Evaluación de resultados del programa Empleo Incluyente implementado por la Secretaría Distrital de Desarrollo Económico

PRODUCTO 1: INFORME METODOLÓGICO
IPSOS NAPOLEÓN FRANCO

Secretaría Distrital de Planeación
Bogotá, D.C.
3 de diciembre de 2024

Contenido

I. INTRODUCCIÓN 4
1. Programa de Empleo Incluyente 4
1.1 Actores involucrados en la intervención del PEI 6
1.2 Normatividad relacionada con el PEI 7
2. Evaluación de resultados al programa 8
3. La evaluación al PEI en el contexto de evaluaciones a programas similares 9
II. Desarrollo metodológico de la evaluación 13
4. Metodologías cuantitativas 13
4.1 Descripción de las fuentes primarias y/o secundarias 13
4.2 Descripción y justificación de las metodologías. 14
4.2.1 Descripción y justificación de las estadísticas descriptivas 14
4.2.2 Descripción y justificación del Análisis de supervivencia 16
4.2.3 Descripción y justificación de los modelos SEM 20
4.2.4 Descripción y justificación de las redes neuronales 23
5. Metodologías cualitativas 27
5.1 Enfoque teórico conceptual 27
5.2 Desarrollo Metodológico 28
5.2.1Desarrollo y justificación del análisis narrativo 30
5.2.2 Desarrollo y justificación de la evaluación realista 34
5.2.3 Estudios de caso 42
III. Diseño de muestra 46
6.Diseño muestral cuantitativo 46
6.1 Descripción de la estrategia muestral cuantitativa 46
6.1.1 Características del diseño muestral 46
6.1.2 Proceso de conformación de estratos 46
6.1.3 Algoritmos de selección 47
6.2 Descripción de la estrategia de estimación de parámetros 48
6.2.1 Probabilidades de inclusión y factores de expansión 48
6.2.2 Estimadores de parámetros de caracterización 49
6.3 Medidas de calidad de parámetros de caracterización 53
6.3.1 Estimación de la varianza 53
6.3.2. Simulación de escenarios 54

6.4 Determinación del tamaño de muestra 55
6.4.1 Expresión matemática para el cálculo del tamaño de muestra 55
6.4.2 Cálculo del tamaño de muestra 56
6.5 Tratamiento de la población TRANS 56
6.6 Mitigación del efecto de la no respuesta no ignorable 57
7. Diseño muestral cualitativo 61
7.1 Descripción de estrategia y tipologías de muestreo 61
7.2 Componente 1: Personas benefiaciarias 63
7.3 Componente 2: Empresas beneficiarias 66
IV. Triangulación 68

I. INTRODUCCIÓN

1. Programa de Empleo Incluyente[1]

El Programa Empleo Incluyente (PEI) es una iniciativa de la Secretaría Distrital de Desarrollo Económico de Bogotá que busca incentivar la contratación laboral de poblaciones que enfrentan barreras de acceso al empleo formal. A través de la entrega de incentivos monetarios a las empresas, acompañado, según la demanda, de servicios de intermediación para el encuentro entre oferta y demanda, el programa pretende cerrar brechas de inclusión laboral y promover prácticas de diversidad, equidad e inclusión en el ámbito empresarial.

Entre las necesidades y problemáticas que motivaron la creación y puesta en marcha del programa se destacan:

1. **Desigualdad en el acceso al empleo formal**: grupos poblacionales como mujeres, jóvenes, personas mayores de 50 años, personas con discapacidad, víctimas del conflicto armado, comunidades étnicas y personas trans, entre otros, enfrentan discriminación y estereotipos que limitan sus oportunidades laborales. Algunas de estas brechas han sido estructurarles y se profundizaron posterior a la pandemia del COVID-19.
2. **Prácticas empresariales excluyentes**: las empresas a menudo no cuentan con políticas inclusivas o desconocen los beneficios de la diversidad en el lugar de trabajo. A su vez, el ecosistema de empleo de la ciudad requiere la adopción de enfoques que promuevan el acceso a oportunidades laborales de forma diferencial.
3. **Desconexión entre oferta y demanda laboral**: existe una brecha significativa entre las habilidades que poseen las personas que enfrentan barreras de acceso al empleo formal y las competencias requeridas por el mercado laboral. A menudo, los mecanismos de intermediación no logran conectar de manera eficiente a las poblaciones que históricamente han sido excluidas, con las empresas que requieren talento, lo que genera ineficiencias tanto para las empresas como para las posibles personas empleadas.

La principal intervención del PEI son los **Incentivos monetarios para empresas**, que consisten en la entrega de sumas de dinero a empresas que contraten a personas históricamente excluidas, con montos variables según el grupo poblacional y el término del contrato. Para la identificación de estos montos se definieron tres tipos de brechas para las poblaciones y dos términos de la relación contractual. En cuanto a las brechas se tienen aquellas para las poblaciones altamente afectadas por estigmatización (personas trans), las que tienen una mayor afectación por altas tasas de desempleo o desempleo de larga duración (personas jóvenes o mayores de 50 años), y, finalmente, en un tercer grupo están las poblaciones afectadas por brechas significativas de empleo por rezagos de educación, falta de experiencia, estereotipos o necesidad de especial protección constitucional (las demás

[1] Algunos textos fueron editados con la asistencia de herramientas de LLLM.

poblaciones priorizadas). Con respecto a los términos de las relaciones contractuales, se considera un monto para los contratos mayores a seis meses y menores a un año, y otro para aquellos de más de un año o a término indefinido.

De otra parte, y con el fin de hacer una distribución equitativa de los recursos entre las diferentes empresas del tejido empresarial de Bogotá, se definió el "modelo de equidad" con tres objetivos: (i) promover la participación de la micro y pequeña empresa, (ii) evitar la concentración de los beneficios en pocas grandes empresas, y (iii) promover la participación de las 19 poblaciones[2].

Adicionalmente, y aunque no se trata de una oferta que se brindó a todas las empresas, desde la SDDE se dispuso de la capacidad instalada de la Agencia Distrital de Empleo para acompañar la implementación del programa, en particular en acciones de **intermediación laboral**, y a su vez se realizaron alianzas con entidades públicas y privadas que pudieran contribuir con su capacidad y experiencia en este mismo rol de acercamiento de la oferta con la demanda. Lo anterior, operó de acuerdo con la demanda y necesidad de las empresas, quienes en todo caso contaron con la autonomía para la gestión de sus vacantes de la manera que consideraron pertinente.

La implementación del programa se desarrolla en tres etapas:

1. **Gestión del empleo y articulación con el ecosistema de empleo de la ciudad**: de acuerdo con la resolución 003 de 2023, su objetivo es lograr la vinculación de las 19 poblaciones objetivo del programa, mediante la promoción del beneficio entre las empresas, la implementación de rutas para la gestión del empleo y la mitigación de las barreras al empleo de la población.
2. **Postulación de las empresas, verificación y validación de los soportes documentales**: en esta etapa, se busca asignar el beneficio a las empresas que cumplan las condiciones establecidas en el Decreto Distrital 399 de 2022.
3. **Giro del beneficio a las empresas que cumplen los requisitos**. Realizada la verificación de la empresa y de las poblaciones, se determina el/los beneficios a entregar por empresa, de acuerdo con la o las categorías de la población que han sido empleadas. Esta etapa se realiza a través de la entidad financiera encargada de la administración de los recursos asignados al programa.

En cuanto a los resultados esperados del PEI, en el corto plazo, se espera un incremento en el número de contrataciones de poblaciones que han tenido mayores dificultades para la colocación, respondiendo de manera inmediata a las barreras de acceso al empleo formal que enfrentan. A medida que el programa avanza hacia el mediano plazo, los esfuerzos se traducirán en mejoras en la permanencia laboral y en los niveles de ingreso de las personas beneficiarias, reflejando una estabilización en su participación en el mercado laboral. Paralelamente, las empresas participantes adoptarán políticas inclusivas, promoviendo

[2] Aunque el decreto 399 de 2022 por medio del cual se adopta el programa, incluye 19 grupos poblacionales, no fue posible establecer un documento que acreditara la condición para la población que realiza actividades sexuales pagadas.

cambios sustanciales en sus prácticas organizacionales y reduciendo de manera tangible las prácticas discriminatorias en sus procesos de selección y gestión del talento.

En el largo plazo, el impacto acumulado del PEI se manifestará en una disminución sostenida de las tasas de desempleo entre las poblaciones con barreras de acceso al mercado laboral, contribuyendo a la reducción de la desigualdad socioeconómica en Bogotá. La transformación esperada no solo abarcará aspectos cuantitativos, como el acceso al empleo y los ingresos, sino también cualitativos, con el establecimiento de una cultura empresarial y un ecosistema del empleo que valora y promueve la diversidad e inclusión como pilares fundamentales para el desarrollo y la competitividad.

Los impactos esperados del PEI sobre las personas beneficiarias abarcan múltiples dimensiones que van más allá de la vinculación laboral. En primer lugar, se proyecta un empoderamiento económico significativo, que se traduce en una mayor autonomía y capacidad de los participantes para tomar decisiones financieras, incrementando su control sobre los recursos y mejorando su estabilidad económica. En paralelo, el programa contribuirá al desarrollo de competencias de las personas beneficiarias, proporcionándoles la experiencia y las habilidades necesarias para acceder a nuevas oportunidades de crecimiento dentro del mercado laboral formal.

Adicionalmente, el PEI generará mejoras en la calidad de vida de los participantes, facilitando su acceso a la seguridad social y proporcionando un entorno de estabilidad económica que repercute positivamente en su bienestar general. Este impacto se complementará con un incremento en la inclusión social, ya que el programa busca reducir la exclusión y estigmatización que enfrentan las poblaciones con difícil colocación, promoviendo un entorno más equitativo y justo para todas las personas beneficiarias.

1.1 Actores involucrados en la intervención del PEI

Para el funcionamiento del programa entre los actores involucrados se destaca, el equipo de la Dirección de Desarrollo Empresarial y Empleo (DDEE) de la Secretaría Distrital de Desarrollo Económico (SDDE), más específicamente el equipo de la Subdirección de Empleo y Formación (SEF) que tuvo a cargo el programa, así como el personal de la Agencia Distrital de Empleo (ADE) que acompaña la ruta de intermediación laboral.

A su vez, el talento humano que hace parte de los tres operadores principales que permiten el funcionamiento del programa: la entidad financiera que administró los recursos y realizó los giros (FIDUCIARIA LA PREVISORA S.A.), el gestor documental que hizo la verificación y validación de los soportes (CORPORACION PARA LA INVESTIGACION EL DESARROLLO SOSTENIBLE Y LA PROMOCION SOCIAL) y la interventoría técnica externa (ITECO CONSULTORES S.A.S BIC).

También es importante mencionar las instancias de coordinación y validación de las orientaciones del programa, como lo es el Comité Interinstitucional del Sistema Distrital para la Mitigación del Impacto Económico, el Fomento y la Reactivación Económica (Comité EMRE) del Distrito.

Finalmente, las empresas que vinculan a las personas de las diferentes poblaciones, y los aliados y prestadores autorizados del Servicio Público de Empleo que facilitaron la gestión de la búsqueda de personas para la vinculación laboral.

1.2 Normatividad relacionada con el PEI

A continuación, se listan los documentos normativos que reglamentan y dan sustento a la intervención:

Del orden internacional:

- Convenio No. 111 de la Organización del Trabajo (OIT) de 1958 (Ratificado por Colombia en 1969): Convenio sobre la discriminación en materia de empleo y ocupación

Del orden nacional:

- Constitución Política de Colombia (Art. 25, 53, 54, 113, 115, 209): artículos relacionados con el derecho al trabajo, los principios mínimos fundamentales del estatuto del trabajo y los deberes del Estado en relación con la ubicación laboral.
- Ley 1257 de 2008: Por la cual se dictan normas de sensibilización, prevención y sanción de formas de violencia y discriminación contra las mujeres, se reforman los Códigos Penal, de Procedimiento Penal, la Ley 294 de 1996 y se dictan otras disposiciones.
- Ley 1448 de 2011 Por la cual se dictan medidas de atención, asistencia y reparación integral a las víctimas del conflicto armado interno.

Del orden distrital:

- Acuerdo 257 de 2006: Por el cual se dictan normas básicas sobre la estructura, organización y funcionamiento de los organismos y de las entidades de Bogotá, Distrito Capital, y se expiden otras disposiciones
- Acuerdo 421 de 2009: Por medio del cual se ordena la creación del Sistema Integral de Protección a Mujeres Víctimas de Violencias – SOFIA -
- Decreto 437 de 2016: Por el cual se modifica la Estructura Organizacional de la Secretaría Distrital de Desarrollo Económico
- Decreto 399 de 2022: Por medio del cual se crea el Programa Empleo Incluyente para el cierre de brechas de población de difícil empleabilidad y se dictan otras disposiciones
- Resolución 003 de 2023: Por medio de la cual se reglamenta el Decreto Distrital 399 de 2022 y se adopta el modelo para la equidad e imparcialidad, en la asignación de beneficios a las empresas participantes en el Programa Empleo Incluyente para el cierre de brechas de población de difícil empleabilidad y se dictan otras disposiciones

Políticas Públicas con enfoque poblacional del orden nacional:

- CONPES 3554 de 2008: Política Nacional de Reintegración Social y Económica para Personas y Grupos Armados Ilegales
- CONPES 3931 de 2018: Política Nacional para la Reincorporación Social y Económica de Exintegrantes de las FARC-EP

Políticas Públicas con enfoque poblacional del orden distrital:

- Decreto 470 de 2007: Política Pública de Discapacidad para el Distrito Capital
- Decreto 380 de 2015: Política Pública de Trabajo Decente y Digno de Bogotá
- Decreto 193 de 2022: Política Pública de Lucha Contra la Trata de Personas en Bogotá Distrito Capital 2022-2031
- CONPES D.C. 11 de 2019 y su respectiva Adenda 1, de mayo del 2022: Política Pública de Actividades Sexuales Pagadas 2020 – 2029
- CONPES D.C. 14 de 2021: Política Pública de Mujeres y Equidad de Género 2020-2030
- CONPES D.C. 16 de 2022: Política Pública LGBTI, identidades de género y orientaciones sexuales en el D.C. 2021-2032
- CONPES D.C. 19 de 2022: Actualización del Plan de Acción de la Política Pública Social para el Envejecimiento y la Vejez 2010-2025
- CONPES D.C. 20 de 2022: Actualización del Plan de Acción de la Política Pública Distrital para el Fenómeno de Habitabilidad en Calle 2015-2025
- CONPES D.C. 25 de 2022 Actualización del Plan de Acción de la Política Pública de Productividad, Competitividad y Desarrollo Socioeconómico 2011-2038

Sentencias/Autos de la Corte Constitucional

- Sentencias T-724 de 2003 y T-291 de 2009. Autos 268 de 2010; 183 de 2011; 189 de 2011; 275 de 2011; T-783 de 2012; 366 de 2014; 118 de 2014 y 587 de 2015: Pronunciamientos de la Corte Constitucional en relación con la actividad de los Recicladores de Oficio en la prestación del servicio público. Les reconoce como sujetos de especial protección respecto de los cuales se deben promover acciones afirmativas.
- Sentencias C-410 de 1994, C-251 de 1996, C-507 de 2008, C-234 de 2009, C-027 de 2016, C-115 de 2017: Criterios redistributivos que contribuyen al bienestar general, impactando de manera directa en la reducción de los índices de desempleo y propenden por la sostenibilidad empresarial en todas sus categorías sin favorecer a aquellas que tienen mayor capacidad para acceder a los beneficios, buscando un equilibrio en el bienestar general como máxima de la Constitución Política de 1991.

2. Evaluación de resultados al programa

Este estudio tiene como objetivo realizar una evaluación de resultados del Programa Empleo Incluyente implementado por la Secretaría Distrital de Desarrollo Económico. Para ello, se tendrá en cuenta el periodo desde el inicio del Programa en enero de 2023, hasta la actualidad. Dicha evaluación se centra sobre los siguientes objetivos específicos:

1. Analizar los resultados del programa en términos de las condiciones de empleabilidad, permanencia y los niveles de ingreso de las personas vinculadas a las empresas.
2. Determinar de qué manera cambia la probabilidad de enganche de las personas vinculadas a las empresas y sus niveles de ingreso después de participar en el programa.
3. Analizar los resultados del programa en cuanto a la calidad de vida de las personas vinculadas a las empresas.

4. Identificar los factores que influyen en que las empresas decidan participar o no del programa.
5. Analizar los cambios en la apropiación de prácticas alrededor de la inclusión, diversidad y equidad al interior de las empresas.
6. Formular recomendaciones para la focalización del programa con el fin de mejorar o mantener los resultados en términos de las poblaciones que enfrentan barreras de ingreso al mercado laboral.

Con base en los objetivos descritos, la evaluación de resultados del Programa de Empleo Incluyente (PEI) es pertinente para comprender los efectos de esta intervención en el mercado laboral formal y en la vida de las personas y empresas beneficiarias. Al centrarse en aspectos clave como la empleabilidad, la permanencia en el empleo, los niveles de ingreso y la calidad de vida de los participantes, este estudio permitirá identificar si el programa está cumpliendo con sus objetivos estratégicos, y proporcionará insumos valiosos para la mejora continua del programa, con recomendaciones enfocadas en la focalización adecuada para beneficiar a las poblaciones que enfrentan mayores barreras para acceder al mercado laboral formal, garantizando así una mayor efectividad y sostenibilidad de la política pública.

3. La evaluación al PEI en el contexto de evaluaciones a programas similares

La presente sección hace una breve revisión de la literatura sobre evaluaciones ya existentes que se han hecho a programas similares al PEI. A nivel internacional, se relacionan algunos estudios que analizan los resultados generales de evaluaciones de subsidios al salario compilando diversos trabajos. Igualmente se presenta algunas experiencias para América Latina. Para Colombia se presentan las evaluaciones que han existido para algunos programas de apoyo al empleo.

Los hallazgos a nivel internacional

La mayoría de resultados de las evaluaciones sobre políticas activas de mercado laboral se puede encontrar en recientes estudios tipo meta análisis. Estos son estudios que utilizan técnicas estadísticas para analizar resultados de un buen número de estudios sobre un tema en particular.

Card et. al. 2015 hacen su análisis sobre 207 evaluaciones de impacto sobre políticas activas del mercado laboral en todo el mundo, incluidos los programas de subsidios al empleo. Encuentran que, junto con los programas de formación o entrenamiento, los programas de subsidios al empleo tienen efectos positivos en el mediano y largo plazo. Igualmente, estos programas tienen un mayor efecto en desempleados de largo plazo.

Otro estudio a nivel mundial reciente es el de ILO and World Bank 2022. Toma como base resultados de evaluaciones de 220 programas de políticas activas del mercado laboral enfocados en personas entre 15 y 35 años, evaluaciones producidas entre 1990 y 2022. De ese total de evaluaciones, 29 corresponden a los subsidios al salario. Encuentra que los subsidios

al salario y la formación tienen más grandes efectos en países de ingreso alto. Igualmente, los subsidios al salario junto con los programas de trabajos públicos son los más costosos para los gobiernos.

Por su parte, OIT 2016 realiza un análisis para América Latina y el Caribe, basado en 44 estudios, aunque solo cuatro fueron relacionados con los subsidios al empleo. De estos cuatro, tres estudios observaron un impacto positivo sobre la empleabilidad de los participantes y no arrojaron diferencias entre hombre y mujeres. Los estudios disponibles demuestran que las subvenciones salariales pueden incrementar las probabilidades de empleo, en particular cuando se asignan directamente a los individuos como un complemento a sus ingresos. Un estudio similar (Escudero y otros 2017) encuentra los mismos resultados dado que el conjunto de evaluaciones sobre subsidios al empleo es el mismo.

De estos meta análisis, se encuentra que la mayor parte de las evaluaciones se han hecho para programas de formación y relativamente son pocas las dirigidas a esquemas de subsidios al empleos como el del PEI. Esto responde, en gran parte, a que en su mayoría las intervenciones son de formación. Por otra parte, se observa que las poblaciones objetivo de estos programas se han circunscrito a unas pocas principalmente a jóvenes, mujeres y personas mayores, lo cual contrasta con las 17 poblaciones objetivo del PEI. A continuación se describen algunas evaluaciones encontradas para América Latina[3].

En Argentina se hizo un experimento con el programa de empleos públicos Trabajar del cual se tomó una muestra para ofrecer subsidios al salario y capacitación. El experimento se denominó ProEmpleo y se llevó acabo entre 1998 y 2000. Se dirigió a personas desempleadas provenientes de familias pobres. El componente de subsidios al salario otorgaba un subsidio de 150 dólares por mes a trabajadores mayores de 45 años y de 100 dólares para trabajadores jóvenes, siendo el salario mínimo de 200 dólares en esa época. La evaluación realizada (Salvia y otros) encontró impacto positivo en cuanto al empleo pero no significante en los ingresos. Los efectos positivos principalmente se encontraron para mujeres y jóvenes.

En Chile existe el programa Subsidio al Empleo Joven -SEJ - el cual fue creado en 2009. Su población objetivo es los jóvenes entre 18 y 24 años de edad del 40% más pobre de la población para los cuales se pretende incrementar sus tasa de empleo e informalidad al igual que sus ingresos. Entrega un subsidio de hasta 30% del salario percibido por el joven, 20% se entrega al trabajador y 10% a la empresa, y se va reduciendo a medida que aumenta el salario del joven (Dema y otros 2015). La evaluación de 2012 (Centro de Microdatos 2012), usando el método de regresión discontinua encuentra que el SEJ generó incrementos en las tasas de ocupación y de participación, tanto para mujeres como para hombres. Sin embargo, los impactos decrecen entre julio de 2009 y diciembre de 2010, en especial para los jóvenes entre 18 y 22 años de edad. Por su parte, la otra evaluación realizada al programa (Sence 2022) encuentra un aumento cercano al 3,2% en la probabilidad de cotizar formalmente de 11% en los ingresos laborales. Esta evaluación utilizó el método diferencias en diferencias con datos que abarcaron el periodo 2010 – 2021. formales y dependientes

Evaluaciones de impacto sobre subsidios al salario en Colombia

[3] Por otra parte, muchos esquema de subsidios al salario surgieron para contrarrestar los efectos del confinamiento del COVID19 pero no se tiene evidencia de sus evaluaciones.

La implementación de programas de subsidio al salario en Colombia ha sido relativamente reciente. Su inicio se puede establecer en 1998 cuando se aprobó un descuento tributario por un monto máximo del quince por ciento (15%) del impuesto neto de renta del respectivo período, según los nuevos empleos generados. Luego en 2003, el gobierno estableció la exención en el pago de aportes parafiscales al SENA, ICBF y cajas de compensación, a empleadores que generaran nuevos puestos de trabajo. Los dos programas no tuvieron muchos efectos (Bohórquez y Merchán 2005) y su diseño operativo difería un poco al del PEI, dado que se trataba de descuentos y exenciones que implicaban ciertos trámites operativos adicionales y el beneficio para el empleador podía tardar varios meses después de generado el puesto de trabajo[4].

El primer programa implementado a nivel nacional con características similares al PEI fue el Programa de Apoyo Directo al Empleo, PADE. El programa inició en mayo de 2004, otorgaba a las empresas un 30% del salario mínimo durante seis meses por cada empleo generado a personas de los niveles I y II del SISBEN y se dirigía a las zonas urbanas. Su evaluación, basada en encuestas a empresas, mostró que el diseño del programa en cuanto al monto del subsidio fue adecuado, la duración de seis meses del subsidio fue muy corta y hubo problemas en la implementación (Olivera 2006).

