

Impacto de la elección de carreras STEM en el razonamiento cuantitativo: análisis diferencial por género y factores individuales y preuniversitarios

Impact of STEM Career Choice on Quantitative Reasoning: Differential Analysis by Gender and Individual and Pre-University Factors

Ana María Barrera Rodríguez*

Faculty of Economic, Administrative and Accounting Sciences, Universidad Libre Seccional Pereira, Pereira, Risaralda, Colombia.

ORCID iD: <https://orcid.org/0000-0003-3701-2278>

Email: anam.barrerar@unilibre.edu.co

Víctor David Jaramillo Mejía

Faculty of Economic, Administrative and Accounting Sciences, Universidad Libre Seccional Pereira, Pereira, Risaralda, Colombia.

ORCID iD: <https://orcid.org/0000-0001-6426-4855>

Email: victord.jaramillom@unilibre.edu.co

Paola Andrea Echeverri Gutiérrez

Faculty of Engineering, Universidad Libre Seccional Pereira, Pereira, Risaralda, Colombia.

ORCID iD: <https://orcid.org/0000-0003-3029-4997>

Email: paolaa.echeverrig@unilibre.edu.co

Recibido / Received: 10/01/2026

Aceptado / Accepted: 07/06/2026

Resumen: Este estudio analiza el impacto de la elección de carreras STEM (Ciencia, Tecnología, Ingeniería y Matemáticas) en el desarrollo del razonamiento cuantitativo y sus diferencias por género, factores preuniversitarios y características individuales. La investigación adopta un enfoque cuantitativo con un diseño ex post facto. Se utilizaron datos secundarios provenientes de las pruebas Saber 11 y Saber Pro del Instituto Colombiano para la Evaluación de la Educación (ICFES), complementados con datos primarios obtenidos a partir de encuestas aplicadas a 653 estudiantes universitarios de Risaralda, Colombia. Para el análisis se aplicaron modelos de regresión por Mínimos Cuadrados Ordinarios (MCO), Propensity Score Matching (PSM) y árboles de decisión (CART) para estimar efectos diferenciales y explorar patrones asociados a la elección de programas STEM. Los resultados indican que cursar una carrera STEM se asocia con un incremento promedio de aproximadamente 7.6 puntos en las competencias de razonamiento cuantitativo frente a programas no STEM, estimación que es consistente con un efecto causal positivo bajo los supuestos de identificación del PSM, pero no constituye evidencia causal definitiva dada la naturaleza observacional de los datos. Adicionalmente, la ganancia asociada a las carreras STEM en razonamiento cuantitativo es significativamente menor para las mujeres que para los hombres, evidenciando una brecha de género en los retornos educativos de estas disciplinas. El análisis de segmentación mediante árbol de decisión identificó el desempeño previo en ciencias, el capital educativo familiar y la zona de residencia como los factores con mayor poder discriminante en la probabilidad de que las mujeres elijan carreras STEM. Los hallazgos subrayan la necesidad de fortalecer políticas educativas orientadas a la equidad de género en STEM, con especial atención al fortalecimiento de competencias científicas en la educación media y a la reducción de brechas territoriales y socioeconómicas en el acceso a trayectorias STEM.

Palabras clave: Educación STEM, razonamiento cuantitativo, brecha de género, elección de carrera, Propensity Score Matching (PSM), Árboles de decisión (CART).

Abstract: This study analyzes the impact of choosing STEM careers (Science, Technology, Engineering, and Mathematics) on the development of quantitative reasoning and its differences by gender, pre-university factors, and individual characteristics. The research adopts a quantitative approach with an ex post facto design. Secondary data from the Saber 11 and Saber Pro assessments conducted by the Colombian Institute for the Evaluation of Education (ICFES) were used, complemented by primary data collected through surveys administered to 653 university students from Risaralda, Colombia. The analysis employed Ordinary Least Squares (OLS) regression models, Propensity Score Matching (PSM), and decision trees (CART) to estimate differential effects and explore patterns associated with the choice of STEM programs. The results indicate that enrolling in a STEM degree is associated with an average increase of approximately 7.6 points in quantitative reasoning competencies compared to non-STEM programs, an estimate that is consistent with a positive causal effect under the identification assumptions of PSM, but does not constitute definitive causal evidence given the observational nature of the data. Additionally, the quantitative reasoning gains associated with STEM careers are significantly smaller for women than for men, evidencing a gender gap in the educational returns of these disciplines. The decision tree segmentation analysis identified prior science performance, family educational capital, and area of residence as the factors with the greatest discriminant power in predicting the likelihood of women choosing STEM careers. The findings underscore the need to strengthen educational policies aimed at gender equity in STEM, with particular attention to building scientific competencies at the secondary education level and reducing territorial and socioeconomic disparities in access to STEM pathways.

Keywords: STEM Education, Quantitative Reasoning, Gender Gap, Career Choice. Propensity Score Matching (PSM), Decision Trees (CART).

1. Introducción

El término STEM hace referencia a las disciplinas que se enfocan en la Ciencia, Tecnología, Ingeniería y Matemáticas (Botella et al., 2019). En los últimos años, estas disciplinas han adquirido un papel cada vez más importante (Zhao y Perez-Felkner, 2022) debido a que son consideradas motores del desarrollo económico, de la innovación y de la competitividad de las organizaciones (Verdugo-Castro et al., 2022). En diversas investigaciones se ha encontrado que la formación en STEM promueve el pensamiento crítico, la capacidad para encontrar soluciones innovadoras y la capacidad para enfrentar retos reales relacionados con la salud, la energía y el medio ambiente (Kayan-Fadlelmula et al., 2022). En ese orden de ideas, estas áreas se han consolidado como unos ejes estratégicos importantes que ayudan a alcanzar ventajas competitivas tanto a nivel regional como internacional. Esto ha generado que se cree un consenso sobre la importancia de fomentar la participación de los jóvenes en estas áreas consideradas como campos de estudios clave para obtener ventajas competitivas a nivel profesional (Hermans et al., 2022). Y que en últimas, se gestó un cambio en los paradigmas educativos de las instituciones de educación superior (Kayan-Fadlelmula et al., 2022). De ahí que el potencial de estas áreas o disciplinas para formar estudiantes con pensamiento crítico y capacidad para proponer soluciones a los problemas que enfrenta la sociedad moderna las convierte en un pilar esencial del desarrollo (García-Peñalvo et al., 2022). En este sentido, los sistemas educativos de todo el mundo han comenzado a establecer como prioridad la preparación de los jóvenes para un entorno laboral dinámico y exigente, caracterizado por avances tecnológicos, una economía globalizada y competitiva, y una creciente diversidad social (López et al., 2023). Es por esto que se ha determinado que la educación STEM no solo es clave para la inserción laboral en un contexto cada vez más competitivo y especializado, sino que también actúa como un motor fundamental

para el desarrollo de capacidades humanas (Kim et al., 2021).

Lo anterior pone de manifiesto que la educación en STEM desempeña un papel único en múltiples ámbitos y explica la alta demanda actual por profesionales, investigadores y técnicos cualificados en estas disciplinas (Hermans et al., 2022). A esto se suma que, en promedio los trabajos en áreas de STEM poseen remuneraciones que duplican a los no STEM (Zhao y Perez-Felkner, 2022). Esta combinación de alta demanda y mejores condiciones laborales ha motivado que, en diversos países, se implementen políticas y estrategias educativas para incentivar el interés de los jóvenes por estas áreas (Zhao y Perez-Felkner, 2022).

Sin embargo, a pesar de las favorables perspectivas laborales y del creciente esfuerzo por promover la educación en STEM, la realidad muestra que el número de estudiantes que optan por este tipo de carreras sigue siendo insuficiente (Miller y Halpern, 2013). Esta situación se traduce en una escasez de personal con las competencias necesarias en estas áreas, lo que ha generado una brecha significativa entre la demanda de profesionales y la oferta existente. Particularmente preocupante es la persistente brecha de género (Riegle-Crumb y King, 2010). A nivel mundial, las mujeres representan apenas alrededor del 35 % de los estudiantes matriculados en programas de educación superior relacionados con STEM (UNESCO, 2024). Aunque su participación ha aumentado lentamente en algunas áreas (Patterson et al., 2020), su presencia sigue siendo muy baja en campos como la ingeniería, la física y las tecnologías de la información. Esta desigualdad no solo refleja la influencia de estereotipos y sesgos culturales, sino también la falta de referentes femeninos y de políticas sostenidas que incentiven su acceso y permanencia en estas disciplinas (Arredondo Trapero et al., 2019).

En América Latina, investigaciones recientes muestran que las brechas de género en STEM continúan siendo un desafío persistente, particularmente en áreas relacionadas con las matemáticas, la ingeniería y la tecnología. En Uruguay, Fernández Soto et al. (2024) encontraron que las desigualdades de género no solo se presentan en el acceso a programas STEM, sino también en las trayectorias académicas y profesionales, evidenciando fenómenos de segregación horizontal y vertical que limitan la permanencia y progresión de las mujeres en estos campos. De manera similar, estudios realizados en Chile identificaron que los ambientes académicos masculinizados, la ausencia de referentes femeninos y determinados estereotipos culturales continúan afectando la experiencia y permanencia de las mujeres en carreras STEM, especialmente en áreas tradicionalmente asociadas a la ingeniería y la construcción (Salazar et al., 2024). Asimismo, investigaciones desarrolladas en Costa Rica y otros contextos iberoamericanos han señalado que factores como la autopercepción de habilidades, la ansiedad matemática, las expectativas familiares y las dinámicas escolares influyen significativamente en la elección de carreras científicas y tecnológicas por parte de las mujeres (Merayo y Ayuso, 2023).

En ese orden de ideas, se ha encontrado que las brechas de género en STEM son fruto de una combinación de factores individuales, familiares, institucionales y socioculturales. Entre los principales factores identificados se encuentran la menor autoeficacia percibida en matemáticas y ciencias, la ansiedad matemática, las expectativas diferenciadas de docentes y familias, y la persistencia de estereotipos asociados a las capacidades científicas y tecnológicas de hombres y mujeres (Esteban Ramiro et al., 2024; Merayo y Ayuso, 2023). Lo anterior sugiere que las brechas de género en STEM son un fenómeno multidimensional que no puede explicarse únicamente por diferencias de habilidades o intereses individuales.

En Colombia, las brechas de género en áreas asociadas a STEM han mostrado comportamientos particularmente relevantes en evaluaciones estandarizadas. Estudios recientes evidencian que en Colombia, los hombres obtienen, en promedio, mejores resultados en matemáticas y razonamiento cuantitativo, mientras que las mujeres suelen presentar mejores desempeños en lectura crítica (Rodríguez de Luque, 2025). Asimismo, investigaciones desarrolladas en la capital del país, Bogotá, encontraron que las brechas en matemáticas y ciencias naturales desfavorables para las mujeres están presentes en la mayoría de instituciones educativas analizadas, evidenciando la persistencia de desigualdades estructurales asociadas al género.

Adicionalmente, la evidencia nacional ha mostrado que variables como el nivel socioeconómico, la ubicación geográfica y las características de las instituciones educativas influyen significativamente en el desempeño académico en áreas cuantitativas, generando diferencias importantes entre grupos poblacionales y territorios (Fajardo Ortiz et al., 2025). En este contexto, comprender los factores asociados a la elección de carreras STEM y su relación con el desarrollo de competencias cuantitativas resulta especialmente relevante para el diseño de estrategias orientadas a reducir las brechas educativas y promover una mayor participación femenina en estas disciplinas.

