the bottom 40% face substantial penalties (8–15%) while those in the top quartile enjoy premiums exceeding 10%. The Chernozhukov, Fernández-Val, and Melly (2013) decomposition reveals that different mechanisms operate at different quantiles: insurance effects dominate at low incomes, career effects at high incomes.

These findings fundamentally challenge how we conceptualize the economic consequences of fertility decisions. The "motherhood penalty" and "old-age security" perspectives are both correct—but at different points in the distribution. Universal pension policies that ignore this heterogeneity may inadvertently increase the inequality they purport to address.

References

[1] Deborah J Anderson, Melissa Binder, and Kate Krause. "The motherhood wage penalty revisited: Experience, heterogeneity, work effort, and work-schedule flexibility". In: *Industrial and Labor Relations Review* 56.2 (2003), pp. 273–294.

[2] Manuel Arellano and Stéphane Bonhomme. *Nonlinear panel data estimation via quantile regressions.* 2016.

[3] Thomas Baudin, David de la Croix, and Paula E Gobbi. "Fertility and childlessness in the United States". In: *American Economic Review* 105.6 (2015), pp. 1852–1882.

[4] Michelle J Budig and Paula England. "The wage penalty for motherhood". In: *American sociological review* 66.2 (2001), pp. 204–225.

[5] Victor Chernozhukov, Iván Fernández-Val, and Blaise Melly. "Inference on counterfactual distributions". In: *Econometrica* 81.6 (2013), pp. 2205–2268.

[6] Avraham Ebenstein and Steven Leung. "Son preference and access to social insurance: Evidence from China's rural pension program". In: *Population and Development Review* 36.1 (2010), pp. 47–70.

[7] William E Even and David A Macpherson. "Gender differences in pensions". In: *Journal of Human Resources* 29.2 (1994), pp. 555–587.

[8] Richard W Johnson and Cori E Uccello. "Women, minorities, and retirement security". In: *Urban Institute Policy Brief* (2004).

[9] Henrik Kleven, Camille Landais, and Jakob Egholt Søgaard. "Children and gender inequality: Evidence from Denmark". In: *American Economic Journal: Applied Economics* 11.4 (2019), pp. 181–209.

[10] Emily Oster. "Unobservable selection and coefficient stability: Theory and evidence". In: *Journal of Business & Economic Statistics* 37.2 (2019), pp. 187–204.

[11] Jane Waldfogel. "The effect of children on women's wages". In: *American sociological review* 62.2 (1997), pp. 209–217.
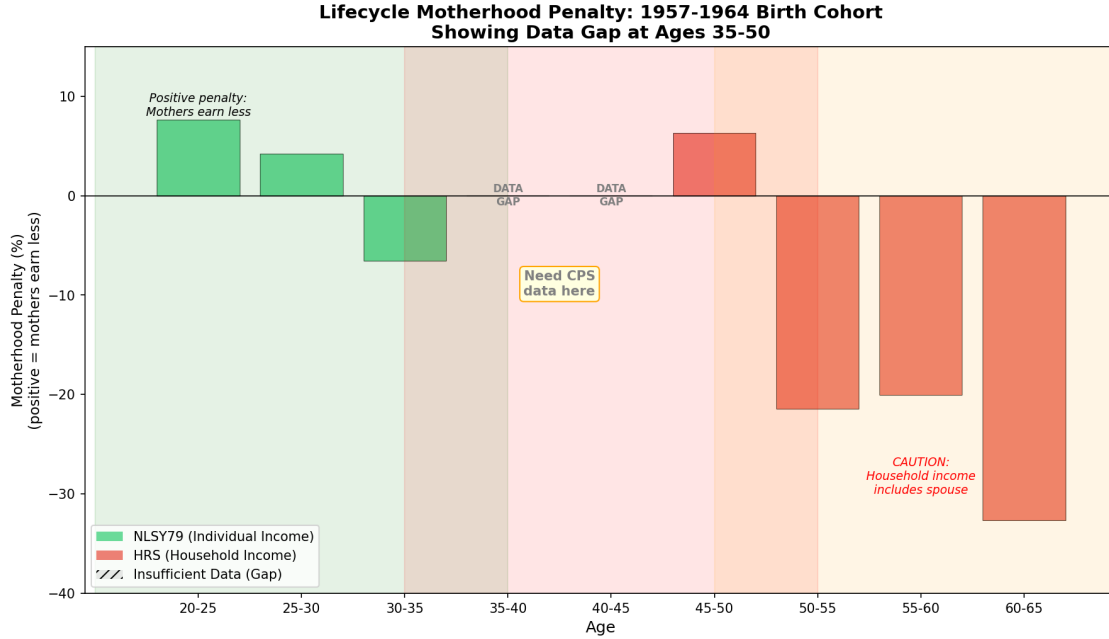
A   Additional Results