El segundo programa de subsidio directo al salario con una evaluación sobre sus resultados fue 40.000 Primeros Empleos. Estuvo dirigido a jóvenes entre 18 y 28 años sin empleo ni experiencia formal relacionada con su área de formación y operó entre 2015 y 2018. Ofrecía un subsidio atractivo para los empleadores. Cubría el salario, las prestaciones sociales y la seguridad social hasta por seis meses con el compromiso de mantener al joven durante 12 meses. Establecía una diferenciación del monto de subsidio al mes según el nivel de formación que ibas desde $978.567 hasta $1.332.272 (Ministerio del Trabajo 2015). La evaluación encontró que el programa generó ingresos y empleo en actividades formales con una focalización acertada en jóvenes con menores posibilidades de engancharse laboralmente (menor nivel de estudios, menor capacitación y experiencia laboral y con menos conocimientos de lenguas extranjeras). Y algunos jóvenes lograron mejorar sus condiciones laborales en su mismo lugar de trabajo y lograr incremento de salario, ascensos y/o puestos más afines con sus intereses.

Un tercer programa con su respectiva evaluación fue el Programa de Apoyo al Empleo Formal, PAEF. Se creó en mayo de 2020 para proteger a trabajadores y empresas formales del país que afectadas por las restricciones de movilidad y operación para contrarresta el COVID19. Otorgaba un subsidio a las empresas de 40% del salario mínimo por cada trabajador empleado, monto que luego aumentó a 50% si los empleados eran mujeres y de los sectores turismo, hostelería, gastronomía y actividades artísticas. El programa parece haber ayudado a mitigar un poco los efectos adversos causados por la pandemia, como lo son la pérdida de empleos y la quiebra de empresas (DNP 2022).

[4] Algo similar pudo ocurrir con las exenciones que se otorgaron a partir de 2011 a raíz de los beneficios que se dieron con la Ley 1429 de 2010.

Del PAEF se derivó el Programa Nuevo Empleo, el cual funcionó entre 2021 y 2023. Sus principales diferencias respecto al PAEF fueron el requisito de generar nuevos empleos, la inclusión de jóvenes como un grupo prioritario y la disminución del monto del subsidio. Sus grupos objetivo fueron jóvenes entre los 18 y 28 años, mujeres mayores de 28 años y hombres mayores de 28 años, por quienes los empleadores recibían un subsidio equivalente a 25%, 15% o 10% de un salario mínimo mensual, respectivamente, por cada empleo. La evaluación del Programa indicó que los jóvenes y las personas con remuneraciones entre 1 y 2 salarios mínimos fueron los más beneficiados. Sin embargo, encontró que el Programa no logró cambiar la tendencia preexistente en el comportamiento de la contratación de nuevos empleados.

Una segunda evolución del Programa Nuevo Empleo tuvo resultados adicionales (Guido 2024) utilizando, entre otras herramientas, la metodología diferencias en diferencias. Entre ellos, Encontró un aumento de 46% en empleo de jóvenes, 58% para hombres y 32% para mujeres. También halló efectos positivos para micro y pequeñas empresas y en empleos entre 1 y 4 salarios mínimos. Sin embargo, 47% de empleos subsidiados no requerían subsidio, siendo una evidencia de peso muerto.

Como continuación a los dos últimos programas mencionados, en 2024 el Ministerio del Trabajo transformó el Programa Nuevo Empleo identificándolo como Empleos para la Vida. Los principales cambios fueron el aumento del subsidio, la inclusión de personas con discapacidad y un diferencial de subsidio por duración del desempleo. Los montos de subsidio respecto al salario mínimo son 30% por jóvenes entre 18 y 28 años, 35% por personas con discapacidad, 20% por mujeres mayores de 28 años, 15% por hombres mayores de 28 años y un 10% adicional si cualquiera de ellos ha estado desempleado durante los últimos 4 meses. Dado su reciente inicio de ejecución, el programa aún no tiene evaluación, pero su desempeño podría servir de referencia para cualquier ajuste del PEI como complemento a esta evaluación.

Cuadro 1 Evaluaciones sobre subsidios al salario en América Latina y Colombia

País	Nombre del programa	Fecha implementación del programa	Estudio o documento de la evaluación	Fecha de la evaluación	Población objetivo	Método de evaluación	Principales resultados
Argentina	Experimento ProEmpleo	1998 - 2000	Salvia et al, 2004.	2004	Trabajadores de bajos salarios provenientes de familias pobres.	Experimento aleatorio utilizando dobles diferencias y variables instrumentales	Impacto positivo en el empleo pero no significante en los ingresos. Efectos positivos principalmente para mujeres y jóvenes.
Chile	Subsidio Empleo Joven	2009 - 2024	Centro de Microdatos. 2012.	2009 - 2010	Jóvenes de 18 a 25 años de edad en el 40% de la población más vulnerable.	Regressión discontinua	Incrementos en tasas de ocupación y de participación mujeres y hombres, con tendencia decreciente entre julio de 2009 y diciembre de 2010, en especial para los jóvenes entre 18 y 22 años de edad.
Chile	Subsidio Empleo Joven	2009 - 2024	Sence 2022	2010 - 2021	Jóvenes de 18 a 25 años de edad en el 40% de la población más vulnerable.	Diferencias en diferencias	Aumento cercano a 3,2% en la probabilidad de cotizar formalmente y de 11% en ingresos laborales formales.
Colombia	Programa de Apoyo Directo al Empleo - PADE	2004 - 2005	Olivera 2006	2006, octubre	Personas desempleadas en niveles I y II del SISBEN, en especial jefes de hogar	Encuestas a empresas beneficiarias del subsidio	El diseño del programa en cuanto al monto del subsidio fue adecuado pero hubo problemas en la implementación.
Colombia	40.000 Primeros Empleos	2015 - 2018	DNP 2020.	2019	Jóvenes de 18 a 28 años, sin empleo ni experiencia formal relacionada con su área de formación.	Metodología cuasiexperimental con emparejamiento exacto.	Generó ingresos y empleo en actividades formales con una focalización acertada en jóvenes con menores posibilidades de engancharse laboralmente.
Colombia	Programa de Apoyo al Empleo Formal, PAEF	2020 - 2021	DNP 2022	2019 - 2021	Cualquier persona al inicio. Luego hubo mayores subsidios por mujeres y para empleadores en turismo, hostelería, gastronomía y actividades artísticas.	Diferencias en diferencias	Mitigó las pérdidas de empleo formal pero no contribuyó a aumentar el empleo.
Colombia	Nuevo Empleo	2021 - 2023	DNP 2024	2023	Jóvenes entre 18 y 28 años, mujeres mayores de 28 años y hombres mayores de 28 años	Análisis descriptivo.	No hubo cambio en la tendencia preexistente de contratación de trabajadores nuevos. Jóvenes y personas entre 1 y 2 salarios mínimos fueron los más beneficiados.
Colombia	Nuevo Empleo	2021 - 2023	Guio Daniela 2024	2024	Jóvenes entre 18 y 28 años, mujeres mayores de 28 años y hombres mayores de 28 años	Diferencias en diferencias	Aumento de 46% en empleo de jóvenes, 58% para hombres y 32% para mujeres. Efectos positivos para micro y pequeñas empresas y en empleos entre 1 y 4 salarios mínimos. 47% de empleos subsidiados no requerían subsidio (peso muerto).

Fuente: Recopilación con base en estudios citados en este capítulo

A continuación, este informe describe el desarrollo metodológico necesario para llevar a cabo la para la evaluación de resultados al programa de Empleo Incluyente implementado por la Secretaría Distrital de Desarrollo Económico. En el capítulo II se describen de manera detallada las metodologías tanto cuantiativas como cualitativas que se emplearán, su descripción, la justificación de su pertinencia para esta evaluación y el PEI y los pasos necesarios para su implementación. Más adelante, en el capítulo III se describe todo lo relacionado con el trabajo de campo que se adelantará, puntualmente se incluye la propuesta de diseño muestral cuantitativo y cualitativo y finalmente se incluye en el capítulo IV una propuesta de triangulación de información y metodologías para la estructuración de hallazgos y formulación de recomendaciones para el Programa.

II. Desarrollo metodológico de la evaluación

4. Metodologías cuantitativas

4.1 Descripción de las fuentes primarias y/o secundarias

El siguiente cuadro resume las variables, sus características, la fuente primaria/secundaria, la frecuencia y el horizonte temporal. Para comparaciones en las variables de empleo usaremos la GEIH.

Variable	Característica	Fuente de Información	Frecuencia	Horizonte Temporal
Tiempo en empleo	Dependiente	Encuesta de evaluación	Única observación	Desde el inicio del desemploe hasta el final del período de estudio
Tiempo de desempleo	Dependiente	Encuesta de evaluación	Única observación	Desde el inicio del desempleo hasta el final del período de estudio
Ingreso	Dependiente	Encuesta de evaluación	Única observación	
Experiencia laboral previa	Covariable	Encuesta de evaluación	Única observación	N/A
Sexo, Edad, Nivel Educativo	Covariable	Encuesta de evaluación	Única observación	N/A
Pertenencia a grupos vulnerables	Covariable	Encuesta de evaluación	Única observación	N/A
Inclusión laboral	Latente	Encuesta de evaluación	Única observación	N/A
Discriminación laboral	Latente	Encuesta de evaluación	Única observación	N/A
Empleabilidad	Latente	Encuesta de evaluación	Única observación	N/A
Calidad del empleo	Latente	Encuesta de evaluación	Única observación	N/A
Calidad de vida	Latente	Encuesta de evaluación	Única observación	N/A
Empoderamiento económico	Latente	Encuesta de evaluación	Única observación	N/A
Adaptación laboral	Latente	Encuesta de evaluación	Única observación	N/A
Enfoque interseccional	Covariable	Encuesta de evaluación	Única observación	N/A
Características sociodemográficas	Latente	Encuesta de evaluación	Única observación	N/A
Variables organizacionales (percepción del encuestado)	Latente	Encuesta de evaluación	Única observación	N/A
Duración del programa	Dependiente	Encuesta de evaluación	Única observación	Desde la inscripción hasta el final del estudio
Tipo de beneficios recibidos	Covariable	Encuesta de evaluación	Única observación	Desde la inscripción hasta el final del estudio
Tasa de reenganche	Dependiente	PILA	Mensual	Últimos 5 años antes del programa y hasta el último mes disponible en 2024
Fecha de inicio del empleo	Para construir Dependiente	PILA	Mensual	Últimos 5 años antes del programa y hasta el último mes disponible en 2024
Fecha de fin del empleo	Para construir Dependiente	PILA	Mensual	Desde la inscripción hasta la desvinculación o censura
Salario (IBC)	Dependiente	PILA	Mensual	Últimos 5 años antes del programa y hasta el último mes disponible en 2024
Estado de cotización en seguridad social	Dependiente	PILA	Mensual	Últimos 5 años antes del programa y hasta el último mes disponible en 2024
Duración del empleo en términos de meses cotizados	Dependiente	PILA	Mensual	Un año antes de la inscripción. Desde la inscripción hasta el final del estudio
Permanencia en el empleo	Dependiente	PILA	Mensual	Desde la inscripción hasta la desvinculación o censura

4.2 Descripción y justificación de las metodologías.

4.2.1 Descripción y justificación de las estadísticas descriptivas

4.2.1.1 Descripción

Utilizaremos estadísticas descriptivas como una metodología fundamental para comprender las características y composiciones de las personas beneficiarias del Programa de Empleo

Incluyente (PEI), así como para analizar las dinámicas de empleo y desempleo en el contexto del programa. A continuación, detallamos los aspectos clave de esta metodología:

i) Principales usos:

- **Resumen de datos:** Facilitar la comprensión de grandes conjuntos de datos mediante la síntesis de información relevante.
- **Identificación de patrones y tendencias:** Detectar comportamientos comunes y tendencias generales en las variables estudiadas.
- **Detección de anomalías:** Identificar valores atípicos o inconsistencias en los datos que puedan influir en los análisis posteriores.
- **Base para análisis avanzados:** Proporcionar una fundamentación sólida para la aplicación de metodologías más complejas, como modelos estadísticos inferenciales.

ii) Requerimientos de información:

- **Datos cuantitativos:** Información numérica sobre variables como salario, edad, duración del empleo/desempleo, nivel educativo, entre otras variables señaladas en el Cuadro 3 al inicio de este capítulo.
- **Bases de datos completas y confiables:** Registros administrativos del programa (marco de lista), encuesta a personas beneficiarias y sus empresas, y datos de fuentes oficiales como la PILA.
- **Variables categóricas y continuas:** Datos estructurados que permitan clasificar y medir las diferentes características de los participantes y empresas (Cuadro 3).

iii) Limitaciones:

- **No permite inferencias causales:** Las estadísticas descriptivas no pueden establecer relaciones de causa y efecto.
- **Dependencia de la calidad de los datos:** Resultados pueden estar sesgados si existen errores o datos faltantes significativos.

iv) Descripción del análisis asociado:

- **Análisis cuantitativo:** Aplicaremos técnicas estadísticas para cuantificar y resumir los datos numéricos y categóricos.
- **Visualización de datos:** Utilizaremos herramientas gráficas para representar la información de manera comprensible.
- **Análisis exploratorio:** Investigaremos las características de los datos sin formular hipótesis previas, identificando patrones y relaciones potenciales.

v) Tipo de resultados que genera:

- **Medidas de tendencia central y dispersión:** Valores como media, mediana, moda, desviación estándar.
- **Distribuciones de frecuencias y proporciones:** Porcentajes y conteos de categorías en variables binarias y categóricas.
- **Gráficos y visualizaciones:** Histogramas, boxplots, diagramas de barras, matrices de correlación.
- **Identificación de relaciones preliminares:** Correlaciones y asociaciones entre variables.

vi) Características de las conclusiones que se pueden obtener:

- **Descriptivas y generales:** Conclusiones sobre cómo se distribuyen y comportan las variables en la muestra analizada.
- **Identificación de perfiles:** Caracterización de grupos específicos dentro de las personas beneficiarias y sus empresas.
- **Orientación para análisis posteriores:** Señalamiento de áreas que requieren análisis más profundos o metodologías adicionales.

4.2.1.2 Justificación

Las estadísticas descriptivas se ajustan directamente a los objetivos específicos de la evaluación y las preguntas orientadoras planteadas. Al proporcionar un resumen detallado de las características de las personas beneficiaras y empresas asociadas al empleado, esta metodología nos permitirá:

- **Objetivo 1:** Analizar las condiciones de empleabilidad, permanencia y niveles de ingreso de las personas beneficiarias.
    - **Aplicación:** Utilizaremos medidas de tendencia central y dispersión para evaluar salarios, duración del empleo/desempleo y tasa de cotización anual.
- **Objetivo 4:** Identificar los factores que influyen en que las empresas decidan participar o no del programa.
    - **Aplicación:** A través de tablas de frecuencias y proporciones, caracterizaremos las empresas según tamaño, sector y ubicación.
- **Objetivo 6:** Formular recomendaciones para la focalización del programa.
    - **Aplicación:** Al identificar patrones y tendencias, podremos señalar qué grupos poblacionales requieren mayor atención.

Las preguntas orientadoras, como "¿Cuáles son las características de las personas vinculadas en cuanto a su formación académica, perfil laboral y características socioeconómicas?" serán respondidas mediante el análisis descriptivo de variables como nivel educativo, experiencia laboral previa y pertenencia a grupos de difícil colocación y todas las variables del Cuadro 3.
El Anexo 5 se presenta la implementación y el abordaje de desafíos metodológicos de la estadísticas descriptivas (secciones 3.1.3 y 3.1.4 respectivamente).

4.2.2 Descripción y justificación del Análisis de supervivencia

4.2.2.1 Descripción

i) Usos principales:

- Analizar el tiempo transcurrido desde un punto de inicio (por ejemplo, la contratación) hasta que ocurre un evento de interés (por ejemplo, desvinculación laboral). Para la información de la encuesta normalmente se pregunta por el tiempo en el empleo y el desempleo. Para la información de la PILA se solicitará información cinco años antes del inicio del programa hasta la fecha disponible (posiblemente hasta tercer o cuarto trimestre de 2024).
- Evaluar los factores asociados de variables explicativas (edad, sexo, experiencia previa, sector económico, etc.) en la probabilidad de que ocurra el evento en un momento determinado.
- Comparar curvas de supervivencia entre diferentes grupos poblacionales.

ii) Requerimientos de información:

- Datos sobre la fecha en la que inició el contrato por el que la persona beneficiaria fue postulada y aceptada en el Programa.
- Tiempo en desempleo anterior a su vinculación (esto se levantará a través de la encuesta).
- Datos longitudinales que incluyan fechas de inicio y fin de cada empleo (y desempleo) para las personas beneficiarias. Esto a partir de la **PILA**.
- Información sobre variables censuradas (personas beneficiarias que aún permanecen empleados/desempleados al final del periodo de observación). Esto a partir de la

encuesta de la evaluación y de la **GEIH** para comparar los datos generales de Bogotá con los de la muestra de personas beneficiarias.

- Datos sobre variables explicativas relevantes (demográficas, socioeconómicas, laborales). Esto para las 3 fuentes de información.

iii) Limitaciones:

- Estos modelos suponen que las censuras[5] son no informativas y que los eventos son independientes entre individuos:
    - La **censura** (cuando no se conoce el tiempo exacto en que ocurre el evento) es **no informativa**, es decir, la razón por la cual se censura a un individuo no está relacionada con el evento en estudio.
    - Los tiempos hasta el evento son **independientes entre los individuos**. Esto significa que lo que le ocurre a un individuo no debe influir en el riesgo o tiempo de ocurrencia del evento para otro individuo.
- Es necesario validar el supuesto de riesgos proporcionales en los modelos de supervivencia.
    - En el modelo de **Cox**, se asume que las **razones de riesgos** (la tasa a la que ocurre el evento para un individuo en relación con otro) son **proporcionales** a lo largo del tiempo. Esto significa que los efectos de las covariables multiplican la tasa de riesgo de manera constante en el tiempo.
- En modelos paramétricos como el **Weibull** o el **Exponencial**, se asume que el **tiempo hasta el evento sigue una distribución específica**.
    - Por ejemplo, en el modelo Weibull, se asume que el riesgo cambia de manera predecible con el tiempo (puede aumentar o disminuir). Por ejemplo, en la duración del desempleo el riesgo aumenta con el tiempo: a lo largo de los meses cada vez es más difícil encontrar empleo.
- Se asume que la **censura es independiente** de las covariables incluidas en el modelo. Esto implica que las variables explicativas no deberían influir en la probabilidad de que un individuo sea censurado.
- Los datos de la PILA deben manejarse como **datos panel** o **historias de eventos**, dado que se observarán periodos de cotización y no cotización que generan trayectorias de duración de empleo y desempleo diferentes para cada individuo.

iv) Análisis asociado:

- **Análisis cuantitativo** mediante:
    - Estimación de funciones de supervivencia utilizando el método de Kaplan-Meier.
    - Modelos de riesgos proporcionales de Cox para evaluar el efecto de variables explicativas.
    - Modelos paramétricos (Weibull, Exponencial, etc.) si se cumplen los supuestos necesarios.

[5] La censura de los datos de empleo y desempleo se refiere a que cuando la persona contesta la encuesta no sabe si el mes siguiente pierde el empleo o consigue un empleo respectivamente. Eso también sucede con los datos de la PILA. Por ejemplo, si los datos conseguidos a través del Ministerio de Salud llegan hasta septiembre de 2024, no sabemos si en octubre de 2024 un empleado pierde el empleo o se observa a alguien que consiguió empleo en octubre (pero se observaba en septiembre como no cotizante).

v) Tipo de resultados que genera:

- Curvas de supervivencia que muestran la probabilidad de continuar empleado o desempleado a lo largo del tiempo.
- Estimaciones de riesgos relativos o hazard ratios asociados a variables explicativas.
- Identificación de factores que aumentan o disminuyen la probabilidad de permanencia en el empleo.

vi) Características de las conclusiones:

- Permite concluir cuáles variables tienen un efecto significativo en la duración del empleo y el desempleo.
- Ayuda a entender patrones de permanencia laboral entre diferentes grupos a través de los meses cotizados en un periodo de la PILA (por ejemplo, en el último año, o desde que inició el programa). Estos se compararán con los patrones antes de la inscripción al programa para mostrar el resultado del programa.
- Proporciona información para diseñar estrategias que mejoren la retención laboral de las personas beneficiarias según sus características, sector económico, tamaño de la empresa.

4.2.2.2. Justificación

Los **modelos de supervivencia** se ajustan de manera óptima a los objetivos y preguntas orientadoras de la evaluación del PEI, ya que permiten analizar en profundidad la **permanencia en el empleo, el desempleo** y los **factores que la afectan**. A continuación, se detalla cómo esta metodología responde a cada una de las preguntas orientadoras pertinentes:

- **Pregunta 1a:**

"Analizar los resultados del programa en términos de las condiciones de empleabilidad, permanencia y los niveles de ingreso de las personas vinculadas a las empresas."

- **Uso del modelo:**
  - Los modelos de supervivencia permiten analizar la **duración en el empleo y el desempleo** de las personas beneficiarias, proporcionando una medida directa de la **permanencia laboral**.
  - Al incluir variables como nivel de ingreso, condiciones de empleo, características socioeconómicas y de la firma se puede evaluar cómo estas influyen en la supervivencia laboral.
- **Tipo de resultados generados:**
  - Estimaciones de la **probabilidad de continuar empleado/desempleado** a lo largo del tiempo para diferentes grupos.
  - Identificación de factores que afectan la **empleabilidad**, **permanencia**, desempleo tales como experiencia laboral previa, nivel educativo y sector económico.
- **Contribución a los objetivos de la evaluación:**
  - Al comparar el antes y después de los beneficiarios es posible obtener evidencia cuantitativa sobre la **efectividad del programa** en mejorar la permanencia laboral y disminuir la duración del desempleo.
  - Utilizando las covariables, es posible identificar áreas de mejora y fortalezas del PEI en términos de empleabilidad.
- **Pregunta 1b:**

"Determinar de qué manera cambia la probabilidad de enganche de las personas vinculadas a las empresas y sus niveles de ingreso después de participar en el programa."

- **Uso del modelo:**

- Al analizar el tiempo hasta la desvinculación laboral en la PILA, se puede inferir cómo el programa influye en la **probabilidad de enganche** y **retención** en el empleo. La probabilidad de enganche se puede medir fácilmente en la PILA teniendo observaciones de cuando deja de cotizar y cuando lo vuelve a hacer (reenganche). Esto es posible gracias a que obtendremos datos de 5 años antes del PEI para poder evaluar el antes y el después de los beneficiarios.
- Comparando curvas de supervivencia (en el empleo y el desempleo) antes y después de la participación en el programa se evalúa el **resultado del PEI**.

- **Tipo de resultados generados:**
  - **Hazard ratios** que muestran el efecto del programa en la probabilidad de continuar empleado.
  - Estimaciones que reflejen cambios en **niveles de ingreso** asociados con la permanencia laboral. Esto comparando curvas de superviviencia entre quienes tuvieron cambios negativos, no cambios, cambios positivos y variables binarios con que reproduzcan estos cambios para incluirlas como covariables en los modelos paramétrico (Cox) y no paramétricos, según las pruebas de proporcionalidad que se describen más adelante.
- **Contribución a los objetivos de la evaluación:**
  - Permite cuantificar el **cambio en la probabilidad de enganche** (antes y después) y vincularlo directamente con la participación en el programa.
  - Ofrece información para ajustar estrategias que aumenten la retención y mejoren los niveles de ingreso por tipos de personas beneficiarias.
- **Pregunta 2:**

"¿De qué forma cambia la probabilidad de enganche de las personas beneficiarias y sus niveles de ingreso después de participar en el programa?"