Superar estas brechas implica comprender mejor los factores que influyen en la elección de carrera entre hombres y mujeres, así como analizar cómo dicha decisión repercute en el desarrollo de competencias clave para la empleabilidad (López et al., 2023). En respuesta a este desafío, diferentes países han implementado estrategias orientadas a fortalecer la educación en STEM, no solo como una forma de impulsar la competitividad económica, sino también como una herramienta para promover la equidad de género y el desarrollo social sostenible.

Aunque la literatura ha avanzado en la comprensión de las brechas de género en STEM y en el análisis del desempeño académico en matemáticas y ciencias, todavía existe evidencia limitada sobre cómo la elección de carreras STEM se relaciona con el desarrollo de competencias de razonamiento cuantitativo en estudiantes universitarios colombianos. Asimismo, son escasos los estudios que analizan simultáneamente los factores preuniversitarios, las características individuales y las diferencias de género asociadas con la probabilidad de elegir este tipo de carreras en contextos regionales específicos.

En este contexto, surge la pregunta de investigación: ¿Cuál es el impacto diferencial de la elección de carreras STEM en el desarrollo de competencias en razonamiento cuantitativo entre estudiantes universitarios en Risaralda, y qué factores pre-universitarios y características individuales se asocian con la probabilidad de que las mujeres elijan dichas carreras?

2. Metodología

2.1. Tipo de investigación y estudio

La presente investigación adopta un enfoque cuantitativo (Saunders et al., 2016) para examinar los factores asociados a la elección de carreras STEM (Ciencia, Tecnología, Ingeniería y Matemáticas) por parte de las mujeres y el impacto de dichas carreras en el desarrollo de competencias. Emplea un diseño no experimental de tipo *ex post facto*, dado que se analizan datos preexistentes y se observan fenómenos que ya han ocurrido sin manipulación directa de las variables por parte de los investigadores. El estudio

combina análisis longitudinales implícitos (al comparar resultados de las pruebas Saber 11 y las pruebas Saber Pro para los mismos individuos) y transversales (al analizar los determinantes de la elección de carrera en un punto específico). Se busca establecer relaciones asociativas y, mediante técnicas cuasiexperimentales, inferir posibles efectos causales. El alcance del estudio es descriptivo en su fase inicial, al caracterizar las diferencias en competencias y la elección de carreras, y analítico-explicativo al intentar determinar el impacto de las carreras STEM y los factores asociados a su elección (Creswell, 2014).

2.2. Fuentes de datos y muestras

Para abordar la pregunta de investigación se recurrió a dos fuentes de información complementarias. Inicialmente, se utilizaron microdatos anonimizados de las pruebas Saber 11 y Saber Pro, facilitados por el Instituto Colombiano para la Evaluación de la Educación (ICFES). Estos datos permitieron construir un panel de estudiantes que presentaron la prueba Saber 11 entre los años 2014-2020 y la prueba Saber Pro entre 2018-2023, posibilitando el seguimiento de su trayectoria académica. Para los análisis comparativos y de impacto con estos datos secundarios, la muestra se focalizó en estudiantes de instituciones de educación superior del departamento de Risaralda, resultando en 1401 observaciones que fueron variando según la naturaleza de modelo.

En segundo lugar, se recolectaron datos primarios a través de 653 encuestas aplicadas a estudiantes de instituciones de educación superior del departamento de Risaralda durante el año 2023. El levantamiento de la información se realizó mediante un diseño de muestreo probabilístico bietápico (Creswell, 2014), en el cual, en la primera etapa, se definieron como Unidades Primarias de Muestreo (UPM) a las Instituciones de Educación Superior (IES) que ofrecen programas de pregrado en el departamento. De un total de 12 IES identificadas, se seleccionaron aleatoriamente 10, asegurando la inclusión de la Universidad Tecnológica de Pereira (UTP) como la única institución pública del departamento y, al mismo tiempo, la de mayor tamaño. Esta selección buscó garantizar una diversidad institucional representativa, incluyendo tanto universidades públicas como privadas.

En la segunda etapa, dentro de cada una de las 10 IES seleccionadas, las Unidades Secundarias de Muestreo (USM) fueron los estudiantes de pregrado. La selección de los 653 estudiantes por IES se realizó mediante un muestreo por conveniencia en puntos de alto tráfico estudiantil (cafeterías, bibliotecas), buscando una representación equilibrada de estudiantes de carreras STEM y No STEM, así como de diferentes semestres. Cabe destacar que, dado el tamaño de población estudiantil de la UTP y la diversidad de programas ofertados, cerca del 19,2% de las encuestas se aplicaron en esta institución educativa. El uso del muestreo por conveniencia permitió acceder de manera ágil a estudiantes pertenecientes a diferentes programas académicos e instituciones de educación superior, estrategia que ha sido ampliamente utilizada en investigaciones sociales y educativas cuando existen limitaciones logísticas y de acceso a la población objetivo (Etikan et al., 2015). No obstante, este tipo de muestreo puede generar posibles sesgos de selección asociados a la disponibilidad y disposición de participación de los encuestados, por lo que los resultados deben interpretarse considerando las características específicas de la muestra analizada.

2.3. Variables del estudio

Las variables empleadas en los diferentes modelos analíticos se definieron y operacionalizaron como se detalla a continuación. La tabla 1 resume las variables clave provenientes de los datos del ICFES, mientras que la tabla 2 describe las variables adicionales o con diferente operacionalización obtenidas de los datos primarios.

Tabla 1: Descripción de variables clave – Datos ICFES (Saber 11 y Saber Pro).

Variable	Descripción explícita	Tipo	Escala / Codificación usada en modelos	Fuente
spro_rc_var	Puntaje en Razonamiento Cuantitativo (Saber Pro)	Cuantitativa	Numérico (continuo, rango [ej. 0-300])	Data ICFES
spro_lc_var	Puntaje en Lectura Crítica (Saber Pro)	Cuantitativa	Numérico (continuo, rango [ej. 0-300])	Data ICFES
stem_dm	Elección de carrera STEM (1=STEM, 0=No STEM)	Binaria	Numérico (0, 1), 0=No STEM, 1=STEM	Data ICFES
s11_mat_var	Puntaje en Matemáticas (Saber 11)	Cuantitativa	Numérico (continuo, rango [ej. 0-100])	Data ICFES
s11_lc_var	Puntaje en Lectura Crítica (Saber 11)	Cuantitativa	Numérico (continuo, rango [ej. 0-100])	Data ICFES
mujer_dm	Género del estudiante	Binaria	Numérico (0, 1), 0=Hombre, 1= Mujer	Data ICFES
fami_estrato vivienda	Estrato socioeconómico del hogar	Ordinal	Factor (Niveles: 1, 2, 3, 4, 5, 6, Sin Estrato=NA)	Data ICFES
cole_caracter	Carácter académico del colegio de egreso	Nominal / Ordinal	Factor (Niveles: 1. Académico, 2. Técnico, 3. Técnico/académico)	Data ICFES
cole_naturaleza	Naturaleza del colegio (Público/Privado)	Binaria	Numérico (0, 1), 0. No oficial, 1. Oficial.	Data ICFES

Tabla 2: Descripción de variables clave – Datos primarios.

Nombre Variable	Descripción Explícita de la Variable	Tipo de Variable	Formato / Codificación	Escala / Niveles o Etiquetas	Notas Adicionales Árbol de decisión
Sexo (usada para género)	Género del estudiante	Binaria (Nominal)	Numérico (1, 0)	0. Mujer, 1. Masculino	Usada para filtrar (mujeres).
STEM (usada para elección de STEM)	Indica si la carrera elegida por el estudiante es STEM o No STEM	Binaria (Nominal)	Numérico (1, 0)	1 = Sí (STEM), 0 = No (No STEM)	Variable dependiente del árbol.
Edu_padre	Nivel educativo más alto alcanzado por el padre	Ordinal	Cualitativa	0. Sin formac, 1. Primaria, 2. Secundaria, 3. Técnico, 4. Tecnólogo, 5. Universitario, 6. Maestría, 7. Doctorado	Considerado predictor.
Edu_madre	Nivel educativo más alto alcanzado por la madre	Ordinal	Cualitativa	0. Sin formac, 1. Primaria, 2. Secundaria, 3. Técnico, 4. Tecnólogo, 5. Universitario, 6. Maestría, 7. Doctorado	Predictor en árbol.
Estrato	Estrato socioeconómico del hogar del estudiante	Ordinal	Cualitativa	1. Bajo (1-2), 2. Medio (3-4), 3. Alto (5-6)	Predictor en árbol.

Nombre Variable	Descripción Explícita de la Variable	Tipo de Variable	Formato / Codificación	Escala / Niveles o Etiquetas	Notas Adicionales Árbol de decisión
Tip_colegio	Tipo de colegio del que egresó el estudiante	Binaria (Nominal)	Numérico (1, 0)	0. Público, 1. Privado	Predictor en árbol.
Actividades_STEM_previas	Participación del estudiante en actividades STEM pre-universitarias	Binaria (Nominal)	Numérico (1, 0)	0. No, 1. Sí	Predictor en árbol.
Puntaje_ciencias_previo	Área de énfasis o mejor desempeño en pruebas previas (ej. Saber 11)	Ordinal	Cualitativa	1. Matemáticas, 2. Ciencias, 3. Artística, 4. Tecnología, 5. Lenguaje, 6. Deporte	Predictor en árbol.
Zona_residencia	Zona de ubicación (probablemente del colegio o residencia)	Binaria (Nominal)	Numérico (1, 0)	0. Rural, 1. Urbana	Predictor en árbol.

La operacionalización de las carreras STEM (Ciencia, Tecnología, Ingeniería y Matemáticas) se fundamentó en el listado de programas de grado STEM de la Universidad de Indiana (Indiana University Indeanapolis, 2012), permitiendo una clasificación consistente de los campos de estudio. Para el análisis de regresión, las variables de control categóricas se introdujeron en los modelos como series de indicadores binarios (variables dummy), designando una categoría de cada variable original como el grupo de referencia omitido, una práctica estándar para la estimación e interpretación de los efectos

2.4. Métodos de análisis estadístico y econométrico

Para abordar los objetivos de la investigación, se aplicó una secuencia de técnicas analíticas. Inicialmente, se estimaron Modelos de Regresión por Mínimos Cuadrados Ordinarios (MCO) para cuantificar la asociación entre la elección de una carrera STEM y los puntajes de los estudiantes en Razonamiento Cuantitativo al finalizar su educación superior (Saber Pro). Estos modelos se controlaron por el desempeño previo del estudiante en la prueba Saber 11 (específicamente, el puntaje en Matemáticas), y se construyeron de forma progresiva, añadiendo secuencialmente bloques de covariables socioeconómicas y del contexto escolar para evaluar la estabilidad y robustez del coeficiente asociado a la variable STEM. La especificación general de estos modelos fue:

$$Y_{pro(i)} = \beta_0 + \beta_1 STEM_i + \beta_2 Y_{saber11(i)} + \mathbf{X}_i' \gamma + \epsilon_i$$

Donde $Y_{pro(i)}$ es el puntaje Saber Pro, $STEM_i$ es la dummy de carrera STEM, $Y_{saber11(i)}$ es el puntaje Saber 11, y $\mathbf{X}_i$ es un vector de controles. Se emplearon errores estándar robustos a heterocedasticidad (White, 1980).

