

Does Receiving a Pension Improve Wellbeing?

Evidence from Propensity Score Matching Among Older Adults in Colombia

Author : Sara Isabel Jiménez Jiménez

Thesis Director EAFIT : Henry Daniel Puerta Álvarez

Thesis Director UCLouvain : Vincent Vandenberghe

Academic Year 2025-2026

In order to obtain the joint degree:

Master 120 en Sciences économiques, Orientation générale, Finalité spécialisée (UCLouvain/UNamur)

and

Magister en Economía Aplicada (EAFIT)

Dedication

To my family, for their love, support, and constant presence throughout this journey.

And especially to my mother, Maria Victoria, whose wellbeing has always been a priority in my life and whose story inspired the central question of this thesis.

Acknowledgements

I would like to express my sincere gratitude to my supervisors, Henry and Vincent, for giving me the opportunity to develop this thesis under their guidance, their trust, feedback, and support throughout has been valuable during this research process.

I also want to thank Universidad EAFIT and UCLouvain/UNamur for welcoming me into their academic communities and for giving me the opportunity to grow personally and professionally through this double-degree Master's program in Applied Economics.

Thanks to my classmate Monica and my work colleagues for their encouragement. Their conversations, advice, and shared experiences made this journey so much more meaningful.

Finally, I want to thank my loved ones for their understanding and emotional support throughout this journey. To my sister, Maria Camila, thank you for being a source of inspiration and joy in my life. To my grandmother, Luz Ofelia, thank you for accompanying me in spirit. And to my boyfriend, Pablo Andres, thank you from the bottom of my heart for standing by me with so much patience, love, and support.

Completing this thesis has been a challenge as an important learning experience, this is not just the end of an academic stage, but also a period of personal development that I will always remember.

Sara I. Jimenez J.

Note: Generative AI tools (Claude, Anthropic) were used as linguistic assistants to review and improve written expression, and as a coding assistant during the development of the R analysis pipeline. All analytical decisions, interpretations, and conclusions are the sole responsibility of the author.

Table of Contents

1. INTRODUCTION.....	9
1.1 Institutional Context: Why pension coverage is low and access is socially stratified	9
1.2 The retirement transition and subjective wellbeing	10
1.3 The empirical gap in Colombia.....	10
1.4 Scope of the Causal Claim	11
1.5 Research question and objectives	12
2. THEORETICAL FRAMEWORK AND EMPIRICAL LITERATURE	13
2.1 Subjective wellbeing: definition and measurement	13
2.2 The economics of retirement and the retirement shock	13
2.3 Pensions and wellbeing in low-coverage contexts: evidence from Latin America.....	13
2.4 Composite indices and the multiple-comparisons problem	14
2.5 Propensity Score Matching as an identification strategy	15
3. DATA AND ANALYTICAL SAMPLE	16
3.1 The ECV 2024.....	16
3.2 Sample construction.....	16
3.3 Treatment variable.....	17
3.4 Outcome variables	17
3.5 Covariates.....	18
4. METHODOLOGY	20
4.1 Identification strategy.....	20
4.2 The propensity score and its logit transformation.....	20
4.3 Propensity score model	20
4.4 Matching procedure.....	21
4.5 Outcome model and standard errors	21
4.6 Sensitivity analysis.....	22
4.7 Heterogeneity analysis.....	22
4.8 Software and reproducibility.....	22
5. DESCRIPTIVE ANALYSIS AND COMPOSITE INDEX CONSTRUCTION.....	23
5.1 Sociodemographic characteristics by pension status.....	23
5.2 Construction of the composite wellbeing index	24
6. RESULTS.....	26

6.1 Propensity score model	26
6.2 Common support: probability scale vs. logit scale	28
6.3 Matching results and post-matching balance	29
6.4 Main result: ATT on the composite wellbeing index	31
6.5 Disaggregated results: twelve wellbeing dimensions	32
6.6 Heterogeneity.....	34
7. SENSITIVITY ANALYSIS	37
7.1 Rosenbaum bounds.....	37
7.2 E-values	37
7.3 Robustness of the composite index to the exclusion of income satisfaction	38
8. DISCUSSION	40
8.1 Magnitude and pattern of the estimated effects.....	40
8.2 Why the Colombian estimate exceeds the European benchmark	40
8.3 Comparison with other middle-income countries	41
8.4 Mechanisms (preliminary).....	42
9. POLICY IMPLICATIONS	43
10. CONCLUSIONS, LIMITATIONS, AND FUTURE RESEARCH	45
10.1 Summary of findings	45
10.2 Limitations.....	46
10.3 Future research	46
10.4 Closing remark.....	46
References.....	47

List of Tables

Table 1. Sample construction	16
Table 2. Key variables used in the analysis.....	18
Table 3. Sociodemographic characteristics by pension status (weighted by FEX_C).....	23
Table 4. Propensity score model: summary statistics	26
Table 5. Logistic regression: propensity score model coefficients.....	27
Table 6. Doubly robust OLS regression: ATT on the composite wellbeing index.....	32
Table 7. ATT estimates by wellbeing dimension	33
Table 8. ATT on composite wellbeing index by subgroup	34
Table 9. Rosenbaum bounds: critical Γ values.....	37
Table 10. Sensitivity of the composite index to alternative specifications	38

List of Figures

Figure 1. Distribution of the composite wellbeing index by pension status (pre-matching).....	25
Figure 2. ROC curve of the propensity score model (AUC = 0.935).	26
Figure 3. Propensity score distribution by group, probability scale.....	28
Figure 4. Distribution of logit(PS) by group: the scale used for matching.	28
Figure 5. Mirror histogram of the propensity score by group.	29
Figure 6. Love plot: covariate balance pre and post matching.	30
Figure 7. Distributional balance for the health insurance regime, pre and post matching	30
Figure 8. Distributional balance for education level (P8587), pre and post matching.	31
Figure 9. Propensity score distribution post-matching: near-complete overlap.	31
Figure 10. Forest plot of the ATT on the composite wellbeing index.	32
Figure 11. Forest plot of the ATT by wellbeing dimension with Bonferroni correction.	33
Figure 12. Heterogeneity by sex: ATT across 13 outcomes (composite + 12 dimensions).	35
Figure 13. Heterogeneity by area of residence: ATT across 13 outcomes.	35
Figure 14. Heterogeneity by education level: only the No educ./Basic subgroup converged....	36
Figure 15. Heterogeneity by socioeconomic stratum: only the No stratum/Low subgroup converged.	36
Figure 16. E-values for outcomes with significant ATT.	38

Abstract

Most adults spend their working lives looking forward to retirement, yet the transition from a structured routine to open-ended days raises an important question: is that long-awaited moment enough when a clear sense of purpose is no longer there?

This study examines whether receiving pension income is associated with greater well-being among older adults in Colombia.

Using data from the 2024 Quality of Life Survey (ECV 2024) and a propensity score matching approach (Rosenbaum & Rubin, 1983), pensioners are compared with observationally similar individuals who do not receive a pension, controlling for differences between the two groups.

A composite well-being index constructed following Kling et al. (2007) shows that receiving a pension is associated with an increase of 0.097 standard deviations. The effect is most evident in the economic dimension: pensioners report 0.664 points higher satisfaction with their income and 0.433 points higher satisfaction with their work situation, on a 0–10 scale. No significant effects are found for health or day-to-day emotional well-being. Effects are larger in rural areas and among men, though these differences are not conclusive. Rosenbaum bounds (2002) suggest moderate robustness to unobserved confounders (critical $\Gamma \approx 1.3$).

Finally, in terms of public policy, complementary programs focused on psychosocial support and active aging are proposed, in order to address the dimensions of well-being, health and emotional welfare, that financial transfers do not reach.

Keywords: pension, subjective wellbeing, propensity score matching, older adults, Colombia, causal inference, cross-sectional design

Resumen

El adulto promedio aspira al momento de obtener su retiro laboral, sin embargo, la transición de una rutina fija a la vida en el hogar plantea una pregunta relevante: ¿es suficiente ese momento tan anhelado cuando ya no existe un propósito claro de vida?

La presente investigación tiene como objetivo conocer si recibir ingresos de pensión está asociado con un mayor bienestar entre los adultos mayores en Colombia.

Con esta finalidad, se utilizaron los datos de la Encuesta de Calidad de Vida 2024 (ECV 2024) y una técnica de emparejamiento por puntaje de propensión (Rosenbaum & Rubin, 1983), que permite comparar pensionados con personas similares que no reciben pensión, controlando por diferencias observables entre los dos grupos.

Construyendo un índice compuesto de bienestar siguiendo a Kling et al. (2007), se encuentra que recibir pensión se asocia con un aumento de 0,097 desviaciones estándar. El efecto es más claro en la esfera económica: quienes reciben pensión reportan 0,664 puntos más de satisfacción con sus ingresos y 0,433 puntos más de satisfacción con su situación laboral (escala 0–10). No se encuentran efectos significativos en salud ni en emociones del día a día. Aunque las diferencias en los efectos no son concluyentes, son más altos en zonas rurales y entre hombres. Los análisis de sensibilidad con cotas de Rosenbaum (2002) indican robustez moderada frente a variables no observadas (Γ crítico $\approx 1,3$).

Finalmente, en términos de política pública, se proponen programas complementarios de acompañamiento psicosocial y envejecimiento activo podrían atender las dimensiones del bienestar, salud y bienestar emocional, que el ingreso pensional no alcanza.

Palabras clave: pensión, bienestar subjetivo, emparejamiento por puntaje de propensión, adultos mayores, Colombia, inferencia causal, diseño transversal

1. INTRODUCTION

The world's population, including Colombia's, is aging rapidly, the Departamento Administrativo Nacional de Estadística DANE (2024b) projected that adults over 60 will grow from 15% of the total population in 2025 to 23% by 2050, which makes it urgent to ask: when people retire and receive a pension, does their well-being actually improve?

The Colombian pension system has, among several, two challenges that must be highlighted, on the one hand, very few people manage to retire with a pension, and those who do tend to be in better conditions than those who do not; therefore, understanding whether genuinely obtaining a pension improves the well-being of older adults, and how and for whom, is essential to creating effective public policies.

A pension might improve how people evaluate their lives without necessarily improving how they feel on a day-to-day basis. It might also benefit those living in urban areas more than those in rural ones, or the other way around, so the policies must adapt accordingly.

1.1 Institutional Context: Why pension coverage is low and access is socially stratified

Colombia's pension system has two tracks (Law 100, 1993), one managed by the government (RPM, pay-as-you-go), one run by private companies (RAIS, individual savings). Both require two things, the first one is reaching retirement age (57 for women, 62 for men) and the second one is accumulating 1,300 weeks of contributions. However, there is a problem with the last one, as most people reach retirement age, but accumulating 1,300 weeks of contributions in a labor market where 58% of workers are informal (DANE, 2024a) is a challenge, as the informal employment generates no contribution weeks, this is an structural constraint, not individual failure.

Fernández and Villar (2017) studied a tax change in 2012, to understand why is informality so high, and found that Colombia lowered employer payroll contributions from 29.5% to 16%, while the informality dropped by 4.8 percentage points. He mentioned a discussion where a high payroll taxes were actively pushing people into informal work, while when taxes were lower, more people entered formal job, concluding that high tax rates were sustaining informal employment, the same informality that blocks pension contributions.

Mejía and Forero (2022) were interested to underservant how many Colombians are trapped in this cycle, and they found that only one in three working-age Colombians (ages 20-59) is contributing to a pension at any given moment. A second opinion that highlighted the fact that the principal reason is informality, not lack of work. Informal workers, the self-employed, agricultural workers, domestic employees, do contribute if they can, but most do not accumulate enough weeks before reaching retirement age, here we are back on the idea that the problem is not individual choice or poor planning, but it is that the structure of their labor market offers no clear path to 1,300 weeks.

While we analyse the informality, there is also an important aspect in formal employment, which is not randomly distributed, as it clusters with education, urban living, and access to the contributory health system, so pension access is not just rare, it is unequally distributed. The people most likely to get pensions are already better-off, when you compare pensioners to non-pensioners, you are not just comparing two groups with different incomes, you are also comparing two groups with entirely different life trajectories: different jobs, different places they lived, different schooling, so a simple comparison between the two groups will mislead you because the differences are so large.

Colombia does have a backup program for people who do not qualify, the Consorcio Colombia Mayor runs two programs, one gives modest cash transfers to vulnerable older adults (Colombia Mayor), another subsidizes contributions for low-income workers. By 2024, Colombia Mayor covered 1.7 million people, where the monthly amount, COP 80,000 to 225,000, is less than one-fifth of a regular contributory pension (COP 1.3 million), however this thesis focuses on contributory pensions, not these subsidies.

It is important to mention that a new reform (Law 2381, 2024) tried to expand coverage through a four-pillar system, however, the Constitutional Court stopped it in mid-2025 while Congress corrected procedural problems, so none of the people in our data were affected by the reform.