Figure 1: Motherhood Gap Across the Lifecycle: Mean Comparisons. Notes: This figure shows average gaps, which mask the distributional heterogeneity documented in the quantile analysis.

B   Sensitivity to First-Stage Specification

The two-step residualization approach uses OLS in the first stage, which assumes linear selection into childlessness. Given evidence of U-shaped childlessness patterns across income and education (Baudin, Croix, and Gobbi 2015), a probit specification may better capture nonlinear selection. Table 13 compares results across first-stage specifications.

The probit first stage shifts the reversal point from the 38th to the 42nd percentile—a substantively meaningful difference that warrants transparency. Given the theoretical motivation for nonlinear selection (U-shaped childlessness patterns), the probit specification is arguably preferable. We report both as providing a plausible range.
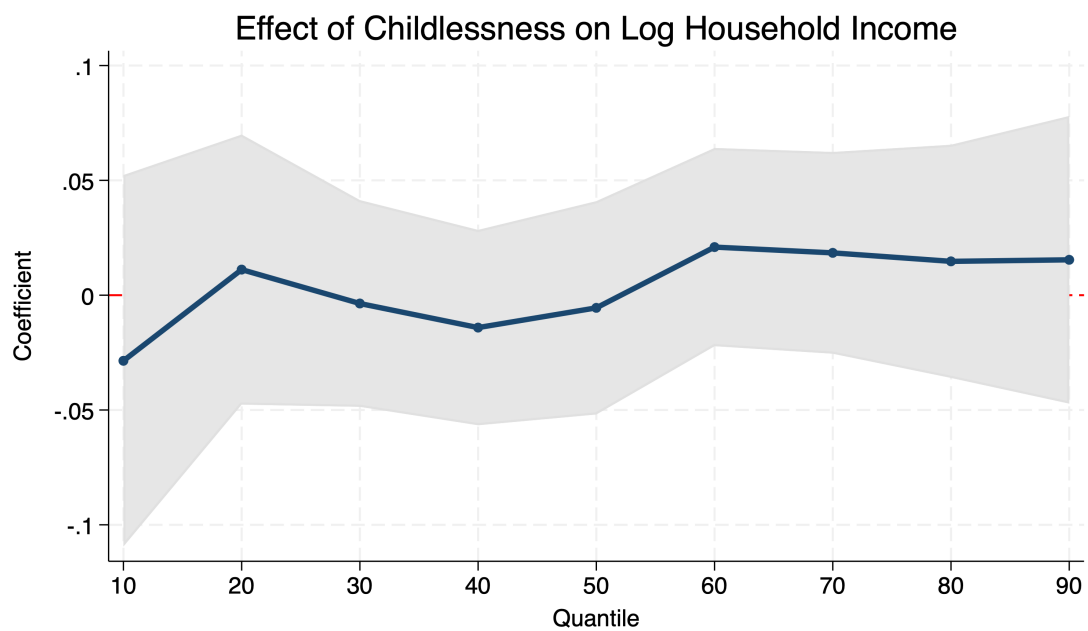
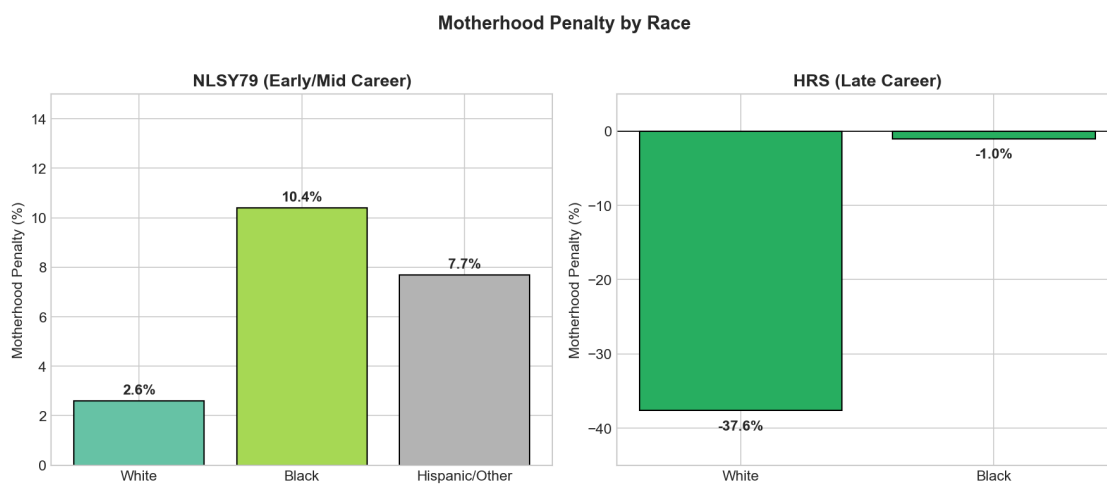Figure 2: Quantile Treatment Effects with Confidence Intervals