Posteriormente, como lo sugiere Becker y Ichino (2002), para mitigar el potencial sesgo de autoselección inherente a la elección de una carrera, se implementó la técnica de Propensity Score Matching (PSM) (Rosenbaum y Rubin, 1983), enfocada en la submuestra de mujeres. Este método busca crear grupos de tratamiento (mujeres en

STEM) y control (mujeres en No STEM) que sean estadísticamente comparables en un conjunto de características observables pre-universitarias. La probabilidad de pertenecer al grupo STEM (propensity score) se estimó mediante un modelo de regresión logística, utilizando como predictores los puntajes de Saber 11, el estrato socioeconómico, la educación de los padres y características del colegio de origen.

$$P(STEM_i = 1 | \mathbf{X}_i) = \Lambda(\alpha_0 + \mathbf{X}_i' \alpha)$$

Donde Λ es la función logística y $\mathbf{X}_i$ son covariables pre-universitarias.

Se aplicó un algoritmo de emparejamiento por el vecino más cercano con reemplazo (Austin, 2011) y un caliper de 0.2 desviaciones estándar del logit del propensity score (Caliendo y Kopeinig, 2008). La calidad del emparejamiento se evaluó mediante la inspección del balance de covariables (Stuart, 2010). De esta forma, se obtuvo una muestra inicial de 449 mujeres en carreras STEM (tratadas) y 358 mujeres en carreras No STEM (control), el proceso de Propensity Score Matching logró emparejar exitosamente a 446 mujeres STEM. Para estos emparejamientos, se utilizaron 225 mujeres No STEM únicas, donde algunas de estas últimas pudieron haber sido emparejadas más de una vez (indicado por un Tamaño Efectivo de Muestra (ESS) de control de 144.98). Como resultado, 3 mujeres STEM y 133 mujeres No STEM no fueron incluidas en la muestra final emparejada utilizada para el análisis del efecto del tratamiento.

Con estos datos, se procede a medir el Efecto Promedio del Tratamiento en los Tratados (ATT) (Imbens y Wooldridge, 2009), estimando una regresión ponderada en la muestra emparejada, controlando por el puntaje Saber 11, con la forma:

$$Y_{pro(i)} = \delta_0 + \delta_1 STEM_i + \delta_2 Y_{saber11(i)} + u_i$$

donde δ_i es el ATT, con errores estándar robustos. Se realizaron análisis de sensibilidad a la especificación del modelo de propensity score y al algoritmo de matching para verificar la robustez del ATT estimado.

No obstante, la validez causal de las estimaciones PSM descansa sobre el supuesto de ignorabilidad condicional o Conditional Independence Assumption (CIA) (Rosenbaum y Rubin, 1983), según el cual, una vez condicionado el conjunto de covariables preuniversitarias observadas, la asignación al tratamiento (elección de carrera STEM) es independiente de los resultados potenciales. Este supuesto, estándar en la literatura de evaluación de programas con datos observacionales (Imbens y Wooldridge, 2009), no puede verificarse empíricamente de forma directa, puesto que implica la ausencia de factores no observados que influyan simultáneamente en la probabilidad de elegir una carrera STEM y en el desempeño posterior en razonamiento cuantitativo.

En el presente estudio, es relevante que variables no observadas como la motivación intrínseca hacia las matemáticas y las ciencias, la autoeficacia académica en áreas cuantitativas (Bandura, 1997; Wigfield y Eccles, 2000), el capital social familiar orientado a trayectorias STEM y la calidad diferencial de la instrucción universitaria recibida en cada programa no estén plenamente capturadas por las covariables incluidas en el modelo de propensity score (puntajes Saber 11, estrato socioeconómico y características del colegio de egreso). Si estas variables no observadas correlacionan positivamente tanto con la elección de STEM como con el desempeño en razonamiento cuantitativo, el ATT estimado podría estar sobreestimando el efecto causal neto. Por el contrario, si operan factores compensatorios no observados (como barreras institucionales o ambientes

académicos menos estimulantes para las mujeres en STEM), el efecto estimado podría estar subestimado.

La estabilidad del coeficiente de la variable STEM a través de las tres especificaciones progresivas del modelo ATT (incluyendo controles adicionales por estrato y carácter de la institución) constituye una señal de robustez ante la omisión de variables de control observables, aunque no garantiza la ausencia de sesgo por variables no observadas. En consecuencia, los hallazgos derivados del PSM deben interpretarse como estimaciones consistentes con un efecto causal positivo, pero no como evidencia causal definitiva, condición inherente a toda evaluación basada en datos observacionales sin asignación aleatoria al tratamiento (Imbens y Wooldridge, 2009).

Para explorar si el impacto de las carreras STEM varía por género, se utilizó un modelo de interacción, conceptualmente similar a un enfoque de Diferencias en Diferencias (DiD) modificado (Angrist y Pischke, 2009). Este modelo, estimado sobre la muestra completa de hombres y mujeres, incluyó un término de interacción (Wooldridge, 2010) entre la variable dummy de carrera STEM y una variable dummy que identifica a las mujeres:

$$Y_{pro(ij)} = \beta_0 + \beta_1 STEM_j + \beta_2 Mujer_i + \beta_3 (STEM_j \times Mujer_i) + \beta_4 Y_{saber11(ij)} + \mathbf{X}_{ij}'\gamma + \epsilon_{ij}$$

El coeficiente β_3 del término de interacción $STEM_j \times Mujer_i$ captura el efecto DiD. Se usaron errores estándar robustos.

Para todos los modelos de regresión, se realizaron pruebas de diagnóstico estándar, incluyendo la prueba RESET de Ramsey para linealidad, el Factor de Inflación de la Varianza (VIF) para multicolinealidad, la prueba de Breusch-Pagan para heterocedasticidad (complementada con el uso de errores estándar robustos), y la prueba de Durbin-Watson para autocorrelación (Wooldridge, 2010).

Enseguida, con el propósito de identificar la estructura jerárquica de los factores preuniversitarios e individuales asociados a la probabilidad de que una mujer elija una carrera STEM, se implementó un Árbol de Decisión de Clasificación con el algoritmo CART (Classification and Regression Trees; Breiman et al., 1984) sobre la submuestra completa de mujeres proveniente de los datos primarios recolectados mediante encuesta (N = 320; 177 STEM, 143 No STEM). A diferencia de los modelos de regresión lineal, que estiman efectos promedio bajo supuestos de linealidad y aditividad, el árbol de decisión permite detectar interacciones no lineales y segmentaciones condicionales entre predictores, ofreciendo una representación visual e interpretable de los perfiles de elección (Loh, 2011).

La variable dependiente fue la elección de carrera STEM (binaria: STEM / No STEM), y los predictores incluyeron: nivel educativo del padre (edu_padre), nivel educativo de la madre (edu_madre), estrato socioeconómico, tipo de colegio, zona de residencia, conocimiento o participación previa en actividades STEM (Conocim_STEM y actividades_stem_previas) y desempeño previo en ciencias (puntaje_ciencias_previo). El modelo se construyó utilizando el parámetro de complejidad $cp = 0.01$ y un mínimo de 20 observaciones por nodo de división ($minsplit = 20$), con semilla aleatoria fijada en 123 para garantizar la reproducibilidad de los resultados.

La validación del modelo se realizó en dos niveles complementarios. En primer lugar, se calcularon métricas de ajuste in-sample sobre la muestra completa de construcción (N = 320), obteniendo una exactitud (accuracy) de 0.722 (IC 95%: 0.669–0.770), un coeficiente Kappa de 0.439, una sensibilidad de 0.740, una especificidad de 0.699, un

F1-Score de 0.746 y un AUC de 0.759, indicando una capacidad discriminante moderada-alta del modelo. En segundo lugar, se realizó una validación cruzada de 10 pliegues (10-fold cross-validation) sobre la muestra completa, obteniendo un AUC máximo de 0.579. La diferencia entre el AUC in-sample (0.759) y el AUC en validación cruzada (0.579) evidencia un sobreajuste moderado, atribuible al tamaño muestral disponible en relación con la complejidad estructural del árbol.

Así, siguiendo las recomendaciones de la literatura (Loh, 2011), el árbol se interpreta prioritariamente como una herramienta de segmentación descriptiva: su valor analítico reside en la jerarquía e interacciones que revela entre los predictores (sintetizadas a través de la importancia relativa de variables) más que en su capacidad predictiva fuera de muestra.

2.5. Software Utilizado

El procesamiento, análisis y visualización de los datos se llevaron a cabo utilizando el software estadístico R (versión 4.4.2) a través del entorno de desarrollo integrado RStudio. Los paquetes de R principales para este estudio incluyeron *dplyr* para la manipulación de datos; *lmtest* y *sandwich* para la estimación de errores estándar robustos y pruebas de diagnóstico en modelos de regresión; *MatchIt* y *cobalt* para la implementación del Propensity Score Matching y la evaluación del balance de covariables; *rpart* y *rpart.plot* para la construcción y visualización de árboles de decisión; *stargazer* para la generación de tablas de resultados de regresión; y *ggplot2* junto con *patchwork* para la creación de visualizaciones estadísticas.

3. Resultados

El análisis descriptivo inicial de los puntajes obtenidos en las pruebas Saber Pro para el año 2023 en el departamento de Risaralda revela diferencias en las distribuciones de desempeño entre hombres y mujeres, particularmente en la competencia de Razonamiento Cuantitativo donde los hombres presentan una distribución de puntajes más favorable.

La Figura 1 muestra diferencias importantes entre mujer y hombre en la competencia de Razonamiento Cuantitativo. La curva de densidad y la distribución para los hombres se presenta una mayor concentración de puntajes en el extremo superior y desplazada hacia la derecha en comparación con la de las mujeres, sugiriendo que, en general, los hombres tienden a alcanzar puntajes superiores en esta competencia. En contraste, en Lectura Crítica, las distribuciones presentan un mayor nivel de solapamiento, lo que indica un desempeño más homogéneo entre géneros en esta área. Estos resultados sugieren que las brechas de desempeño se manifiestan con mayor intensidad en competencias asociadas al razonamiento matemático y analítico.

La Figura 2 complementa estos resultados mediante diagramas de caja y nubes de puntos para la muestra general y la submuestra STEM. En ambos casos, se observa que los hombres presentan medianas superiores en Razonamiento Cuantitativo, lo que confirma la existencia de las diferencias de género en esta competencia. El panel superior, que considera a todos los estudiantes corrobora visualmente la tendencia observada en los gráficos de densidad. Asimismo, dentro de las disciplinas STEM se evidencia que, aunque las distribuciones de puntajes son relativamente altas para ambos grupos, los hombres continúan exhibiendo una ligera ventaja respecto a las mujeres.

Figura 1: Funciones de densidad estimadas de los puntajes Saber Pro en Razonamiento Cuantitativo y Lectura Crítica, según sexo (Risaralda, 2023).