El cambio en la probabilidad de enganche requiere el uso de la PILA. Requiere medir el Hazard rate antes y después de la intervención.

- **Uso del modelo:** A través de la medición del antes y el después se medirá el cambio en la probabilidad de reenganche. Este ejercicio es posible de hacer para aquellos que perdieron el empleo, pues de otra forma no hay reenganche.
- **Tipo de resultados generados:** Estimaciones que relacionan el cambio en la probabilidad de reenganche después de la intervención.
- **Contribución a los objetivos:** Responde directamente al objetivo 2.
- **Requiere armar un panel de datos como se explicará en la sección 3.2.3.4.**
- **Pregunta 4 de empresas:**

"Analizar los cambios en la apropiación de prácticas alrededor de la inclusión, diversidad y equidad al interior de las empresas."

- **Uso del modelo:**
  - Al incluir variables de la encuesta de la evaluación relacionadas con **prácticas empresariales** de inclusión y equidad, se puede evaluar cómo estas afectan la **duración del empleo** de las personas beneficiarias. La conformación de estas variables se hará a través de modelos de Análisis Factorial Confirmatorio que se explica en la sección 3.3.
  - Permite comparar la permanencia laboral en empresas con diferentes niveles de **apropiación de prácticas inclusivas**.
  - La inclusión de variables como las prácticas empresariales, obtenidas a partir de las respuestas de los beneficiarios, como covariables en un modelo de supervivencia, se basa en la premisa de que estas variables reflejan

características relevantes del entorno laboral que pueden influir en la duración del empleo. Aunque la información proviene directamente de los beneficiarios, estas variables no se consideran exclusivamente atributos individuales, sino que sirven como aproximaciones (o "proxies") de características organizacionales.
    - Por ejemplo, cuando un beneficiario reporta prácticas inclusivas o políticas de apoyo en la empresa que lo contrató, esto proporciona información indirecta sobre las condiciones y dinámicas organizacionales de esa empresa. Al incluir esta variable en el modelo de supervivencia, se evalúa cómo estas prácticas organizacionales, percibidas por los beneficiarios, afectan la probabilidad de que un contrato de trabajo se mantenga durante un cierto período.
    - El uso de esta estrategia es similar al caso del tamaño de la empresa, donde la información reportada por el beneficiario (como el número de empleados o el sector económico) también se utiliza como una covariable que captura un aspecto organizacional que puede influir en la estabilidad laboral. Esto es lo que se hace tradicionalmente con el uso de variables de la empresa de la GEIH que responden los trabajadores encuestados.
    - Aunque la fuente de datos sea el beneficiario, el objetivo es modelar el efecto de las características organizacionales, en este caso prácticas empresariales, sobre la duración del empleo, entendiendo que estas variables se refieren a las condiciones laborales y organizativas en las que se desenvuelven los beneficiarios. De nuevo, no se puede generalizar a las empresas (ni es el objetivo), sino analizar si las políticas de la empresa del trabajador tienen un efecto sobre las variables dependientes.
- **Tipo de resultados generados:**
    - Comparaciones de curvas de supervivencia entre empresas con y sin prácticas inclusivas.
    - Análisis del efecto de las políticas de inclusión en la **retención de talento, duración del empleo y estabilidad laboral**.
- **Contribución a los objetivos de la evaluación:**
    - Proporciona evidencia sobre la **efectividad de las prácticas inclusivas** en mejorar la permanencia laboral.
    - Informa sobre el **impacto de las intervenciones del programa** en la cultura organizacional de las empresas.

El Anexo 5 se presenta la implementación y el abordaje de desafíos metodológicos de la estadísticas descriptivas (secciones 3.2.3 y 3.2.4 respectivamente).

4.2.3 Descripción y justificación de los modelos SEM

4.2.3.1 Descripción

i) **Usos principales:**

Los **Modelos de Ecuaciones Estructurales (SEM)** son una metodología estadística que permite analizar relaciones complejas entre variables observables y latentes. En el contexto de la evaluación del Programa de Empleo Incluyente (PEI), los SEM se utilizarán para:

- **Modelar variables latentes** como inclusión laboral, discriminación laboral, empleabilidad, calidad del empleo, capital humano, empoderamiento económico y adaptación laboral, que no son directamente observables, pero se infieren a partir de indicadores medibles.

- **Analizar relaciones estructurales** entre estas variables latentes y otras variables observables, permitiendo comprender cómo diferentes factores influyen en los resultados del programa.
- **Explorar efectos de mediación y moderación**, identificando mecanismos a través de los cuales el PEI impacta a las personas beneficiarias y cómo variables como el enfoque interseccional pueden moderar estas relaciones.

ii) **Requerimientos de información:**

- **Datos primarios** recolectados a través de encuestas aplicadas a las personas beneficiarias, que incluyan indicadores observables para las variables latentes.
- **Variables observables** relacionadas con prácticas de inclusión, percepción de discriminación, capacidades y competencias adquiridas, condiciones laborales, entre otros.
- **Información demográfica y socioeconómica** de las personas beneficiarias: edad, sexo, nivel educativo, experiencia laboral previa, pertenencia a grupos vulnerables, etc.
- **Variables de resultado:** empleo/desempleo, calidad de vida, empoderamiento económico, adaptación laboral, calidad del empleo.
- **Otras variables descritas en la sección 2.1.**

iii) **Limitaciones:**

- **Supone normalidad multivariante** de los datos, lo cual puede no cumplirse en todos los casos.
- Se asume que no hay **multicolinealidad** severa entre las **variables independientes** o **indicadores**.
- Se asume que los **errores de medición** asociados con los indicadores de las variables latentes son **independientes** entre sí, es decir, los errores no deben correlacionarse.
- **Requiere muestras de tamaño adecuado** para obtener estimaciones estables y confiables.
- **Es sensible a la especificación del modelo**, por lo que un modelo mal especificado puede conducir a resultados erróneos.
- **Dificultad para capturar relaciones no lineales complejas**, que podrían estar presentes en los datos.

iv) **Análisis asociado:**

- **Análisis cuantitativo** que involucra:
  - **Análisis factorial confirmatorio (CFA)** para validar las variables latentes a partir de indicadores observables.
  - **Estimación de modelos estructurales** para analizar relaciones entre variables latentes y observables.
  - **Evaluación de mediación y moderación** para identificar efectos directos e indirectos.
- **Usaremos software estadístico especializado**, como **STATA** o paquetes en **R** y **Python**.

v) **Tipo de resultados que genera:**

- **Estimaciones de cargas factoriales** que indican la relación entre variables latentes y sus indicadores observables.
- **Coeficientes de regresión estructurales** que muestran la influencia de unas variables sobre otras.
- **Identificación de efectos directos, indirectos y totales**, proporcionando una comprensión detallada de las relaciones.

- **Índices de ajuste del modelo**, como RMSEA, CFI y TLI, que indican qué tan bien el modelo se ajusta a los datos.

vi) **Características de las conclusiones que se pueden obtener:**

- **Comprensión de los factores clave** que influyen en los resultados del programa, como la empleabilidad (1/0), la calidad del empleo, la adaptación laboral, la inclusión laboral y el empoderamiento económico.
- **Identificación de mecanismos de mediación**, mostrando cómo ciertas variables median o moderan los efectos del PEI.
- **Información para la toma de decisiones**, proporcionando evidencia para mejorar y focalizar las estrategias del programa.
- **Reconocimiento de diferencias entre grupos poblacionales**, permitiendo adaptar intervenciones según las necesidades específicas.

4.2.3.2 Justificación

Los **Modelos SEM** se ajustan de manera óptima a los objetivos y preguntas orientadoras de la evaluación del PEI, ya que permiten analizar relaciones complejas y multidimensionales entre variables que son esenciales para entender el impacto del programa.

- **Pregunta 3:**

  "Analizar los resultados del programa en cuanto a la calidad de vida de las personas vinculadas a las empresas."
    - **Uso del modelo:** Los SEM permiten modelar la **calidad de vida** como una variable latente compuesta por múltiples indicadores (salud, bienestar emocional, satisfacción laboral, etc.), y analizar cómo el PEI influye en ella.
    - **Tipo de resultados generados:** Estimaciones que muestran la relación entre la participación en el programa y la mejora en la calidad de vida, incluyendo efectos directos e indirectos.
    - **Contribución a los objetivos:** Proporciona una comprensión integral del impacto del PEI en la vida de las personas beneficiarias, más allá de los aspectos laborales.
- **Pregunta 4 de empresas:**
- *"Analizar los cambios en la apropiación de prácticas alrededor de la inclusión, diversidad y equidad al interior de las empresas."*
    - **Uso del modelo:** Permite evaluar cómo las empresas han adoptado prácticas inclusivas y cómo estas prácticas afectan otros resultados organizacionales.
    - **Tipo de resultados generados:** Relaciones entre la variable latente **prácticas de inclusión** y otras variables como retención de talento y clima organizacional.
    - **Contribución a los objetivos:** Identifica el efecto del PEI en promover culturas organizacionales más inclusivas.
- **Pregunta 2:**
- *"¿De qué forma cambia la probabilidad de enganche de las personas beneficiarias y sus niveles de ingreso después de participar en el programa?"*
- El cambio en la probabilidad de enganche requiere el uso de la PILA. Requiere medir en un periodo anterior al programa la probabilidad de enganche, es decir, si se observa un período de no cotización, cuando tarda en reengancharse, antes y después de ser persona beneficiaria en el PEI.
    - **Uso del modelo:** A través de la medición del antes y el después se medirá el cambio en la probabilidad de reenganche. Este ejercicio es posible de hacer para aquellos que perdieron el empleo, pues de otra forma no hay reenganche.

- **Tipo de resultados generados:** Estimaciones que relacionan el cambio en la probabilidad de reenganche después de la intervención.
- **Contribución a los objetivos:** Responde directamente al objetivo 2.

- Para esta pregunta también es necesario analizar los cambios en la sostenibilidad financiera y económica de las personas beneficiarias.
  - **Uso del modelo:** A través de la variable latente **empoderamiento económico**, se puede analizar cómo el PEI influye en la autonomía financiera y acceso a recursos de las personas beneficiarias.
  - **Tipo de resultados generados:** Estimaciones que relacionan la participación en el programa con mejoras en la sostenibilidad financiera.
  - **Contribución a los objetivos:** Proporciona evidencia sobre el impacto económico del PEI en las personas beneficiarias.
- **Objetivo 6:**
- *"Formular recomendaciones para la focalización del programa..."*
  - **Uso del modelo:** Al identificar cómo diferentes características de las personas beneficiarias y las empresas influyen en los resultados, se pueden hacer recomendaciones específicas para mejorar la focalización.
  - **Tipo de resultados generados:** Análisis de mediación y moderación que muestran qué grupos se benefician más y en qué condiciones.
  - **Contribución a los objetivos:** Informa decisiones estratégicas para optimizar el impacto del programa.

El Anexo 5 se presenta la implementación y el abordaje de desafíos metodológicos de la estadísticas descriptivas (secciones 3.3.3 y 3.3.4 respectivamente).

4.2.4 Descripción y justificación de las redes neuronales

4.2.4.1 Descripción

i) **Principales usos:**

Las **redes neuronales artificiales** son modelos computacionales inspirados en el cerebro humano, diseñados para reconocer patrones complejos y relaciones no lineales entre variables. En el contexto de la evaluación del PEI, se utilizarán para:

- **Modelar relaciones no lineales y complejas** entre variables sociodemográficas, políticas empresariales de inclusión, discriminación y factores interseccionales (género, etnia, discapacidad, edad).
- **Capturar interacciones complejas** y patrones ocultos que las metodologías tradicionales podrían no detectar.
- **Predecir con alta precisión** las probabilidades de inclusión laboral, cambios en la probabilidad de enganche, permanencia en el empleo y mejoras en la calidad del empleo de las personas beneficiarias, según las características observables y variables latentes de las personas beneficiarias y las empresas.
- **Manejar datos de alta dimensionalidad**, permitiendo analizar simultáneamente numerosas variables y factores.

ii) **Requerimientos de información:**

- **Datos cuantitativos** recolectados de encuestas aplicadas a personas beneficiarias y sus empresas incluyendo variables observables y latentes.

- **Variables independientes (X):** Características sociodemográficas, nivel educativo, experiencia laboral, políticas de inclusión empresarial, prácticas de diversidad, percepciones de discriminación, entre otras.
- **Variables dependientes (Y):** Resultados laborales como inclusión laboral, calidad del empleo, empleabilidad, permanencia en el empleo y empoderamiento económico.
- **Conjunto de datos amplio y representativo** para entrenar y validar el modelo, asegurando suficientes casos para evitar el sobreajuste y garantizar la generalización.

iii) **Limitaciones:**

- **Necesidad de grandes volúmenes de datos** para entrenar eficazmente las redes neuronales y evitar el sobreajuste.
- **Interpretabilidad limitada:** Requiere de ejercicios adicionales para identificar los pesos y conexiones de las variables independientes.
- **Complejidad computacional:** Requiere recursos computacionales significativos y experiencia técnica para su implementación y ajuste.
- **Riesgo de sobreajuste (overfitting):** Si no se manejan adecuadamente, las redes pueden ajustarse demasiado a los datos de entrenamiento y no generalizar bien a nuevos datos.

iv) **Análisis asociado:**

- **Análisis cuantitativo avanzado** utilizando técnicas de aprendizaje automático supervisado.
- **Entrenamiento y validación del modelo** mediante métodos como división de datos en conjuntos de entrenamiento (80%-70%) y prueba (20%-30%), validación cruzada y uso de métricas de rendimiento.
- **Preprocesamiento de datos**, incluyendo normalización, codificación de variables categóricas y manejo de datos faltantes.
- **Optimización de hiperparámetros**, como el número de capas, neuronas por capa, funciones de activación, tasa de aprendizaje y técnicas de regularización.

v) **Tipo de resultados que genera:**

- **Predicciones precisas** sobre variables de resultado, como la probabilidad de permanencia en el empleo, probabilidad de empleo, desempleo, reenganche, mejoras en la calidad del empleo, nivel de ingreso, empoderamiento económico y sus diferencias basadas en sus características individuales y de la empresa.
- **Identificación de patrones ocultos** y relaciones no evidentes entre variables.
- **Modelos capaces de manejar interacciones complejas** y efectos no lineales entre múltiples factores.
- **Métricas de rendimiento** que evalúan la exactitud, precisión, sensibilidad y capacidad de generalización del modelo.

vi) **Características de las conclusiones que se pueden obtener:**

- **Clasificación** de individuos y empresas según los resultados.
- **Comprensión profunda** de cómo múltiples factores interactúan para influir en los resultados del programa.
- **Identificación de combinaciones específicas** de características individuales y empresariales que conducen a mejores resultados.
- **Información para la toma de decisiones** y formulación de políticas más efectivas, al entender qué estrategias son más exitosas para diferentes grupos poblacionales.
- **Recomendaciones basadas en datos** para mejorar la focalización y eficacia del PEI.

4.2.4.2 Justificación

Las **redes neuronales** se ajustan de manera óptima a los objetivos y preguntas orientadoras de la evaluación del PEI, ya que permiten abordar la complejidad inherente al programa y capturar relaciones no lineales y patrones ocultos entre múltiples variables. A continuación, se detalla cómo esta metodología responde a las necesidades específicas de cada pregunta orientadora, teniendo en cuenta las preguntas orientadoras:

Pregunta 1:

"¿Cuáles son las características de las personas vinculadas en cuanto a su formación académica, perfil laboral y características socioeconómicas por cada uno de los 18 tipos de poblaciones priorizadas por el programa?"

- **Uso del modelo:** Utilizaremos las redes neuronales para **clasificar a las personas** según sus brechas, barreras y características. La red neuronal analizará múltiples variables sociodemográficas y laborales para identificar patrones y segmentar a las personas beneficiarias en función de sus necesidades y desafíos específicos.
- **Tipo de resultados generados:** Generación de perfiles detallados de las personas beneficiarias, identificando grupos con características similares y las barreras que enfrentan, lo que permitirá una comprensión profunda de las 18 poblaciones priorizadas.
- **Contribución a los objetivos:** Ayuda a **personalizar las estrategias del programa** y a **focalizar las intervenciones en función de las características y necesidades de cada grupo**, mejorando la efectividad del PEI.

Pregunta 2:

"¿Cuáles son las condiciones de empleabilidad, permanencia y los niveles de ingreso de las personas beneficiarias después de haber participado del programa? ¿De qué forma cambia la probabilidad de enganche de las personas beneficiarias y sus niveles de ingreso después de participar en el programa?"

- **Uso del modelo:** Las redes neuronales **predecirán cambios en empleabilidad y permanencia** basados en las características individuales y empresariales. Se modelarán las relaciones complejas entre variables para estimar cómo la participación en el programa afecta estas condiciones.
- **Tipo de resultados generados:** Predicciones precisas sobre la probabilidad de enganche, niveles de ingreso y permanencia en el empleo, considerando las interacciones entre múltiples factores.
- **Contribución a los objetivos:** Permite evaluar el impacto del programa en términos cuantificables y entender qué factores influyen más en la mejora de la empleabilidad y permanencia, informando decisiones para optimizar el programa.
- **Objetivo 6:** Permite identificar **qué características y estrategias** son más efectivas para mejorar la empleabilidad y permanencia de las poblaciones vulnerables, informando la **focalización del programa** hacia aquellos grupos donde puede generar mayor impacto.

Pregunta 3:

"¿Cuáles son las características de la calidad de vida de los beneficiarios después de haber participado en el programa?"

- **Uso del modelo:** La red neuronal **predecirá cambios en la calidad de vida** según las características previas de las personas beneficiarias y su participación en el programa. Esto incluye aspectos como bienestar económico, satisfacción laboral y equilibrio entre vida laboral y personal.

- **Tipo de resultados generados:** Estimaciones sobre mejoras en la calidad de vida, identificando patrones y factores que contribuyen a dichos cambios.
- **Contribución a los objetivos:** Proporciona una visión integral del impacto del PEI en la vida de las personas beneficiarias, más allá de los aspectos laborales, esencial para evaluar la efectividad global del programa.

Pregunta 4:

"¿Se observan resultados diferenciales según la experiencia laboral previa, sexo, tipos de poblaciones priorizadas por el programa, grupos de edad, nivel de educación, sector de la economía, entre otros?"

- **Uso del modelo:** Las redes neuronales **predecirán diferencias de resultados** basados en características individuales. Podremos identificar cómo distintos factores afectan los resultados del programa en diferentes subgrupos.
- **Tipo de resultados generados:** Análisis detallados que muestran variaciones en los resultados según variables como experiencia laboral previa, género, edad, nivel educativo y sector económico.
- **Contribución a los objetivos:** Ayuda a adaptar las estrategias del PEI para abordar las necesidades específicas de cada grupo, asegurando una mayor equidad y eficacia en la intervención.

Pregunta 5:

"¿Cuáles son los cambios en el empoderamiento económico de las personas vinculadas a las empresas a partir del programa?"

- **Uso del modelo:** Utilizaremos las redes neuronales para **predecir cambios en el empoderamiento económico** basados en características personales y laborales. Esto incluye analizar cómo factores como ingresos, estabilidad laboral y acceso a recursos influyen en el empoderamiento de las personas beneficiarias.
- **Tipo de resultados generados:** Estimaciones sobre mejoras en el empoderamiento económico, identificando los factores que más contribuyen a este cambio.
- **Contribución a los objetivos:** Permite evaluar cómo el PEI contribuye al empoderamiento económico y qué estrategias pueden potenciar este efecto, informando políticas para mejorar la autonomía económica de las personas beneficiarias.

Pregunta 6:

"¿Qué percepción tienen las personas vinculadas acerca de su proceso de inclusión en las empresas que las contrataron? ¿Cuáles son las experiencias que se identifican en el proceso de adaptación de las personas vinculadas a través del programa?"

- **Uso del modelo:** Las redes neuronales **clasificarán a las empresas** según las percepciones de las personas beneficiarias y las variables de resultado. Analizaremos cómo diferentes prácticas empresariales y culturas organizacionales afectan la percepción de inclusión y la adaptación de los empleados.
- **Tipo de resultados generados:** Identificación de patrones en las percepciones de inclusión y adaptación, correlacionados con características específicas de las empresas.
- **Contribución a los objetivos:** Informa a las empresas y al programa sobre las prácticas que favorecen una mejor inclusión y adaptación, permitiendo mejorar los procesos de integración y satisfacción laboral.

Pregunta 7:

"¿Cuáles deberían ser las poblaciones beneficiarias del programa teniendo en cuenta el objetivo de incentivar la contratación laboral de las poblaciones que enfrentan brechas de acceso al empleo formal?"

- **Uso del modelo:** Las redes neuronales **predecirán qué poblaciones tienen mayores probabilidades de beneficiarse** del programa, identificando grupos que enfrentarían mayores mejoras al participar en el PEI.
- **Tipo de resultados generados:** Recomendaciones sobre focalización de poblaciones, basadas en predicciones de impacto y análisis de brechas y barreras al empleo formal.
- **Contribución a los objetivos:** Ayuda a optimizar la asignación de recursos del PEI, enfocándose en las poblaciones que más lo necesitan y donde el programa puede generar mayor impacto, alineándose con el objetivo de incentivar la contratación de poblaciones vulnerables.

Resumen:

La utilización de redes neuronales en la evaluación del PEI permite:

- **Abordar cada pregunta orientadora** con herramientas capaces de manejar la complejidad y diversidad de variables involucradas.
- **Generar resultados precisos y detallados** que proporcionan perspectivas profundas sobre el impacto del programa en diferentes dimensiones y grupos poblacionales.
- **Informar decisiones estratégicas**, adaptando y mejorando el PEI para maximizar su efectividad y equidad.

De esta manera, las redes neuronales se alinean perfectamente con los objetivos y necesidades de la evaluación, ofreciendo una metodología robusta y flexible para obtener resultados significativos.

El Anexo 5 se presenta la implementación y el abordaje de desafíos metodológicos de la estadísticas descriptivas (secciones 3.4.3 y 3.4.4 respectivamente).

5. Metodologías cualitativas

5.1 Enfoque teórico conceptual

El marco teórico para la evaluación del Programa de Empleo Incluyente (PEI) subraya la relevancia de entender las barreras que enfrentan las poblaciones vulnerables en el acceso al empleo formal. Teorías como la discriminación laboral de Becker y Arrow proporcionan un contexto para identificar y caracterizar las necesidades de estos grupos en el mercado laboral. Además, la teoría del capital humano de Becker es empleada para evaluar cómo las habilidades, educación y experiencia laboral afectan la empleabilidad y el valor de las personas en el mercado laboral, buscando potenciar su inclusión y desarrollo.

En cuanto a las empresas, la evaluación del PEI también se apoya en teorías como la responsabilidad social corporativa de Carroll y la adopción de innovaciones de Rogers para entender los factores que influyen en la adopción de prácticas inclusivas. Estas teorías ayudan a explorar cómo las empresas pueden integrar a poblaciones vulnerables en sus equipos de trabajo y cómo la diversidad y la inclusión pueden convertirse en una ventaja competitiva. Asimismo, la teoría de la diversidad en el lugar de trabajo de Ely y Thomas explica cómo gestionar eficazmente la diversidad para mejorar los procesos organizacionales y fomentar un entorno laboral más inclusivo.