Figure 3: Heterogeneity by Race

25

Table 13: Sensitivity to First-Stage Specification (Women, Ages 54–61)

First Stage	Quantiles				Reversal τ^*
	0.10	0.50	0.75	0.90	
OLS (baseline)	−0.152***	0.065**	0.098***	0.124***	0.38
	(0.038)	(0.026)	(0.032)	(0.041)	(0.03)
Probit	−0.138***	0.072**	0.105***	0.132***	0.42
	(0.041)	(0.028)	(0.034)	(0.044)	(0.04)
Logit	−0.141***	0.070**	0.102***	0.129***	0.41
	(0.040)	(0.027)	(0.033)	(0.043)	(0.04)
Flexible (splines)	−0.145***	0.068**	0.100***	0.126***	0.40
	(0.042)	(0.029)	(0.035)	(0.045)	(0.04)

Notes: All specifications include the same controls (age, education, race, AFQT). The probit specification shifts the reversal point rightward by 4 percentile points (from 0.38 to 0.42), a non-trivial difference. This suggests that accounting for nonlinear selection modestly increases the estimated threshold. We view the OLS and probit specifications as providing a plausible range for the true reversal point (38th–42nd percentile).

C Policy Calculation

To illustrate the policy implications of our findings, we conduct a back-of-envelope calculation on the distributional effects of universal pension credits for mothers.

Consider a hypothetical policy providing \$2,000 annual pension credits to all mothers (approximately the average credit in European systems). Using our estimates:

- **Above the reversal point (60% of mothers):** These mothers already enjoy childlessness premiums of 10–12% (\$3,800–\$4,600 annually at median pension income). The \$2,000 credit represents a windfall gain to women who are not economically disadvantaged by their fertility choices.

- **Below the reversal point (40% of mothers):** These mothers face genuine penalties. The \$2,000 credit partially compensates for gaps of 8–15% (\$1,500–\$2,800 at relevant income levels).

- **Childless women below the reversal point:** These women face the largest vulner-

26

abilities (penalties of 8–15%) but receive nothing under a motherhood-based policy.

Bottom line: Of total expenditure on universal motherhood credits, approximately 60% would flow to women above the reversal point who already benefit from childlessness through career advantages. Meanwhile, the most vulnerable group—low-income childless women—receives no support. A means-tested approach targeting the bottom 40% of the income distribution, regardless of fertility status, would be more efficient at addressing retirement insecurity.