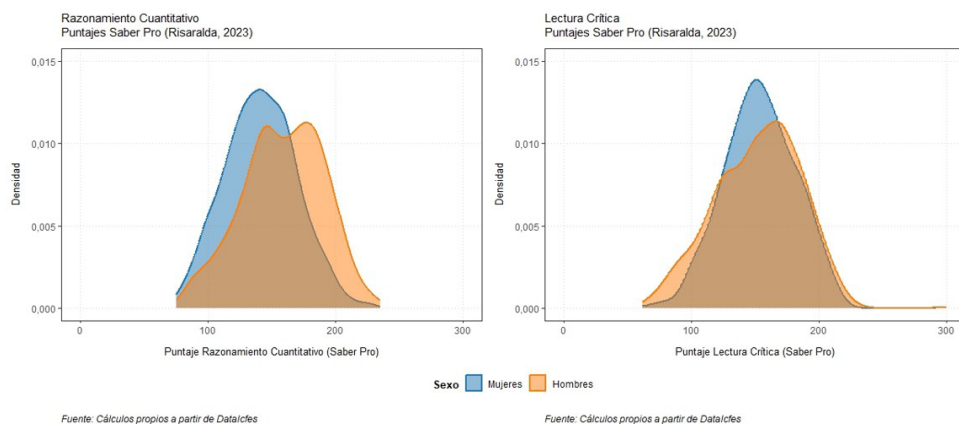
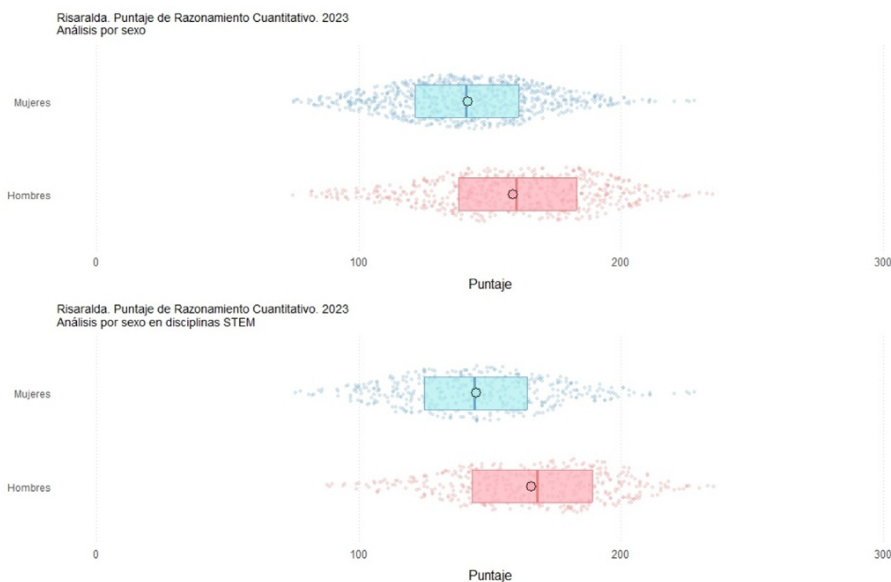


Figura 2: Distribución en diagrama de cajas de puntajes de Razonamiento Cuantitativo Saber Pro por sexo: muestra general y submuestra STEM (Risaralda, 2023).



Cuando el análisis se restringe exclusivamente a los estudiantes matriculados en disciplinas STEM (Figura 2, panel inferior), la diferencia en el desempeño en Razonamiento Cuantitativo entre géneros persiste. Aunque ambos grupos (hombres y mujeres en STEM) muestran distribuciones de puntajes relativamente altas, la distribución para los hombres en STEM sigue posicionada ligeramente a la derecha de la de las mujeres en STEM.

Analizar el efecto de cursar un pregrado STEM en el rendimiento cuantitativo de

un estudiante posibilita medir el impacto específico de estos programas en la mejora de habilidades analíticas, y también identificar si dicho impacto es homogéneo entre hombres y mujeres, y qué factores contextuales o individuales podrían estar mediando esta relación. Con este propósito, se exploró la asociación entre el rendimiento académico al finalizar la educación media y las competencias desarrolladas al concluir la educación superior. Para ello se generaron gráficos de dispersión comparando los puntajes de las pruebas Saber 11 y Saber Pro.

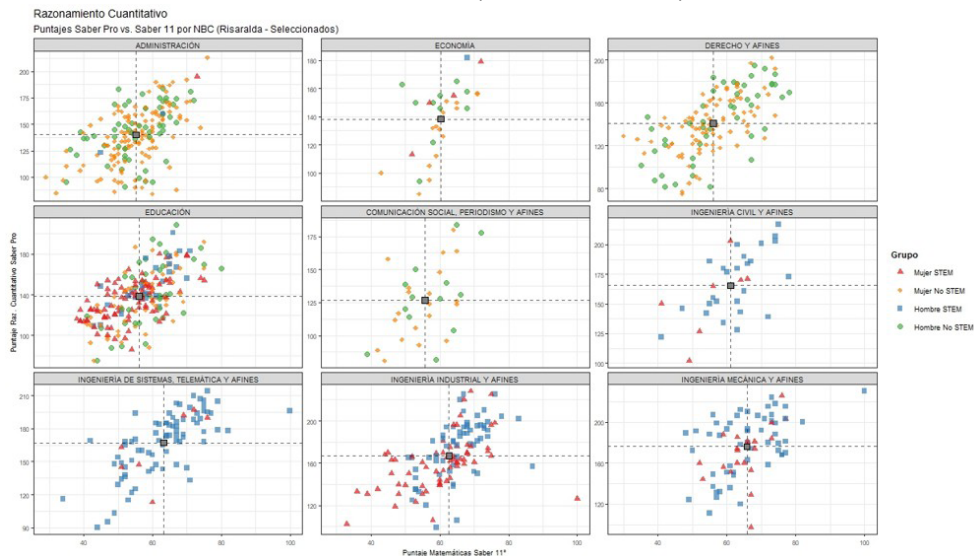
La Figura 3 presenta la relación entre los puntajes obtenidos en las pruebas Saber 11 y Saber Pro para las competencias de Razonamiento Cuantitativo y Lectura Crítica, diferenciando a los estudiantes según género y elección de carrera STEM. En ambos casos, se observa una asociación positiva entre el desempeño previo y el desempeño al finalizar la educación superior, indicando que los estudiantes con mejores resultados en Saber 11 tienden a obtener mayores puntajes en Saber Pro. Asimismo, los grupos Hombre STEM y Mujer STEM muestran una mayor concentración en niveles altos de desempeño, especialmente en Razonamiento Cuantitativo, aunque persiste una importante variabilidad individual entre estudiantes.

Figura 3: Relación entre Puntajes Saber 11 y Saber Pro por Competencia y Grupo Género-STEM (Risaralda, 2023).



Con el fin de profundizar en el análisis de estas relaciones, los gráficos de dispersión fueron desagregados por Núcleos Básicos de Conocimiento (NBC), para las competencias de Razonamiento Cuantitativo y Lectura Crítica. La Figura 4 evidencia que la concentración de estudiantes Hombre STEM en los cuadrantes de alto rendimiento es particularmente notable en programas de ingeniería como. En contraste, programas pertenecientes a áreas no STEM presentan una distribución más heterogénea entre grupos de género y niveles de desempeño. La posición del punto de promedio central (cuadrado negro) varía entre los NBC, indicando diferencias en el rendimiento promedio tanto de entrada (Saber 11) como de salida (Saber Pro) para cada área.

Figura 4: Relación entre Puntajes de Matemáticas (Saber 11) y Razonamiento Cuantitativo (Saber Pro) por Núcleo Básico de Conocimiento (NBC) y Grupo Género-STEM (Risaralda - 2023).



De manera similar, la Figura 5 muestra que la relación entre desempeño previo y desempeño posterior en Lectura Crítica también presenta variaciones según el campo de estudio. Los resultados evidencian que la composición de los grupos género-STEM y los niveles de desempeño difieren considerablemente entre NBC, sugiriendo que el contexto disciplinar desempeña un papel relevante en el desarrollo de competencias académicas.

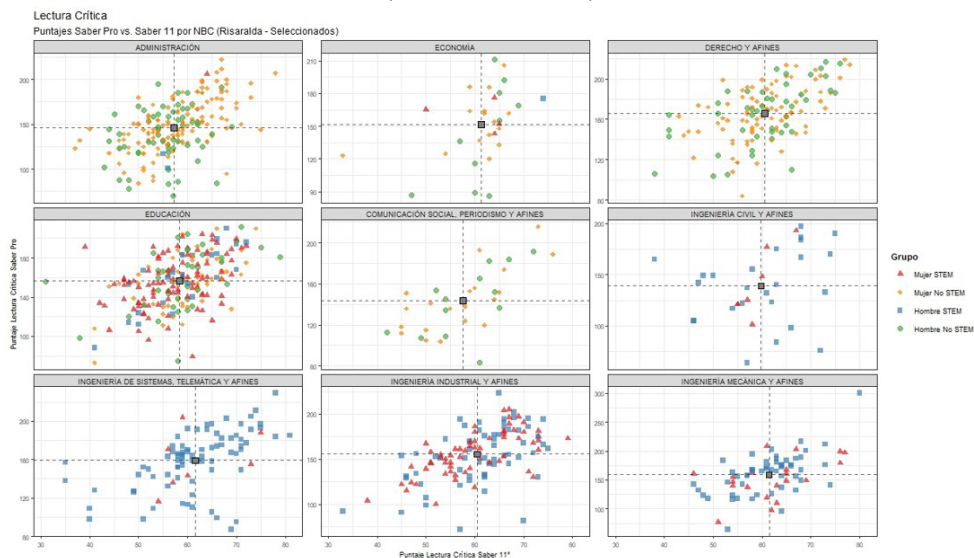
Estos análisis de dispersión subrayan la importancia de considerar tanto el punto de partida de los estudiantes (Saber 11) como el contexto disciplinar (NBC) al evaluar el desarrollo de competencias y las posibles brechas entre grupos.

Así, si bien la exploración visual de los datos sugiere la existencia de diferencias en el desempeño y posibles asociaciones con la elección de carreras STEM, estas observaciones no controlan simultáneamente por múltiples factores que podrían influir en dichos resultados, como el rendimiento académico previo y las características socioeconómicas. Para obtener una cuantificación más rigurosa de la asociación entre la elección de una carrera STEM y el desarrollo de competencias en Razonamiento Cuantitativo, y para evaluar la estabilidad de esta relación ante la inclusión de diversos controles, se procedió a la estimación de Modelos de Regresión por Mínimos Cuadrados Ordinarios (MCO).

Para una cuantificación más rigurosa de la asociación entre la elección de una carrera STEM y el desarrollo de competencias en Razonamiento Cuantitativo, se estimaron tres modelos de regresión por Mínimos Cuadrados Ordinarios (MCO) de forma progresiva. Todos los modelos utilizan el puntaje de Razonamiento Cuantitativo de la prueba Saber Pro como variable dependiente y controlan por el desempeño previo del estudiante en la prueba Saber 11 de Matemáticas. El Modelo 1 constituye la especificación base; el Modelo 2 introduce el estrato socioeconómico del estudiante

como covariable; y el Modelo 3 incorpora adicionalmente el carácter de la institución educativa de egreso. Los resultados, empleando errores estándar robustos a la heterocedasticidad (HC1), se detallan en la Tabla 3.

Figura 5: Relación entre Puntajes de Lectura Crítica (Saber 11) y Lectura Crítica (Saber Pro) por Núcleo Básico de Conocimiento (NBC) y Grupo Género-STEM (Risaralda - 2023).