En el Anexo 6, marco interpretativo, se encuentran desarrollados más ampliamente estos enfoques teóricos.

5.2 Desarrollo Metodológico

Como metodologías cualitativas para la evaluación de resultados del PEI se empleará **Análisis Narrativo** y **Evaluación Realista (Configuraciones CMO).** Se utilizará **análisis narrativo** para profundizar en las experiencias y percepciones de las personas beneficiarias del programa, así como de las empresas participantes, sobre los procesos de inclusión, diversidad y equidad en el entorno laboral. Por su parte, **la evaluación realista (configuraciones CMO)** se usará para entender no solo si el programa de empleo incluyente ha sido efectivo, sino cómo y en qué contextos ha generado los resultados deseados. Este enfoque permite identificar los mecanismos subyacentes que explican por qué el programa funciona para algunos grupos poblacionales o en ciertas empresas, y no en otros. Al centrarnos en las interacciones entre contextos específicos, mecanismos de cambio, y resultados observados, podremos ofrecer recomendaciones más precisas y ajustadas a las condiciones locales.

Adicionalmente, se usará Análisis automatizado de texto (NLP) como técnica de análisis complementaria. Puntualmente, se usará para analizar las descripciones de las vacantes laborales y evaluar cómo las empresas participantes están alineando sus ofertas con los principios de inclusión, diversidad y equidad. Se harán ejercicios de NLP: i) Análisis de frecuencia para identificar palabras clave como 'inclusión' o 'diversidad' en vacantes; ii) Extracción de Entidades Nombradas (NER) para detectar menciones de poblaciones vulnerables; iii) Modelos de Tópicos (LDA) para explorar temas latentes como inclusión o flexibilidad laboral en descripciones de vacantes.

El Cuadro 4.1 presenta las metodologías cualitativas propuestas para responder a las preguntas orientadoras.

Cuadro 4.1 Preguntas orientadoras y metodologías de análisis cualitativo

Preguntas Orientadoras	Análisis Narrativo	Evaluación Realista
1. ¿Cuáles son las características de las personas vinculadas en cuanto a su formación académica, perfil profesional y características socioeconómicas por cada uno de los 18 tipos de poblaciones priorizadas por el programa?	Los participantes describirán cómo sus antecedentes educativos y socioeconómicos afectan su acceso al empleo y sus expectativas, permitiendo identificar las particularidades de cada grupo poblacional.	
2. ¿Cuáles son las condiciones de empleabilidad, permanencia y los niveles de ingreso de los beneficiarios después de haber participado del programa? ¿De qué forma cambia la probabilidad de enganche de los beneficiarios y sus niveles de ingreso después de participar en el programa?	Los testimonios personales reflejarán cambios en las percepciones de estabilidad laboral y mejoras en los ingresos, aportando un entendimiento cualitativo sobre cómo los participantes viven su transición post-programa.	Se analizarán los mecanismos que explican el aumento en la probabilidad de enganche y los cambios en ingresos, evaluando cómo el programa influye en la permanencia laboral de distintos grupos.
3. ¿Cuáles son las características de la calidad de vida de los beneficiarios después de haber participado en el programa?	A través de los relatos de los beneficiarios, se explorarán cómo ha cambiado su bienestar emocional, financiero y social, capturando una perspectiva integral de su calidad de vida.	Se identificarán los mecanismos que han influido en la mejora o deterioro de la calidad de vida de los beneficiarios, evaluando qué factores del programa han sido determinantes en dichos cambios.
4. ¿Se observan resultados diferenciales según la experiencia laboral previa, sexo, tipos de poblaciones priorizadas por el programa, grupos de edad, nivel de educación, sector de la economía, entre otros?	Los relatos mostrarán cómo distintas características personales (edad, sexo, experiencia laboral) afectan el impacto del programa en cada grupo poblacional, revelando diferencias percibidas por los mismos participantes.	Permitirá comprender por qué y cómo el programa tiene efectos diferenciales según el contexto específico de cada grupo poblacional, analizando qué mecanismos propician mejores resultados en ciertos subgrupos.
5. ¿Cuáles son los cambios en el empoderamiento económico de las personas vinculadas a las empresas a partir del programa?	Los participantes describirán en qué medida el programa ha mejorado su capacidad de tomar decisiones económicas autónomas, reflejando una evolución en su poder adquisitivo y seguridad financiera.	Se analizarán los mecanismos que explican el empoderamiento económico, investigando cómo y por qué el programa favorece la independencia económica en ciertos contextos.
6. ¿Qué percepción tienen las personas vinculadas acerca de su proceso de inclusión en las empresas que las contrataron? ¿Cuáles son las experiencias que se identifican en el proceso de adaptación de las personas vinculadas a través del programa?	Los testimonios de los participantes presentarán sus experiencias de inclusión y adaptación en las empresas, permitiendo identificar posibles barreras y facilitadores en el proceso de integración laboral.	Se examinarán los mecanismos que facilitan o dificultan la inclusión de los beneficiarios en las empresas, evaluando cómo y por qué la experiencia de adaptación varía según el contexto empresarial y el perfil del empleado.
7. ¿Cuáles deberían ser las poblaciones beneficiarias del programa teniendo en cuenta el objetivo de incentivar la contratación laboral de las poblaciones que enfrentan brechas de acceso al empleo formal?	Los beneficiarios describirán las barreras específicas que enfrentaron, ayudando a identificar qué otras poblaciones podrían beneficiarse del programa, a partir de su experiencia personal.	Se analizarán los mecanismos que explican el éxito o fracaso del programa en reducir las brechas laborales, proponiendo la inclusión de nuevas poblaciones en función de estos hallazgos.
8. ¿Cuáles son las características de las empresas participantes del programa en cuanto a tamaño, actividad económica, ubicación, entre otras variables?	Los testimonios de los directivos de las empresas permitirán explorar cómo perciben su participación en el programa en relación con su tamaño, actividad y ubicación, reflejando cómo estos factores influyen en su decisión de participar.	
9. ¿Cuáles son los factores que influyen en que las empresas decidan participar o no del programa? ¿Cuáles son los factores que influyen en que las empresas participen y no logren recibir el beneficio?	A través de sus relatos, los empresarios podrán señalar los incentivos o barreras percibidas para participar en el programa, revelando sus motivaciones y dificultades.	Se identificarán los mecanismos que explican la participación empresarial y las razones por las cuales algunas empresas no logran recibir el beneficio, analizando las condiciones que afectan su participación.
10. ¿Cuáles son los factores que inciden en que los aliados (Agencias de Empleo, gremios, alcaldías locales) promuevan la participación de las empresas en el programa?	Los testimonios de los actores clave permitirán conocer su perspectiva sobre la promoción del programa, detallando los factores que influyen en su apoyo o desmotivación para incentivar la participación empresarial.	Se analizarán los mecanismos que impulsan o limitan la promoción del programa por parte de aliados estratégicos, identificando cómo y por qué ciertos factores contextuales afectan su nivel de compromiso.
11. ¿Cómo han cambiado la apropiación de prácticas alrededor de la inclusión, diversidad y equidad al interior de las empresas? Hacer énfasis en los procesos de reclutamiento y selección.	Las experiencias narradas por los directivos empresariales reflejarán cómo han cambiado sus prácticas en términos de inclusión, diversidad y equidad, destacando los cambios observados en sus procesos de reclutamiento y selección.	Se analizará cómo el programa ha influido en los cambios organizacionales mediante la identificación de mecanismos (como capacitaciones o incentivos) y contextos favorables.
12. ¿Qué cambios culturales se observan en la inclusión, diversidad y equidad de las empresas relacionados con la población que enfrenta barreras de acceso en Bogotá? ¿Hay prácticas de discriminación o exclusión al interior de la empresa hacia los tipos de poblaciones priorizadas por el programa?	A través de los relatos de los beneficiarios y directivos, identificarán percepciones sobre los cambios culturales en las empresas en relación con la inclusión y equidad, así como experiencias de discriminación o exclusión.	Permitirá identificar los contextos en los que las prácticas de inclusión han generado cambios culturales, evaluando los mecanismos que explican la persistencia o eliminación de prácticas discriminatorias en las empresas.

5.2.1Desarrollo y justificación del análisis narrativo

El análisis narrativo del Programa de Empleo Incluyente es una herramienta metodológica clave para evaluar sus resultados en las oportunidades de empleo, el empoderamiento económico y la adaptación de las personas beneficiarias en sus nuevos entornos laborales. Este enfoque metodológico se alinea con los objetivos de inclusión laboral de poblaciones vulnerables, promoviendo la equidad y diversidad. A través de relatos detallados, se exploran experiencias y percepciones de todos los actores involucrados, revelando tanto las experiencias directas en el entorno laboral como el impacto en la vida personal de las personas beneficiarias. El análisis narrativo es crucial para identificar patrones en las características de las personas beneficiarias y los resultados en términos de empleabilidad e inclusión, además de evidenciar cómo las prácticas empresariales y el apoyo influyen en su estabilidad laboral. También informa el fortalecimiento de la autonomía económica y cómo la inclusión laboral contribuye al empoderamiento económico y social de las personas beneficiarias, mientras evalúa la efectividad de las políticas de inclusión para mitigar la discriminación laboral.

Así, el análisis narrativo aplicado a la evaluación del Programa de Empleo Incluyente permite una comprensión integral de cómo los objetivos del programa se traducen en experiencias concretas y significativas para las personas beneficiarias y las empresas.

El **análisis narrativo** es una metodología cualitativa que implica la recolección y análisis de relatos personales para entender cómo las personas experimentan y perciben su entorno. Estos métodos se desarrollan a partir de la hipótesis de que a través de las historias y narraciones que las personas construyen, le dan sentido a sus experiencias y al mundo que las rodea. Según Riessman (2008) en 'Narrative Methods for the Human Sciences', el análisis narrativo implica examinar la estructura de la historia, el contenido y el contexto en el que se cuenta, para entender cómo las personas construyen sus identidades y realidades sociales. Además, Chase (2005) enfatiza que las narrativas no son meramente personales, sino que están moldeadas por narrativas culturales, que pueden influir en cómo las personas entienden y relatan sus experiencias.

Así, las narrativas pueden develar elementos relevantes sobre los contextos culturales, sociales y personales y la manera en que son experimentadas por cada quien en la construcción de su subjetividad. Al escucharlas atentamente, es posible comprender los significados que las personas atribuyen a sus experiencias y cómo estos están influenciados por su entorno social. En el contexto del **PEI**, esta metodología es particularmente pertinente porque permite capturar en profundidad las experiencias y percepciones de las personas beneficiarias, los empleadores y otros actores clave sobre las oportunidades laborales, el empoderamiento económico y la adaptación en sus nuevos entornos de trabajo.

De igual manera, esta metodología permite comprender cómo personas de grupos poblacionales que han sido históricamente excluidos del mercado laboral o que enfrentan barreras para emplearse, construyen su relación con el trabajo, con la búsqueda de oportunidades, con otras personas en entornos laborales, entre otros aspectos. Es decir, permite entender cómo características personales como el género, la edad o el lugar de origen condicionan la experiencia de un colectivo de personas frente al mercado laboral y al trabajo y cómo las narrativas individuales se diferencian o refuerzan a partir de estas experiencias colectivas. Así mismo, permite indagar cómo las experiencia y relaciones con el mercado laboral y el trabajo puede ser diferente para diferentes grupos poblacionales, comprendiendo lo particular de cada uno. Esto es esencial no solamente para identificar los resultados del

Programa sino también las barreras o situaciones específicas vividas por las poblaciones priorizadas que facilitan o dificultan dichos resultados.

Así mismo, para los directivos y equipos de recursos humanos de las empresas que participaron en el Programa, el enfoque narrativo nos permite entender cómo fue su experiencia y cómo dieron sentido a esta experiencia. Así, es el enfoque idóneo para indagar con detalle cuáles fueron sus motivaciones para registrarse en el programa, cómo conocieron del mismo, cómo fue la experiencia del registro y la postulación de las personas que vincularon hasta recibir el beneficio y finalmente, cómo dan sentido a los cambios e implicaciones que pudo tener para la empresa la inclusión de personas de difícil colocación o que hacen parte de grupos históricamente vulnerados.

En comparación con otras metodologías que podrían aplicarse a los objetivos de la evaluación como técnicas etnográficas en los lugares de trabajo o en los procesos de búsqueda de trabajo, las metodologías narrativas son más ventajosas porque al centrarse en historias personales, la indagación narrativa empodera a los participantes, que ya hacen parte de grupos poblacionales marginados o subrepresentados, al darles voz en el proceso de evaluación poniendo como elemento central su perspectiva.

El anexo 7.1 describe las categorías de análisis que guiaran esta metodología cualitativa en su aplicación al PEI.

Análisis de datos

Investigación narrativa y el modelo de análisis de Labrov

Para realizar el análisis narrativo de los resultados del PEI, se seguirá una metodología basada en la investigación narrativa que permitirá comprender en profundidad las experiencias de los participantes. En un primer paso, se procederá a la **identificación de narrativas** mediante una revisión de las transcripciones de entrevistas, grupos focales, talleres participativos y estudios de caso. Este proceso buscará extraer historias que reflejen las experiencias laborales y de inclusión de las personas beneficiarias, capturando relatos sobre cómo la participación en el PEI ha impactado su vida cotidiana, estabilidad económica e integración en el mercado laboral. Este proceso se complementará con técnicas avanzadas de procesamiento de lenguaje natural (NLP) como se describe en el Anexo 7.3.

El segundo paso consiste en la **codificación de las narrativas** recolectadas, empleando el modelo estructural de Labov (Labov & Waletzky, 1967; Labov, 1972: Parks, 2023). Este modelo nos permite descomponer cada historia en componentes clave, revelando su estructura narrativa y el significado subyacente que los participantes otorgan a sus experiencias con el programa de empleo incluyente.

Las etapas del modelo de Labov que aplicaremos son las siguientes:

1. **Abstract (Resumen)**: es la introducción de la historia que establece el propósito de la narración. Aquí identificaremos por qué el narrador decide compartir su experiencia y qué espera comunicar. Esto nos ayuda a entender el enfoque y las expectativas del participante desde el inicio.

2. **Orientation (Orientación)**: Proporciona detalles contextuales esenciales como el tiempo, lugar, personajes y situación inicial. Esta información contextualiza la historia y permite situar al investigador en el escenario de los eventos narrados, facilitando la comprensión de las circunstancias en las que se desarrollan.

3. **Complicating Action (Acción Complicante)**: Describe las secuencias de eventos significativos que constituyen el núcleo de la historia. En esta etapa, identificamos los acontecimientos clave que impulsan la narrativa y que son cruciales para entender las experiencias y desafíos enfrentados por los participantes.

4. **Evaluation (Evaluación)**: Destaca el significado y la importancia que el narrador atribuye a los eventos descritos. Esta sección es esencial para comprender las percepciones y emociones de los participantes respecto al programa, revelando cómo interpretan y valoran sus experiencias.

5. **Resolution (Resolución)**: Presenta el desenlace de la historia, mostrando cómo se resolvieron las situaciones o cuáles fueron los resultados finales de los eventos narrados. Esto nos permite entender las consecuencias directas de las experiencias vividas.

6. **Coda (Coda)**: Ofrece un cierre a la narración, conectando la experiencia pasada con el presente o con futuros cambios esperados. La coda permite al narrador reflexionar sobre el impacto duradero de su experiencia y cómo ha influido en su situación actual o expectativas futuras.

Al aplicar este modelo al análisis de las narrativas de los participantes del PEI, buscamos:

I) **Comprender en profundidad** cómo los participantes perciben y experimentan el programa, identificando aspectos positivos y áreas de mejora. II) **Identificar patrones comunes y diferencias** en las experiencias de los participantes, lo cual puede estar relacionado con variables las características de la empresa o características sociodemográficas del individuo. III) **Revelar los resultados** que el programa tiene en las vidas de las personas. Iv) **Informar recomendaciones** para el mejoramiento del programa, basadas en las experiencias y evaluaciones expresadas por los participantes.

En el anexo 7.2 se presentan ejemplos concretos de la alicación del modelo propuesto para posibles narrativas asociadas a personas beneficiarias del PEI.

Importancia del Modelo de Labov

El modelo de Labov es una herramienta ampliamente reconocida en el análisis cualitativo de narrativas por su capacidad para estructurar y profundizar en el significado de las historias personales. Al **descomponer** las narrativas en sus componentes fundamentales, podemos c**apturar la complejidad** de las experiencias individuales, entendiendo no solo lo que sucedió, sino también cómo y por qué los eventos son significativos para los narradores y a**nalizar la construcción del significado**, explorando cómo los participantes interpretan sus experiencias y cómo estas influyen en su percepción del programa.

Desde una segunda se **contextualizan las historias** dentro de marcos culturales y sociales más amplios, reconociendo la influencia de factores externos en las narrativas personales. A partir de este análisis, las narrativas serán organizadas bajo **temas y subtemas** comunes. Los temas principales incluirán la inclusión laboral, el empoderamiento económico y la permanencia en el empleo. Los subtemas explorarán aspectos como las percepciones sobre la discriminación, la equidad, los cambios en la calidad de vida y las oportunidades de desarrollo laboral.

El **análisis de las narrativas** se llevará a cabo tanto a nivel individual con las personas participantes a través de entrevistas y también grupal en las empresas y aliados (grupos focales y talleres). Se examinarán las historias individuales en conjunto, observando cómo los participantes validan o refutan narrativas compartidas sobre la inclusión laboral y las políticas de diversidad implementadas por las empresas. Esto permitirá identificar patrones comunes, divergencias y tendencias que aporten una comprensión más completa de las dinámicas presentes en el programa.

Para complementar el análisis, se procederá a la **contabilización de temas y subtemas**, registrando la frecuencia con la que ciertos aspectos, como mejoras en la calidad del empleo, reducción de la discriminación o aumento de la autonomía económica, aparecen en las historias. Este paso facilitará la relación de los hallazgos cualitativos con tendencias cuantitativas, ofreciendo una perspectiva más integrada del impacto del PEI. Este proceso se llevará a cabo con técnicas de NLP que se presentan en el Anexo 3 (ver la primera sección sobre Análisis de frecuencia y nube de palabras).

Posteriormente, se dedicará especial atención a la **reflexión y co-construcción del significado**, analizando el rol del investigador en la interpretación de las narrativas y la dinámica social durante las sesiones grupales y participativas. Esto es fundamental para asegurar una interpretación auténtica y ética de las experiencias narradas, reconociendo posibles sesgos y garantizando que las voces de los participantes sean representadas con fidelidad y respeto.

Finalmente, se realizará una **síntesis y presentación final** de los hallazgos, agrupando las narrativas bajo categorías temáticas claras. Se utilizarán ejemplos representativos y citas textuales que ilustren el impacto del PEI en la inclusión laboral y el empoderamiento económico de las personas beneficiarias. Esta presentación integrará los diversos elementos del análisis, vinculándolos con el marco teórico, las percepciones de las empresas y proporcionando una comprensión profunda y coherente de cómo el programa ha influido en las vidas de los participantes.

En conjunto, este enfoque estructurado del análisis narrativo no solo permite capturar la profundidad y la complejidad de las experiencias vividas por los participantes, sino que también garantiza que el análisis sea riguroso, validado por las propias voces de las personas beneficiarias y centrado en los resultados clave que el jurado calificador considerará esenciales para una evaluación exitosa.

Para llevar a cabo estos pasos, los datos narrativos recolectados serán transcritos y analizados utilizando técnicas de análisis de contenido temático. Esto implicará:

- **Codificación inicial:** Se dará una lectura exhaustiva de las transcripciones y la información de vacantes extraída con web scraping (ver Anexo 2) para familiarizarse con los datos.
- **Generación de códigos:** Se identificarán temas y patrones recurrentes que surgen de las narrativas. Este paso será complementado por análisis automatizado de texto (ver Anexo 3, tercera sección: modelo de tópicos).
- **Agrupación temática:** Se organizarán los códigos relacionados en temas más amplios que respondan directamente a las preguntas de la evaluación.
- **Interpretación:** Se relacionarán los hallazgos con las condiciones y objetivos del PEI para formular conclusiones y recomendaciones.

Como se ha mencionado, en el marco de nuestra evaluación cualitativa, además de emplear técnicas tradicionales de análisis de contenido temático y codificación manual, incorporaremos técnicas avanzadas de Procesamiento de Lenguaje Natural (NLP), no como un sustituto de la codificación manual sino como un complemento que agilice todo el proceso y le de mayor precisión. El NLP es un campo de la inteligencia artificial que se enfoca en la interacción entre los computadores y el lenguaje humano, y tiene la capacidad de analizar grandes volúmenes de texto de manera eficiente y efectiva. Esta integración tiene como objetivo enriquecer nuestro entendimiento de los datos narrativos y optimizar el proceso de análisis.

En nuestro enfoque de análisis de datos, combinaremos la codificación manual tradicional con la codificación asistida por técnicas de NLP. Utilizaremos NLP en las etapas iniciales para obtener una visión general de los datos, identificar rápidamente temas y conceptos recurrentes en las narrativas que son significativos para responder a las preguntas orientadoras y guiar el enfoque del análisis cualitativo (etapa exploratoria). Esto servirá como punto de partida para la codificación manual, ahorrando tiempo y asegurando que se consideren los patrones emergentes en el corpus completo. Posteriormente emplearemos los hallazgos de NLP para complementar y enriquecer el análisis manual, identificando áreas que requieran mayor atención. La codificación manual es crucial para asegurar que comprendemos el significado detrás de las experiencias personales relatadas.

Consideraciones éticas

- **Consentimiento informado**: Los participantes serán informados sobre los objetivos de la investigación y se obtendrá su consentimiento antes de participar.
- **Confidencialidad**: Se garantizará la confidencialidad de los datos personales y empresariales y se protegerá la identidad de los participantes mediante pseudónimos (por ejemplo, la empresa X…).
- **Acción sin daño**: Se cuidará queque, en el desarrollo de las entrevistas, talleres y otros instrumentos no se generen efectos adversos inesperados en las personas o grupos que van a participar de la evaluación. Este enfoque es relevante en términos éticos porque trabajaremos con personas de grupos poblacionales que experimentan discriminación en el contexto cultural de Bogotá, así que es fundamental cuidar que desde la evaluación no se repliquen acciones que refuercen o amplien daños causados por dicha discriminación. Esto implica, por ejemplo, cuidar que no se cuestionen o juzguen sus experiencias de vida o decisiones laborales y que no se les trate con condecendencia o desde “el lugar del salvador”.

5.2.2 Desarrollo y justificación de la evaluación realista

La evaluación de programas sociales complejos como el PEI requiere enfoques metodológicos que permitan comprender no solo si el programa funciona, sino también cómo, para quién y en qué contextos produce sus efectos. En este sentido, la **evaluación realista** propuesta por Pawson y Tilley (1997) ofrece un marco adecuado para responder a las preguntas orientadoras de la evaluación del PEI.

La evaluación realista se basa en la premisa de que los programas son intervenciones que generan cambios a través de **mecanismos** específicos que operan en ciertos **contextos** para producir **resultados**. Este enfoque se centra en identificar y comprender las relaciones entre **contexto (C)**, **mecanismo (M)** y **resultado (O)**, conocidas como configuraciones CMO. Al aplicar esta metodología al PEI, es posible explorar cómo las diversas estrategias del programa

influyen en las personas beneficiarias y las empresas participantes, considerando las particularidades de cada entorno.