1.2 The retirement transition and subjective wellbeing

Retirement is not just an economic event, as it disrupts daily routine, social connections, and sense of purpose, things that matter for wellbeing independently of money (Ekerdt, 1987), this is also known as "retirement shock."

In one part of the planet, Europe, the evidence is mixed, the economists Bonsang and Klein (2012) tracked 12 countries with SHARE data and found that retirement hurt life satisfaction in the short term, especially for people who valued work highly, while Börsch-Supan and Jürges (2006) saw similar patterns in Germany, however, more recent studies using sharper research designs found effects near zero, retirement happened, and wellbeing stayed roughly the same (Coe & Zammaro, 2011), so we can assume that in Europe, the retirement shock is real but small.

On this part of the world, Latin America, it is evidence a different story, as when Mexico rolled out the 70 y Más program, Galiani, Gertler, and Bando (2016) found mental health and life satisfaction improved among rural older adults, also, Peru's Pensión 65 showed similar results, reaching pension eligibility age improved wellbeing (Bando, Galiani, & Gertler, 2020).

1.3 The empirical gap in Colombia

Very little research in Colombia has explored whether pensions improve wellbeing, there are researches about general subjective wellbeing, those say that job security matters a lot, and job loss has a stronger negative effect on wellbeing than equivalent income changes, as well as access to health services, pension plans, and safer communities is closely linked to higher

wellbeing (Graham & Krauss, 2013), and there also focusing on the number of people covered or the poverty levels of pensioners (Bosch et al., 2013; Sánchez & Mesa, 2018), or Mejía and Forero (2022) provide the most thorough recent analysis of old-age protection in Colombia, as they demonstrate how informality leads to pension exclusion and poverty but do not measure wellbeing. The World Bank (2023) confirms that the coverage gap for older adults remains around 25%, a low figure by regional standards but unevenly distributed, mentioning that the rural, indigenous, and Afrodescendant populations are often left behind.

So, while Colombia has data and literature on coverage gaps and poverty, and with them presents analyses and policies on proposed reforms (as Salazar-Saenz & Robayo, 2020 have done), there is little evidence at the individual level on whether receiving a pension actually changes the way older adults feel or evaluate their lives, that is why the 2024 ECV survey is a good source of information as this one asks both, Do you receive a pension? AND How satisfied are you with your life, health, income, and security? (following OECD, 2013 standards).

1.4 Scope of the Causal Claim

The thesis asks a conditional question, for people with similar observable characteristics, does receiving a pension predict higher wellbeing? This is not the same as a causal effect, as pensions are not assigned randomly; they are given to individuals with more education, formal jobs, better health insurance, and urban residence. If you just compare pensioners to non-pensioners, you are comparing two different groups rather than measuring the pension's true effect.

In order to address this, it is used the Propensity Score Matching (PSM), as the idea is find non-pensioners who resemble pensioners (in terms of education, job history, and health insurance) and compare them, if they are genuinely similar, the difference between them should closely represent a real causal effect.

PSM identification is based on the conditional independence assumption (CIA), where all the variables influence the pension access and wellbeing observed (Rosenbaum & Rubin, 1983).

There is one more limitation, the survey asks, "Did you receive pension income last month?" It does not specify the type of pension, whether it is old-age, survivor's, or disability. These types differ and might operate differently, as someone receiving a disability pension might have worse health than someone with an old-age pension, so when we group them together, we measure an average of different effects, which may obscure the true dynamics at play.

The timing is also important, as education and sex clearly precede receiving a pension and do not change, however, health regime, current job, and reported illness are measured at the same time as the pension question, because it is used as indicators of past employment history, which seems reasonable, but it is not perfect. If receiving a pension alters someone's job situation (for example, they begin volunteering upon retirement), it might control for something that the pension itself could influence.

1.5 Research question and objectives

The main research question is: Among older adults in Colombia with similar observed characteristics, is receiving a pension associated with higher wellbeing?

To answer it, the thesis does the following:

1. Characterize the sociodemographic profile of older adults in Colombia by pension status, identifying observable differences in age, sex, education, health regime, main activity, socioeconomic stratum, ethnicity, and area of residence.
2. Construct a composite subjective wellbeing index that integrates evaluative, hedonic, and eudaimonic dimensions available in the ECV 2024, following the OECD (2013) and Diener et al. (1999) frameworks and using the equal-weight aggregation method of Kling et al. (2007).
3. Estimate a logistic propensity score model that captures the probability of receiving pension income as a function of observable characteristics, verifying the common support assumption and assessing post-matching balance through standardized mean differences and Love plots.
4. Estimate the ATT of pension receipt on the composite wellbeing index and, in disaggregated analyses, on each of the twelve component dimensions, applying Bonferroni correction for multiple comparisons.
5. Assess the robustness of the main result to hidden bias using Rosenbaum bounds (Rosenbaum, 2002) and E-values (VanderWeele & Ding, 2017), and examine heterogeneity patterns across sex, area, education, and socioeconomic stratum.
6. Derive policy implications consistent with the empirical findings, with explicit attention to the limitations imposed by the cross-sectional design and the partial scope of the CIA.

2. THEORETICAL FRAMEWORK AND EMPIRICAL LITERATURE

2.1 Subjective wellbeing: definition and measurement

Subjective wellbeing includes several aspects of how people experience their lives, while the hedonic approach, rooted in classical economics, measures this through emotional states like happiness, worry, sadness, and daily satisfaction (Diener, 1984; Kahneman, Diener, & Schwarz, 1999), the eudaimonic methods focus on a different question, not just whether you feel good, but whether your life has purpose and meaning (Ryff, 1989).

There are three components that the OECD (2013) proposes under the Guidelines on Measuring Subjective Wellbeing, the evaluative (overall judgments about life), the hedonic (emotional and experiential states), and the eudaimonic (sense of purpose), as by looking just at one single indicator, will offer a limited view of wellbeing, risking misinterpretation, while the full picture is more complicated.

The ECV 2024 uses this framework in its wellbeing module, the evaluative items include overall life satisfaction and the Cantril ladder (OECD, 2013), while hedonic indicators assess happiness, worry, and sadness experienced the day before, a eudaimonic item gauges whether life feels worthwhile. The survey also inquires about satisfaction in specific areas such as income, work, free time, health, and security, this broad coverage enables the creation of a composite index that avoids the oversimplified view that would result from using a single indicator.

2.2 The economics of retirement and the retirement shock

Life-cycle models in labor economics have long examined how individuals plan savings and consumption around retirement (Modigliani & Brumberg, 1954; Gruber & Wise, 1999), as the simplest version assumes pension income replaces wage income, a substitution that would leave wellbeing unchanged, yet evidence from high-income, formal-labor contexts tells a more complicated story.

The retirement shock introduced by Ekerdt (1987), is used to describe what some individuals experience, including deterioration in subjective wellbeing when paid work ends, along with its structure, routine, identity, social networks, and sense of purpose.

Through the introduction, Europe was mentioned, as what emerges from the European evidence is heterogeneity and often small average effects, in formal labor markets with high pension coverage, the association between retirement and subjective wellbeing is small, inconsistent across studies, and varies significantly across population subgroups.

2.3 Pensions and wellbeing in low-coverage contexts: evidence from Latin America

Research on pension programs in developing economies reports stronger welfare effects than those observed in high-income countries, Galiani, Gertler, and Bando (2016) study

Mexico's 70 y Más program using a difference-in-differences design that exploits geographic rollout timing, they find substantial improvements in mental health (depression measures decrease by approximately 0.10 standard deviations) and subjective wellbeing, alongside reductions in paid work hours, a pattern explained by income effects. While Peru's Pensión 65, analyzed by Bando, Galiani, and Gertler (2020) using regression discontinuity around the age cutoff, shows effects on subjective wellbeing of 0.08 to 0.12 standard deviations, concentrated in economic security and life satisfaction domains.

Colombia's existing literature documents structural constraints on pension access, Sánchez and Mesa (2018) show that RPM subsidies accrue disproportionately to higher-income workers who accumulate the required contribution weeks, while low-income workers contribute irregularly and rarely qualify, while Bosch, Melguizo, and Pagés (2013) classify the Colombian system as one with high benefits for the covered minority and low coverage overall, as the coverage gap reflects structural labor market features, Mejía and Forero (2022) find that the probability of completing the 1,300 contribution weeks required under the RPM falls below 20% for workers in the bottom two income quintiles, not due to insufficient work effort but due to the structural informality of available employment.

On other aspect, Fernández and Villar (2017) analyzed the 2012 Colombian tax reform, reducing non-wage labor costs by 13.5 percentage points reduced informality by 4.8 points, revealing the tax sensitivity of contribution patterns, the present thesis contributes by providing the first matching-adjusted estimate of the association between pension receipt and multidimensional subjective wellbeing in Colombia using a nationally representative sample, filling a gap that microsimulation studies like Salazar-Saenz and Robayo (2020) have not addressed.

2.4 Composite indices and the multiple-comparisons problem

Testing 12 separate wellbeing outcomes without statistical correction creates a practical problem, with 12 independent tests at $\alpha = 0.05$, the probability of finding at least one false positive by chance is $1 - (0.95)^{12} \approx 0.46$, approaching one in two, a single composite indicator avoids this inflation, so a composite index also reflects a conceptual reality, no single wellbeing indicator captures the full construct.

The OECD (2013) emphasizes that evaluative and hedonic measures each miss what the other captures, and neither reflects how satisfied someone is with specific life domains, an index that combines all twelve can represent the full multidimensional picture that theory and policy require, that is why, the equal-weight z-score aggregation method, formalized by Kling, Liebman, and Katz (2007) in their Moving to Opportunity analysis, has become standard in development economics when multiple outcomes are tested (Anderson, 2008).

The method standardizes each indicator using the pre-treatment sample mean and standard deviation, then averages the standardized scores, with this approach no assumption are required about latent factor structure, which matters when twelve indicators are correlated but not reducible to a single dimension, as it gives each conceptual domain equal voice rather

than allowing the most variable indicator to dominate, it is also robust to changes in the indicator set in ways that principal-components weighting is not, finally, Anderson (2008) discusses inverse-covariance weighting as a more efficient alternative, but it is less interpretable and more sensitive to outlier indicators.

2.5 Propensity Score Matching as an identification strategy

Propensity Score Matching, developed by Rosenbaum and Rubin (1983), is a quasi-experimental method that addresses selection bias by reducing the dimensionality of the conditioning problem, the propensity score is the conditional probability of treatment given observed covariates:

$$p(X) = \Pr(D = 1 | X)$$

It is a balancing score, conditioning on $p(X)$ achieves the same covariate balance as conditioning on the full vector of covariates. Under two assumptions, the conditional independence assumption (CIA) and overlap, the average treatment effect on the treated (ATT) is identified as the difference in mean outcomes between treated units and matched controls (Rosenbaum & Rubin, 1983; Imbens, 2004).

Meanwhile the elimination of the selection bias is conditioned by the CIA on complete covariate measurement, as all variables together influencing treatment and outcome must be observed, the propensity model achieves a pseudo- R^2 of 0.479 and AUC of 0.935, indicating that included covariates record a significant systematic change in the treatment condition.

A high AUC of 0.935 reveals that the model performs well and also shows that pensioners and non-pensioners are structurally different groups. This places a practical constraint on the matching, as the estimate will be based on a subset of pensioners for whom it is possible to find a comparable non-pensioner; the latter will be smaller than the total population. It is important to note that the complete history of contributions is the main unobserved concern, and the ECV does not measure it.

3. DATA AND ANALYTICAL SAMPLE

3.1 The ECV 2024

The data source is the 2024 Encuesta de Calidad de Vida, ECV, conducted by Colombia's National Administrative Department of Statistics (DANE) and released in 2025, the ECV is a cross-sectional household survey with national, urban, and rural representativeness. It is Colombia's official source of subjective wellbeing data and the only nationally representative survey that combines a multidimensional wellbeing module with detailed socioeconomic information and a pension-receipt question, the 2024 wave includes 228,688 individual observations across 557 variables, with a calibrated expansion factor (FEX_C) that enables population-level inference.

3.2 Sample construction

The analytical sample is restricted to older adults eligible for an old-age pension under Law 100 of 1993, Article 33, women aged 57 and over and men aged 62 and over. These thresholds correspond to the minimum retirement ages of the Colombian contributory pension system. The pension eligibility age, rather than the conventional cutoff of 60 used in much of the gerontology literature, is chosen because the study aims to capture the population that can legitimately access the contributory system, applying these age restrictions yields an eligible sample of 41,835 individuals (25,122 women and 16,713 men).

We dropped three types of observations with missing data. First, 169 people did not answer the pension question (P8642). Second, 619 had no response on health regime (P6100), this was the single covariate with the largest imbalance. Finally, 234 did not report their socioeconomic stratum (P8520S1A1). The last two categories accounted for most of the exclusions. The final analytical sample contains 40,813 individuals, retaining 97.6% of the eligible sample the 14.1% pensioner share (5,753 individuals) aligns with OECD (2023) estimates when adjusted for age and rural composition of the eligible sample.