De esta forma, los modelos MCO indican una asociación positiva y estadísticamente significativa entre cursar una carrera STEM y los puntajes en Razonamiento Cuantitativo Saber Pro. En el Modelo 1, que solo ajusta por el puntaje previo en matemáticas, ser estudiante de un programa STEM (Dummy STEM) se asocia con un incremento promedio de 6.058 puntos ($p < 0.001$) en el puntaje de la prueba Saber Pro, en comparación con estudiantes de programas No STEM. El puntaje de matemáticas en Saber 11 también muestra una fuerte asociación positiva ($\beta = 1.834$, $p < 0.001$).

Al incorporar el Estrato socioeconómico en el Modelo 2, el coeficiente estimado para la variable Dummy STEM se mantiene robusto y altamente significativo, situándose en 5.908 puntos ($p < 0.001$). Esta ligera disminución sugiere que, si bien el estrato puede tener alguna influencia, la asociación positiva de las carreras STEM persiste notablemente.

Luego, el Modelo 3, que incluye tanto el Estrato como el Carácter de la Institución Educativa de egreso, arroja un coeficiente para Dummy STEM de 5.881 puntos ($p < 0.001$). La estabilidad de este coeficiente a través de las diferentes especificaciones subraya la fuerte asociación positiva entre la formación en áreas STEM y un mayor rendimiento en razonamiento cuantitativo, incluso tras controlar por estas características contextuales y el desempeño académico previo. A lo largo de los tres modelos, el Puntaje matemáticas prueba Saber 11 conserva su poder predictivo con un coeficiente consistentemente alrededor de 1.8-1.9 ($p < 0.001$).

Tabla 3: Modelos de regresión progresivos para Razonamiento Cuantitativo (Saber Pro).

Dependent variable: Puntaje Saber Pro			
	Modelo 1	Modelo 2	Modelo 3
	(1)	(2)	(3)
<i>Dummy STEM (1=Stem)</i>	6.058	5.908	5.881
	(1.553)	(1.568)	(1.573)
	t = 3.900	t = 3.768	t = 3.738
	p = 0.0001***	p = 0.0002***	p = 0.0002***
Puntaje matemáticas prueba Saber 11	1.834	1.864	1.864
	(0.103)	(0.105)	(0.105)
	t = 17.859	t = 17.752	t = 17.735
	p = 0.000***	p = 0.000***	p = 0.000***
Constant	35.568	31.984	36.234
	(5.781)	(8.492)	(24.604)
	t = 6.153	t = 3.766	t = 1.473
	p = 0.000***	p = 0.0002***	p = 0.141
Covariables			
Estrato	x	✓	✓
Carácter de la Institución Educativa	x	x	✓
Pruebas del modelo			
Linealidad (Prueba RESET: p>0.05)	✓	x	x
Multicolinealidad (VIF < 5)	✓	x	x
Normalidad (JB: p>0.05)	✓	✓	✓
Heterogeneidad (BP_Test: p>0.05)	✓	✓	✓
Observations	807	807	807
R ²	0.380	0.385	0.385
Adjusted R ²	0.379	0.378	0.375
Residual Std. Error	22.11	22.13	22.17
F Statistic	246.816***	55.329***	38.234***
Durbin-Watson	2.02	2.01	2.01
Note: *p<0.05; **p<0.01; ***p<0.001 Errores estándar robustos (HC1). Dummies creadas para variables categóricas.			

Las pruebas de diagnóstico aplicadas a los modelos (resumidas en la Tabla 3) indican que los tres modelos cumplen con los supuestos de Normalidad de los residuos (prueba Jarque-Bera, $p > 0.05$) y Homocedasticidad (prueba Breusch-Pagan, $p > 0.05$), y el estadístico Durbin-Watson (cercano a 2) no sugiere problemas de autocorrelación. Sin embargo, los Modelos 2 y 3 presentan indicios de No Linealidad (Prueba RESET, $p < 0.05$) y Multicolinealidad ($VIF > 5$). Si bien la multicolinealidad podría afectar la precisión de los estimadores de las variables de control introducidas, la consistencia del coeficiente de la variable Dummy STEM es un hallazgo central. La indicación de no linealidad en los modelos más complejos sugiere que podrían existir relaciones no lineales no capturadas o la necesidad de explorar interacciones adicionales, aunque el efecto principal de STEM parece ser robusto.

Los modelos MCO, si bien controlan por un conjunto de características observables, no pueden descartar por completo el sesgo de autoselección inherente a la decisión de los estudiantes de optar por una carrera STEM. Es probable que existan diferencias sistemáticas preexistentes entre quienes eligen programas STEM y quienes no, más allá de las variables incluidas en los modelos de regresión, que podrían estar

correlacionadas tanto con la elección de carrera como con el desempeño académico posterior. Para abordar esta potencial limitación y obtener una estimación del efecto del tratamiento menos susceptible a este tipo de sesgo, se implementó la técnica de Propensity Score Matching (PSM), enfocada en la submuestra de mujeres. Este método busca crear grupos de tratamiento (mujeres en STEM) y control (mujeres en No STEM) que sean estadísticamente más comparables en un amplio rango de características observables pre-universitarias.

Previo al proceso de emparejamiento, se identificaron diferencias iniciales entre las mujeres que eligieron carreras STEM y aquellas matriculadas en programas No STEM. En particular, las estudiantes STEM presentaban, en promedio, puntajes ligeramente superiores en Razonamiento Cuantitativo en la prueba Saber 11, lo que sugiere la existencia de diferencias preuniversitarias entre ambos grupos y justifica la implementación del PSM. La muestra inicial para este análisis, antes del matching, consistió en 449 mujeres en carreras STEM y 358 mujeres en carreras No STEM. Estas diferencias preexistentes, junto con posibles disparidades en otras covariables no mostradas en esta tabla resumen (pero detalladas en la tabla de balance posterior), subrayan la importancia de aplicar PSM para obtener una comparación más equitativa.

Tabla 4: Características de la Muestra de Mujeres Antes del Emparejamiento (Puntaje Raz. Cuantitativo Saber 11).

Grupo de Tratamiento	Media Razonamiento Cuantitativo	S.D Razonamiento Cuantitativo	N
No STEM (Control)	55,43	8,7	358
STEM (Tratado)	56,28	9,6	449

El proceso de emparejamiento se realizó utilizando variables relacionadas con el desempeño académico previo, el estrato socioeconómico y las características de la institución educativa de egreso. Para ello, se implementó un algoritmo de vecino más cercano (1:1) con reemplazo y restricción de caliper, buscando conformar grupos de mujeres STEM y No STEM estadísticamente comparables.

La efectividad del emparejamiento se evaluó mediante el análisis del balance de covariables antes y después del matching. La Tabla 5 resume el balance logrado para las principales covariables en la muestra emparejada. Como se observa, la distancia del propensity score (distance) entre los grupos tratado y control es mínima después del matching (Diferencia de Medias Estandarizada, SMD = 0,0003). Para las variables continuas clave, como el Puntaje Matemáticas Saber 11 y el Puntaje Lectura Crítica Saber 11, las SMD son notablemente bajas (0,0385 y -0,0034, respectivamente), indicando un excelente equilibrio en estas importantes medidas de habilidad previa.

De manera similar, las diferencias de medias estandarizadas para las diversas categorías de estrato y características del colegio (Colegio No Oficial, Colegio Académico, etc.) se redujeron considerablemente, situándose en su mayoría por debajo del umbral comúnmente aceptado de 0,1, lo que sugiere que el proceso de matching fue exitoso en la creación de grupos comparables en términos de estas características observables. La muestra emparejada final consistió en 671 observaciones (446 mujeres STEM y 225 mujeres No STEM únicas, con algunas de estas últimas utilizadas más de una vez debido al matching con reemplazo).

Tabla 5: Tabla de balance del PSM.

Variable	Type	Media Tratado	Media Control	Diferencia de Medias	Distancia Media eCDF	Máxima Distancia eCDF
Distancia del Propensity Score	<i>Distance.</i>	0,5686	0,5685	0,0003	0,0021	0,0179
Puntaje Matemáticas	<i>Contin.</i>	56,2309	55,861	0,0385	0,0271	0,0650
Puntaje Lectura Crítica	<i>Contin.</i>	58,2242	58,2511	-0,0034	0,0130	0,0471
Estrato (Sin especificar)	<i>Binary</i>	0,0135	0,0179	-0,0391	0,0045	0,0045
Estrato 1	<i>Binary</i>	0,1592	0,1345	0,0676	0,0247	0,0247
Estrato 2	<i>Binary</i>	0,3924	0,4417	-0,1009	0,0493	0,0493
Estrato 3	<i>Binary</i>	0,3117	0,2892	0,0484	0,0224	0,0224
Estrato 4	<i>Binary</i>	0,0762	0,0717	0,017	0,0045	0,0045
Estrato 5	<i>Binary</i>	0,0426	0,0404	0,0111	0,0022	0,0022
Estrato 6	<i>Binary</i>	0,0045	0,0045	0,0000	0,0000	0,0000
Colegio No Oficial	<i>Binary</i>	0,2018	0,213	-0,0278	0,0112	0,0112
Colegio Oficial	<i>Binary</i>	0,7982	0,787	0,0278	0,0112	0,0112
Colegio Académico	<i>Binary</i>	0,4664	0,4731	-0,0135	0,0067	0,0067
Colegio Técnico	<i>Binary</i>	0,2399	0,2556	-0,0368	0,0157	0,0157
Colegio Técnico/Académico	<i>Binary</i>	0,2937	0,2713	0,0493	0,0224	0,0224

Una vez conformada la muestra emparejada, se estimó el Efecto Promedio del Tratamiento en los Tratados (ATT) mediante modelos de regresión ponderados, controlando siempre por el puntaje de Matemáticas en Saber 11. La Tabla 6 presenta los resultados de tres especificaciones progresivas del modelo ATT.

Tabla 6: Modelos ATT Progresivos (Muestra Emparejada por PSM).

	Variable Dependiente: Puntaje Raz. Cuantitativo Saber Pro		
	ATT 1	ATT 2	ATT 3
<i>Dummy</i> STEM (1=Stem)	7.648*** (2.137)	7.654*** (2.141)	7.587*** (2.109)
Puntaje matemáticas prueba Saber 11	1.946*** (0.097)	1.962*** (0.099)	1.984*** (0.099)
Constant	27.889*** (5.681)	25.028** (8.314)	19.761* (8.649)
Covariables			
Estrato	×	✓	✓
Carácter de la Institución Educativa	×	×	✓
Observaciones (emparejadas)	671	671	671
R Cuadrado (ponderado)	0.396	0.399	0.407
R Cuadrado Ajustado (pond.)	0.394	0.392	0.397
Observations	671	671	671
R ²	0.396	0.399	0.407
Adjusted R ²	0.394	0.392	0.397
Notas: *p<0.05; **p<0.01; ***p<0.001 Errores estándar robustos (HC1) entre paréntesis. Modelos estimados en muestra emparejada por PSM usando pesos.			

El Modelo ATT 1, que ajusta únicamente por el puntaje de Matemáticas Saber 11, muestra que la elección de una carrera STEM (*Dummy* STEM (1=Stem)) se asocia con un incremento promedio de 7.648 puntos ($p < 0.001$) en el puntaje de Razonamiento Cuantitativo Saber Pro, en comparación con mujeres en carreras No

STEM con características pre-universitarias similares.