El uso de la evaluación realista es especialmente adecuado/justificado para el PEI por varias razones. En primer lugar, permite una comprensión profunda de los **mecanismos de cambio**. El PEI busca mejorar la inclusión laboral de grupos vulnerables que enfrentan barreras para acceder al empleo formal. A través de la evaluación realista, se puede analizar cómo los mecanismos del programa—, como incentivos a empresas, acciones con aliados del ecosistema de empleo para la inclusión de poblaciones y generación de información—generan cambios en la empleabilidad y el empoderamiento económico de las personas beneficiarias. Esto se alinea con las teorías de empleabilidad de Hillage y Pollard (1998) y de capital humano de Becker (1964), que enfatizan el desarrollo de habilidades y competencias para mejorar la inserción laboral.

En segundo lugar, este enfoque permite un análisis detallado del **contexto y su influencia**. Las experiencias de discriminación y exclusión laboral varían según el contexto socioeconómico y cultural. Las teorías de discriminación de Becker (1957) y Arrow (1973) destacan cómo los prejuicios y estereotipos afectan las oportunidades laborales. La evaluación realista permite examinar cómo estos contextos específicos influyen en la efectividad del PEI y cómo el programa puede adaptarse para abordar diferentes formas de discriminación, incorporando el enfoque interseccional de Crenshaw (1989) para comprender las múltiples dimensiones de la desigualdad.

Además, la evaluación realista facilita la **evaluación de resultados diferenciales**. Dado que el PEI atiende a diversas poblaciones priorizadas, es crucial entender qué funciona para quién. Este enfoque permite analizar cómo los resultados del programa pueden variar entre grupos como mujeres, personas con discapacidad, migrantes, jóvenes y personas mayores de 50 años. Así, es posible identificar configuraciones CMO específicas que expliquen los resultados diferenciales, alineándose con el énfasis en la inclusión laboral y la calidad del empleo definida por la OIT (1999).

El análisis del **rol de las empresas** es otro aspecto clave donde la evaluación realista aporta valor. Las empresas juegan un papel fundamental en la implementación del PEI. Esta metodología permite explorar cómo factores como la responsabilidad social corporativa (Carroll, 1991), la adopción de innovaciones (Rogers, 2003) y la gestión de la diversidad (Ely y Thomas, 2001) influyen en la decisión de las empresas de participar en el programa y adoptar prácticas inclusivas. Esto es esencial para responder a preguntas sobre los factores que motivan o dificultan la participación empresarial y los cambios culturales en las organizaciones.

Finalmente, la evaluación realista proporciona una base sólida para **informar mejoras en el programa**. Al comprender los mecanismos y contextos que conducen al éxito o fracaso del PEI, es posible ajustar estrategias para mejorar la adaptación laboral de las personas beneficiarias (Dawis y Lofquist, 1984), fortalecer el empoderamiento económico (Kabeer, 1999) y promover prácticas laborales decentes (OIT, 1999). Este enfoque permite que la evaluación no solo sea un ejercicio de rendición de cuentas, sino también una herramienta para el aprendizaje y la optimización continua del programa.

En conclusión, la metodología de evaluación realista es adecuada para el PEI porque permite una comprensión profunda y contextualizada de cómo el programa impacta a diferentes actores en diversos entornos. Al centrarse en las configuraciones de contexto, mecanismo y

resultado, esta metodología facilita la identificación de los mecanismos efectivos y las condiciones que potencian o limitan el éxito del programa, proporcionando una base sólida para la toma de decisiones informadas y la mejora continua del PEI.

Antes de abordar la explicación sobre la implementación, es fundamental destacar que las configuraciones **Contexto-Mecanismo-Resultado (CMR)** son intrínsecas a la **Evaluación Realista (ER)**. Estas configuraciones son esenciales porque buscan comprender **cómo** y **por qué** los programas producen determinados resultados en diversos contextos. La premisa central de la ER es que los **mecanismos** (es decir, los procesos internos que explican cómo funciona el programa) no operan de manera aislada; su efectividad depende profundamente del **contexto** en el que se implementan. Como afirman Pawson y Tilley (1997), "*los resultados causales se derivan de mecanismos que actúan en contextos*".

Las configuraciones CMR permiten explorar y explicar las relaciones dinámicas entre el contexto, los mecanismos activados por la intervención y los resultados obtenidos. Los autores proponen la configuración CMR como herramienta clave para evaluar cómo y por qué los programas funcionan bajo diferentes circunstancias. Mediante la formulación de hipótesis sobre posibles configuraciones CMR, es posible entender en profundidad los impactos de las intervenciones en políticas públicas. Estas hipótesis se validan empíricamente a través del trabajo de campo, lo que implica una **interacción iterativa entre teoría y evidencia**.

La **Evaluación Realista** (ER) es una metodología cualitativa y teóricamente informada que se centra en comprender los procesos causales subyacentes a la implementación de un programa, identificando para quién funciona, en qué condiciones y de qué manera (Pawson y Tilley, 1997). Dada la complejidad y multidimensionalidad del PEI, la ER es especialmente pertinente y viable para responder a las preguntas orientadoras de la evaluación.

La aplicabilidad de la ER al contexto del PEI **se justifica por tres razones principales**: primero, la ER permite identificar los **mecanismos subyacentes** que generan los resultados del programa. Esto es crucial para entender cómo el PEI influye en la empleabilidad, la calidad del empleo y la calidad de vida de las personas beneficiarias, entre otras categorías. Por ejemplo, podemos explorar cómo la capacitación, el acompañamiento y los incentivos económicos actúan como mecanismos que mejoran las habilidades y la confianza de los participantes, facilitando su inserción laboral.

Segundo, la ER analiza cómo los **contextos específicos** afectan la implementación y los resultados del programa. El PEI opera en un entorno con diversas poblaciones y empresas, cada una con características particulares. La ER nos permite examinar cómo factores contextuales como el género, la edad, la etnia, el nivel educativo de las personas beneficiarias, así como el sector económico, el tamaño y la cultura organizacional de las empresas, influyen en la efectividad de los mecanismos del programa.

Tercero, la ER proporciona perspectivas y hallazgos para **optimizar el programa**, permitiendo adaptar estrategias según las necesidades y circunstancias de los diferentes actores y grupos de interés. Al comprender las configuraciones CMR específicas que conducen al éxito o limitan la eficacia del PEI, es posible formular recomendaciones precisas para mejorar su diseño e implementación.

Implementación de la metodología

La implementación de la ER en la evaluación del PEI se **adaptará a las diferentes fases de la evaluación**, permitiendo un análisis flexible y contextualizado:

1. **Diseño de instrumentos**: el diseño de instrumentos parte de las configuraciones CMR presentadas en el Anexo 4 y se explican más adelante.

2. **Recolección de datos**: se emplearán técnicas cualitativas como entrevistas semiestructuradas, grupos focales y talleres participativos, dirigidos tanto a personas beneficiarias como a representantes de empresas y otros actores clave. Las herramientas de recolección se ajustarán para captar información relevante sobre los contextos específicos y los mecanismos percibidos, facilitando la identificación de configuraciones CMR. Se garantizará la diversidad en la selección de participantes, abarcando diferentes grupos poblacionales y tipos de empresas (ver la sección del diseño de la muestra cualitativa).

3. **Análisis de datos**: se utilizará un enfoque analítico basado en la identificación y articulación de configuraciones CMR (las cuales se presentan a continuación). Mediante el análisis temático (categorías de análisis presentadas en los Cuadros 4.2 a 4.7) y la codificación de las narrativas recogidas, se explorarán los mecanismos que impulsan los resultados del programa en diferentes contextos. Se aplicarán técnicas como el **análisis de patrones y la comparación constante** para detectar regularidades y variaciones en las experiencias de los participantes. El análisis de patrones y la comparación constante se realizará con el apoyo de técnicas de NLP que se presentaron en la sección 4.3 y se explican en el Anexo 3.

4. **Validación y refinamiento de hipótesis**: las hipótesis iniciales sobre las configuraciones CMR se contrastarán con los datos empíricos obtenidos. **Este proceso iterativo** permitirá refinar las teorías sobre cómo funciona el PEI, identificando **factores contextuales críticos y mecanismos efectivos**. Se buscará corroborar las hipótesis mediante la triangulación de fuentes (análisis narrativo y resultados cuantitativos) y la contrastación con la literatura existente de nuestro marco teórico.

5. **Presentación de resultados**: los hallazgos se integrarán para ofrecer una visión comprensiva y práctica que informe la toma de decisiones y la mejora continua del programa. Se elaborarán recomendaciones específicas basadas en las configuraciones CMR identificadas, orientadas a **optimizar los mecanismos efectivos y a adaptar el programa a los contextos identificados y los grupos poblacionales focalizados**. Se presentarán casos ilustrativos que ejemplifiquen las configuraciones CMR más relevantes.

Configuraciones CMO para cada pregunta orientadora

En el Anexo 4, se presentan las configuraciones CMO que se utilizarán **para cada pregunta orientadora** que lidera esta metodología cualitativa (un cuadro por cada pregunta), las cuales están sombreadas en gris en el Cuadro 4.1 al comienzo del capítulo (preguntas 4, 7, 9 10, 11 y 12). Estas configuraciones son las hipótesis iniciales que deben ser revisadas y adaptadas como se explica posteriormente. Los Cuadros del Anexo 4 presentan e**n negrilla categorías de análisis** en cada celda de los cuadros que presentan a continuación las configuraciones CMO para las preguntas orientadoras que liderará la evaluación realista propuesta.

[ACA ANEXO 4]

Este avance en la identificación de las configuraciones CMO representa un paso crucial en el diseño de los instrumentos de recolección de datos, ya que estas configuraciones constituyen las hipótesis que serán verificadas en campo. No obstante, la propuesta va más allá: consiste en un **enfoque iterativo** que implica ir a campo, analizar los datos obtenidos y luego regresar con las configuraciones refinadas y los instrumentos ajustados, basados en los primeros hallazgos. Esta metodología garantiza que el proceso de evaluación se adapte continuamente a la realidad observada, permitiendo un análisis más profundo y preciso de los mecanismos subyacentes.

En la práctica, **el proceso iterativo entre teoría y evidencia** en la evaluación realista se desarrolla de manera sencilla y estructurada. Inicialmente, formulamos hipótesis basadas en el marco teórico y en nuestro conocimiento previo sobre cómo debería funcionar el programa en diferentes contextos (Teoría del Cambio). Estas hipótesis plantean relaciones entre el contexto, los mecanismos del programa y los resultados esperados y se han expresado en los seis cuadros anteriores. Con estas configuraciones se diseñan los instrumentos de recolección.

Una vez establecidas estas hipótesis, procedemos al trabajo de campo, donde recopilamos evidencia empírica a través de entrevistas, grupos focales, talleres y observaciones. Esta información nos permite ver cómo operan realmente los mecanismos del programa en contextos específicos y si los resultados observados coinciden con lo que anticipamos.

Al analizar los datos recopilados, podemos encontrar que algunas hipótesis se confirman, mientras que otras requieren ajustes. Por ejemplo, podríamos descubrir que un mecanismo no funciona como esperábamos en un determinado grupo poblacional debido a factores contextuales no considerados inicialmente.

Este hallazgo nos lleva a refinar nuestras hipótesis, incorporando los nuevos conocimientos adquiridos sobre los contextos y mecanismos en juego. Con estas hipótesis revisadas, regresamos al campo para recopilar más datos que nos permitan validar estos nuevos planteamientos. Durante la evaluación, se acordará con el CTS un porcentaje de instrumentos de recolección para la primera ronda de validación de las hipótesis iniciales (recomendamos, un 60%), con el fin de volver a campo con las hipótesis revisadas y poder expandir el conocimiento.

Este ciclo de formular hipótesis, recolectar evidencia, analizar y ajustar se repite una vez. Es un proceso dinámico donde la teoría y la evidencia se alimentan mutuamente. A través de esta interacción continua, logramos una comprensión más profunda y precisa de cómo y por qué el programa produce ciertos resultados en diferentes situaciones.

Al final, este proceso iterativo nos permite construir una explicación sólida y detallada del funcionamiento del programa, adaptada a los diversos contextos en los que se implementa. Esto es esencial para poder hacer recomendaciones informadas, mejorar la efectividad del programa y asegurar que realmente beneficie a las personas y empresas involucradas.

Adaptación de la Evaluación Realista en dos etapas

Como se mencionó, la Evaluación Realista se implementará en dos fases diferenciadas y complementarias, con el propósito de profundizar en los mecanismos causales que explican los resultados observados en el PEI. Este enfoque permitirá ir más allá de una descripción estática de los resultados y, en cambio, identificar y probar hipótesis sobre **cómo** y **por qué** el programa funciona de manera diferente para los distintos grupos poblacionales y contextos.

Etapa 1: Exploración de mecanismos clave

En la **primera fase**, se llevarán a cabo entrevistas, grupos focales y análisis de datos cualitativos preliminares centrados en las preguntas orientadoras. En esta etapa, el enfoque será verificar los **mecanismos causales emergentes** que se activan en ciertos contextos para generar los resultados observados. Es decir que se explorarán como las configuraciones CMO de los cuadros 4.2 a 4.7 tienen validez en la realidad. A través de la recopilación de narrativas y relatos de las personas beneficiarias y empresas, se obtendrán hallazgos sobre **qué configuraciones de contexto (C)** están interactuando con los **mecanismos (M)** subyacentes para generar los **resultados (R)**.

Algunos mecanismos que pueden emerger incluyen:

- **Adaptación a la cultura organizacional** en las empresas por parte de las personas beneficiarias del programa.
- **Barreras de discriminación estructural** en ciertos sectores que limitan el acceso a empleo de calidad para grupos vulnerables.
- **Dinámicas de apoyo laboral** que refuerzan el empoderamiento económico y la permanencia en el empleo.

Etapa 2: Refinamiento y profundización de hipótesis

Con base en los hallazgos obtenidos en la etapa 1, la **segunda fase** del levantamiento cualitativo tendrá un enfoque más **profundo y específico**, donde se **reformularán** o **ajustarán las hipótesis** iniciales sobre los mecanismos emergentes. Esta fase buscará confirmar y precisar nuevas configuraciones **Contexto-Mecanismo-Resultado (CMO)** que fueron detectadas en la etapa 1.

En esta etapa, los instrumentos de recolección de información se **ajustarán** para reflejar estas nuevas hipótesis. Por ejemplo:

- Se incorporarán preguntas más **específicas** sobre los mecanismos emergentes, como cómo se percibe y aborda la discriminación en empresas de determinados sectores.
- Se crearán instrumentos ajustados para indagar más a fondo sobre las condiciones contextuales que permiten o inhiben el empoderamiento económico en las mujeres o jóvenes personas beneficiarias.

Asimismo, se llevarán a cabo entrevistas y talleres adicionales con los actores involucrados en los contextos clave donde se han identificado resultados particulares (ver Capítulo 6). Estos esfuerzos adicionales no solo profundizarán en áreas críticas, sino que también permitirán ajustar la evaluación para abordar nuevos temas emergentes que no fueron considerados en la primera etapa, ya que no se habían manifestado inicialmente. Esta segunda fase garantiza una mayor flexibilidad y capacidad de adaptación, asegurando que la evaluación cubra todas las dinámicas relevantes que puedan surgir.

Uso de los resultados de la primera fase para guiar la segunda

El diseño iterativo de la evaluación realista asegura que la segunda fase no solo profundice en los mecanismos detectados en la primera, sino que también permita comprobar si esos mecanismos se activan de forma similar en otros contextos o grupos poblacionales. Además,

la **retroalimentación constante** entre las dos fases facilita que la evaluación no se limite a un enfoque estático, sino que evolucione conforme emergen nuevos hallazgos. Esto refuerza la validez del análisis cualitativo, ya que asegura que las configuraciones CMO que explican los resultados del programa estén basadas en evidencia empírica robusta y contextualizada. En una evaluación realista tradicional, el trabajo de campo se centra en probar si las hipótesis planteadas exante son válidas o no. Sin embargo, este enfoque no permite explorar nuevas hipótesis subyacentes que puedan surgir durante el proceso y que refuercen o complementen los hallazgos obtenidos.

En resumen, la estructura en dos fases de la **Evaluación Realista** permite una **exploración iterativa** y un refinamiento de las hipótesis, ajustando los instrumentos de recolección y profundizando en los mecanismos causales. Esto garantiza que la evaluación no solo capture resultados, sino que también ofrezca explicaciones sólidas sobre los **factores subyacentes** que generan los efectos observados en **diferentes grupos y contextos**, cumpliendo así con los objetivos de la evaluación del PEI.

El proceso de iteración en el levantamiento de información se hará tanto para las personas como para las empresas beneficiarias. Como, como se describe en el capítulo 6 (Diseño de muestra cualitativo). A continuación, se describe más específicamente cómo se diseñan la primera y segunda etapa de recolección y el ajuste de instrumentos, y algunas precisiones particulares para la distribución de instrumentos entre las dos etapas de levantamiento para actores específicos con quienes se levantará información:

De manera general se seguirán los siguientes pasos:

1. **Primer levantamiento:** Siguiendo lo propuesto metodológicamente con la incorporación de técnicas narrativas y de evaluación realista, se aplicarán entrevistas individuales, entrevistas grupales, grupos focales o talleres.
2. **Análisis, ajuste de hipótesis e identificación factores a profundizar:** En esta etapa se tendrá como insumo las transcripciones de las entrevistas realizadas y las notas y percepciones de moderadores y equipos de campo. Tomando en consideración que el tiempo para el campo es limitado, no se pretende esperar a tener toda la información transcrita, codificada manualmente y analizada para hacer los ajustes necesarios para el segundo levantamiento. Así, con las transcripciones se adelantarán el análisisel análisis automatizado de texto para obtener hallazgos preliminares y estos serán revisados en una sesión de trabajo con los profesionales que participaron del levantamiento, quienes con sus notas e impresiones de campo podrán complementar los hallazgos para corregir las hipótesis planteadas preliminarmente y se ajustarlas según lo encontrado. Así mismo, se identificará si hay alguna indagación que sea pertinente incluir o algún aspecto a profundizar y las características puntuales de los actores con quienes sea relevante hacer el segundo levantamiento.
   Ahora bien, de acuerdo con lo encontrado en la prueba piloto de la evaluación, se revisará si el levantamiento en dos etapas es inviable desde el punto de vista operativo para alguno de los actores planteados dentro del diseño de muestra. Por ejemplo, si el contacto con las empresas se dificulta, se evaluará levantar la información de directivos de empresas en una sola etapa. Estas modificaciones al diseño de muestra serán presentadas al CTS con el informe de prueba piloto.

3. **Ajuste de instrumentos:** Con los hallazgos preliminares se ajustarán los instrumentos[6].
4. **Segundo levantamiento:** De acuerdo con el diseño muestral realista, la elección de los actores con quienes ser hará este segundo levantamiento no se especifica desde el inicio de la evaluación. Al contrario, será definido según los hallazgos preliminares encontrados.
5. **Análisis de la información:** Las nuevas transcripciones de analizarán de manera integrada con toda la información levantada, de acuerdo con los planteamientos metodológicos descritos en las secciones anteriores.

Beneficiarias

Personas beneficiarias con alta participación en el programa: Personas jóvenes, mujeres, personas mayores de 50 años y personas en condición de pobreza.

Primer levantamiento	Se aplicarán entrevistas individuales: dos entrevistas para cada grupo poblacional, buscando la máxima variabilidad entre ellas.
Segundo levantamiento	Se levantarán cuatro entrevistas adicionales con personas cuyas características sean definidas de acuerdo con lo encontrado en la primera etapa. Así, puede ser que cada entrevista adicional sea dirigda a una persona de cada uno de los cuatro grupos pero este no será necesariamente el caso, si se encuentra interés en dos o incluso un grupo específico, se concentrarán estos instrumentos en ellos.

Personas beneficiarias de grupos pobalcionespobalcionespobalcionespoblacionales de interés con baja participación en el Programa: personas con pertenencia étnica, personas con discapacidad, personas migrantes, víctimas del conflicto armado, personas transgénero y habitantes de zonas rurales. Primer levantamiento	Se harán entrevistas grupales para cada población en donde se buscará tener la máxima variación, como se describe en el capítulo 6. Allí se buscará identificar rasgos comunes a su experiencia y la construcción de narrativas colectivas alrededor del mercado laboral y su relación con el programa y el empleo facilitado por él. Dentro de esta etapa se identificarán las características de las personas con las que se quiera profundizar.
Segundo levantamiento	Se harán entrevistas a profundidad en donde se levantará la historia de vida de una persona para cada uno de los grupos definidos según las características identificadas en la etapa anterior. Así mismo, se contemplan dos entrevistas a profundidad adicionales que con otras poblaciones que no se hayan abordado o

[6] Siguiendo lo especificado en el anexo técnico de la evaluación, este ajuste tendrá que ser revisado y aprobado por el CTS, por lo cuálcual se necesitará de una gestión articulada entre el equipo consultor y el Comité para que el ajuste, revisión y aprobación de instrumentos se haga en un máximo de 1 semana y de esta manera no se afecte el desarrollo del trabajo de campo.

	con alguna de este grupo en la que se vea la necesidad de más de una historia de vida para complementar el estudio de caso. La elección de esas dos entrevistas será informada no solamente por las entrevistas grupales sino también, idealmente, por la información de resulte del primer levantamiento con las empresas.

Aliados: Cajas de compensación, Agencia Distrital de Empleo, SENA, bolsas de empleo, Universidades. Para este grupo se revisará con el CTS la pertinencia de dos levantamientos o si se hacen dos grupos focales en un solo levantamiento separando por las características de los aliados, por ejemplo, públicos y privados o fondos de empleo y otros.

Empresas

Primer levantamiento	Se hará un primer levantamiento según lo establecido en el diseño de muestra.
Segundo levantamiento	El segundo levantamiento no presenta diferencias en el tipo de instrumento o actor, se buscarán ajustes en las guías y se aplicarán cuando se consideren pertinentes.

5.2.3 Estudios de caso

Acorde con lo establecido en el Anexo Técnico, para las poblaciones que son de interés y están focalizadas por el programa pero para las que no se ha conseguido una suficiente participación en el mismo, se hará un estudio de caso por grupo poblacional que permita una mejor comprensión de las barreras de acceso específicas que enfrentan así como su experiencia en relación con el mercado laboral y con el Programa, todo ello con miras a recomendar ajustes al programa que permitan aumentar su participación. Así, se harán estudios de caso para las personas con discapacidad, migrantes, con pertenencia étnica, habitantes de la ruralidad, víctimas del conflicto y personas transgénero[7].

Un estudio de caso es una metodología de investigación que se enfoca en el análisis descriptivo y contextualizado de un fenómeno, individuo, grupo, organización o evento específico. Este enfoque permite una comprensión focalizada de la complejidad y contexto del sujeto, en este caso un grupo poblacional, proporcionando información valiosa que puede no ser fácilmente obtenible a través de métodos de investigación más generales. De acuerdo con Robert Stake (1995), un estudio de caso se caracteriza por su particularidad y su capacidad para ofrecer una perspectiva integral del fenómeno estudiado, utilizando múltiples fuentes de evidencia como entrevistas, observaciones, y análisis de documentos. Esta metodología es especialmente útil en situaciones donde es necesario considerar la variabilidad y diversidad de experiencias individuales y grupales dentro de un contexto particular.

[7] Cómo se explica en el capítulo 6 se podrá incluir un grupo poblacional más que sea identificado por el equipo consultor como de especial interés.