Table 1. Sample construction

Step	N
Full ECV 2024 dataset	228,688
Eligible sample (women ≥ 57 , men ≥ 62)	41,835
Women ≥ 57 years	25,122 (60.1%)
Men ≥ 62 years	16,713 (39.9%)
Exclusions for DK/NR (P8642, P6100, P8520S1A1)	- 1,022
Final analytical sample	40,813 (97.6% retained)
Pensioners (P8642 = 1)	5,753 (14.1%)
Non-pensioners (P8642 = 2)	35,060 (85.9%)

3.3 Treatment variable

The treatment variable is binary and is derived from question 48 of the labor force module (variable P8642): "Last month, did you receive any income from retirement pension, survivor's pension, disability pension, or old-age pension?" An individual is classified as treated (pension = 1) if the response is "Yes" and as control (pension = 0) if the response is "No."

The treatment variable conflates distinct pension types (old-age, survivor's, disability), a key design constraint, the treatment is heterogeneous by design, the ECV 2024 does not allow disaggregation by pension type, so old-age, retirement, survivor's, and disability pensions are all coded as treatment. These pensions are not equivalent, as a survivor's pension may be linked to widowhood, a disability pension may be linked to worse baseline health, and an old-age pension may be linked to a long formal labor trajectory.

Grouping is necessary because of the data, but this implies that the estimate should be understood as a relationship with receiving pension income, rather than being the result of a specific type of pension. Furthermore, the treatment takes into account the receipt of the pension in the previous month, not the pension status as a phase of the life cycle, therefore independent validation is necessary using P6920 ("Are you currently contributing to a pension fund?" with response 3 = "Already a pensioner") identified 665 discordant cases where individuals report not having received pension income in the previous month but report being pensioners, therefore, according to the operational definition, they are considered non-pensioners, since their presence may indicate temporary interruptions in payments or measurement errors, and both factors would reduce the estimated ATT to zero.

This limitation is especially relevant for the interpretation of health and emotional wellbeing outcomes, as if some treated individuals receive disability pensions, their baseline health may be worse than that of comparable non-pensioners, and this could attenuate the estimated association between pension receipt and health.

3.4 Outcome variables

The primary outcome is a composite subjective wellbeing index constructed from 12 indicators in the ECV 2024 wellbeing module, these 12 indicators cover the three dimensions of the OECD (2013) framework, life satisfaction, Cantril ladder, happiness, worry, sadness, life worthwhileness, income satisfaction, work satisfaction, free time satisfaction, health satisfaction, security satisfaction, and self-reported general health.

As a secondary and exploratory analysis, the ATT is estimated separately for each of the twelve individual indicators, with Bonferroni correction for multiple comparisons, this disaggregated analysis identifies which wellbeing dimensions are more strongly associated with pension receipt, and which remain without significant effect, providing suggestive evidence on possible mechanisms.

3.5 Covariates

Covariates are characteristics theoretically associated with both pension access and subjective wellbeing, for some, clearly precede pension receipt, such as sex, age, education, literacy, ethnicity, and area of residence, for others are measured at the same time as pension receipt, such as health regime, main activity, chronic illness, and socioeconomic stratum, so these variables are included because they proxy long-run socioeconomic and labor-market trajectories, but they should not be interpreted as perfectly pre-treatment.

Literacy and education reveal the path of human capital toward formal employment, while health insurance coverage highlights the divide between formal and informal labor markets. Furthermore, area of residence, employment status, and socioeconomic level reflect current economic circumstances and demographic factors gather data on life cycle aspects and identity, and considering chronic illness as an indirect indicator of baseline health, although its causal relationship is uncertain. Pensioners may be more likely to receive a formal diagnosis through the contributory health system, leading to an asymmetry in measurements rather than a confounding effect.

Table 2. Key variables used in the analysis

Variable	Label in results	Role	Literal question (ECV 2024)	Response
P8642		Treatment	Last month, did you receive any pension income?	1 = Yes, 2 = No
P6040		Covariate	How old are you?	57-98 (continuous)
P6020		Covariate	Sex at birth	1 = Male, 2 = Female
P5502		Covariate	Marital status	1-6
P6080		Covariate	Ethnic self-identification	Indígena / Afro / None / Other
P6100		Covariate	Health insurance regime	Contrib. / Special / Subsidized
P6240		Covariate	Main activity last week	Working / Seeking / Household / Other
P1930		Covariate	Diagnosed with chronic illness (12 months)	1 = Yes, 2 = No
P8520S1A1		Covariate	Socioeconomic stratum (utility bill)	0-6, recoded
P6160		Covariate	Can read and write	1 = Yes, 2 = No
P8587		Covariate	Highest education level attained	1-13 ordinal
CLASE		Covariate	Area classification (design variable)	1 = Urban, 2 = Rest
P1895	Life satisfaction	Outcome	How satisfied are you with your life?	0-10
P1927	Cantril ladder	Outcome	Cantril Ladder: current step	0-10
P1905	Worthwhile	Outcome	Things in your life are worthwhile	0-10
P1901	Happiness yesterday	Outcome	How happy did you feel yesterday?	0-10

P1903	Worry (inverted)	Outcome	How worried were you yesterday? (inverted)	0-10
P1904	Sadness (inverted)	Outcome	How sad were you yesterday? (inverted)	0-10
P1896	Income satisfaction	Outcome	Satisfaction with income (99 = NA)	0-10
P1897	Health satisfaction	Outcome	Satisfaction with health	0-10
P1898	Security satisfaction	Outcome	Satisfaction with security	0-10
P1899	Work/activity satisfaction	Outcome	Satisfaction with work/activity	0-10
P3175	Free time satisfaction	Outcome	Satisfaction with free time	0-10
P6127	Self-reported health	Outcome	General health status (inverted)	1-4
FEX_C		Design	Calibrated expansion factor	Continuous

Source: own elaboration based on the ECV 2024 questionnaire and the DDI documentation (DANE, 2026). Inverted outcomes: P1903, P1904, P6127 recoded so that higher values indicate greater wellbeing.

4. METHODOLOGY

4.1 Identification strategy

To identify the effect, we use Propensity Score Matching (Rosenbaum & Rubin, 1983), the specific quantity we estimate is called the Average Treatment Effect on the Treated ATT, define it as:

$$ATT = E[Y(1) - Y(0) | D = 1]$$

where $Y(1)$ and $Y(0)$ denote the potential outcomes under treatment and control, respectively, and D is the treatment indicator ($pension = 1$).

Identification of the ATT under PSM rests on two key assumptions. The conditional independence assumption:

$$(Y(0), Y(1)) \perp D | X.$$

The CIA is not directly testable; And the overlap or common support assumption is:

$$0 < Pr(D = 1 | X) < 1 \text{ for all } X \text{ in the support.}$$

4.2 The propensity score and its logit transformation

Matching on the propensity score directly creates two practical problems, the visual assessment fails because non-pensioners spike near zero, the control distribution is compressed into $[0, 0.05]$ while the treated distribution spans the full unit interval, making overlap appear much worse than it is. And the caliper problem is more subtle on the probability scale, the same numerical distance of 0.05 spans a 250% range in underlying probability at the low end but only 10% near the center, so matching precision is inconsistent across the distribution, so the logit transformation resolves both at once.

The logit transformation maps the bounded $(0, 1)$ interval to the full real line $(-\infty, +\infty)$:

$$\text{logit}(p) = \log\left(\frac{p}{(1 - p)}\right)$$

The overlap region becomes visually assessable, and the caliper maintains uniform meaning across the distribution. Following Austin (2011), who demonstrates through simulation that a width of 0.2 standard deviations of the logit of the propensity score performs well across a wide range of sample sizes and treatment prevalence levels, all overlap diagnostics and matching in this thesis are implemented on the logit scale.

4.3 Propensity score model

The propensity score is estimated via logistic regression with eleven observed covariates, age, sex, marital status, ethnicity, health regime, main activity, chronic illness, socioeconomic stratum, literacy, education level, and area of residence, so the model includes 34 parameters when categorical variables are expanded using standard dummy-coding. The reference

categories are, married (marital status), no ethnic identification (ethnicity), subsidized health regime, household duties, no stratum, and urban area.

$$\begin{aligned} \text{logit}(\text{Pr}(\text{pension} = 1 | X)) \\ = \beta_0 + \beta_1 \cdot \text{age} + \beta_2 \cdot \text{sex} + \beta_3 \cdot \text{marital} + \beta_4 \cdot \text{ethnicity} + \beta_5 \\ \cdot \text{health_regime} + \beta_6 \cdot \text{activity} + \beta_7 \cdot \text{chronic} + \beta_8 \cdot \text{stratum} + \beta_9 \\ \cdot \text{literacy} + \beta_{10} \cdot \text{education} + \beta_{11} \cdot \text{area}. \end{aligned}$$

The model is estimated by maximum likelihood using the iteratively reweighted least squares (IRLS) algorithm, the contribution history variable P6920 is excluded from the propensity model because its pre-matching SMD of 2.730 would induce quasi-perfect separation in the logit; it is retained only as a post-matching validation variable.

4.4 Matching procedure

Matching is performed by nearest-neighbor 1:1 without replacement on the logit of the propensity score, with a caliper of 0.2 standard deviations of the logit (Austin, 2011). The algorithm is implemented in the R package MatchIt (Ho et al., 2011). Pensioners for whom no control unit lies within the caliper are excluded from the matched sample, avoiding extrapolation into regions of poor overlap. These excluded units represent pensioners who are systematically different from any non-pensioner in the eligible sample; excluding them is necessary to avoid comparing incomparable groups.

This exclusion rule has two implications, one improves internal validity by ensuring comparisons occur in the region of common support where causal inference is more defensible, and the other one, reduces external validity, the matched-sample ATT characterizes only those pensioners for whom a similar non-pensioner exists, not all pensioners in Colombia.

4.5 Outcome model and standard errors

The ATT is estimated by doubly robust OLS on the matched sample:

$$Y_i = \alpha + \beta D_i + \gamma X_i + \varepsilon_i$$

where Y_i is the outcome variable, D_i is the treatment indicator, X_i contains the same covariates as the propensity model, and β is the estimated ATT. Including the covariates in the outcome regression on the matched sample provides double robustness, the estimator is consistent if either the propensity score model or the outcome model is correctly specified (Ho et al., 2007; Imbens & Wooldridge, 2009). Standard errors are clustered at the matched-pair (subclass) level using the cluster-robust covariance estimator `vcovCL`, which accounts for the within-pair dependence induced by the matching procedure (Abadie & Spiess, 2022).

As a supplementary diagnostic, the proportional bias reduction for each wellbeing dimension is computed as:

$$1 - \left(\frac{ATT_{\{PSM\}}}{\Delta_{\{raw\}}} \right)$$

where $ATT_{\{PSM\}}$ is the doubly robust matched estimate and $\Delta_{\{raw\}}$ is the unweighted pre-matching difference in means between pensioners and non-pensioners for that outcome.

4.6 Sensitivity analysis

Two complementary sensitivity tools are applied to assess robustness to hidden bias, one is Rosenbaum bounds (Rosenbaum, 2002) which quantify how much an unobserved confounder would need to alter the odds of treatment, holding observed covariates fixed, to nullify the statistical significance of the estimated effect, and the critical Γ is the value of the odds-ratio at which the upper-bound p-value reaches 0.05. E-values (VanderWeele & Ding, 2017) report the minimum strength of association, on the risk-ratio scale, that an unmeasured confounder would need to have with both treatment and outcome to fully explain away the observed estimate.

4.7 Heterogeneity analysis

Heterogeneity in the ATT is examined across four pre-specified subgroups, sex, area of residence, education level, and socioeconomic stratum. For each subgroup, an independent PSM is performed on the relevant subset, excluding the stratifying variable from the propensity model within the subgroup. Subgroups in which the propensity model fails to converge due to insufficient variance in the covariates are reported as non-estimable. Since they are based on independent matching procedures and not on a formal interaction model estimated on the full matched sample, differences across groups should be interpreted as suggestive rather than conclusive.

4.8 Software and reproducibility

All evaluations are performed in R. The cobalt package is used for balance diagnostics; MatchIt 4.x (Ho et al., 2011) for the matching algorithm; lmtest and sandwich (Zeileis, 2004) for doubly robust standard errors; rbounds for Rosenbaum limits; and pROC for ROC and AUC diagnostics. The workflow is structured in script sequences (01_merge.R to 08_heterogeneity.R), in addition to an exploratory script (phase8_spillover.R).