Para evaluar la robustez de este hallazgo, el Modelo ATT 2 incorpora adicionalmente el Estrato socioeconómico como variable de control en el modelo de resultado. El coeficiente para la variable STEM se mantiene prácticamente inalterado en 7.654 puntos ($p < 0.001$). Finalmente, el Modelo ATT 3 añade también el carácter de la Institución Educativa de egreso. Incluso con este conjunto más amplio de controles post-matching, el efecto estimado de las carreras STEM permanece fuerte y significativo, con un coeficiente de 7.587 puntos ($p < 0.001$).

La notable estabilidad del coeficiente de Dummy STEM (1=Stem) a través de estas tres especificaciones (aproximadamente +7.6 puntos) sugiere que, después de equilibrar los grupos en una serie de características pre-universitarias mediante PSM, la asociación positiva entre la formación en áreas STEM y un mayor desarrollo de competencias en razonamiento cuantitativo para las mujeres es robusta. El Puntaje matemáticas prueba Saber 11 consistentemente se muestra como un predictor fuerte del desempeño posterior, con un coeficiente cercano a 1.95-1.98 en todos los modelos ATT. Los modelos explican aproximadamente el 40% de la varianza en los puntajes de Razonamiento Cuantitativo Saber Pro (R^2 ponderado ~ 0.396 - 0.407) en la muestra emparejada.

Aunque los resultados del PSM evidencian un efecto positivo de las carreras STEM sobre el desempeño cuantitativo de las mujeres, resulta relevante analizar si dicho impacto difiere respecto al observado en los hombres. Para explorar esta posible heterogeneidad por género, se estimó un modelo de interacción, conceptualmente alineado con un enfoque modificado de Diferencias en Diferencias (DiD). Este modelo permite no solo controlar por el desempeño previo y otras covariables relevantes, sino también identificar y aislar el efecto diferencial asociado a ser mujer y cursar una carrera STEM.

Tabla 7: Modelos DiD Progresivos: Efecto STEM en Mujeres sobre Razonamiento Cuantitativo.

	DIFF-DIFF (1)	DIFF-DIFF (2)	DIFF-DIFF (3)
Dummy STEM (1=Stem)	12.822*** (2.056)	12.591*** (2.067)	12.476*** (2.071)
Dummy Mujer (1=Si)	-2.616** (2.085)	-2.754** (2.102)	-2.804** (2.099)
Puntaje matemáticas prueba Saber 11	1.924*** (0.071)	1.941*** (0.073)	1.947*** (0.073)
STEM x Mujer (Interacción DiD)	-6.842** (2.594)	-6.792** (2.607)	-6.788** (2.609)
Constant	33.185*** (4.532)	28.996*** (6.163)	36.627* (16.722)
Covariables			
Estrato	×	✓	✓
Carácter de la Institución Educativa	×	×	✓
Observaciones	1401	1401	1401
Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ Errores estándar robustos (HC1). El coeficiente de interacción es el efecto DiD.			

Para tal fin, se utilizó la muestra completa de estudiantes ($N = 1.401$) con información disponible de las pruebas Saber 11 y Saber Pro. El modelo incluye como variables

explicativas la elección de una carrera STEM, el género del estudiante (dummy Mujer, donde 1=Sí), el puntaje previo en Matemáticas (Saber 11), y, de manera crucial, un término de interacción entre la elección de una carrera STEM y el género femenino. La Tabla 7 presenta los resultados de tres especificaciones progresivas de este modelo, estimados con errores estándar robustos (HC1).

El coeficiente de la variable Dummy STEM (1=Stem) representa el efecto de cursar una carrera STEM para el grupo de referencia de género, es decir, los hombres. En el Modelo DiD (1), que solo controla por el puntaje de Saber 11, los hombres en carreras STEM obtienen, en promedio, 12,822 puntos más ($p < 0,001$) en Razonamiento Cuantitativo Saber Pro que los hombres en carreras No STEM. Este efecto se mantiene robusto y altamente significativo al añadir controles por Estrato (Modelo DiD 2: $\beta = 12,591$, $p < 0,001$) y Carácter de la Institución Educativa (Modelo DiD 3: $\beta = 12,476$, $p < 0,001$).

La variable Dummy Mujer (1=Sí) indica la diferencia en el puntaje Saber Pro entre mujeres y hombres dentro de las carreras No STEM, controlando por los demás factores del modelo. En el Modelo DiD (1), las mujeres en carreras No STEM obtienen en promedio 2.616 puntos menos que los hombres en carreras No STEM ($p < 0.01$). Esta diferencia se mantiene estadísticamente significativa y de magnitud similar al añadir controles por estrato (Modelo DiD 2: coef. = -2.754, $p < 0.01$) y carácter de la institución (Modelo DiD 3: coef. = -2.804, $p < 0.01$). Esto sugiere una brecha de género base en el desempeño en Razonamiento Cuantitativo dentro de las disciplinas No STEM, desfavorable para las mujeres, incluso después de controlar por el puntaje previo en matemáticas.

El coeficiente de interés principal para este análisis DiD es el término de interacción STEM $\times$ Mujer (Interacción DiD). Este coeficiente en el Modelo DiD (1) es de -6.842 y es estadísticamente significativo al nivel del 1% ($p < 0.01$). Esto indica que el “efecto STEM” es significativamente diferente para las mujeres en comparación con los hombres. Específicamente, el impacto positivo de una carrera STEM en el Razonamiento Cuantitativo es aproximadamente 6.842 puntos menor para las mujeres que para los hombres. Este hallazgo se mantiene notablemente estable y significativo al incluir controles por Estrato (Modelo DiD 2: coef. = -6.792, $p < 0.01$) y Carácter de la Institución Educativa (Modelo DiD 3: coef. = -6.788, $p < 0.01$).

Por lo tanto, si bien las carreras STEM se asocian con un mayor rendimiento en Razonamiento Cuantitativo para ambos géneros en comparación con sus pares en carreras No STEM (controlando por el puntaje Saber 11), el “impulso” o “valor añadido” estimado de una carrera STEM es significativamente menor para las mujeres que para los hombres en la muestra analizada. Para calcular el efecto neto de STEM para las mujeres, se suma el coeficiente de Dummy STEM (efecto para hombres) y el coeficiente de STEM $\times$ Mujer. Por ejemplo, utilizando el Modelo DiD (3), el efecto neto de STEM para las mujeres sería de $12.476 - 6.788 = 5.688$ puntos. Este efecto neto sigue siendo positivo y estadísticamente significativo, pero es considerablemente inferior al efecto de 12.476 puntos estimado para los hombres. El Puntaje matemáticas prueba Saber 11 es un predictor consistentemente fuerte y positivo en todos los modelos.

Estos resultados sugieren que, aunque las carreras STEM contribuyen positivamente al fortalecimiento de competencias cuantitativas para ambos géneros, las mujeres no

parecen beneficiarse en la misma magnitud que los hombres. Esta diferencia podría estar asociada a factores estructurales y socioculturales ampliamente documentados en la literatura, como la persistencia de estereotipos de género, menores niveles de autoeficacia en áreas cuantitativas y ambientes académicos históricamente masculinizados dentro de algunas disciplinas STEM. En este sentido, los hallazgos evidencian que el acceso de las mujeres a programas STEM no garantiza por sí solo la eliminación de las brechas en el desarrollo de competencias analíticas.

Los análisis precedentes, basados en datos secundarios del ICFES, han cuantificado el impacto de la elección de carreras STEM en el desarrollo de competencias, revelando un efecto positivo, aunque diferencial por género. Sin embargo, para complementar esta perspectiva, resulta crucial comprender los factores que subyacen a la decisión inicial de las mujeres de optar por estas trayectorias académicas. Factores individuales, familiares, y experiencias educativas previas, no siempre disponibles en bases de datos administrativas, pueden jugar un rol determinante. Para explorar esta dimensión, la investigación se apoyó en datos primarios recolectados mediante encuestas a estudiantes universitarias en Risaralda, con el objetivo de identificar, a través de un modelo de árbol de decisión, los perfiles y las características más influyentes en la probabilidad de que una mujer elija una carrera STEM.

El árbol de decisión CART construido sobre la submuestra de 320 mujeres identificó seis variables con participación activa en la construcción de las divisiones del árbol: desempeño previo en ciencias, nivel educativo del padre, zona de residencia, nivel educativo de la madre, estrato socioeconómico y conocimiento previo en STEM. La Tabla 8 presenta la importancia relativa de cada variable, calculada como la contribución porcentual acumulada a la reducción del índice de Gini a lo largo de todas las divisiones.

Tabla 8: Importancia relativa de variables en el árbol de decisión CART.

Variable	Importancia relativa (%)
Desempeño previo en ciencias	47,8
Nivel educativo del padre	17,2
Zona de residencia	10,9
Nivel educativo de la madre	9,2
Estrato socioeconómico	7,4
Conocimiento previo en STEM	3,5
Tipo de colegio	3,4
Actividades STEM previas	0,6
Total	100
Nota: Importancia calculada como reducción acumulada del índice de Gini atribuible a cada predictor en todas las divisiones del árbol (algoritmo CART; cp = 0.01, minsplit = 20, N = 320 mujeres, semilla = 123). Métricas de ajuste in-sample: accuracy = 0.722 (IC 95%: 0.669–0.770), Kappa = 0.439, sensibilidad = 0.740, especificidad = 0.699, F1 = 0.746, AUC = 0.759. AUC en validación cruzada de 10 pliegues = 0.579, indicando sobreajuste moderado atribuible al tamaño muestral disponible.	

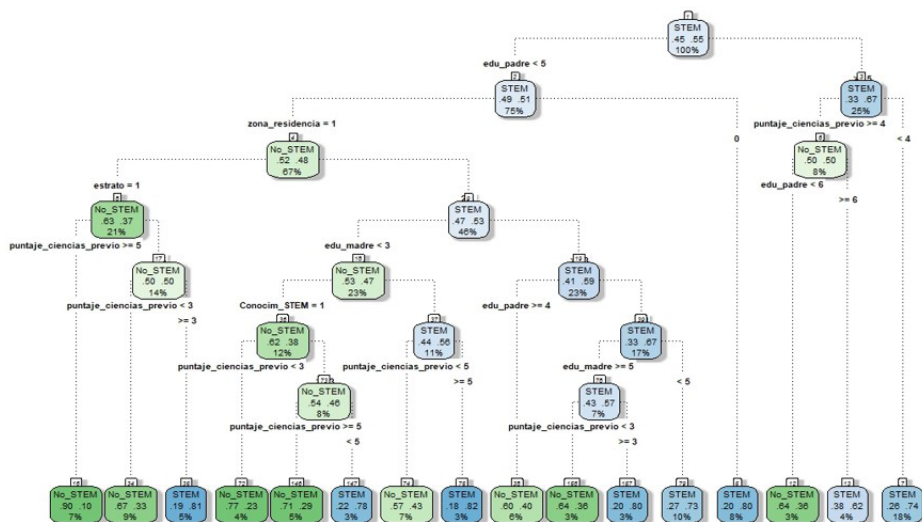
El desempeño previo en ciencias (puntaje_ciencias_previo) emerge como el factor con mayor poder discriminante, concentrando el 47.8% de la importancia total del modelo. Este resultado indica que la orientación y el rendimiento académico en áreas científicas durante la educación media constituye el predictor individual más diferenciador entre las mujeres que eligen carreras STEM y las que no, consistentemente con la literatura sobre el rol del rendimiento previo en la elección

de carrera científica (Wigfield y Eccles, 2000; Zhao y Perez-Felkner, 2022). Le sigue el nivel educativo del padre (17.2%), la zona de residencia (10.9%), el nivel educativo de la madre (9.2%) y el estrato socioeconómico (7.4%).