El proceso de llevar a cabo un estudio de caso a menudo implica la recopilación de datos a través de diversas técnicas cualitativas, como entrevistas a profundidad y grupos focales. Para esta evaluación, se emplearán entrevistas grupales y entrevistas a profundidad en las que se indagará por historias de vida. Estas técnicas permiten a los investigadores explorar las experiencias, percepciones y motivaciones de los participantes, proporcionando una visión detallada del fenómeno en cuestión. Yin (2018) destaca que la fortaleza de un estudio de caso radica en su capacidad para abordar preguntas de investigación de "cómo" y "por qué", centrándose en fenómenos contemporáneos dentro de su contexto de la vida real. Además, el uso de múltiples fuentes de datos contribuye a la triangulación, aumentando la validez y fiabilidad de los hallazgos.

Un aspecto fundamental de los estudios de caso es su aplicación en la investigación de grupos poblacionales específicos, como es el caso de esta evaluación. Estos estudios permiten a los investigadores comprender las barreras y desafíos únicos que enfrentan, así como los factores que influyen en su éxito o fracaso en determinados contextos, como el mercado laboral o la participación en programas sociales. Merriam (1998) resalta la importancia de considerar el contexto socio-cultural y económico de los participantes, ya que estos factores a menudo juegan un papel crucial en la configuración de sus experiencias y resultados.

Para cada población, la información cualitativa será también complementada por la información cuantitativa que se tenga para estos grupos. Si bien la muestra cuantitativa no es representativa para estos segmentos de población, el análisis descriptivo de la información cuantitativa secundaria del programa será triangulado con la información cualitativa encontrada para complementar y contrastar y así tener una mejor comprensión de cada grupo. Esta comprensión, sin duda, debe ser informada también por la experiencia de las empresas, por lo que será fundamental en los instrumentos dirigidos a las empresas beneficiarias o postulantes preguntar de manera específica por las motivaciones o dificultades que tuvieron para incluir personas con discapacidad, migrantes, víctimas, habitantes rurales, personas trans y con pertenencia étnica. Ello será fundamental para identificar, por ejemplo, si los estereotipos dirigidos a estas poblaciones dentro de las empresas son más agudos y persistentes o si, por ejemplo, las empresas no tienen la capacidad de hacer las adaptaciones necesarias de cara a su inclusión.

Ahora bien, es importante aclarar que los estudios de caso no se presentan como una metodología independiente de las metodologías narrativa y de evaluación realista planteadas en las secciones anteriores de este capítulo. Al contrario, estas dos metodologías guiarán la exploración de cada caso para el entendimiento particular del mismo de cara a las barreras que enfrentan en el mercado laboral y la capacidad del programa para disminuir o no las brechas de inclusión laboral que cada una enfrenta.

Como dijo anteriormente, cada estudio contempla la construcción de una historia de vida, que es una técnica narrativa y por ello responde a lo descrito metodológicamente al inicio de este capítulo. Esta metodología se centra en la reconstrucción detallada de la vida de los participantes a través de sus propios relatos y perspectivas, lo que permite explorar cómo los individuos interpretan y dan sentido a sus experiencias en contextos específicos. Según Denzin y Lincoln (2018), las historias de vida ofrecen una ventana única hacia la subjetividad de los individuos, revelando no sólo sus acciones, sino también sus emociones, pensamientos y motivaciones subyacentes.

Por otro lado, se busca también indagar para cada grupo los mecanismos de Contexto-Mecanismo – Resultado que se plantean dentro de la evaluación realista. Para los estudios de caso, el contexto estará configurado en una buena medida por la pertenencia a ese grupo poblacional específico.

5.2.3 Herramientas de análisis complementario e incorporación del enfoque diferencial en el análisis

Como se menciona en el marco normativo propuesto para la evaluación, los enfoques diferenciales poblacionales y de género así como la perspectiva interseccional tienen un papel dentro de la evaluación teniendo en cuenta que hacen parte también de la formulación del Programa y las poblaciones a las que este se dirigen que están subrepresentadas en el mercado laboral y que han sufrido discriminación y exclusión social de manera histórica. Para la incorporación efectiva de estos enfoques se hace necesario complementar las metodologías cualitativas presentadas con elementos de análisis de investigación feminista y economía feminista. Puntualmente, se incorporará en el diseño de instrumentos, desarrollo de campo y análisis de información elementos del análisis de discurso feminista.

El análisis de discurso feminista es una herramienta cualitativa que examina cómo el lenguaje y las prácticas discursivas reflejan, perpetúan o desafían las desigualdades de género y otras formas de opresión. Según Michelle Lazar (2005) en su libro "Feminist Critical Discourse Analysis: Gender, Power and Ideology in Discourse", esta herramienta busca desentrañar las relaciones de poder y las ideologías de género que se manifiestan en diferentes contextos discursivos, considerando cómo el lenguaje actúa como una herramienta para mantener o desafiar las estructuras de poder.

Una característica clave del análisis de discurso feminista es su enfoque en el contexto, que implica situar los discursos dentro de las estructuras socioculturales y económicas más amplias. Tal como señala Judith Baxter (2003) en "Positioning Gender in Discourse: A Feminist Methodology", el análisis de discurso feminista no solo estudia el contenido del discurso, sino también las interacciones y las posiciones de los hablantes, explorando cómo las identidades de género son negociadas y representadas en diferentes escenarios, como el lugar de trabajo.

Así mismo, el análisis de discurso feminista se interesa por las voces marginadas y busca dar visibilidad a las experiencias y perspectivas de las mujeres y otros grupos subrepresentados.

Concretamente, dentro de la Metodología de análisis narrativo se buscará incorporar una mirada crítica en cómo las relaciones de género o de poder están presentes en las narrativas tanto de las personas como las empresas beneficiarias. Por ejemplo, desde el análisis de discurso feminista, en una entrevista con una joven contratada a través para un empleo formal, se pueden observar las relaciones de género en cómo ella describe su proceso de contratación y las expectativas laborales. La joven podría mencionar sentirse evaluada de manera diferente debido a su género, o cómo las dinámicas de poder en la empresa influyen en su integración. Además, puede expresar si las políticas de inclusión son vistas como genuinas o simbólicas, y cómo esto afecta su percepción de empoderamiento y reconocimiento en el lugar de trabajo. Este análisis revela cómo el lenguaje y las prácticas discursivas reflejan y desafían las desigualdades de género en el ámbito laboral.

La incorporacion de esta herramienta se hará de la siguiente manera:

Aquí tienes los pasos para aplicar el análisis de discurso feminista en entrevistas e historias de vida:

1. Diseño de instrumentos para la recolección de datos:

- Revisión de lenguaje inclusivo dentro de los instrumentos de recolección de información (instrumentos, consentimientos etc).

- Incorporación de preguntas específicas para explorar experiencias de género, poder y agencia.

2. Transcripción y Organización:

- Inclusión de categorías específicas relacionas con género y poder dentro del libro de códigos. Por ejemplo: roles de género, experiencias de discriminación como microagresiones o lenguaje violento, narrativas de empoderamiento en situaciones de discriminación o desventaja, negociación de las necesidades propias en el entorno laboral, conciliación con trabajo de cuidado etc.

- Identificaicón de patrones en el uso del lenguaje que reflejen normas de género, desigualdades de poder o narrativas de resistencia y empoderamiento.

- Identificación de cómo se describen las interacciones de poder y las dinámicas de género en las historias.

4. Análisis Crítico:

- Analisis de cómo las personas entrevistadas narran sus experiencias y cómo el lenguaje refleja las estructuras de poder de género y otras categorías como edad, origen, ocupación etc.

- Análisis de presencia o refuerzo de estereotipos persistentes en las narraciones y su internalización por parte de las poblaciones.

5. Interpretación y Reflexión:

- Interpreta los hallazgos en relación con las teorías de economía feministas, considerando cómo los discursos identificados refuerzan o desafían las normas de género.

- Reflexiona sobre propios sesgos dentro de la investigación y cómo pueden influir en la interpretación de los datos.

6. Presentación de Resultados:

- Comunicación de los hallazgos de manera destacando la diversidad de experiencias.

- Recomendaciones orientadas a la profundización de los enfoques diferenciales en los casos en los que se encuentre pertiente de acuerdo con los resultados.

III. Diseño de muestra

6.Diseño muestral cuantitativo

6.1 Descripción de la estrategia muestral cuantitativa

6.1.1 Características del diseño muestral

Para la elaboración de un muestreo probabilístico es esencial disponer de un marco de muestreo adecuado. En este caso el marco está constituido por un conjunto de datos que incluye a todas las personas vinculadas a empresas a través del Programa Empleo Incluyente. Tanto el universo de estudio como la población objetivo son las personas pertenecientes al Programa.

Con el fin de obtener información relevante sobre esta población se propone una encuesta mediante muestreo probabilístico. El diseño propuesto es un muestreo estratificado, aplicando muestreo aleatorio simple dentro de cada estrato (EST-MAS).

Para la población trans, que cuenta con un tamaño poblacional de 91 individuos, los cuales se encuentran incluidos en el grupo "un grupo poblacional sin incluir mujeres ni jóvenes", se realizará una sobremuestra dirigida a este grupo minoritario, en las siguientes secciones se desarrollarán más los detalles técnicos.

Tanto las unidades de muestreo como unidades de observación, así como las unidades de análisis están representadas por los registros individuales del grupo de personas vinculadas a empresas a través del Programa Empleo Incluyente.

Para la elaboración de este diseño muestral se tomó de base la información del marco muestral anonimizada, provista en el insumo llamado "Postulados-aprobados.xlsx".

La fuente principal de información a utilizar para contacto y recolección de información será el marco muestral de personas beneficiarias del Proyecto de Empleo Incluyente, que será provisto por la SDDE, dado que en el anexo técnico se manifiesta que: "*Los proponentes tendrán a su disposición una base de datos (marco de lista) que contiene información que permite la identificación de las personas beneficiarias de las distintas poblaciones y otras variables para la realización del diseño muestral probabilístico.*"

6.1.2 Proceso de conformación de estratos

Los estratos están dados por las variables "Grupo.poblacional.1" y "Grupo.poblacional.2" del archivo insumo "Postulados-aprobados.xlsx". La decisión frente a esta elección de variable de estratificación ha sido de carácter administrativo. Existe un total de cinco estratos, los cuales se encuentran definidos según los siguientes criterios: i) mujeres, ii) jóvenes de 18 a 28 años, iii) otros pertenecientes a una sola población diferentes a mujeres y jóvenes, iv) personas pertenecientes a dos grupos poblacionales incluyendo pobreza según SISBEN y v) personas

pertenecientes a dos grupos poblacionales, exceptuando pobreza según SISBEN. El cálculo del tamaño de muestra se efectuará al interior de cada uno de estos estratos.

Tabla 6.1. Descripción y tamaño poblacional por estrato

Tipo de población	Total	Porcentaje
1. Una población: Mujer	4076	18.1%
2. Una población: jóvenes de 18 a 28 años	5515	24.5%
3. Una población: Otros (No jóvenes, No mujeres)	1871	8.3%
4. Dos poblaciones: Combinación incluye pobreza	2900	12.9%
Dos poblaciones: combinación no incluye pobreza	8138	36.2%
Total	22500	100%

Fuente: Bases de datos anexas a los pliegos de condiciones

6.1.3 Algoritmos de selección

El algoritmo de selección a utilizar en cada uno de cinco estratos corresponderá al algoritmo de *coordinado negativo*, puesto que, en primer lugar, este algoritmo produce las probabilidades acordes del diseño MAS, y en segundo lugar, este algoritmo arroja un ordenamiento de los elementos de la población lo cual facilita la labor de la selección de los posibles reemplazos. La aplicación de dicho algoritmo en el estrato h ,$(h = 1,2,3,4,5)$, donde se seleccionan n_h de los N_h unidades es como sigue:

1. Para cada elemento de la población del estrato h generar una realización de una variable aleatoria con distribución uniforme sobre el intervalo (0,1).
2. Ordenar la lista de estos N_h elementos de forma descendente (o ascendente) con respecto a los números generados en el paso anterior.
3. Seleccionar los primeros n_h elementos del anterior ordenamiento. Estos elementos corresponden a la muestra aleatoria simple seleccionada en el estrato h.

Teniendo en cuenta que la selección de la muestra depende de la semilla utilizada en el proceso de generación descrito en el paso 1, en caso se ser adjudicado, se acordará la semilla utilizada junto con la Secretaría Distrital de Desarrollo Económico, con el fin de garantizar la transparencia y la replicabilidad de la muestra.

Estos procesos de selección de unidades muestrales se efectuarán con la función *strata* del paquete *sampling*[8] del lenguaje de programación estadístico R.[9] disponible para su descarga y uso en el sitio web de *CRAN-R Project*[10], que es el repositorio público de las bibliotecas del software. En el Apéndice 1 Códigos computacionales en R para la selección de una muestra EST-MAS con el algoritmo coordinado negativo, se encuentran los códigos computacionales donde se toma el insumo “Base_Datos_Postulados-aprobados.xlsx”, se crea la variable estrato y se aplica el algoritmo de coordinado negativo a cada estrato, y como resultado se obtiene

[8] Versión 2.10.
[9] R versión 4.2.3 (2023-03-15).
[10] https://cran.r-project.org/web/packages/available_packages_by_name.html

una base de datos "marco.muestra" que corresponde a la misma base de datos de postulados-aprobados con una variable adicional "en_muestra" cuyo valor es 1 si el individuo es seleccionado en la muestra y 0 si no. De esta forma se lleva a cabo el proceso de selección de la muestra.

6.2 Descripción de la estrategia de estimación de parámetros

6.2.1 Probabilidades de inclusión y factores de expansión

Como el diseño muestral propuesto es un EST-MAS (muestreo estratificado con muestreo aleatorio simple dentro de cada estrato), se tiene que la probabilidad de inclusión π_k para el individuo k perteneciente al estrato h está dada por:

$$\pi_k = \frac{n_h}{N_h}$$

Donde n_h denota el tamaño muestral del estrato h, y N_h denota el tamaño poblacional de este mismo estrato. De lo anterior, se tiene que el factor de expansión general para el individuo k perteneciente al estrato h está dado por:

$$f_k = \frac{1}{\pi_k} = \frac{N_h}{n_h}$$

En la siguiente tabla, se muestran las probabilidades de inclusión y los factores de expansión para los 5 estratos:

Estrato	Tamaño poblacional del estrato (N_h)	Tamaño muestral del estrato (n_h)	Probabilidad de inclusión para cualquier elemento del estrato	Factor de expansión para cualquier elemento del estrato
Una Población: Mujer	4076	259	$\frac{259}{4076}$	$\frac{4076}{259}$
Una Población: Joven de 18 a 28 años	5515	263	$\frac{263}{5515}$	$\frac{5515}{263}$
Una Población: Otros (NO jóvenes. NO mujeres)	1871	241	$\frac{241}{1871}$	$\frac{1871}{241}$
Dos Poblaciones: Combinación Incluye Pobreza SISBEN	2900	252	$\frac{252}{2900}$	$\frac{2900}{252}$
Dos Poblaciones: Combinación No Incluye Pobreza SISBEN	8138	267	$\frac{267}{8138}$	$\frac{8138}{267}$

Tabla 5.2. Probabilidades de inclusión y factores de expansión para cada estrato.

6.2.2 Estimadores de parámetros de caracterización

Los potenciales indicadores claves que resuelven el objetivo general de la evaluación son de la forma de proporciones, totales, promedios y razones poblacionales, que son definidos a continuación.

- ***Proporción poblacional:***

La proporción de la población de estudio para cierta característica de interés está expresada como:

$$P = \frac{1}{N}\sum_{k\in U} y_k$$

donde U denota la población de estudio, y_k vale 1 si el individuo k tiene cierta característica y vale 0 si no.

- ***Total poblacional:***

El total de una variable de interés en la población de estudio está expresado como:

$$t_y = \sum_{k\in U} y_k$$

donde y_k denota el valor de la variable de interés y observada en el individuo k.

Promedio poblacional:

El valor promedio de una variable de interés en la población de estudio está expresado como:

$$\bar{Y} = \frac{1}{N}\sum_{k\in U} y_k$$

donde U denota la población de estudio, y_k denota el valor de la variable de interés y observada en el individuo k y N denota el tamaño poblacional.

- ***Razón poblacional:***

La razón poblacional es la comparación entre dos totales poblacionales por medio de un cociente, y está expresada como:

$R = \frac{t_y}{t_z} = \frac{\sum_{k\in U} y_k}{\sum_{k\in U} z_k}$donde U denota la población de estudio, y_k y z_k denota el valor de la variable de interés y y z observada en el individuo k. Un caso particular de una razón poblacional son los promedios y proporciones de tal forma que los ejemplos de preguntas orientadoras citados anteriormente corresponderían a la estimación de razones poblacionales.

- ***Diferencia de dos promedios***

Como herramienta para la cuantificación de efecto, se considera la diferencia de dos medias poblacionales definidas como:

$$\bar{Y}_1 - \bar{Y}_2 = \frac{1}{N}\sum_{k\in U} y_{1,k} - \frac{1}{N}\sum_{k\in U} y_{2,k}$$

Donde $y_{1,k}$ y $y_{2,k}$ denotan el valor de dos variables y_1 y y_2 sobre el individuo k, U denota la población de estudio y N denota el tamaño poblacional. Un ejemplo de un parámetro de esta naturaleza puede ser:

- Diferencia entre el salario promedio antes y después del programa.

Tenga en cuenta que en la anterior expresión, cuando las variables $y_{1,k}$ y $y_{2,k}$ corresponde a una variable indicadora para un mismo fenómeno de interés antes y después del programa, la diferencia de dos promedios se convierte en una diferencia de proporciones.

Ahora, uno de los dos términos que aparecen en la diferencia de promedios será conocido ya que en el estudio solo se llevará a cabo un levantamiento de información, en el caso de estimar la diferencia entre el salario promedio antes y después del programa, el valor del salario promedio antes del programa se obtendrá cruzando el marco muestral de personas con la información de PILA y corresponderá a una valor conocido.

Otro tipo de diferencia de promedios que es de interés estimar corresponde a la diferencia en el valor promedio entre dos subgrupos poblacionales. Suponga que U_1 y U_2 son dos subconjuntos del universo U, por ejemplo U_1 y U_2 pueden corresponder al subgrupo **Mujeres (una población)** y **Jóvenes de 18 a 28 años** (una población) indicados en la Tabla 4 del anexo técnico, y es de interés conocer la diferencia de salario promedio entre estos dos grupos, el parámetro de interés está dado por:

$$\Delta = \frac{1}{N_1}\sum_{k\in U_1} y_k - \frac{1}{N_2}\sum_{k\in U_2} y_k$$

- Donde N_1 y N_2 son los tamaños poblacionales conocidos de U_1 y U_2, y y_k es el valor de la característica de interés en el individuo k.***Coeficientes de regresión***

  En la fórmula del modelo de regresión se definen los parámetros de interés en la estimación de coeficientes de regresión como $\beta_1, \beta_2, \ldots, \beta_q$, en un modelo de la forma:

$$y_k = \beta_1 z_{1k} + \beta_2 z_{2k} + \ldots + \beta_q z_{qk} + e_k,$$

  Si se considera la regresión con intercepto $z_{1k} = 1$ para todo $k \in U$.

La característica poblacional para estimar es el vector de coeficientes de regresión $\beta = (\beta_1, \beta_2, \ldots, \beta_Q)'$.A continuación, se presentan las expresiones teóricas de los estimadores de los parámetros poblacionales:

- ***Estimador de un total poblacional:***

El estimador insesgado de mayor popularidad para un total poblacional es el estimador de Horvitz-Thompson, dada por:

$$\hat{t}_{y,\pi} = \sum_{k \in s} \frac{y_k}{\pi_k}$$

donde s denota la muestra probabilística y π_k denota la probabilidad de inclusión del individuo k. La ventaja de este estimador insesgado frente a otros estimadores insesgados es que es el único estimador hiperadmisible en la clase de los estimadores insesgados generalizados.

Teniendo en cuenta el diseño EST-MAS de esta propuesta, el estimador Horvitz-Thompson de un total poblacional queda expresado como:

$$\hat{t}_{y,\pi} = \sum_{h=1}^{5} \sum_{k \in s_h} \frac{y_k}{\pi_k} = \sum_{h=1}^{5} \frac{N_h}{n_h} \sum_{k \in s_h} y_k$$

Donde s_h denota la muestra seleccionada del estrato h, n_h denota el tamaño muestral del estrato h, y N_h denota el tamaño poblacional de este mismo estrato.

- ***Estimador de una proporción poblacional:***

El estimador de una proporción P está dado por

$$\hat{P} = \frac{1}{N} \sum_{k \in s} \frac{y_k}{\pi_k}$$

donde y_k es la variable indicadora para la presencia de la característica de interés en el k-ésimo individuo.

Teniendo en cuenta el diseño EST-MAS de esta propuesta, el estimador Horvitz-Thompson de un total poblacional queda expresado como:

$$\hat{P} = \frac{1}{N} \sum_{k \in s} \frac{y_k}{\pi_k} = \frac{1}{N} \sum_{h=1}^{5} \sum_{k \in s_h} \frac{y_k}{\pi_k} = \frac{1}{N} \sum_{h=1}^{5} \frac{N_h}{n_h} \sum_{k \in s_h} y_k$$

Donde s_h denota la muestra seleccionada del estrato h, n_h denota el tamaño muestral del estrato h, N_h denota el tamaño poblacional de este mismo estrato y N el tamaño poblacional.

- ***Estimador de un promedio poblacional:***

En cuanto a la estimación de un promedio poblacional, cuando el tamaño poblacional N es conocido, el estimador insesgado del promedio $\bar{Y}$ viene dado por

$$\hat{\bar{Y}} = \frac{1}{N} \hat{t}_{y,\pi}$$

donde $\hat{t}_{y,\pi}$ denota el estimador de Horvitz-Thompson del total poblacional t_y presentado anteriormente.

- ***Estimador de una razón poblacional:***

El estimador de uso popular para la razón poblacional $R = \frac{t_y}{t_z}$ está dada por

$$\hat{R} = \frac{\hat{t}_{y,\pi}}{\hat{t}_{z,\pi}} = \frac{\sum_{k\in s} y_k/\pi_k}{\sum_{k\in s} z_k/\pi_k} = \frac{\sum_{h=1}^{5} \frac{N_h}{n_h} \sum_{k\in s_h} y_k}{\sum_{h=1}^{5} \frac{N_h}{n_h} \sum_{k\in s_h} z_k}$$

Si bien este estimador es sesgado, por medio de desarrollos de linealización de Taylor, se comprueba que $\hat{R}$ es un estimador aproximadamente insesgado para R.

- ***Estimador de una diferencia de dos promedios***

El estimador insesgado para una diferencia de promedios corresponde a la diferencia entre el estimador de Horvitz-Thompson para cada uno de los promedios, el cual está dado por:

$$\bar{Y}_1 - \widehat{\bar{Y}}_2 = \frac{1}{N}\sum_{k\in U} y_{1,k} - \frac{1}{N}\hat{t}_{y2,\pi}$$

Donde $\hat{t}_{y2,\pi}$ corresponde al estimador de Horvitz-Thompson del total poblacional de la variables y_2, tal como se definió anteriormente. Ahora, el primer término $\frac{1}{N}\sum_{k\in U} y_{1,k}$ no será estimado ya que se obtendrá para todas las personas del marco muestral a partir de informaciones auxiliares o registros administrativos tal como la PILA.