5. DESCRIPTIVE ANALYSIS AND COMPOSITE INDEX CONSTRUCTION

5.1 Sociodemographic characteristics by pension status

Table 3 presents weighted descriptive statistics for the analytical sample by pension status. The use of standardized mean differences (SMD) rather than significance tests reflects a deliberate methodological choice. With $N = 40,813$, virtually any difference, however small in magnitude, would be statistically significant at conventional levels. The SMD measures imbalance in standardized units and is independent of sample size. The threshold $|SMD| > 0.1$, following Rubin (2001), is standard in the propensity score matching literature for identifying covariates with meaningful imbalance. For continuous variables, the SMD is calculated as the difference in group means divided by the pooled standard deviation; for binary variables, the denominator uses the binomial expression $p(1 - p)$.

Table 3. Sociodemographic characteristics by pension status (weighted by FEX_C)

Variable	Total	Pensioners	Non-pens.	SMD
Age (years)	69.46 (8.44)	71.44 (8.02)	69.15 (8.52)	0.296
Education (1-13)	3.70 (2.52)	6.06 (3.65)	3.33 (2.06)	0.924
Subsidized regime	69.6%	4.9%	79.7%	2.236
Contributory regime	28.2%	85.6%	19.2%	1.719
Middle stratum (3-4)	7.6%	23.4%	5.1%	0.561
Activity: Working	29.8%	11.7%	32.6%	0.522
Literacy (yes)	84.1%	97.0%	82.1%	0.508
Area: Urban	30.3%	50.8%	27.2%	0.499
Sex: Male	39.4%	45.4%	38.5%	0.157
N (unweighted)	40,813	5,753	35,060	

Notes: Weighted proportions are shown for categorical variables, and the mean (standard deviation) is shown for continuous variables. SMD = normalized difference of means, calculated by Rubin (2001). The variable contrib_status (P6920), which has an SMD of 2.730, is omitted from the propensity model to prevent near-perfect separation. Source: own calculations on ECV 2024 (DANE, 2025).

Thirty-one of the thirty-four covariate-parameters showed meaningful imbalance ($SMD > 0.1$). Health regime had the largest gap ($SMD = 2.236$), almost all pensioners were in the contributory regime, almost no non-pensioners were. This pattern reflects the institutional reality of the Colombian pension system, access requires formal employment, which generates both contributory health regime affiliation and pension contribution weeks. Education followed ($SMD = 0.924$), then stratum, activity, literacy, and urban residence (all between 0.5 and 0.6).

5.2 Construction of the composite wellbeing index

All indicators are first oriented so that higher values denote greater wellbeing. Worry (P1903) and sadness (P1904) are inverted as $worried_inv = 10 - P1903$ and $sad_inv = 10 - P1904$, and self-reported health (P6127) as $srh = 5 - P6127$. Each indicator is then standardized using the mean and standard deviation of the full analytical sample ($N = 40,813$) before any matching, so that the index is interpretable in standard-deviation units relative to the pre-treatment population. The index is the simple arithmetic mean of available standardized scores:

$$I_i = \left(\frac{1}{K_i}\right) \sum_k z_{ki}$$

where K_i is the number of available indicators for individual i . Of the 40,813 observations, 34,824 (85.3%) have all twelve indicators ($K = 12$); 5,551 (13.6%) have $K_i = 11$ because they answered P1896 = 99 ("no income to evaluate"), which is recoded as missing.

The asymmetric missingness on P1896 is mechanically driven, 17.3% of matched non-pensioners have missing income-satisfaction data because they report no income to evaluate, compared to 1.5% of pensioners. This compositional asymmetry is addressed through two robustness checks, the ATT is re-estimated using an 11-item index that excludes P1896, and the sample is restricted to observations with complete data on all twelve indicators, these both checks confirm the stability of the main result.

Measured using Cronbach's alpha coefficient (Cronbach, 1951), the index has an internal consistency of 0.885, exceeding the conventional limits of 0.7 (acceptable) and 0.8 (good). All twelve indicators exhibit a consistent underlying structure, the equal weighting method is preferred over other options, such as principal component analysis. Unlike PCA weighting, equal weighting does not require any assumptions about the underlying factor structure and is resilient to variations in the set of indicators, qualities that are relevant due to the correlations among these factors.

The distribution of the composite index before matching is shown in Figure 1, where shows that the gross difference decreased by 65%, from 0.279 SD to 0.097 SD after matching. This demonstrates that the determining factor in much of the observed gap is the pensioner's selection, not the pension amount itself.

Distribution of the composite wellbeing index by pension status

Pre-matching raw difference: 0.279 SD | Cronbach alfa = 0.885



Figure 1. Distribution of the composite wellbeing index by pension status (pre-matching).

Note: 17.3% of non-pensioners have missing values in P1896 (income satisfaction, coded as 99 in the ECV) and contribute to this distribution with an average of 11 items, instead of the total index of 12 items. The distribution presented here is robust when the sample is limited to complete cases ($K = 12$ only).

6. RESULTS

6.1 Propensity score model

The logistic propensity model is estimated on 40,786 observations (27 dropped for missing values on education, P8587). The model converges via the standard iteratively reweighted least squares (IRLS) algorithm. Goodness-of-fit indicators are summarized in Table 4. The McFadden pseudo- R^2 is 0.479, the Nagelkerke R^2 is 0.579, and the AUC-ROC is 0.935. An AUC of 0.935, meaning the model correctly distinguishes pensioners from non-pensioners 93.5% of the time, is well above the 0.7–0.8 range typically considered acceptable (Austin, 2011). The eleven covariates effectively capture the structural determinants of pension access in Colombia, who had formal employment, their residential area, and education level. The common support region is [0.001, 0.967] in probability units, containing 98.9% of the sample.

Table 4. Propensity score model: summary statistics

Indicator	Value
N observations	40,786
N pensioners ($D = 1$)	5,750 (14.1%)
Covariates	11 variables (34 parameters)
Log-likelihood (null)	- 16,177.4
Log-likelihood (model)	- 8,420.7
AIC	16,911.4
McFadden pseudo- R^2	0.479
Nagelkerke R^2	0.579
AUC-ROC (C-statistic)	0.935
Common support	[0.001, 0.967] - 98.9% of sample

Source: own calculations on ECV 2024 (DANE, 2026).

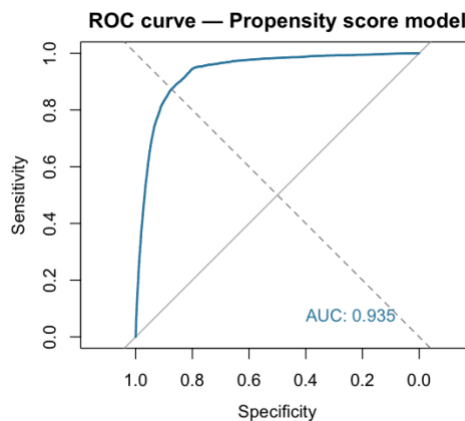


Figure 2. ROC curve of the propensity score model (AUC = 0.935).

Table 5 reports the odds ratios from the logistic regression. The strongest predictor is health regime. Compared to someone in the subsidized regime, a person in the contributory regime is 56.8 times more likely to have a pension (OR = 56.8), as this makes sense institutionally, being in the contributory health regime means you had formal employment, and formal employment is the only real path to a pension in Colombia, the education (OR = 1.12 per scale unit), urban residence (rural areas have lower odds, OR = 0.70), and middle and high stratum (OR = 1.92 and 2.15 respectively) round out the strong predictors. The chronic-illness coefficient (OR = 1.21) is interpreted with caution, it most likely reflects reverse causality, pensioners have better access to the contributory health regime and therefore a higher probability of receiving a formal diagnosis rather than a true higher prevalence of disease.

Table 5. Logistic regression: propensity score model coefficients

Covariate	Odds Ratio [95% CI]
Age (P6040)	1.013*** [1.006, 1.019]
Male (ref: Female)	1.504*** [1.353, 1.672]
Marital: Widowed (ref: Married)	0.700*** [0.612, 0.800]
Marital: Separated	0.827** [0.720, 0.949]
Marital: Single	0.620*** [0.524, 0.733]
Ethnicity: Afrodescendant (ref: None)	0.771** [0.615, 0.968]
Ethnicity: Ethnic minority	0.555*** [0.391, 0.787]
Health regime: Contributory (ref: Subsidized)	56.82*** [48.36, 66.76]
Health regime: Special	43.89*** [32.75, 58.82]
Activity: Working (ref: Household)	0.486*** [0.424, 0.556]
Activity: Job-seeking	0.523*** [0.337, 0.812]
Activity: Disabled	0.750* [0.567, 0.992]
Activity: Other	0.583*** [0.496, 0.684]
Chronic illness: Yes (ref: No)	1.211*** [1.095, 1.340]
Stratum: Low (ref: No stratum)	1.346*** [1.171, 1.547]
Stratum: Middle	1.924*** [1.577, 2.348]
Stratum: High	2.152*** [1.562, 2.966]
Illiterate (ref: Literate)	0.568*** [0.456, 0.707]
Education (P8587, 1-13)	1.121*** [1.096, 1.147]
Area: Rural/populated center (ref: Urban)	0.703*** [0.613, 0.807]
Area: No info/rural dispersed	0.847** [0.744, 0.964]

*Notes: *p<0.1; **p<0.05; ***p<0.01. Source: own calculations on ECV 2024 (DANE, 2026). N = 40,786. Coefficients are odds ratios.*

6.2 Common support: probability scale vs. logit scale

Figure 3 presents the propensity score distribution on the probability scale. The non-pensioners form a dense spike near zero, with the control distribution compressed into $[0, 0.05]$ while the treated distribution spans $[0, 1]$. This visual representation is misleading for assessing overlap.

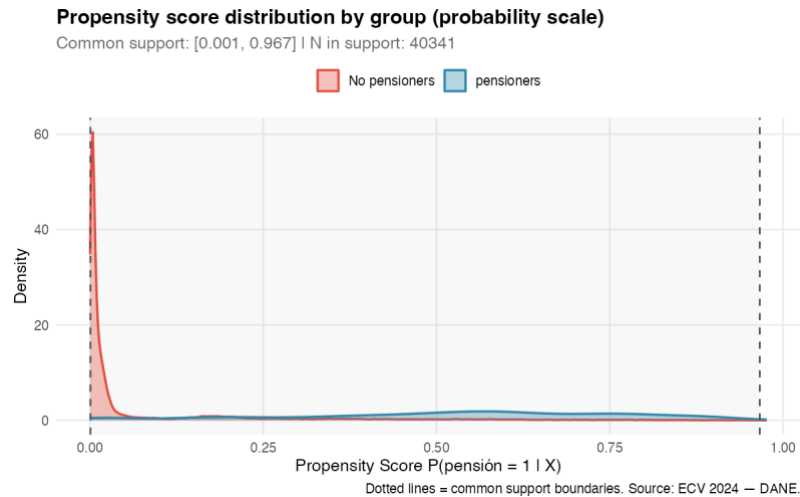


Figure 3. Propensity score distribution by group, probability scale.

The same data, but on the logit scale, are shown in Figure 4. The logit scale reveals two distributions that have significant overlap in the range of $[-2, +2]$ and a significant distance. People who do not receive a pension cluster around a logit of approximately -3.8 (with a probability of about 0.02), while those who do receive a pension cluster around a logit of approximately 0 (with a probability of about 0.5). There is overlap in the interval $[-2, +2]$, and the logit scale is used to determine the caliber of 0.2 standard deviations and provides an appropriate measure for assessing common support.

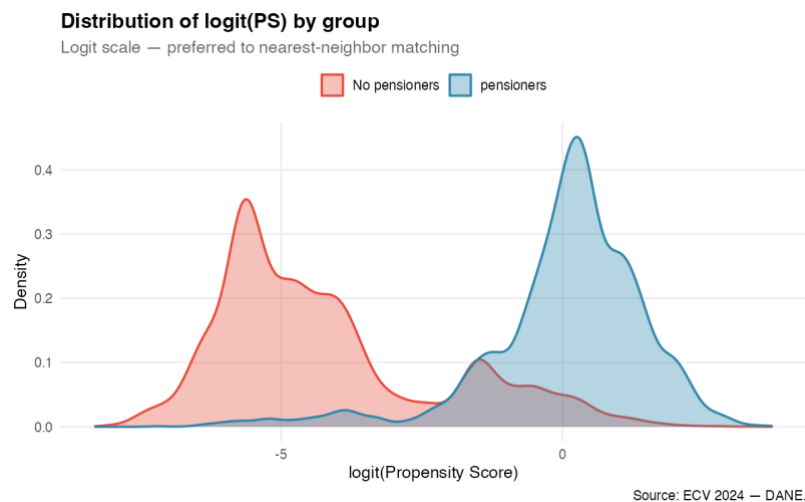


Figure 4. Distribution of logit(PS) by group: the scale used for matching.