El peso conjunto del capital educativo familiar (sumando la contribución del padre y la madre) alcanza el 26.4% de la importancia total, subrayando la relevancia del entorno familiar como factor estructurante de la trayectoria académica femenina en línea con la teoría del capital cultural (Bourdieu, 1986). El conocimiento previo en STEM (Conocim_STEM) y el tipo de colegio aportan importancias menores (3.5% y 3.4% respectivamente), mientras que la participación en actividades STEM previas presenta la contribución más baja (0.6%).

La Figura 6 presenta la estructura jerárquica completa del árbol, que permite identificar las reglas de segmentación condicional a través de las cuales estos factores se combinan para definir perfiles de elección. En el nodo raíz, la primera división se produce sobre el nivel educativo del padre ($\text{edu_padre} < 5$), diferenciando a las mujeres según el capital educativo paterno. En la rama izquierda —que agrupa al 75% de la muestra con padre de menor nivel educativo—, la segunda división se produce sobre la zona de residencia, donde las mujeres de zona rural ($\text{zona_residencia} = 0$) presentan una mayor proporción No STEM (63%–37%). La rama derecha (que agrupa al 25% de la muestra con padre de mayor nivel educativo) se subdivide según el desempeño previo en ciencias ($\text{puntaje_ciencias_previo} \geq 4$), donde las mujeres con mayor rendimiento científico previo muestran una probabilidad STEM del 67%.

Figura 6: Árbol de Decisión: Probabilidad de Elección de Carrera STEM por Mujeres (Risaralda, 2023).



Estas reglas de segmentación revelan que el capital educativo familiar y el rendimiento científico previo no operan de forma independiente, sino que configuran perfiles de elección donde la combinación de factores familiares, territoriales y académicos determina de manera condicional la probabilidad de elección STEM.

Esta perspectiva, inaccesible desde los modelos de regresión estándar, permite identificar subgrupos específicos con diferente propensión a la elección STEM, información directamente útil para el diseño de políticas focalizadas de promoción de la participación femenina en estas áreas.

4. Conclusiones

El conjunto de análisis realizados, empleando diversas metodologías y fuentes de datos, converge en hallazgos complementarios que permiten una comprensión más integral del fenómeno estudiado, al tiempo que evidencia la importancia de triangular estrategias empíricas cuando se trabaja con datos observacionales.

Los Modelos de Mínimos Cuadrados Ordinarios (MCO) evidenciaron una asociación positiva y estadísticamente significativa entre la elección de carreras STEM y el desempeño en razonamiento cuantitativo. En el modelo más completo, dicha asociación se estimó en aproximadamente 5.88 puntos en el puntaje Saber Pro para las mujeres, controlando por el puntaje Saber 11 y características del estrato y tipo de colegio. No obstante, dado que los modelos MCO con datos observacionales no eliminan la posibilidad de sesgo por autoselección, esta estimación debe interpretarse como una asociación parcialmente ajustada antes que como un efecto causal neto.

Con el objetivo de mitigar el potencial sesgo de autoselección en variables observables, se implementó el Propensity Score Matching (PSM). Los resultados del Efecto Promedio del Tratamiento en los Tratados (ATT), estimado sobre la muestra emparejada, indicaron una asociación positiva de las carreras STEM de aproximadamente 7.59 puntos en el puntaje de razonamiento cuantitativo. Esta estimación es ligeramente superior a la obtenida con MCO, lo que podría sugerir que los modelos de regresión sin emparejamiento subestiman moderadamente la asociación para las mujeres que efectivamente eligen STEM, o que el PSM logró conformar grupos de comparación más equilibrados en variables observables, aislando de mejor forma la diferencia asociada al tratamiento. La robustez de este hallazgo a través de diferentes especificaciones del modelo ATT post-matching refuerza la conclusión de un beneficio tangible. No obstante, esta estimación está condicionada al supuesto de ignorabilidad condicional (esto es, la ausencia de confusión por variables no observadas), por lo que debe interpretarse como consistente con un efecto causal positivo antes que como evidencia causal definitiva.

El modelo de interacción con término STEM $\times$ Mujer ofreció una perspectiva comparativa entre géneros. Este modelo estimó un efecto base de las carreras STEM para los hombres de aproximadamente 12.48 puntos. Crucialmente, el término de interacción “STEM $\times$ Mujer” fue negativo y significativo (aprox. -6.79 puntos), indicando que el “impulso” de una carrera STEM en razonamiento cuantitativo es significativamente menor para las mujeres en comparación con los hombres. El efecto neto de STEM para las mujeres, derivado de este modelo, fue de 5.688 puntos (12.476 - 6.788), el cual, aunque positivo y significativo, es inferior tanto al efecto estimado para los hombres en el mismo modelo como al ATT estimado mediante PSM (7.59 puntos). Esta diferencia entre el ATT del PSM (enfocado exclusivamente en mujeres comparadas con sus pares No STEM más similares en observables) y el efecto neto para mujeres del modelo de interacción (que incorpora a los hombres como referencia para estimar el diferencial de género) es esperable dado que ambos métodos responden a preguntas

y poblaciones de referencia distintas. No obstante, los dos resultados convergen en señalar un beneficio positivo y significativo de las carreras STEM para las mujeres en términos de razonamiento cuantitativo, siendo este beneficio significativamente menor en magnitud relativa al estimado para los hombres.

Este hallazgo resulta consistente con investigaciones internacionales y latinoamericanas que han documentado que, aunque las mujeres logran acceder cada vez más a programas STEM, continúan enfrentando barreras estructurales y socioculturales que afectan diferencialmente sus trayectorias académicas y el desarrollo de competencias cuantitativas. Factores como la persistencia de estereotipos de género, menores niveles de autoeficacia en matemáticas, la ausencia de referentes femeninos y los ambientes académicos históricamente masculinizados podrían contribuir a explicar por qué las ganancias asociadas a las carreras STEM son menores para las mujeres en comparación con los hombres (Fernández Soto et al., 2024; Merayo y Ayuso, 2023; Salazar et al., 2024).

Por su parte, el Árbol de Decisión CART, construido con datos primarios sobre la submuestra de 320 mujeres, identificó un conjunto de factores jerárquicamente importantes en la segmentación de perfiles de elección STEM. El desempeño previo en ciencias emergió como el predictor con mayor poder discriminante, concentrando el 47.8% de la importancia total del modelo, seguido por el nivel educativo del padre (17.2%), la zona de residencia (10.9%), el nivel educativo de la madre (9.2%) y el estrato socioeconómico (7.4%). El capital educativo familiar en su conjunto (sumando la contribución de ambos progenitores) alcanza el 26.4% de la importancia total, subrayando el peso del entorno familiar como factor estructurante de la trayectoria académica femenina hacia las carreras STEM.

Es relevante notar que varias de estas variables (estrato socioeconómico, zona de residencia y tipo de colegio) también fueron incorporadas como controles en los modelos MCO y en la estimación del propensity score para el PSM. Si bien en los modelos de regresión estas variables no mostraron coeficientes individualmente significativos después de controlar por el puntaje Saber 11 y la condición STEM, el árbol de decisión sugiere que sí desempeñan un rol relevante en la segmentación de perfiles de elección, particularmente cuando interactúan de forma condicional con el rendimiento científico previo y el capital educativo familiar. Este resultado ilustra una complementariedad metodológica importante: los modelos de regresión estiman efectos promedio ajustados bajo supuestos de linealidad y aditividad, mientras que el árbol de decisión revela reglas de clasificación condicional y perfiles de elección que emergen de las interacciones no lineales entre predictores. Ambas aproximaciones, lejos de ser contradictorias, iluminan dimensiones distintas del mismo fenómeno y se refuerzan mutuamente en la comprensión de los determinantes de la elección femenina de carreras STEM.

No obstante, es preciso señalar que el árbol presentó sobreajuste moderado evidenciado por la diferencia entre el AUC in-sample (0.759) y el AUC en validación cruzada de 10 pliegues (0.579), por lo que sus hallazgos deben interpretarse con carácter exploratorio y descriptivo, como insumo para la generación de hipótesis y el diseño de intervenciones focalizadas, antes que como evidencia predictiva definitiva.

La triangulación de los resultados obtenidos mediante las distintas metodologías aplicadas permite establecer dos conclusiones centrales. En primer lugar, la elección

de carreras STEM se asocia con un desarrollo positivo y estadísticamente significativo de las competencias de razonamiento cuantitativo en mujeres de Risaralda, con estimaciones que oscilan entre 5.69 y 7.59 puntos según el método empleado, y que son consistentes con un efecto causal positivo bajo los supuestos de identificación de cada estrategia. En segundo lugar, este beneficio es significativamente menor en magnitud que el estimado para los hombres, lo que evidencia una brecha de género en los retornos de la educación STEM en competencias cuantitativas que merece atención tanto desde la investigación como desde la política pública. Adicionalmente, la elección de carreras STEM por parte de las mujeres está determinada por una interacción compleja de factores académicos previos, capital educativo familiar y condiciones territoriales y socioeconómicas, donde el desempeño previo en ciencias constituye el factor individual más diferenciador.

En este sentido, los resultados sugieren que la elección de carreras STEM no depende exclusivamente de las capacidades académicas individuales, sino también de factores asociados al entorno familiar, al capital cultural y a las oportunidades de exposición temprana a experiencias científicas y tecnológicas. Esto refuerza la necesidad de comprender la participación femenina en STEM como un fenómeno multidimensional, influenciado por dinámicas educativas, sociales y culturales que trascienden el contexto universitario.

Esta investigación, si bien emplea metodologías robustas para estimar asociaciones y mitigar sesgos de selección observable, no está exenta de limitaciones. La principal radica en la imposibilidad de controlar completamente por factores no observados (como la motivación intrínseca, la autoeficacia académica o el capital social familiar orientado a trayectorias STEM) que podrían influir tanto en la autoselección hacia carreras STEM como en el posterior desarrollo de competencias cuantitativas. Asimismo, la generalizabilidad de los hallazgos está circunscrita al contexto del departamento de Risaralda y a la operacionalización de las competencias mediante los puntajes de las pruebas Saber Pro, los cuales, si bien constituyen una medida estandarizada de amplio uso en el sistema educativo colombiano, no capturan la totalidad del aprendizaje universitario ni las competencias específicas desarrolladas en cada campo STEM. Adicionalmente, los modelos MCO más complejos presentaron indicios de no linealidad y multicolinealidad que, aunque el coeficiente principal de STEM se mostró robusto a través de las especificaciones, sugieren cautela en la interpretación de los coeficientes de las variables de control.

En conjunto, los resultados evidencian que las carreras STEM constituyen un espacio relevante para el fortalecimiento de competencias analíticas y cuantitativas. Sin embargo, también muestran que las brechas de género persisten incluso dentro de estas disciplinas, lo que sugiere que el acceso de las mujeres a programas STEM, aunque necesario, no resulta suficiente por sí solo para garantizar condiciones plenamente equitativas en el desarrollo académico.