En cuanto a la estimación de la siguiente diferencia de promedios entre dos subgrupos poblacionales (estratos):

$$\Delta = \frac{1}{N_1}\sum_{k\in U_1} y_k - \frac{1}{N_2}\sum_{k\in U_2} y_k$$

Considerando que el diseño muestral dentro de cada estrato es un muestreo aleatorio simple, un estimador insesgado para Δ está dado por

$$\hat{\Delta} = \frac{1}{n_1}\sum_{k\in s_1} y_k - \frac{1}{n_2}\sum_{k\in s_2} y_k$$

donde s_1 y s_2 denotan las muestras de los dos estratos, y n_1 y n_2 los tamaños muestrales respectivos.

- ***Estimación de coeficientes de regresión***

Para la estimación de los coeficientes de regresión definido previamente como $\beta = (\beta_1, \beta_2, \ldots, \beta_q)'$. Se utiliza la expresión de (Särndal, Swensson, & Wretman, 1992, pág. 193) definida como:

$$\hat{\beta} = (Z_s \Pi_s^{-1} Z_s^T)^{-1} Z_s \Pi_s^{-1} y_s$$

Donde,

- Z_s es la matriz de todas las columnas de covariables del modelo en la muestra s

- Π_s es una matriz diagonal cuya dimensión corresponde al tamaño de la muestra sobre la población de interés, donde la diagonal contiene a las probabilidades de inclusión de los elementos $k \in s$
- y_s es el vector con los valores observados en la muestra para la variable de interés.

6.3 Medidas de calidad de parámetros de caracterización

6.3.1 Estimación de la varianza

La estimación de la varianza para los diferentes estimadores como total, promedio, proporción, razón, diferencia de promedios, coeficientes de regresión, entre otros, bajo el diseño muestral propuesto puede ser complejo a través de fórmulas explícitas, teniendo en cuenta esto se suelen usar metodologías de remuestreo y permutaciones como Jackknife o Bootstrap para estimar la varianza. Valliant y Dever (2017) exponen que Bootstrap es más eficiente que Jackknife para la estimación de parámetros lineales y no lineales, Cepal (2023) presenta la metodología Bootstrap como el método basado en réplicas más versátil para el cálculo de errores estándar en muestras complejas. A continuación, se presenta la estimación de la varianza del estimador $\hat{\theta}$, donde θ corresponde a los diferentes parámetros a estimar presentados en el capítulo anterior.

Jackknife

Para la aplicación de esta metodología se tienen los siguientes pasos:

1. Se divide la muestra de tamaño n en A grupos de igual tamaño de la siguiente manera $m = n/A$.
2. Para cada grupo a, con $a = 1, 2, \dots, A$, se calcula el estimador $\hat{\theta}_{(a)}$ del parámetro de interés θ luego de eliminar los elementos muestrales que pertenecen al grupo a.
3. Para cada grupo a se define un pseudovalor de θ como $\hat{\theta}_a = A\hat{\theta} - (A-1)\hat{\theta}_{(a)}$ donde $\hat{\theta}$ es el estimador calculado con la muestra completa.
4. Luego de calcular el estimador $\hat{\theta}_{(a)}$ y el pseudovalor $\hat{\theta}_a$ para cada grupo a, el estimador Jackknife se calcula como

$$\hat{\theta}_{JK} = \frac{1}{A}\sum_{a=1}^{A} \hat{\theta}_a$$

5. El estimador de la varianza se expresa como

$$\widehat{VAR}_{JK}(\hat{\theta}) = \frac{1}{A(A-1)}\sum_{a=1}^{A}\left(\hat{\theta}_a - \hat{\theta}_{JK}\right)^2$$

Bootstrap

Para aplicar el método Bootstrap se plantea un proceso en bucle con $l = 1, 2, \dots, L$ iteraciones. En cada iteración se tienen los siguientes pasos:

1. Tomar una muestra aleatoria simple con reemplazamiento de la muestra completa s, obteniendo de esta forma la submuestra Bootstrap s_{BS}.
2. Calcular el estimador Bootstrap $\hat{\theta}_{BS}$ de manera análoga como se calcula el estimador con la muestra completa, sin embargo, en este caso se usa la submuestra s_{BS}.

Luego de ejecutar las L iteraciones, la varianza Bootstrap queda definida como:

$$\widehat{VAR}_{BS}(\hat{\theta}) = \frac{1}{L}\sum_{l=1}^{L}\left(\hat{\theta}_{BS} - \hat{\theta}\right)^2.$$

De igual modo, paquetes estadísticos como R tienen librerías que permiten estimar la varianza por estos métodos de remuestreo y usan correcciones basadas en aproximaciones asintóticas de matrices de varianzas y de covarianzas. En la literatura se encuentran diferentes variaciones de los métodos descritos, teniendo en cuenta esto y el comportamiento de los datos muestrales, al finalizar la recolección de la muestra se evaluará cuál es el mejor método de estimación.

6.3.2. Simulación de escenarios

En la siguiente tabla se muestran los resultados de simulación de escenarios para la estimación de una proporción poblacional considerando diferentes valores de la proporción en los diferentes estratos. Los valores de p_i denotan diferentes valores del parámetro poblacional, de manera que determinan una amplia gama de 10 diferentes escenarios posibles; la columna me indica los valores del margen de error global para cada uno de los escenarios planteados; por su parte los valores en las columnas me_i denotan el margen de error al interior de cada estrato. Se evidencia que, tanto a nivel global como a nivel de los estratos, el margen de error toma valores pequeños, inferiores al valor máximo de 7% requerido por el anexo técnico.

p_1	p_2	p_3	p_4	p_5	me	me_1	me_2	me_3	me_4	me_5
0,1	0,4	0,6	0,6	0,2	2,47	3,5	5,8	5,8	5,8	4,7
0,3	0,2	0,1	0,8	0,6	2,67	5,4	4,7	3,5	4,7	5,8
0,3	0,8	0,6	0,7	0,3	2,61	5,4	4,7	5,8	5,4	5,4
0,5	0,8	0,3	0,5	0,5	2,8	5,9	4,7	5,4	5,9	5,9
0,5	0,2	0,9	0,4	0,7	2,63	5,9	4,7	3,5	5,8	5,4
0,5	0,5	0,5	0,5	0,5	2,94	5,9	5,9	5,9	5,9	5,9
0,8	0,6	0,8	0,5	0,5	2,82	4,7	5,8	4,7	5,9	5,9
0,8	0,2	0,7	0,9	0,1	2,04	4,7	4,7	5,4	3,5	3,5
0,9	0,2	0,3	0,7	0,2	2,31	3,5	4,7	5,4	5,4	4,7
0,9	0,6	0,4	0,6	0,9	2,2	3,5	5,8	5,8	5,8	3,5

Tabla 5.3. Margen de error (%) global y para cada estrato para diferentes escenarios simulados.

Los códigos utilizados para obtener la anterior tabla se encuentran en el Apéndice 2 Códigos computacionales en R para el cálculo del error estándar y margen de error para diferentes valores de proporciones.

6.4 Determinación del tamaño de muestra

6.4.1 Expresión matemática para el cálculo del tamaño de muestra

Teniendo en cuenta que el diseño muestral propuesto es un diseño estratificado con muestreo aleatorio simple dentro de cada estrato, e tamaño de muestra se determina de la siguiente manera: (1) se determina el tamaño muestral n_h para cada estrato h, y (2) se calcula el tamaño muestral como $n = \sum_h n_h$.

Dentro de cada estrato, se busca garantizar que el estimador de un porcentaje poblacional P_h esté cercano al parámetro poblacional con un nivel de confianza de $1-\alpha$ y un determinado margen de error denotado por me. Esto es, se debe garantizar que $1-\alpha \leq \Pr\left(\left|\widehat{P_h} - P_h\right| < me\right)$. Desarrollos algebráicos arrojan que el tamaño muestral n_h debe cumplir la siguiente condición:

$$n_h \geq \frac{\frac{z_{1-\frac{\alpha}{2}}^2}{me^2} * P_h * (1-P_h) * Deff}{1 + \frac{z_{1-\frac{\alpha}{2}}^2}{N_h * me^2} * P_h * (1-P_h) * Deff}$$

Donde

- N_h es el tamaño poblacional para el estrato h.
- $z_{1-\frac{\alpha}{2}}$ denota el percentil $1-\frac{\alpha}{2}$ de la distribución normal estándar (para esta evaluación se toma $\alpha = 5\%$ equivalente a un nivel de confianza del 95%, es decir, $z_{1-\frac{\alpha}{2}} = 1.96$)
- P_h es la proporción poblacional la cual se fija en el valor de 0.5 pues este maximiza la varianza del estimador de P_h, con el cual se logra mayor tamaño de muestra, y se garantiza el margen de error mínimo para diferentes indicadores del estudio. En este caso se asume el escenario más conservador de la varianza de P_h para cada estrato.
- $Deff$ denota el efecto del diseño definido como el cociente de la varianza del estimador de P_h bajo el diseño muestral propuesto y el diseño de muestreo aleatorio simple. Dado que el muestreo en cada estrato es el muestreo aleatorio simple, se tiene que $Deff = 1$ para todos los estratos.
- me denota el margen de error absoluto que para esta evaluación se toma el valor del 5.9%, inferior al valor máximo de 7% especificado por el anexo técnico, y con el cual se obtiene tamaños muestrales superiores a los exigidos por el anexo técnico.
- El tamaño total de muestra n se calcula al sumar los tamaños de muestra calculados en cada estrato: $n = \sum_{h=1}^{H} n_h$.

6.4.2 Cálculo del tamaño de muestra

En la siguiente tabla se presentan los resultados de aplicación de la fórmula presentada en la sección anterior con los tamaños poblacionales presentados en la Tabla 5.1.

Grupo poblacional	Una Población: Mujer	Una Población: Joven de 18 a 28 años	Una Población: Otros (NO jóvenes. NO mujeres)	Dos Poblaciones: Combinación Incluye Pobreza SISBEN	Dos Poblaciones: Combinación No Incluye Pobreza SISBEN	Tamaño muestran total
Tamaño muestral	259	263	241	252	267	**1282**

Tabla 5.4. Tamaño de muestra para los 5 estratos.

Para el cálculo de estos tamaños de muestra se usó el siguiente *script* de R. Los códigos correspondientes se encuentran en el Apéndice 3 Códigos computacionales en R para calcular el tamaño muestral total y por estrato. Dado que estos tamaños muestrales fueron obtenidos usando el valor de 5.9% del margen de error, se puede concluir que estos tamaños cumplen con las medidas de precisión y calidad establecidas en el anexo técnico.

6.5 Tratamiento de la población TRANS

De acuerdo con la información suministrada en el anexo técnico, se desea contar con resultados estadísticamente válidos para la población trans que tiene un tamaño poblacional de 91 personas, y todas ellas están incluidas en el grupo poblacional "Una (1) Población: Otros (NO jóvenes. NO mujeres)". Teniendo en cuenta que en el anexo técnico se decidió no agregar un grupo específico para esta subpoblación, al seleccionar la muestra siguiendo el diseño muestral EST-MAS con los 5 estratos definidos en la tabla 5 del anexo técnico, el número de personas de la población TRANS seleccionado en la muestra no es fijo. Se realizó un estudio de simulación, que consistió en recrear el universo de 1871 personas para el estrato de "Una (1) Población: Otros (NO jóvenes. NO mujeres)" donde 91 de ellas pertenecen a la población TRANS, y se llevó a cabo un muestreo aleatorio de tamaño de 241. El proceso anterior se repitió 1000 veces y se registró el número de personas TRANS que quedaron seleccionadas en la muestra. Se encontró que de estas 1000 simulaciones, mínimo se seleccionaron 3 personas TRANS (un 3% de la población TRANS), máximo 24 personas (un 26% de la población TRANS), y el número promedio de personas TRANS seleccionadas corresponden a 12 (un 13% de la población TRANS). Los códigos computacionales en el *software* R se encuentran en el "Apéndice 4 Códigos computacionales en R relacionados con el tamaño muestral de la población TRANS".

Para garantizar obtener resultados estadísticamente válidos para la población TRANS, se propone que se logren realizar 30 encuestas en esta población, ya que tal como se muestra en

el mismo apéndice 4, el margen de error para dicha población con 30 encuestas es de 14.65%. Si bien es superior al margen de error de 7% requerido para los 5 grupos poblacionales, es un valor aceptable teniendo en cuenta el tamaño reducido de la población TRANS, con lo cual se pueden obtener resultados estadísticamente válidos.

Ahora, en la simulación de 1000 iteraciones, se encontró que en ninguna de estas iteraciones se logra un tamaño muestral de 30 para la población TRANS, por lo cual se propone realizar un sobre muestreo dirigido específicamente a esta población de la siguiente forma: si en el muestreo aleatorio simple en el grupo poblacional "Una (1) Población: Otros (NO jóvenes. NO mujeres)" el número de personas de la población TRANS iguala o supera a 30, entonces no es necesario realizar ningún procedimiento adicional; de lo contrario, suponga que si n_t personas TRANS quedan seleccionadas con $n_t < 30$, entonces se debe realizar un sobre muestreo seleccionando con un muestrea de tamaño $30 - n_t$ de la población TRANS que no fueron seleccionados, y de esta forma se completan las 30 encuestas de dicha población.

6.6 Mitigación del efecto de la no respuesta no ignorable

Selección de reemplazos

En caso de necesitar seleccionar reemplazos para las personas, se utiliza el ordenamiento arrojado por el algoritmo de coordinado negativo del estrato correspondiente, seleccionando el número de reemplazos necesarios como los elementos que siguen inmediatamente después del último elemento seleccionado en la muestra inicial.

Tratamiento de ausencia de respuesta no ignorable y ajuste de factor de expansión

Generalmente en este tipo de estudios se presentan inconvenientes en cuanto a conseguir una tasa de respuesta adecuada por parte de las unidades de muestreo o, como es en este caso, en cuanto a la no respuesta por parte de los encuestados. Dado que cada unidad seleccionada tiene un factor de expansión, se deben realizar ajustes para reducir sesgos por no-respuesta. Para el ajuste de los factores de expansión, se seguirá el procedimiento establecido en (Valliant, Dever, & Kreuter, 2018, p. 324) presentado en la Figura 5.2.

Al final de la recolección se tendrán registradas las incidencias de cada registro seleccionado con el fin de llevar a cabo el proceso de ajuste de los factores. De acuerdo con Valliant y Dever (2017), las unidades de observación se dividen según la incidencia presentada, y se codifican así:

- ER: elegibles respondientes, hace referencia a las encuestas efectivas.
- ENR: elegibles no respondientes, hace referencia a las unidades que deciden no responder la encuesta.
- NEC: no elegible conocidos, es decir, elementos en el marco muestral que no cumplen con las características de la población objetivo y que lograron ser contactados.
- ED: elegibilidad desconocida, son los elementos que no logran ser contactados durante el trabajo de campo.

Al tener la codificación de cada una de las unidades seleccionadas según la clasificación expuesta (ER, ENR, ENC, ED), el siguiente paso es calcular el factor de expansión básico f_k, definido anteriormente. Luego se procede a realizar un ajuste a los factores de expansión por la no respuesta no-ignorable, el procedimiento se expone a continuación.

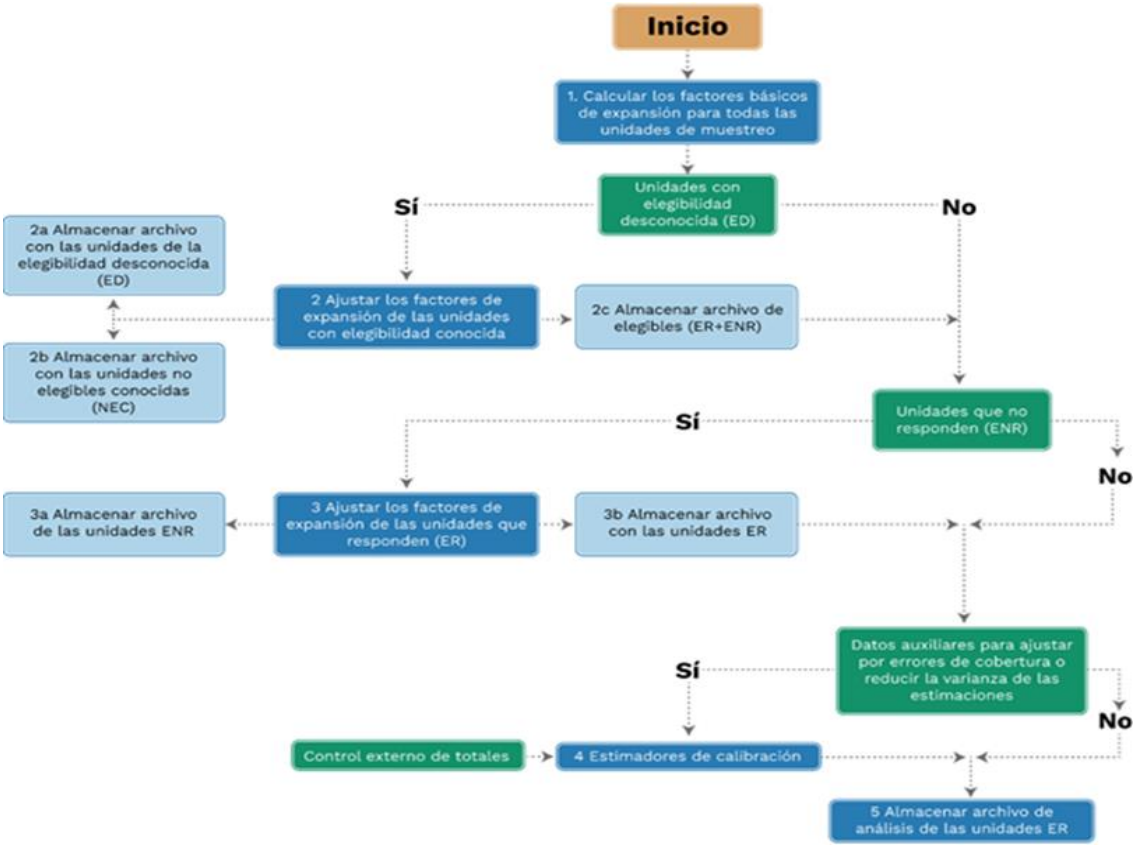


Figura 5.2. Flujo de tratamiento de no respuesta no ignorable y ajuste de factores de expansión.

Ajuste por elegibilidad desconocida

Para el ajuste por no elegibilidad se propone usar el ajuste propuesto por Valliant, Dever y Kreuter (2018), en este caso se define las siguientes clases:

- Sea S_b:$\{ER, ENR, NEC, ED\}$ el conjunto de los subconjuntos de las clases de incidencias que se presentaron.
- $S_{b_{KN}}$:$\{ER, ENR, NEC\}$ es el conjunto de los subconjuntos de elementos con elegibilidad conocida

El ajuste por elegibilidad desconocida es:

$a_1 = \frac{\sum_{k \in s_b} f_k}{\sum_{k \in s_{b_{KN}}} f_k}$, donde $f_k = \frac{N_h}{n_h}$, si el elemento k pertenece al estrato h.

Las unidades no elegibles son descartadas y se aplica el factor de ajuste al resto de unidades, para ello el factor de expansión se ajusta como:

$$f_{1k} = \begin{cases} a_1 f_k & si\ k \in S_{b_{KN}} \\ 0 & si\ no \end{cases}$$

Ajuste por no respuesta

El ajuste por no respuesta no ignorable se puede realizar clasificando los individuos en clases, dada alguna variable auxiliar que se conozca para todos los individuos seleccionados, se define el factor de ajuste para cada clase c como:

$$a_{2,c} = \frac{\sum_{k \in \{ER,ENR\}} f_{1k}}{\sum_{k \in \{ER\}} f_{1k}}$$

El factor de expansión para las unidades respondientes ajustado por la no respuesta se define como $f_{2,k} = a_{2c} f_{1k}$

Otra forma de realizar el ajuste es por medio de la estimación de la probabilidad de respuesta de cada uno de los elementos seleccionados, teniendo en cuenta la información auxiliar que se tenga para cada unidad. Gutiérrez & Rojas (2013) sugiere ajustar la forma funcional del estimador utilizando la metodología de propensity score. El procedimiento es: se denota la muestra de los respondientes como s_A, para cada individuo $k \in s$ se observa la variable R_k (la variable indicadora para el evento "individuo k es respondiente"), así R_k tiene distribución Bernoulli con parámetro $\varphi_k = \Pr(k \in s_A)$. Por otro lado, se debe contar con un vector de información auxiliar conocido $\boldsymbol{z}_k$ para todo $k \in s$; así, se puede estimar φ_k por medio de una regresión logística, esto es,

$$\widehat{\varphi_k} = \frac{e^{\boldsymbol{z}_k{}'\widehat{\boldsymbol{\beta}}}}{1 + e^{\boldsymbol{z}_k{}'\widehat{\boldsymbol{\beta}}}}$$

donde $\boldsymbol{\beta}$ es el vector de coeficientes de la regresión logística. Finalmente, el estimador de Horvitz-Thompson ajustado queda expresado como

$$\hat{t}^a_{y,\pi} = \sum_{k \in s_A} \frac{y_k}{\pi_k \widehat{\varphi_k}}$$

La varianza de este estimador (ver Jae Kwang Kim & Minsun Kim Riddles (2012)) está dada por

$$\hat{V}\left(\hat{t}^a_{y,\pi}\right) = \frac{1}{N^2} \sum_{k \in s_A} \frac{(1 - \widehat{\varphi_k})}{\pi_k \widehat{\varphi_k}^2} (y_k - \widehat{\varphi_k} \boldsymbol{z}_k{}' \widehat{\boldsymbol{\gamma}})^2$$

Y los estimadores para una proporción, un promedio o una razón se define en términos de $\hat{t}^a_{y,\pi}$. Dado lo anterior, es claro que el factor de expansión ajustado por la ausencia de respuesta viene dado por

$$f_{2k} = \frac{f_{1k}}{\widehat{\varphi_k}}$$

para todo $k \in s_A$.

Calibración con información auxiliar

Teniendo en cuenta el diseño muestral EST-MAS, es claro que la suma de los factores de expansión teóricos f_k reproducen los tamaños poblacionales tanto general como por estrato. Sin embargo, esto no se cumple para los factores ajustados por no respuesta f_k^a, por lo que se propone realizar los ajustes por calibración recomendados por Särndal y Lundström (2005) quienes afirman que "cuando los estudios por muestreo están afectados por la falta de respuesta, es deseable que la estructura inferencial que sustenta la encuesta favorezca estimadores con sesgo pequeño o nulo y con errores estándar pequeños" (ver CEPAL (2023)). La calibración busca que los factores de expansión reproduzcan con exactitud totales poblacionales para un conjunto de variables de calibración. Específicamente busca que los factores calibrados f_k^c cumplan con la restricción $\sum_{k\in s} f_k^c x_k = t_x$ donde x_k es el vector de variables de calibración en el individuo k y t_x el vector de totales poblacionales de las variables de calibración. Cuando se definen las variables de calibración como variables que indican la pertenencia de los individuos a los diferentes estratos, se logra que la suma de los factores de expansión calibrados reproduzcan los tamaños poblacionales de los estratos. El proceso computacional para llevar cabo la calibración se puede llevar a cabo en el paquete *survey* usando la función *calibrate* en el software *R*.