The symmetrical histogram shown in Figure 5 places pensioners (5750) above the y-axis, while non-pensioners (35036) are below it. This skewness is characteristic of the low baseline treatment rate in this group and explains why matching preserves only 70.7% of pensioners: the right-hand side of the pensioner distribution lies outside the logit interval of 0.2 standard deviations and lacks equivalent elements among non-pensioners.

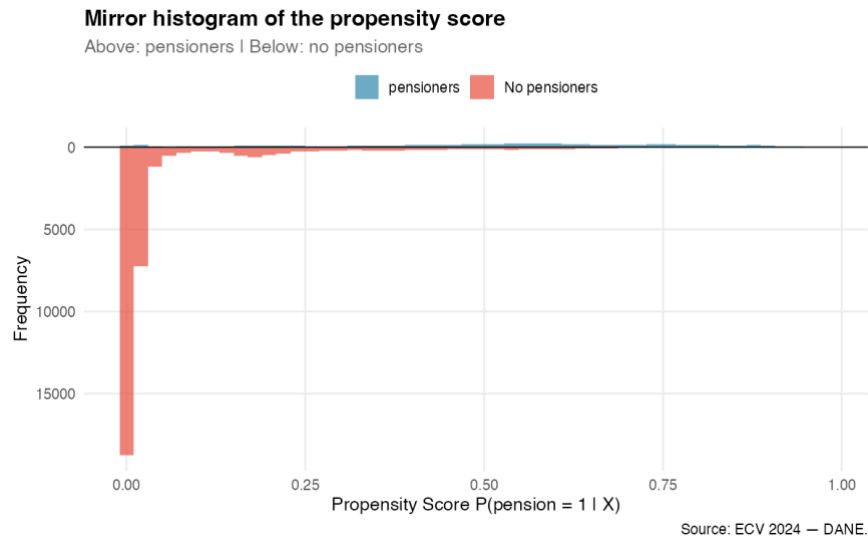


Figure 5. Mirror histogram of the propensity score by group.

6.3 Matching results and post-matching balance

The 1:1 nearest neighbor comparison without replacement, using a logit interval of 0.2 standard deviations, successfully matches 4066 of the 5750 pensioners (a retention rate of 70.7%), resulting in a matched sample of 4066 pairs were observed (8132 observations). The 1684 unmatched pensioners are those whose propensity scores are present in areas that have poor overlap with the control distribution; their exclusion is an essential methodological expense to prevent extrapolation.

Figure 6 displays the Love plot comparing pre- and post-matching standardized mean differences for all 34 covariate-parameters in the propensity model. Pre-matching, the seven most imbalanced covariates exceed $\text{SMD} = 0.3$, with education and the subsidized health regime exceeding 0.7. Post-matching, 32 of the 34 covariates fall below the $\text{SMD} = 0.1$ threshold; the two remaining covariates have post-matching SMDs marginally above the threshold but are controlled for in the doubly robust outcome regression.

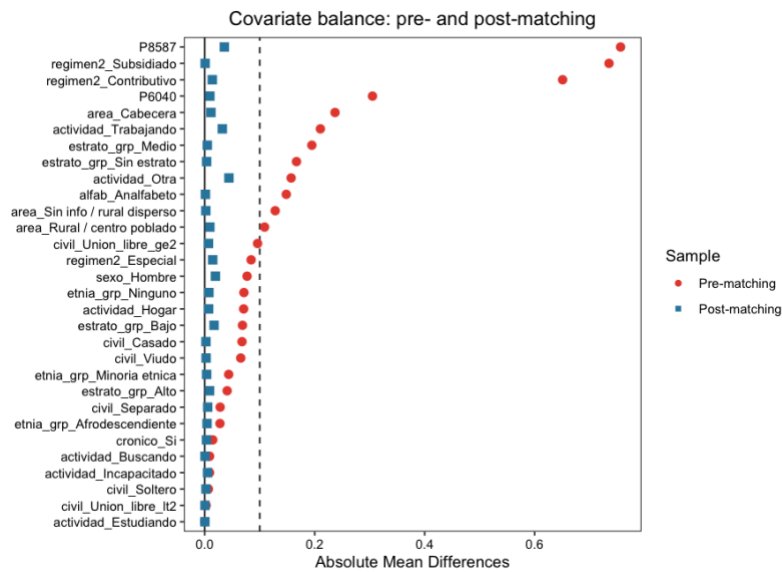


Figure 6. Love plot: covariate balance pre and post matching.

Dashed line: $SMD = 0.10$ threshold (Rubin, 2001).

Figures 7 and 8 reveal the distributional balance of the two covariates that showed the greatest imbalance before matching, the healthcare system and education. In both situations, disjoint distributions are visually observed before matching pensioners and non-pensioners: the majority of pensioners are in the contributory system and have higher levels of education, while non-pensioners predominate in the subsidized system and have lower levels of education. After matching, the two distributions are, they are practically identical, demonstrating that the matching process has been able to equalize the groups of retired and non-retired individuals in these fundamental dimensions.

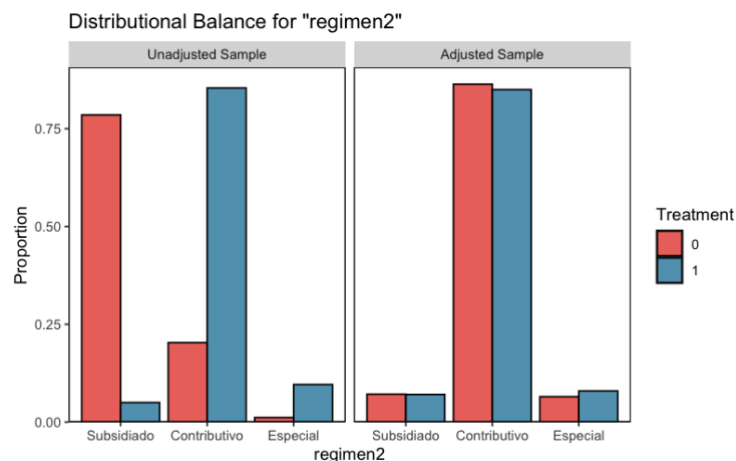


Figure 7. Distributional balance for the health insurance regime, pre and post matching

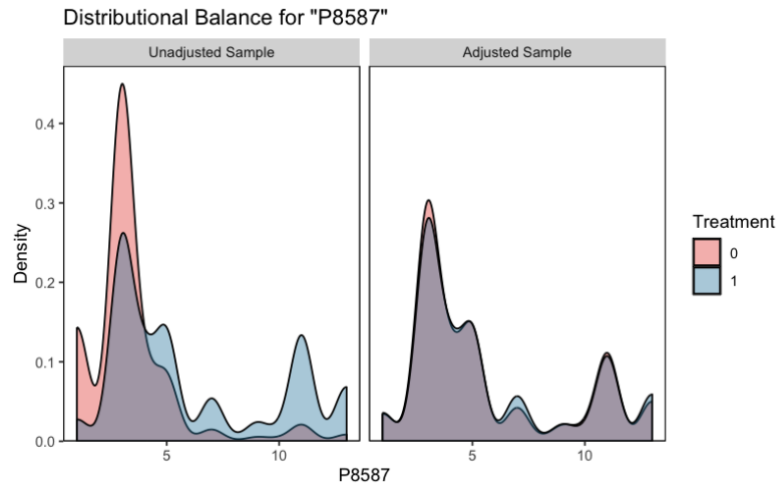


Figure 8. Distributional balance for education level (P8587), pre and post matching.

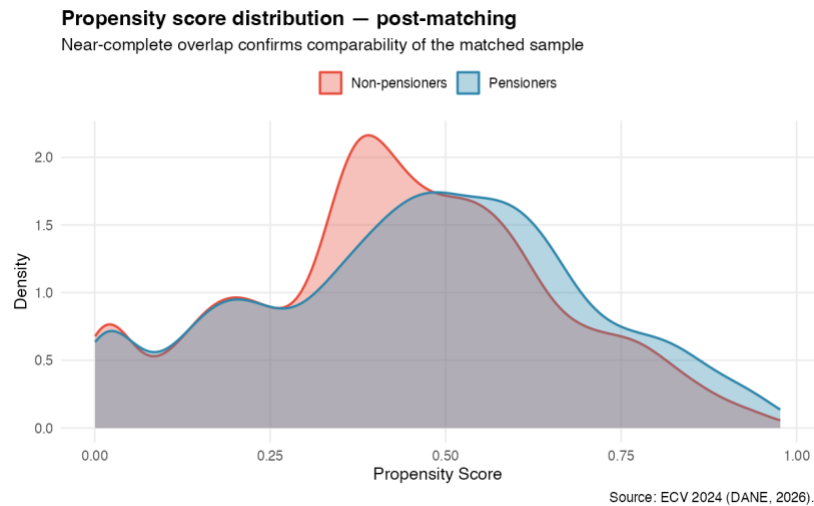


Figure 9. Propensity score distribution post-matching: near-complete overlap.

6.4 Main result: ATT on the composite wellbeing index

Table 6 reports the main result. The doubly robust OLS regression of the composite wellbeing index on the treatment indicator and the propensity-model covariates, with standard errors clustered by matched pair, yields an estimated ATT of +0.097 standard deviations (95% CI: [0.072, 0.123], $p < 0.001$).

The raw outcome gap between pensioners and non-pensioners before matching is 0.279 SD. Matching reduces this to 0.097 SD. The 65% reduction in the gap reflects selection on observable characteristics, pensioners had higher education, worked in formal employment, and lived in urban areas more often than non-pensioners, what remains after accounting for these differences is the 0.097 SD figure, which is more plausibly attributable to pension receipt itself under the conditional independence assumption.

Table 6. Doubly robust OLS regression: ATT on the composite wellbeing index

	Composite Wellbeing Index (z-score)
Pension (D = 1)	0.097*** (0.013) [0.072, 0.123]
Covariates	All PS model covariates
Matching	NNM 1:1, caliper = 0.2 SD logit(PS)
Standard errors	Clustered by matched pair (vcovCL)
Observations	8,132
Matched pairs	4,066
Raw difference (pre-PSM)	+ 0.279 SD
Bias reduction	65.2%

Notes: *** $p < 0.01$. Robust standard errors in parentheses, 95% CI in brackets. Source: own calculations on ECV 2024 (DANE, 2026).

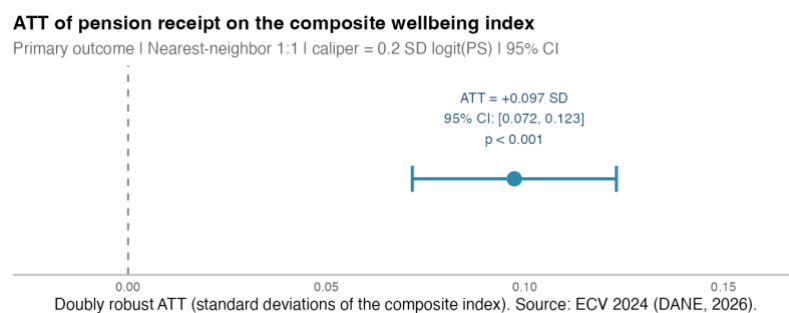


Figure 10. Forest plot of the ATT on the composite wellbeing index.

6.5 Disaggregated results: twelve wellbeing dimensions

Income satisfaction shows the largest effect (+0.664 points), followed by work satisfaction (+0.433 points); both remain significant after Bonferroni correction for twelve simultaneous tests. Life evaluation improves across multiple dimensions: the Cantril ladder, life satisfaction, and happiness all show positive associations. By contrast, day-to-day emotional dimensions (worry, sadness) show smaller effects that are nominally significant at $p < 0.05$ but do not survive Bonferroni correction. Two dimensions show no significant effect: security satisfaction and health.

Income transfers change how people assess their economic situation and their lives overall, they do not measurably affect emotional experience or physical health status.

Table 7. ATT estimates by wellbeing dimension

Outcome	ATT (original scale)	SE	p-value	Bonferroni
Income satisfaction	0.664	0.050	< 0.001	Sig.
Work/activity satisfaction	0.433	0.049	< 0.001	Sig.
Free time satisfaction	0.212	0.038	< 0.001	Sig.
Happiness yesterday	0.209	0.043	< 0.001	Sig.
Cantril ladder	0.181	0.035	< 0.001	Sig.
Life satisfaction	0.171	0.034	< 0.001	Sig.
Health satisfaction	0.165	0.042	< 0.001	Sig.
Worthwhile (eudaimonia)	0.110	0.033	< 0.001	Sig.
Worry (inverted)	0.144	0.063	0.023	p < 0.05
Sadness (inverted)	0.126	0.060	0.035	p < 0.05
Security satisfaction	0.032	0.046	0.487	n.s.
Self-reported health	0.020	0.013	0.111	n.s.

Notes: Robust standard errors clustered by matched pair. Bonferroni-corrected significance threshold: $\alpha = 0.05/12 = 0.0042$. Sig. = survives Bonferroni correction; $p < 0.05$ = nominally significant only; n.s. = not significant. Source: own calculations on ECV 2024 (DANE, 2026).