La recolección de datos primarios para el análisis de los factores de elección, aunque metodológicamente enriquecedora al permitir incorporar variables no disponibles en los datos secundarios, incluyó una etapa de muestreo por conveniencia que podría afectar la representatividad completa de esa submuestra y limitar la inferencia probabilística estricta a la población de referencia. Adicionalmente, la clasificación de las carreras como STEM, aunque basada en un referente externo validado, no

permite un análisis detallado de la heterogeneidad existente dentro de los diversos campos que componen estas disciplinas, lo cual podría enmascarar dinámicas específicas de subáreas particulares, por ejemplo, diferencias entre ingenierías, ciencias básicas y tecnologías de la información.

Los hallazgos y limitaciones identificados abren diversas líneas de investigación futura. Por una parte, sería relevante complementar los presentes resultados cuantitativos con estudios cualitativos que exploren en profundidad las vivencias, percepciones y barreras que enfrentan las mujeres en sus trayectorias STEM en Risaralda, particularmente en lo que respecta a variables psicosociales como la autoeficacia matemática y la identidad STEM, cuya influencia en la elección de carrera está bien documentada en la literatura internacional (Bandura, 1997; Wigfield y Eccles, 2000) pero no pudo ser medida directamente en el presente estudio. Otra línea de investigación podría enfocarse en la heterogeneidad de efectos dentro de los campos STEM, analizando si el impacto en competencias cuantitativas y los factores de elección varían entre ingenierías, ciencias básicas y áreas tecnológicas.

Además, se sugiere la evaluación del impacto de intervenciones específicas (programas de mentoría temprana, talleres de fomento vocacional en educación media, incentivos a la elección de carreras STEM) diseñadas para actuar sobre los factores identificados como más influyentes: el rendimiento científico previo, el capital educativo familiar y las condiciones territoriales. El desarrollo de estudios longitudinales con datos de panel permitiría aplicar estrategias de identificación más robustas (como efectos fijos individuales o variables instrumentales) para un mejor control de la heterogeneidad no observada.

A partir de estos hallazgos, se recomienda que las instituciones educativas y los responsables de política pública fortalezcan estrategias orientadas a promover una mayor participación y permanencia de las mujeres en áreas STEM. Esto incluye el desarrollo de programas de mentoría femenina, iniciativas de orientación vocacional temprana, espacios de acercamiento a la ciencia y la tecnología desde etapas escolares, así como acciones institucionales dirigidas a reducir estereotipos de género y fortalecer la autoeficacia de las estudiantes en competencias cuantitativas. Asimismo, resulta relevante promover ambientes académicos más inclusivos dentro de las disciplinas STEM, especialmente en áreas tradicionalmente masculinizadas.

Por último, investigar el rol de los factores institucionales de las universidades (cultura organizacional, políticas de inclusión, representación femenina en el cuerpo docente STEM) y el desarrollo de instrumentos para medir variables psicosociales latentes (autoeficacia, interés) enriquecería la comprensión de este complejo fenómeno.

Referencias

- Angrist, J. D. y Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press. doi: <https://doi.org/10.1515/9781400829828>
- Arredondo Trapero, F. G., Vázquez Parra, J. C. y Velázquez Sánchez, L. M. (2019). STEM y Brecha de Género en Latinoamérica. *Revista de El Colegio de San Luis*, 9(18), pp. 137–158. doi: <https://doi.org/10.21696/rcsl9182019947>
- Austin, P. C. (2011). Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharmaceutical Statistics*, 10(2), pp. 150–161. doi: <https://doi.org/10.1002/pst.433>

- Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. W H Freeman/Times Books/ Henry Holt & Co.
- Becker, S. O. y Ichino, A. (2002). Estimation of Average Treatment Effects Based on Propensity Scores. *The Stata Journal*, 2(4), pp. 358–377. doi: <https://doi.org/10.1177/1536867X0200200403>
- Botella, C., Rueda, S., López-Iñesta, E. y Marzal, P. (2019). Gender Diversity in STEM Disciplines: A Multiple Factor Problem. *Entropy*, 21(1), pp. 30. doi: <https://doi.org/10.3390/e21010030>
- Bourdieu, P. (1986). The Forms of Capital. En J. G. Richardson (Ed.), *Handbook of Theory and Research for the Sociology of Education* (pp. 241–258). Greenwood Press. Recuperado de <https://publish.illinois.edu/crittheory/files/2023/01/Bourdieu-The-Forms-of-Capital.pdf>
- Breiman, L., Friedman, J., Olshen, R. y Stone, C. (1984). *Classification and Regression Trees* (1st ed.). Chapman and Hall/CRC. doi: <https://doi.org/10.1201/9781315139470>
- Caliendo, M. y Kopeinig, S. (2008). Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys*, 22(1), pp. 31–72. doi: <https://doi.org/10.1111/j.1467-6419.2007.00527.x>
- Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative and Mixed Methods Approaches* (4th ed.). Thousand Oaks, CA: Sage.
- Esteban Ramiro, B., Moreno López, R. y Marí Ytarte, R. M. (2024). ¿De ciencias o de letras? Brechas de género en el espacio educativo STEM. *European Public & Social Innovation Review*, 9, pp. 1–20. doi: <https://doi.org/10.31637/epsir-2024-1462>
- Etikan, I., Musa, S. A. y Alkassim, R. S. (2015). Comparison of Convenience Sampling and Purposive Sampling. *American Journal of Theoretical and Applied Statistics*, 5(1), pp. 1–4. doi: <https://doi.org/10.11648/j.ajtas.20160501.11>
- Fajardo Ortiz, E. J., Romero, H. y Lozano, H. (2025). Brechas de género en los resultados académicos de los estudiantes de educación secundaria en Colombia. *Interciencia*, 50(9), pp. 491–496. Recuperado de https://www.interciencia.net/wp-content/uploads/2025/10/03_7308_A_Fajardo_v50n9_6.pdf
- Fernández Soto, M., Galván, E., Robaina, S., Tenenbaum, V. y Tomassini, C. (2024). Brechas de género en las trayectorias académicas en Uruguay: Formación, producción y acceso a cargos. *Revista Iberoamericana de Ciencia, Tecnología y Sociedad - CTS*, 20(58), pp. 65–89. doi: <https://doi.org/10.52712/issn.1850-0013-510>
- García-Peñalvo, F. J., García-Holgado, A., Dominguez, A. y Pascual, J. (2022). *Women in STEM in Higher Education: Good Practices of Attraction, Access and Retainment in Higher Education*. Springer. doi: <https://doi.org/10.1007/978-981-19-1552-9>
- Hermans, S., Gijsen, M., Mombaers, T. y van Petegem, P. (2022). Gendered patterns in students' motivation profiles regarding iSTEM and STEM test scores: a cluster analysis. *International Journal of STEM Education*, 9(1), pp. 67. doi: <https://doi.org/10.1186/s40594-022-00379-3>
- Imbens, G. W. y Wooldridge, J. M. (2009). Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature*, 47(1), pp. 5–86. doi: <https://doi.org/10.1257/jel.47.1.5>

- Indiana University Indianapolis. (2012). *STEM-Designated Degree Program List*. Indiana: Universidad de Indiana. Recuperado de <https://international.indianapolis.iu.edu/doc/visas/STEM-Designated-Degree-Program-List.pdf>
- Kayan-Fadlelmula, F., Sellami, A., Abdelkader, N. y Umer, S. (2022). A systematic review of STEM education research in the GCC countries: trends, gaps and barriers. *International Journal of STEM Education*, 9(1), pp. 2. doi: <https://doi.org/10.1186/s40594-021-00319-7>
- Kim, J., Jung, J. y Mlambo, Y. A. (2021). Institutional Selectivity and Occupational Outcomes for STEM Graduates: A Generational Comparison. *The Journal of Higher Education*, 92(3), pp. 435–464. doi: <https://doi.org/10.1080/00221546.2020.1819945>
- Loh, W.-Y. (2011). Classification and regression trees. *WIREs Data Mining and Knowledge Discovery*, 1(1), pp. 14–23. doi: <https://doi.org/10.1002/widm.8>
- López, P., Simó, P. y Marco, J. (2023). Understanding STEM career choices: A systematic mapping. *Heliyon*, 9(6), pp. e16676. doi: <https://doi.org/10.1016/j.heliyon.2023.e16676>
- Merayo, N. y Ayuso, A. (2023). Analysis of barriers, supports and gender gap in the choice of STEM studies in secondary education. *International Journal of Technology and Design Education*, 33(4), pp. 1471–1498. doi: <https://doi.org/10.1007/s10798-022-09776-9>
- Miller, D. I. y Halpern, D. F. (2013). Can spatial training improve long-term outcomes for gifted STEM undergraduates? *Learning and Individual Differences*, 26, pp. 141–152. doi: <https://doi.org/10.1016/j.lindif.2012.03.012>
- Patterson, L., Varadarajan, D. S. y Saji Salim, B. (2020). Women in STEM/SET: gender gap research review of the United Arab Emirates (UAE) – a meta-analysis. *Gender in Management: An International Journal*, 36(8), pp. 881–911. doi: <https://doi.org/10.1108/GM-11-2019-0201>
- Riegle-Crumb, C. y King, B. (2010). Questioning a White Male Advantage in STEM: Examining Disparities in College Major by Gender and Race/Ethnicity. *Educational Researcher*, 39(9), pp. 656–664. doi: <https://doi.org/10.3102/0013189X10391657>
- Rodríguez de Luque, J. J. (2025). Educational Gender Gaps in Bogotá and the Right to Education. *Revista Colombiana de Educación*, (95). doi: <https://doi.org/10.17227/rce.num95-18793>
- Rosenbaum, P. R. y Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), pp. 41–55. doi: <https://doi.org/10.1093/biomet/70.1.41>
- Salazar, L. A., Olivari, V., Katherine, O. y Araya, F. (2024). Diagnosis of gender gaps in higher education: a case study in the civil construction program at Federico Santa María Technical University. *Revista Ingeniería de Construcción*, 39(3), pp. 1–15. doi: <https://doi.org/10.7764/RIC.00123.21>
- Saunders, M., Lewis, P. y Thornhill, A. (2016). *Research Methods for Business Students* (7th ed.). Pearson.
- Stuart, E. A. (2010). Matching Methods for Causal Inference: A Review and a Look Forward. *Statistical Science: A Review Journal of the Institute of Mathematical Statistics*, 25(1), pp. 1–21. doi: <https://doi.org/10.1214/09-STS313>

- UNESCO. (2024). *Changing the Equation: Securing STEM Futures for Women*. Recuperado de <https://unesdoc.unesco.org/ark:/48223/pf0000391384>
- Verdugo-Castro, S., García-Holgado, A. y Sánchez-Gómez, M. C. (2022). The gender gap in higher STEM studies: A systematic literature review. *Heliyon*, 8(8), pp. e10300. doi: <https://doi.org/10.1016/j.heliyon.2022.e10300>
- White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica: Journal of the Econometric Society*, 48(4), pp. 817–838. doi: <https://doi.org/10.2307/1912934>
- Wigfield, A. y Eccles, J. S. (2000). Expectancy–Value Theory of Achievement Motivation. *Contemporary Educational Psychology*, 25(1), pp. 68–81. doi: <https://doi.org/10.1006/ceps.1999.1015>
- Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). MIT Press. Recuperado de <https://www.jstor.org/stable/j.ctt5hhcfr>
- Zhao, T. y Perez-Felkner, L. (2022). Perceived abilities or academic interests? Longitudinal high school science and mathematics effects on postsecondary STEM outcomes by gender and race. *International Journal of STEM Education*, 9(1), pp. 42. doi: <https://doi.org/10.1186/s40594-022-00356-w>