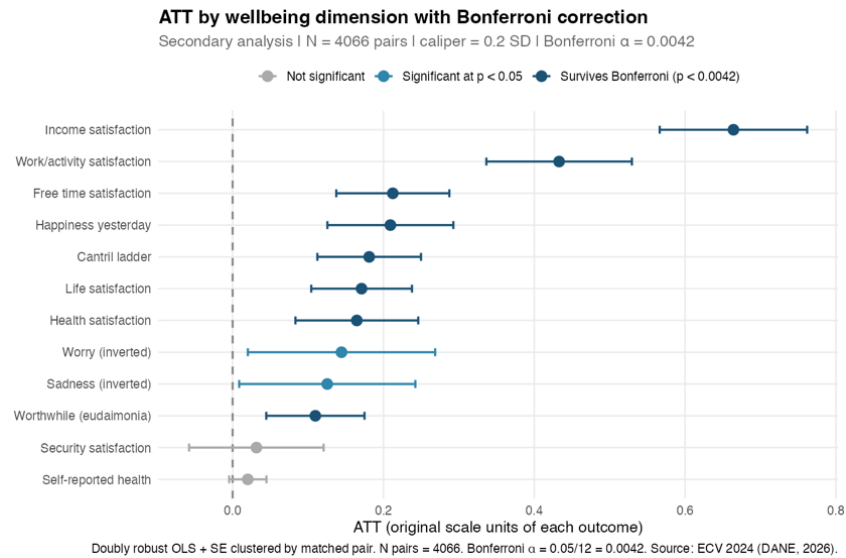


Figure 11. Forest plot of the ATT by wellbeing dimension with Bonferroni correction.

For security satisfaction and self-reported health, matching eliminates 82.7% and 75.3% of the raw gap, this bias reduction magnitude reflects selection on observables rather than pension effects, by contrast, on income and work satisfaction, bias reduction reaches only 45%, leaving substantial differences plausibly attributable to pension receipt.

6.6 Heterogeneity

Table 8 summarizes the heterogeneity analysis on the composite index. The specification converges for sex, area, and the low-education and low-stratum groups. The higher-education and middle-to-high-stratum subgroups do not converge due to insufficient within-subgroup variance: there are essentially no non-pensioners with tertiary education or middle-to-high stratum in the eligible sample, making within-subgroup matching infeasible. Coarsened exact matching (Iacus, King, & Porro, 2012) is recommended for future work to address this constraint.

Table 8. ATT on composite wellbeing index by subgroup

Subgroup	N pairs	ATT (SD)	SE	95% CI
Global	4,066	0.097	0.013	[0.072, 0.123]
Men	1,520	0.134	0.022	[0.091, 0.177]
Women	2,415	0.077	0.017	[0.044, 0.110]
Urban	1,817	0.092	0.019	[0.055, 0.129]
Rural / No info	2,217	0.113	0.018	[0.078, 0.148]
No / Basic education	2,101	0.114	0.019	[0.077, 0.151]
No stratum / Low	3,073	0.098	0.015	[0.069, 0.127]

Notes: Each subgroup PSM is estimated independently with the stratifying variable removed from the propensity model. Source: own calculations on ECV 2024 (DANE, 2026).

The most salient feature is the gender gap: the composite index shows +0.134 SD for men and +0.077 SD for women, which equates to a ratio of about 1.7; the confidence intervals partially overlap; the difference is not definitive at conventional significance levels and the differential effect reflects differences in initial work experience. Men who were eligible to receive a pension at age 62 were more likely to have kept a formal job with a strong professional identity. Women who had the possibility of accessing it at age 57 more frequently reported that their career paths were interrupted, which was determined by care responsibilities and chronic lack of income stability. These different starting points entail different processes: for men, receiving the pension eliminates the disruption of losing their work identity; For women, it guarantees a minimum income. This pattern is consistent with this analysis, as the male advantage is particularly marked in happiness (+0.311 vs. +0.128) and on the Cantril scale (+0.305 vs. +0.109). A formal interaction model in the full matched sample would provide more rigorous evidence of these differentiating mechanisms.

Rural pensioners report almost twice the income satisfaction of their urban counterparts (+0.851 points compared to +0.475). This disparity shows that, in rural areas with fewer economic options and lower baseline incomes, a fixed transfer has a higher marginal benefit. Lower socioeconomic strata and subgroups with lower levels of education, which constitute the majority of the sample, exhibit similar or slightly higher ATT scores than the overall estimate.

This demonstrates that the link between pension perception and well-being is at least as strong among the most economically vulnerable as it is in the population as a whole.

Figures 12 through 14 display heatmaps of the ATT across all twelve dimensions for each subgroup, showing that the patterns of heterogeneity are consistent across multiple outcome domains.



Figure 12. Heterogeneity by sex: ATT across 13 outcomes (composite + 12 dimensions).

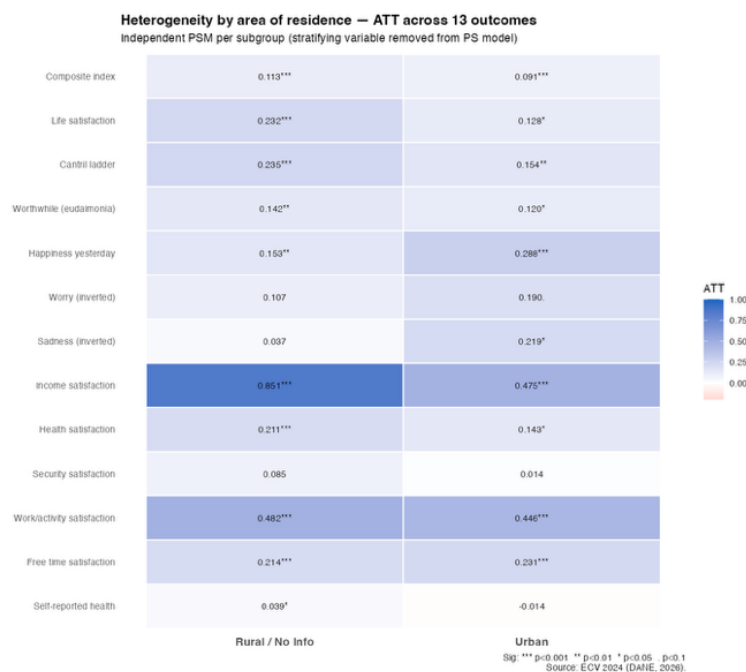


Figure 13. Heterogeneity by area of residence: ATT across 13 outcomes.

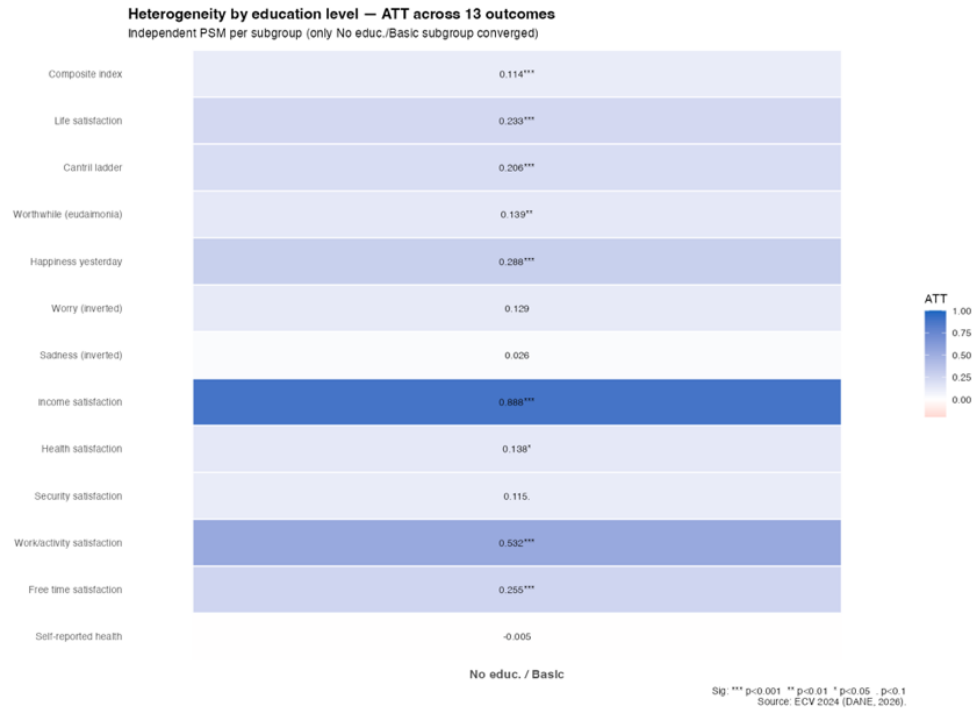


Figure 14. Heterogeneity by education level: only the No educ./Basic subgroup converged.

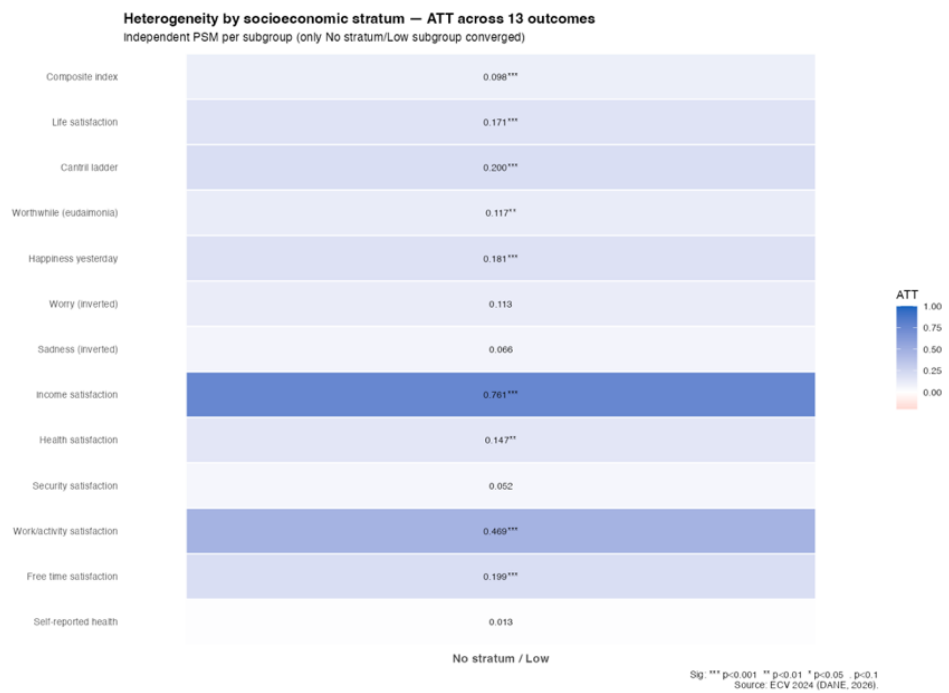


Figure 15. Heterogeneity by socioeconomic stratum: only the No stratum/Low subgroup converged.

7. SENSITIVITY ANALYSIS

7.1 Rosenbaum bounds

Rosenbaum (2002) bounds quantify how much an unobserved binary confounder would need to change the odds of treatment, holding observed covariates fixed, before the statistical significance of the estimated ATT would be lost. The critical Γ is the value at which the upper-bound p-value reaches 0.05. Table 9 reports critical Γ values for the composite index and the main individual dimensions.

Table 9. Rosenbaum bounds: critical Γ values

Outcome	ATT	Critical Γ
Composite index	0.097	1.3
Income satisfaction	0.664	1.6
Work satisfaction	0.433	1.4
Free time satisfaction	0.212	1.2
Happiness yesterday	0.209	1.2

Notes: Critical Γ = value at which the upper-bound p-value first exceeds 0.05. Only outcomes with statistically significant ATTs are reported. Source: own calculations using the *rbounds* R package on ECV 2024 (DANE, 2026).

The critical Γ for the composite index is 1.3. A confounder increasing the odds of pension receipt by 30% above what the covariates capture would nullify the result. A strong omitted variable, like complete formal employment history or accumulated wealth, could still alter the conclusion. However, two factors limit this concern. The propensity model already accounts for 93.5% of the differences between pensioners and non-pensioners (AUC = 0.935), which narrows the plausible range of large omitted factors, the individual components also provide a clearer picture: income satisfaction and work satisfaction have critical Γ values of 1.6 and 1.4 respectively, which are higher than the composite index because the index combines dimensions with different sensitivities.

7.2 E-values

The E-value (VanderWeele & Ding, 2017) is a complementary measure of robustness to unmeasured confounding. Figure 16 presents E-values for all outcomes with statistically significant ATTs. Only income satisfaction (E = 2.62) and work satisfaction (E = 2.10) clear the moderate-robustness threshold of E = 2.0 (VanderWeele & Ding, 2017). The composite index sits just below at E = 1.93.

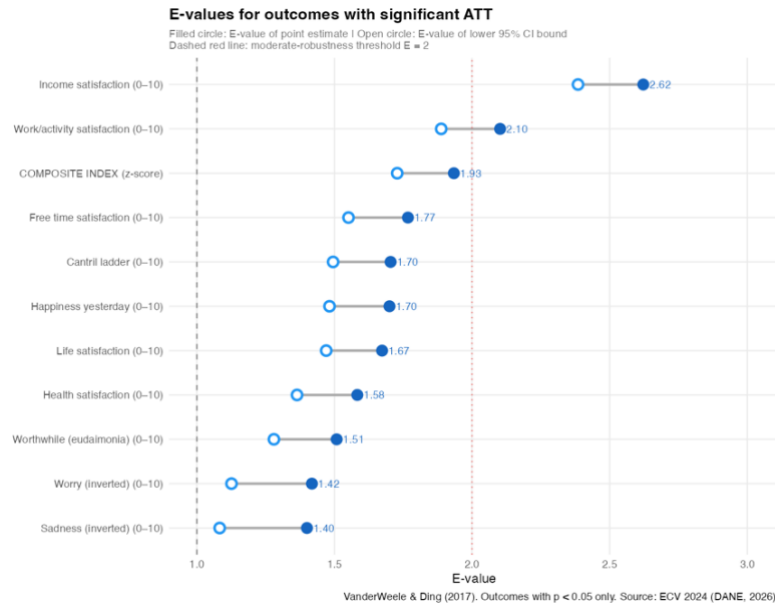


Figure 16. E-values for outcomes with significant ATT.

Filled circles: E-value of the point estimate. The dashed red line represents the moderate robustness threshold ($E = 2$), and the empty circles indicate the E-value of the lower limit of the 95% CI.

7.3 Robustness of the composite index to the exclusion of income satisfaction

Income satisfaction (P1896), which exhibits considerable compositional asymmetry, is included in the composite index. While 1.5% of pensioners have missing data in P1896 (code 99 for "no income"), 17.3% of matched controls lack such data; this represents a difference of approximately 15.8 percentage points. This asymmetry arises mechanically, as many non-pensioners report having no income to assess, whereas pensioners, by definition, have a pension to be evaluated. This concern is addressed with two robustness checks. The first re-estimates the ATT using an index composed of 11 items that does not include P1896. The second is limited to the sample of observations that have complete data for all twelve indicators; both checks corroborate the main result.

Table 10. Sensitivity of the composite index to alternative specifications

Specification	ATT (SD)	95% CI	N
12 items (main)	0.097	[0.072, 0.123]	8,132
11 items (excl. P1896)	0.081	[0.055, 0.107]	8,132
12 items, K = 12 only	0.098	[0.071, 0.125]	7,365

Notes: Main specification uses all 12 wellbeing indicators. 11-item index excludes income satisfaction (P1896). K = 12 only restricts to observations with complete data on all twelve indicators. Source: own calculations on ECV 2024 (DANE, 2026).

The 17% decrease in the estimate (from 0.097 to 0.081 SD) when omitting P1896 is the expected size because income satisfaction has the highest individual ATT (+0.664). It is worth noting that, despite the absence of this dimension, pension perception shows a significant positive correlation of +0.081 SD with the other eleven dimensions. The relationship between the indices of 11 and 12 items within the paired sample is 0.996. This sensitivity check confirms that the main result is not determined by a single component within the set of covariates analyzed.

8. DISCUSSION

8.1 Magnitude and pattern of the estimated effects

The main finding is that receiving a pension is tied to an increase in wellbeing by 0.097 standard deviations, as this is modest but significant and holds up in various tests ($\Gamma \approx 1.3$, $E = 1.93$). However, the raw gap is 0.279 SD, adjusting for health status, education, and socioeconomic stratum alone eliminates 65% of that gap.

Into perspective, 0.097 SD represents a small yet meaningful difference, as while it might not drastically change wellbeing, it is substantial enough to be noticeable in a national survey, comparatively, this effect is larger than typical findings in the European retirement literature, which range from 0.00 to 0.05 SD.

The effects are largest on income and work satisfaction, both survive Bonferroni correction, the life evaluation improves, but daily emotional experience (worry, sadness) does not survive correction, the minor effects on security satisfaction and self-reported health are not significant, and this decreasing trend from economic to emotional wellbeing reflects the Easterlin paradox (Easterlin, 1974; Kahneman & Deaton, 2010), income has a stronger correlation with cognitive life judgments than with emotional responses, while the lack of significant results for security satisfaction makes sense in Colombia, where feelings of insecurity are influenced by local conditions such as crime and conflict, which pension income cannot meaningfully change, the lack of significant results for self-reported health is more complicated; it may partly result from including disability pensions in the treatment variable, since these pensions are tied to worse baseline health, their inclusion could skew the health estimate toward zero.

The limited range of the 4-point health scale may also lower sensitivity, a similar pattern appears in the emotional aspects of wellbeing, worry and sadness, which are statistically significant at $p < 0.05$ but do not pass Bonferroni correction, while the absence of negative emotional effects in Colombia stands in contrast to the European retirement shock, the institutional difference accounts for this, Colombian pensioners with years of informal labor do not suffer from the loss of work identity because their work history was not centered around their identity, any retirement shock that may exist is lessened by the limited work attachment of most in the matched sample and the positive income impact that takes precedence in emotional responses.

8.2 Why the Colombian estimate exceeds the European benchmark

Before interpreting the comparison between the Colombian ATT of +0.097 SD and the European estimates close to zero, it is necessary to clarify a concept, the mechanism by which the perception of pensions works varies significantly between the two contexts, which can make a direct quantitative comparison misleading.

In Europe, a contributory pension replaces a wage. Most workers spend their careers in formal employment, accumulate the required contributions, and retire into a pension that substitutes their prior income stream, the coverage exceeds 80% across OECD countries (OECD,

2023), making pension receipt the expected outcome of a normal working life, it can be interpreted as when something is the norm, receiving it carries no additional signal beyond income continuity.

In Colombia, the pension does not replace income, for most recipients, it creates stable income for the first time, as just the 14.1% of the eligible sample receives a pension, and nationally fewer than 25% of adults over 60 do (OECD, 2023), this scarcity is structural, as with 58% of workers in informal employment (DANE, 2024a), most older Colombians never accumulate the 1,300 contribution weeks required under Law 100. Fernández and Villar (2017) demonstrated that high payroll tax rates actively sustained informality, reducing them cut informal employment by 4.8 percentage points, and Mejía and Forero (2022) found that fewer than one in three working-age Colombians contributes to a pension at any given moment, against this backdrop, receiving a pension signals the completion of a formal labor trajectory that most peers could not achieve, that signal, and the income security it carries, likely amplifies the wellbeing response beyond what pure income substitution would produce.

The largest influence (+0.664 points) is income satisfaction, which aligns with the perception that pension income represents a genuine new economic security, rather than a change in the source of payment. This idea is reinforced by the disparity between urban and rural areas: rural retirees report income satisfaction of +0.851 points, compared to +0.475 for their urban counterparts. This demonstrates that in environments with fewer economic opportunities and lower reference incomes, there is greater marginal utility, as Kahneman and Deaton (2010) and Easterlin (1974) argue regarding diminishing returns on income when evaluating life.

The hedonic dimensions, worry and sadness, do not survive multiple-comparison correction, which is consistent with the observation that retirement disrupts daily routine regardless of income level. The loss of structure and daily purpose that Ekerdt (1987) describes as the retirement shock is not eliminated by pension income; it operates as a separate mechanism. Pension receipt appears to shift how older adults evaluate their lives economically without fully resolving the adjustment in daily experience that retirement itself entails.

8.3 Comparison with other middle-income countries

The estimated ATT of +0.097 SD is consistent with results obtained in other middle-income countries. Galiani, Gertler, and Bando (2016) report that the Mexican program 70 y Más has had an impact of approximately 0.10 SD on subjective well-being and mental health indices. Bando, Galiani, and Gertler (2020) found an effect of 0.08 SD on subjective well-being from Peru's Pensión 65 program, with this effect being most significant in reducing depression. Case (2004) studied the household-level outcomes of expanding pensions for the elderly in South Africa and found similar effect sizes for nutritional and health indicators, which, although not directly comparable, were qualitatively similar. Despite different identification assumptions, the current matching estimator, regression discontinuity (Bando et al., 2020), and geographical deployment (Galiani et al., 2016) generate effect sizes ranging from 0.08 to 0.12 SD.

The Mexican and Peruvian studies focus on non-contributory pensions aimed at vulnerable groups, while this study examines the connection of contributory pensions among a broader

recipient sample. Therefore, the treated populations vary. Additionally, the Colombian study estimates the ATT for those who actually received a pension and matched a comparable control group (4,066 out of 5,750 pensioners). This limits its external validity with respect to the most economically privileged pensioners who could not be matched. These differences should be noted when comparing results, but they do not diminish the qualitative consistency of the regional pattern.

8.4 Mechanisms (preliminary)

The income mechanism provides the most explanation for the income-satisfaction effect, as the estimated ATT of +0.664 points for this one is the largest single estimate in this analysis, and its distribution matches the predictions of income-focused theory, the effects are more pronounced in areas with higher marginal utility, such as, rural areas (+0.851 points) exhibit almost double the urban effect (+0.475), and groups with lower education levels show effects at or above the overall average. For many pensioners in this sample, 79.7% of non-pensioners belong to the subsidized health regime, a reliable sign of informal work history, so, the pension likely represents the first stable, predictable monthly income of their adult lives, this interpretation aligns with structural evidence from Mejía and Forero (2022) about contribution density in informal labor markets and with results from comparative Latin American studies (Galiani et al., 2016; Bando et al., 2020).

The work satisfaction effect needs careful interpretation, as the ECV asks about any activity the respondent did last month, most pensioners in the sample do not work (OR = 0.49), so 'activity' may include volunteer work, household tasks, or hobbies, so the 0.433 point increase could show a shift in how pensioners view their time, as without pension income, work is driven by necessity; with it, people have discretionary time. This aligns with eudaimonic theories of wellbeing, which emphasize autonomy and self-determination (Ryff, 1989). The cross-sectional design does not allow for separating different explanations, it is unclear whether the effect shows real changes in preferences or if pensioners who continue to work report higher satisfaction because financial pressure has lessened.

The difference in composite ATT (+0.134 SD for men compared to +0.077 SD for women) suggests that different mechanisms operate across groups, rather than differing absolute welfare gains, while, men at pension age in this group tended to have steady formal employment histories and a strong sense of occupational identity, this strong work attachment is what Bonsang and Klein (2012) identify as a key factor in the retirement-wellbeing relationship in Europe. Women without pensions in this sample, on the other hand, often have interrupted job histories shaped by caregiving duties and ongoing financial instability, rather than a loss of occupational identity. A pension helps with financial instability by providing a stable income, but the pathway is different, it is about gaining income security rather than restructuring work identity.

9. POLICY IMPLICATIONS

The estimates presented are based on the assumption of conditional independence, which is not directly verifiable, and sensitivity analyses indicate moderate robustness against hidden biases (critical $\Gamma \approx 1.3$). These are the directions suggested by the data, so translating these estimates into resource allocation decisions would require additional longitudinal evidence and a formal cost-effectiveness analysis that is outside the scope of this study.

The disaggregated results of the ATT indicate that pension perception operates primarily through economic and evaluative channels, while income satisfaction and job satisfaction show the strongest associations (+0.664 and +0.433 points, respectively), both exceeding the Bonferroni correction, life satisfaction, happiness, and the Cantril scale also increase significantly. These results suggest that formal income transfers do reach the dimensions of well-being most directly linked to economic security, and that expanding contributory or semi-contributory pension coverage would likely produce quantifiable improvements in well-being.

Neither satisfaction with security nor self-reported health shows a significant association with pension perception in this analysis. In the case of satisfaction with security, the lack of effect is interpreted from an institutional perspective. Perceptions of security in Colombia are determined by territorial factors, crime rates, armed conflict, and displacement, factors that income transfers cannot significantly address, since pension perception modifies a person's economic situation but not the social and physical environment in which they live.

The absence of significant results in the self-assessment of health requires caution before drawing substantial conclusions, given that the treatment variable encompasses old-age, retirement, survivor, and disability pensions, and the latter are associated, by definition, with a worse initial health status, therefore, this compositional problem likely attenuates the estimated effect on health, bringing it close to zero, and renders this analysis an unreliable basis for health policy recommendations.

What the hedonic results suggest, worry and sadness, are nominally positive, but they do not withstand scrutiny, as pension income does not fully resolve the experiential adjustment that retirement entails, the transition from full-time work to retirement represents an abrupt change in daily structure, routine, and sense of purpose that income alone cannot address. These data highlight the importance of well-structured retirement transition programs that allow older adults to gradually move from full-time to part-time work and ultimately to full retirement, rather than facing a sudden and complete exit from working life. Developing such a program in Colombia would require collaboration between employers, the pension system, and public health entities.

Colombia is approaching a demographic transition in which the older adult population will grow substantially while the care infrastructure around it remains underdeveloped, and the results presented in this thesis are one piece of that broader picture: low pension coverage, high informality, and limited access to psychosocial support at the retirement transition constitute disadvantages that pension income alone cannot fully resolve, which suggests that addressing the wellbeing of older adults requires building the ecosystem around pension coverage, mental

health programs, social accompaniment, and modern care facilities, rather than treating income transfers as a sufficient response to what is fundamentally a multidimensional challenge.

10. CONCLUSIONS, LIMITATIONS, AND FUTURE RESEARCH

10.1 Summary of findings

This thesis estimates whether the perception of pension income is associated with greater subjective well-being among older adults in Colombia. The answer is affirmative, as the analysis decomposes this association into selection effects (differences prior to pension receipt) and the residual association with pension receipt itself, after adjusting for observable covariates.

The propensity score matching method was used in the 2024 ECV (N = 40,813, with 14.1% receiving pension income). The average association effect (ATT) on a composite well-being index was estimated to be +0.097 standard deviations (95% CI: [0.072, 0.123]). The difference between pensioners and non-pensioners before matching is 0.279 standard deviations, which is almost three times more.

The disaggregated findings are the central focus of the thesis, where the repercussions are most significant with respect to job satisfaction (+0.433) and income satisfaction (+0.664). Both are surpassed after Bonferroni correction for twelve simultaneous tests. Furthermore, improvements are observed in several aspects of life evaluation: both happiness and life contentment, as well as the Cantril scale, show significant positive correlations. However, daily emotional dimensions (such as sadness or worry) have smaller effects that do not prevail after correction; while the two dimensions that do not reflect a significant effect are satisfaction with health status and self-reported security.

This distinction is important because pensions provide income security, but this is directly related to the broader security that people perceive based on territorial and contextual factors that go beyond income. Furthermore, pensions do not address health through biological mechanisms. Therefore, complementary policies are suggested to achieve results that income alone cannot attain.

The sensitivity analysis indicates that the main result is true under moderate hidden bias, although it is not resistant to strong hidden confounding; meanwhile, Rosenbaum's limits reveal a critical Γ value close to 1.3 for the composite index. The E values are 2.10 for job satisfaction, 2.62 for income satisfaction, and 1.93 for the composite. Only job satisfaction and income exceed the conventional moderate robustness threshold of $E = 2.0$.

Men show larger composite ATTs than women (+0.134 vs. +0.077 SD), and rural residents show larger effects than urban residents (+0.113 vs. +0.092 SD). These patterns are derived from independent matching procedures rather than a formal interaction model and should be treated as hypothesis-generating rather than conclusive. Meanwhile, the higher education and middle-to-upper-class subgroups could not be estimated due to insufficient overlap within subgroups. A structural limitation of the nearest-neighbor PSM design that made exact matching less precise (Iacus, King, and Porro, 2012) could be addressed in future work.

10.2 Limitations

The most significant limitation is the assumption of conditional independence, which cannot be directly verified, as sensitivity analyses indicate moderate, but not strong, robustness against hidden biases (critical $\Gamma \approx 1.3$ for the composite index). The most plausible source of hidden bias is the lack of a complete employment and CVD contribution history up to 2024, while health insurance and education level serve as indirect indicators of formal employment trajectory, though these are not perfect.

The treatment variable groups together disability, retirement, old age and survivor pensions (P8642 does not differentiate them), which could mitigate the estimated impacts on health, given that disability pensions are related to a worse initial health condition. The cross-sectional design, by capturing only a moment in time and not trajectories, makes it impossible to estimate the changing effects of the transition to retirement, including temporal adaptation; for this, longitudinal designs would be required going forward.

The paired ATT is used for 4066 of the 5750 pensioners; the 1684 unmatched pensioners are those with the greatest economic privileges, and the estimates do not describe their experience. Survey expansion weights are used for descriptive statistics but not for matching. Therefore, the ATT describes the paired sample rather than providing a population-representative estimate, and comparisons of subgroup heterogeneity are based on independent PSMs and do not constitute formal evidence of interaction, an interaction model for the entire paired sample is deferred to future research.

10.3 Future research

Using the 4,066 matched pairs, the well-being of cohabiting household members can be examined as an additional outcome, since South African evidence (Case, 2004; Duflo, 2003) documents that pension transfers extend to household members. Therefore, estimating this indirect effect with the same rigorous identification applied to the individual-level analysis is a natural extension.

Once the new 2024 pension reform (Law 2381) is implemented, I would propose a difference-in-differences design that leverages the implementation in different municipalities or contribution week thresholds. This could validate and potentially strengthen the matching-based estimates presented here.

10.4 Closing remark

Colombia is undergoing a demographic transition in which the well-being of older adults is an increasingly central issue in public policy. Pensions, the primary instrument of protection for the elderly, have effects conditioned by the institutional structure that determines who receives them and under what conditions.

The institutional context, low coverage, high informality, and decades of precarious work, is not just background information for this study; it is what gives meaning to pensions for those who receive them and what justifies this estimation.

References

- Abadie, A., & Spiess, J. (2022). Robust post-matching inference. *Journal of the American Statistical Association*, 117(538), 983-995. <https://doi.org/10.1080/01621459.2020.1840383>
- Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association*, 103(484), 1481-1495. <https://doi.org/10.1198/016214508000000841>
- Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, 46(3), 399-424. <https://doi.org/10.1080/00273171.2011.568786>
- Bando, R., Galiani, S., & Gertler, P. (2020). The effects of non-contributory pensions on material and subjective well-being. *Economic Development and Cultural Change*, 68(4), 1233-1255. <https://doi.org/10.1086/702859>
- Bonsang, E., & Klein, T. J. (2012). Retirement and subjective well-being. *Journal of Economic Behavior & Organization*, 83(3), 311-329. <https://doi.org/10.1016/j.jebo.2012.06.002>
- Börsch-Supan, A., & Jürges, H. (2006). Early retirement, social security and well-being in Germany. NBER Working Paper No. 12303. National Bureau of Economic Research. <https://doi.org/10.3386/w12303>
- Bosch, M., Melguizo, Á., & Pagés, C. (2013). Mejores pensiones, mejores trabajos: Hacia la cobertura universal en América Latina y el Caribe. Inter-American Development Bank. <https://doi.org/10.18800/economia.201302.009>
- Case, A. (2004). Does money protect health status? Evidence from South African pensions. In D. A. Wise (Ed.), *Perspectives on the Economics of Aging* (pp. 287-312). University of Chicago Press.
- Coe, N. B., & Zamarro, G. (2011). Retirement effects on health in Europe. *Journal of Health Economics*, 30(1), 77-86. <https://doi.org/10.1016/j.jhealeco.2010.11.002>
- Congreso de la República de Colombia. (1993). Ley 100 de 1993, por la cual se crea el sistema de seguridad social integral y se dictan otras disposiciones. *Diario Oficial* No. 41.148.
- Congreso de la República de Colombia. (2024). Ley 2381 de 2024, por medio de la cual se establece el sistema de protección social integral para la vejez, invalidez y muerte de origen común, y se dictan otras disposiciones. *Diario Oficial* No. 52.819.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334. <https://doi.org/10.1007/BF02310555>
- DANE. (2024a). Gran Encuesta Integrada de Hogares, Boletín técnico. Departamento Administrativo Nacional de Estadística.

- DANE. (2024b). *Proyecciones de población nacional por área, sexo y edad*. Departamento Administrativo Nacional de Estadística.
- DANE. (2025). Encuesta Nacional de Calidad de Vida ECV 2024: Microdatos y documentación DDI. Departamento Administrativo Nacional de Estadística. <https://www.dane.gov.co/index.php/estadisticas-por-tema/salud/calidad-de-vida-ecv/encuesta-nacional-de-calidad-de-vida-ecv-2024>
- Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, 95(3), 542-575. <https://doi.org/10.1037/0033-2909.95.3.542>
- Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin*, 125(2), 276-302. <https://doi.org/10.1037/0033-2909.125.2.276>
- Duflo, E. (2003). Grandmothers and granddaughters: Old-age pensions and intrahousehold allocation in South Africa. *World Bank Economic Review*, 17(1), 1-25. <https://doi.org/10.1093/wber/lhg013>
- Easterlin, R. A. (1974). Does economic growth improve the human lot? Some empirical evidence. In P. A. David & M. W. Reder (Eds.), *Nations and Households in Economic Growth* (pp. 89-125). Academic Press.
- Ekerdt, D. J. (1987). Why the notion persists that retirement harms health. *The Gerontologist*, 27(4), 454-457. <https://doi.org/10.1093/geront/27.4.454>
- Fernández, C., & Villar, L. (2017). The impact of lowering the payroll tax on informality in Colombia. *Economía*, 18(1), 125-155. <https://doi.org/10.31389/eco.53>
- Galiani, S., Gertler, P., & Bando, R. (2016). Non-contributory pensions. *Labour Economics*, 38, 47-58. <https://doi.org/10.1016/j.labeco.2015.11.003>
- Graham, C., & Krauss, A. (2013, October 23). Subjective well-being in Colombia: Some insights on vulnerability, job security, and relative incomes. Brookings Institution. <https://www.brookings.edu/articles/subjective-well-being-in-colombia-some-insights-on-vulnerability-job-security-and-relative-incomes/>
- Gruber, J., & Wise, D. A. (Eds.). (1999). *Social Security and Retirement around the World*. University of Chicago Press.
- Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, 15(3), 199-236. <https://doi.org/10.1093/pan/mpl013>
- Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, 42(8), 1-28. <https://doi.org/10.18637/jss.v042.i08>
- Iacus, S. M., King, G., & Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Political Analysis*, 20(1), 1-24. <https://doi.org/10.1093/pan/mpr013>

- Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *The Review of Economics and Statistics*, 86(1), 4-29. <https://doi.org/10.1162/003465304323023651>
- Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1), 5-86. <https://doi.org/10.1257/jel.47.1.5>
- Kahneman, D., & Deaton, A. (2010). High income improves evaluation of life but not emotional well-being. *Proceedings of the National Academy of Sciences*, 107(38), 16489-16493. <https://doi.org/10.1073/pnas.1011492107>
- Kahneman, D., Diener, E., & Schwarz, N. (Eds.). (1999). *Well-Being: The Foundations of Hedonic Psychology*. Russell Sage Foundation.
- Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1), 83-119. <https://doi.org/10.1111/j.1468-0262.2007.00733.x>
- Ministerio del Trabajo de Colombia. (2024). *Informe de gestión del programa Colombia Mayor*. Bogotá: Ministerio del Trabajo.
- Modigliani, F., & Brumberg, R. (1954). Utility analysis and the consumption function: An interpretation of cross-section data. In K. K. Kurihara (Ed.), *Post-Keynesian Economics* (pp. 388-436). Rutgers University Press.
- Mejía, L. F., & Forero, D. (2022). ¿Qué hacer en protección económica para la vejez? *Libros Fedesarrollo*, Fedesarrollo, number 20249, June.
- OECD. (2013). *OECD Guidelines on Measuring Subjective Well-being*. OECD Publishing. <https://doi.org/10.1787/9789264191655-en>
- OECD. (2023). *Pensions at a Glance 2023: OECD and G20 Indicators*. OECD Publishing. <https://doi.org/10.1787/678d8da1-en>
- Rosenbaum, P. R. (2002). *Observational Studies* (2nd ed.). Springer. <https://doi.org/10.1007/978-1-4757-3692-2>
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55. <https://doi.org/10.1093/biomet/70.1.41>
- Rubin, D. B. (2001). Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, 2(3-4), 169-188. <https://doi.org/10.1023/A:1020363010465>
- Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, 57(6), 1069-1081. <https://doi.org/10.1037/0022-3514.57.6.1069>

- Salazar-Saenz, M., & Robayo, M. (2020). Microsimulación del sistema pensional colombiano: Cobertura, equidad y sostenibilidad. *Coyuntura Económica: Investigación Económica y Social*, 50(1), 25-58.
- Sánchez, F., & Mesa, J. (2018). Subsidios implícitos en el sistema pensional colombiano: Estimación y análisis distributivo. Documentos CEDE, 2018-21. Universidad de los Andes.
- VanderWeele, T. J., & Ding, P. (2017). Sensitivity analysis in observational research: Introducing the E-value. *Annals of Internal Medicine*, 167(4), 268-274. <https://doi.org/10.7326/M16-2607>
- World Bank. (2023). Integrated people-centred care for an aging population: Colombia case study. World Bank Group. <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/099638503072441223>
- Zeileis, A. (2004). Econometric computing with HC and HAC covariance matrix estimators. *Journal of Statistical Software*, 11(10), 1-17. <https://doi.org/10.18637/jss.v011.i10>