Con el fin de monitorear y controlar la representatividad de la muestra, se propone usar los indicadores de representatividad (ver Barry Schouten, Natalie Shlomo & Chris Skinner (2011)) que miden hasta qué punto la composición de la respuesta a una encuesta se desvía de la muestra original. En particular, el indicador de calidad R se puede estimar por medio de:

$$\hat{R} = 1 - 2\sqrt{\frac{1}{N-1}\sum_{k=1}^{n}\frac{(\hat{\varphi}_k - \hat{\bar{\varphi}})^2}{\pi_k}}$$

Donde $\hat{\varphi}_k$ denota la probabilidad estimada de que el k-ésimo individuo sea respondiente, y $\hat{\bar{\varphi}} = \frac{1}{N}\sum_{k=1}^{n}\frac{\hat{\varphi}_k}{\pi_k}$. El valor del indicador $\hat{R}$ se encuentra entre 0 y 1, siendo que el valor 1 indica altos niveles de representatividad, mientras que menor es el valor, más se desvía la composición de la respuesta de la composición de la muestra (ver CEPAL (2023)).

Estimación en muestras cuasi - probabilísticas

Cuando la tasa de reemplazo llega a valores mayores al 200% se puede caer en un muestreo inducido por lo tanto no es posible garantizar la aleatoriedad de la muestra, esta problemática ha ido aumentando en los últimos años por lo que instituciones oficiales de estadística como la CEPAL han optado por cambiar los métodos tradicionales de estimación basados en el diseño e incorporar nuevas metodologías usando información auxiliar que permitan realizar una inferencia válida cuando se tienen tasas de reemplazo altas o incluso cuando la muestra no es probabilística.

Rao en 2021, recopiló diferentes métodos para realizar inferencia válida integrando encuestas e información auxiliar, por medio de la inferencia basada en el modelo. Haremos énfasis en la

metodología de estimadores doblemente robustos expuesta en Chen, Li y Wu (2020)[11], se presenta el detalle de esta técnica en el anexo 8 estimación doblemente robusta.

7. Diseño muestral cualitativo

7.1 Descripción de estrategia y tipologías de muestreo

Como se ha descrito en las secciones metodológicas, en el desarrollo de la evaluación el levantamiento de información cualitativa aportará una visión contextualizada que permitirá profundizar sobre los resultados del PEI, posibilitando también una comprensión descriptiva de los mecanismos de trasmisión de los mismos. Para ello, la participación de actores clave, tanto del lado de la oferta como de la demanda laboral, es esencial para capturar una amplia gama de experiencias y comprender cómo operan los mecanismos de inclusión laboral en diversos entornos.

La estrategia de muestreo cualitativo desarrollada aquí tiene como objetivo explorar de manera detallada las vivencias de las personas beneficiarias y otros actores clave en su participación en el Programa. Este enfoque utiliza un muestreo intencionado, donde se seleccionan deliberadamente casos específicos que abordan estratégicamente los objetivos de la evaluación, como indica Patton (2015). A través de este muestreo, se describe y analiza las experiencias de grupos con características comunes, comparándolas y contrastándolas con las de otros grupos. Este análisis no solo se enfoca en las experiencias individuales, sino que busca un entendimiento más amplio y sistemático de los diferentes actores involucrados. Aunque este muestreo no pretende ser estadísticamente representativo, su objetivo es seleccionar casos que ofrezcan una visión diversa y contrastada, permitiendo obtener resultados más integrales y significativos. Este tipo de muestreo se alinea con metodologías que capturan la complejidad contextual y la diversidad de efectos de una intervención, como la Evaluación Realista y el Análisis Narrativo que se desarrollarán en el estudio.

Como complemento al muestreo intencionado descrito se adoptará también un enfoque de muestreo realista, alineado con el enfoque metodológico de Evaluación Realista que sustenta el análisis cualitativo de los resultados del PEI. En el muestreo realista,

> *la explicación e interpretación de datos empíricos y casos seleccionados prueban y refinan la teoría. Las elecciones de muestreo buscan ejemplos de mecanismos en acción para poder decir algo explicativo sobre sus capacidades causales. Así, el muestreo es tanto pre-especificado como emergente, y se impulsa a través del compromiso con lo que ya se conoce sobre lo que se está investigando y las ideas catalizadas a través del compromiso con relatos empíricos (Emmel, 2005).*

El muestreo realista, como lo indica Emmel (2005), permite seleccionar casos que profundizan en las capacidades causales del programa, probando y refinando hipótesis de la Teoría de Cambio mediante un enfoque flexible y emergente, lo que favorece la evaluación del PEI. La recolección cualitativa se realiza en dos etapas: inicialmente, se busca la máxima variabilidad de casos para probar hipótesis

[11] Chen, Y., Li, P., & Wu, C. (2020). Doubly robust inference with nonprobability survey samples. Journal of the American Statistical Association, 115(532), 2011-2021.

teóricas; luego, se ajustan estas hipótesis con datos empíricos, permitiendo un análisis detallado de experiencias y resultados diferenciales.

El proceso iterativo de recolección de información se desarrollará en dos etapas:

6. **Primer levantamiento:** Diseño de instrumentos acorde con las preguntas orientadoras e hipótesis formuladas en la teoría de cambio.
7. **Análisis, ajuste de hipótesis e identificación factores a profundizar:** Se adelantarán el análisis automatizado de texto para obtener hallazgos preliminares y estos serán revisados en una sesión de trabajo con los profesionales que participaron del levantamiento. Con ello, se identificará si hay alguna indagación que sea pertinente incluir o algún aspecto a profundizar y las características puntuales de los actores con quienes sea relevante hacer el segundo levantamiento[12].
8. **Ajuste de instrumentos:** Con los hallazgos preliminares se ajustarán los instrumentos[13].
9. **Segundo levantamiento:** De acuerdo con el diseño muestral realista, la elección de los actores con quienes ser hará este segundo levantamiento no se especifica desde el inicio de la evaluación. Al contrario, será definido según los hallazgos preliminares encontrados.

Por su parte, el análisis narrativo se hace en paralelo aprovechando la segunda recolección para profundizar sobre los cambios culturales dentro de las empresas y las percepciones de los trabajadores sobre su inclusión.

Considerando estos enfoques intencionado y realista, para la selección de las fuentes de información de los instrumentos cualitativos se implementarán seis tipos de criterios muestrales:

De inclusión forzada:	Población específicada en el Anexo Técnico de los Términos de Referencia.
Por criterio de pertinencia	Esta tipología refiere actores que cumplen perfiles determinados y están preestablecidos según la relevancia estratégica de su rol para la evaluación. Todas las unidades seleccionadas deben reunir una o varias características establecidas previamente.
Por equilibrio:	Procura una distribución equitativa en el número de actores o de casos identificados a propósito de un tema o de una categoría específica, para recolectar información suficiente e igualmente comparable.
Máxima variación:	Procura incluir un espectro diverso de una o diferentes categorías dentro del fenómeno para capturar elementos heterogéneos de éste.
Homogéneo:	esta tipología reúne actores o informantes que tienen características o rasgos en común, lo cual contribuye a obtener evidencia de manera

[12] De acuerdo con lo encontrado en la prueba piloto de la evaluación, se revisará si el levantamiento en dos etapas es inviable desde el punto de vista operativo para alguno de los actores planteados dentro del diseño de muestra. Por ejemplo, si el contacto con las empresas se dificulta, se evaluará levantar la información de directivos de empresas en una sola etapa. Estas modificaciones al diseño de muestra serán presentadas al CTS con el informe de prueba piloto.

[13] Siguiendo lo especificado en el anexo técnico de la evaluación, este ajuste tendrá que ser revisado y aprobado por el CTS, por lo cuálcual se necesitará de una gestión articulada entre el equipo consultor y el Comité para que el ajuste, revisión y aprobación de instrumentos se haga en un máximo de 1 semana y de esta manera no se afecte el desarrollo del trabajo de campo.

	focalizada. Esta tipología se usará para la selección de participantes de las entrevistas grupales para poblaciones con menor participación.
Características de grupo:	Se eligen a los casos para crear un grupo específico rico en información que pueda revelar patrones importantes dentro del grupo (Patton, 2005). Este criterio es especialmente relevante para las personas beneficiarias en los que se buscará identificar patrones similares entre grupos poblacionales específicos, especialmente en aquellos donde más allá de ser sectores sociales, hay movimientos sociales o iniciativas que los agrupan con comunidades que abogan de manera conjunta por sus derechos o inclusión social, por ejemplo, personas con discapacidad, étnias particulares y sectores lgbtiq.

En las siguientes tablas, se establece para cada población incluida en el muestreo cualitativo, los criterios de selección que se emplearán en cada una de las etapas. Adicionalmente, con el objetivo de incorporar la perspectiva interseccional y enfoques de género y poblacional que guían conceptualmente la evaluación, en el componente de personas beneficiarias se buscará escuchar a personas con experiencias de vida en las que se cruza más de una característica priorizada por el programa para comprender las características particulares de dicha intersección. Así, por ejemplo, se buscará de manera intencionada escuchar a mujeres jóvenes o mayores de 50 años, jóvenes migrantes, mujeres con discapacidad, entre otros.

7.2 Componente 1: Personas benefiaciarias

Tipo de actor	Actor	Tipología	Criterio de selección y justificación	Instrumento	No
Personas beneficiarias con alta participación en el Programa	Mujeres[14]	**Primera etapa** ***Por defecto:*** Actor solicitado en el Anexo Técnico ***Teórico y de pertinencia:*** *El segundo grupo poblacional con más participación en el programa 61%*[15] ***Por equilibrio:*** *Teniendo en cuenta que se hará entrevistas a personas postuladas con otros criterios poblacionales de edad y pobreza, estas dos entrevistas se buscarán hacer con personas que solamente hayan sido postuladas por un criterio "mujer".* ***Máxima variación:*** Se buscará diversidad en nivel educativo con los niveles de mayor frecuencia en las mujeres con postulaciones aceptadas por este criterio: 1) entrevista a mujer con nivel técnico o tecnológico y 1 entrevista con mujer bachiller. Se busca diversidad en tipo de contrato: una con término indefinido y otra con término fijo.		Entrevista individual	2
	Jóvenes	**Primera etapa** ***Por defecto:*** Actor solicitado en el Anexo Técnico. ***Teórico y de pertinencia:*** *El grupo poblacional con más participación en el programa 63%*[16]		Entrevista individual	2

[14] Si bien al considerar la categoría "Mujer" de manera amplia se incluye también a las mujeres trans, como hay una categoría de muestra específica para ellas, en esta categoría se incluirán solamente mujeres cisgénero.
[15] Especificaciones del Anexo Técnico
[16] Especificaciones del Anexo Técnico

		Por equilibrio: : Se hará mínimo una entrevista a una persona de género masculino y una persona de género femenino[17]. ***Máxima variación:*** Se buscará diversidad en nivel educativo (una persona bachiller y otra con técnico tecnólogo).		
	Personas en condición de pobreza	**<u>Primera etapa</u>** ***Por defecto:*** Actor solicitado en el Anexo Técnico ***Teórico y de pertinencia:*** *El tercer grupo poblacional con más participación en el programa 15%*[18] ***Por equilibrio:*** Se hará mínimo una entrevista a una persona de género masculino y una persona de género femenino. Una entrevista a personas con bachiller, primaria o secundaria y otra a una persona con técnico, tecnólogo o pregrado. Se buscará también variación en tipo de contrato.	Entrevista Individual	2
	Personas mayores de 50 años	**<u>Primera etapa</u>** ***Por defecto:*** Actor solicitado en el Anexo Técnico ***Teórico y de pertinencia:*** *El cuarto grupo poblacional con más participación en el programa 6%*[19] ***Por equilibrio:*** Se hará mínimo una entrevista a una persona de género masculino y una persona de género femenino. ***Máxima variación:*** Se hará una entrevista a una persona con especialización, maestría o pregrado y otra a una persona con tecnico, tecnologo o bachiller. Se buscará también variación en tipo de contrato.	Entrevista Individual	2
	Segunda etapa para todas las poblaciones (mujeres, jóvenes, pobreza, mayores de 50)	En la segunda etapa se harán 4 entrevistas a personas que cumplan uno o más de estos criterios. Los criterios para la selección de estas entrevistas serán establecidos por el equipo de investigación una vez se haya surtido la primera etapa y con aprobación del Comité Técnico de la evaluación. Estos podrán ser para incluir características no consideras o cruce de categorías.	Entrevista individual	4
Estudios de caso				
Personas beneficiarias con baja colocación	Migrantes	**<u>Primera etapa</u>** ***Por defecto:*** Actor solicitado en el Anexo Técnico ***Máxima variación:*** Se buscará tener diversidad edad, género, nivel de formación, tiempo que lleva en la ciudad y características del contrato y localidad.	Entrevista grupal	1
		<u>Segunda etapa</u> Acorde con los resultados obtenidos en la primera etapa, en la segunda etapa se buscará profundizar en la <u>historia de vida</u> de una persona en particular que haga parte del grupo poblacional para el estudio de caso. Si durante la primera etapa se identificó que hay alguna intersección entre características poblacionales especialmente relevante o	Entrevista a profundidad	1

[17]Al hablar de personas de género femenino hacemos referencia tanto a mujeres cisgénero como a mujeres transgénero, y al hablar de personas de género masculino hacemos referencia también a hombres cis y trans.
[18] Especificaciones del Anexo Técnico
[19] Especificaciones del Anexo Técnico

		sobre la que se tenga menos información o se quiera profundizar (por ejemplo, mujeres migrantes) se buscará en lo posible hacer la historia de vida a una persona que reuna estas características.		
	Víctimas del conflicto armado	**Primera etapa** ***Por defecto:*** Actor solicitado en el Anexo Técnico De acuerdo con la base sumistrada en los Pliegos, este grupo poblacional tuvo únicamente 5 personas con postulaciones aceptadas y vinculadas a empresas. Así, en la entrevista grupal se buscará tener a las cinco personas y se hará la segunda etapa profundizando sobre la historia de vida de una de ellas.	Entrevista grupal	1
		Segunda etapa Acorde con los resultados obtenidos en la primera etapa, en la segunda etapa se buscará profundizar en la historia de vida de una persona en particular que haga parte del grupo poblacional para el estudio de caso. Si durante la primera etapa se identificó que hay alguna intersección entre características poblacionales especialmente relevante o sobre la que se tenga menos información o se quiera profundizar se buscará en lo posible hacer la historia de vida a una persona que reuna estas características.	Entrevista a profundidad	1
	Transgénero	**Primera etapa** ***Por defecto:*** Actor solicitado en el Anexo Técnico ***Máxima variación:*** Se buscará tener diversidad en edad, género (hombres y mujeres trans)[20] y tipo de contrato.	Entrevista grupal	1
		Segunda etapa Acorde con los resultados obtenidos en la primera etapa, en la segunda etapa se buscará profundizar en la historia de vida de una persona en particular que haga parte del grupo poblacional para el estudio de caso. Si durante la primera etapa se identificó que hay alguna intersección entre características poblacionales especialmente relevante o sobre la que se tenga menos información o se quiera profundizar se buscará en lo posible hacer la historia de vida a una persona que reuna estas características.	Entrevista a profundidad	1
	Que habitan la ruralidad	**Primera etapa** ***Por defecto:*** Actor solicitado en el Anexo Técnico ***Máxima variación:*** Se buscará incluir a las 5 personas con postulaciones aprobadas en esta primera etapa. Si la base es actualizada con más registros se buscará variedad en género y edad.	Entrevista grupal	1
		Segunda etapa Acorde con los resultados obtenidos en la primera etapa, en la segunda etapa se buscará profundizar en la historia de vida de una persona en particular que haga parte del grupo poblacional para el estudio de caso. Si durante la primera etapa se identificó que hay alguna intersección entre características poblacionales especialmente relevante o sobre la que se tenga menos información o se quiera profundizar se buscará en lo posible hacer la historia de vida a una persona que reuna estas características.	Entrevista a profundidad	1
	Con discapacidad	**Primera etapa** ***Por defecto:*** Actor solicitado en el Anexo Técnico	Entrevista grupal	1

[20] Buscando también incluir experiencias de hombres trans y otras identidades de género como personas no binarias. |

		Máxima variación: Se buscará diversidad en edad, género, tipo de discapacidad (física, cognitiva, psicosocial), localidad, nivel educativo, tipo de contrato (duración de contrato).		
		Segunda etapa Acorde con los resultados obtenidos en la primera etapa, en la segunda etapa se buscará profundizar en la historia de vida de una persona en particular que haga parte del grupo poblacional para el estudio de caso. Si durante la primera etapa se identificó que hay alguna intersección entre características poblacionales especialmente relevante o sobre la que se tenga menos información o se quiera profundizar se buscará en lo posible hacer la historia de vida a una persona que reuna estas características.	Entrevista a profundidad	1
	Con pertenencia étnica	**Primera etapa** ***Por defecto:*** Actor solicitado en el Anexo Técnico ***Máxima variación:*** Se buscará diversidad en edad, género, pertenencia étnica (indígenes, comunidades negras o afrocolombianas, raizales) localidad, nivel educativo, tipo de contrato (duración de contrato).	Entrevista grupal	1
		Segunda etapa Acorde con los resultados obtenidos en la primera etapa, en la segunda etapa se buscará profundizar en la historia de vida de una persona en particular que haga parte del grupo poblacional para el estudio de caso. Si durante la primera etapa se identificó que hay alguna intersección entre características poblacionales especialmente relevante o sobre la que se tenga menos información o se quiera profundizar se buscará en lo posible hacer la historia de vida a una persona que reuna estas características.	Entrevista a profundidad	1
	Otro	**Segunda etapa** El último estudio de caso se eligirá después del primer levantamiento de información con otras poblaciones beneficiarias y con empresas teniendo en cuenta criterios de **pertinencia técnica** y **equilibrio.** Este caso se elegirá entre las poblaciones focalizadas por el programa que no han sido incluídas como personas pospenadas, en proceso de reintegración, en habitabilidad de calle, en reconversión laboral como vendedores ambulantes o recicladores, entre otras. Se realizarán dos historias de vida.	Entrevista a profundidad	2

7.3 Componente 2: Empresas beneficiarias

Tipo de actor	Actor	Tipología	Criterio de selección y justificación	Instrumento	No
Intermediación	Cajas de compensación, Agencia Distrital de Empleo, SENA, bolsas de empleo, Universidades.	***Por defecto:*** Actor solicitado en el Anexo Técnico Se harán dos grupos focales con características similares, uno en cada etapa de levantamiento. Se buscará dentro de los grupos tener máxima variación en el tipo de actor convocado: Público (SENA y ADE) y privado (cajas, universidades, otras bolsas de empleo). Nota: Si después del primer levantamiento se ve la necesidad de separar el tipo de agencia, por ejemplo teniendo un grupo focal con entidades dirigidas o		Grupo focal	2

		especializadas a personas con formación profesional o ténica (universidades y SENA) y otros dirigidos de manera más general a la población (ADE, cajas de compensación, otras bolsas de empleo, se adicionará un grupo focal para esta segunda etapa de levantamiento.		
Empresas beneficiarias	Directivos y personal de la empresa encargado de la postulación	**<u>Primera etapa</u>** En la primera etapa de levantamiento se harán 9 entrevistas grupales a empresas beneficiarias teniendo en cuenta características de tamaño y sector de la siguiente manera: - Grandes empresas: (3 en sectores de comercio, industria y servicios) - Medianas y pequeñas: (3 en sectores de comercio, industria y servicios) - Micro empresas: (3 en sectores de comercio, industria y servicios) Teniendo en cuenta que en los lineamientos del programa se define el tamaño de la empresa por nivel de ingresos pero que el número de empleados también puede ser una variable de interés y determinante en la participación del programa se buscará, en lo posible transversal a esta muestra diversidad en el número de empleados de las empresas.	Entrevistas grupales	9
		<u>Segunda etapa</u> Según los resultados de la primera etapa, se seleccionarán cuatro entrevistas grupales con empresas de diferente tamaño y sector.	Entrevistas grupales	4
	Personal de empresas seleccionadas	***Por defecto:*** Actor solicitado en el Anexo Técnico Se harán 3 talleres participativos por tamaño de empresa (grande, pequeña y mediana, micro) en la primera etapa de recolección de información. Para la segunda etapa se harán dos talleres participativos adicionales que serán seleccionados según los resultados de la primera etapa.	Taller participativo	5
Empresas que postularon pero no fueron seleccionadas	Personal de la empresa involucrado en la postulación	***Por defecto:*** Actor solicitado en el Anexo Técnico ***Equilibrio:*** Se harán 3 entrevistas por tamaño de empresa (grande, pequeña y mediana, micro). Para la segunda etapa, si se considera relevante, se hará una entrevista más para profundizar con 1 empresa.	Entrevista individual	4
Empresas registradas sin postulaciones	Personal de la empresa involucrado en la postulación	***Por defecto:*** Actor solicitado en el Anexo Técnico ***Equilibrio:*** Se harán 3 entrevistas por tamaño de empresa (grande, pequeña y mediana, micro). Para la segunda etapa, si se considera relevante, se hará una entrevista más para profundizar con 1 empresa.	Entrevista individual	4

IV. Triangulación

La integración de metodologías cuantitativas y cualitativas en la evaluación del Programa de Empleo Incluyente (PEI) se basa en un enfoque mixto, que permite no solo la triangulación de resultados, sino también la construcción de una comprensión más profunda y matizada de los fenómenos observados. Dado el diseño del proyecto, la recolección y análisis de los datos cuantitativos y cualitativos se realizarán en paralelo, con una fase final de integración que maximice las fortalezas complementarias de ambas aproximaciones.

Relación entre metodologías cuantitativas y cualitativas

El enfoque cuantitativo se centrará en la caracterización y cuantificación de los efectos del programa, mientras que el cualitativo explorará los mecanismos subyacentes a estos efectos, enriqueciendo la interpretación de los hallazgos. La combinación de estas metodologías proporciona no solo una validación cruzada (corroborativa), sino también una mayor profundidad explicativa.

Las metodologías cuantitativas, tales como estadísticas descriptivas, análisis de supervivencia, SEM y redes neuronales, se utilizarán para abordar las preguntas orientadoras que requieren cuantificación de efectos y modelado de relaciones entre variables. Por su parte, los métodos cualitativos —análisis narrativo y evaluación realista— ayudarán a contextualizar los hallazgos cuantitativos, identificar barreras y facilitadores de inclusión laboral, y proporcionar una visión detallada de las dinámicas complejas que no son capturadas por los análisis numéricos.

Secuencialidad y complementariedad de los procesos

La implementación simultánea de los análisis cuantitativos y cualitativos permite aprovechar la riqueza de ambos enfoques sin retrasos en la recolección de información. Sin embargo, los resultados cualitativos recolectados en la primera etapa influirán en la etapa posterior de análisis cualitativo y complementarán los resultados de los primeros análisis cuantitativos. De esta forma, se garantizará una retroalimentación continua entre ambos enfoques.

En la fase de integración, se llevará a cabo una triangulación de los hallazgos para complementar los resultados cuantitativos con las percepciones y experiencias recogidas cualitativamente. Por ejemplo, si los análisis de redes neuronales identifican diferencias significativas en la empleabilidad entre subgrupos, se podrán contrastar estos resultados con las narrativas obtenidas en las entrevistas, para explicar las razones detrás de estas diferencias y cómo operan los mecanismos de exclusión o inclusión.

Acciones específicas para la integración

- **Fase I (Simultánea)**: Durante la primera fase, tanto los análisis cuantitativos como cualitativos se realizarán en paralelo. Los análisis cuantitativos, como las estadísticas descriptivas y los modelos de supervivencia, proporcionarán una visión inicial sobre las características sociodemográficas y laborales de las personas beneficiarias, así

como las tendencias generales de permanencia en el empleo. Simultáneamente, las entrevistas y talleres cualitativos explorarán las percepciones de las personas beneficiarias y las empresas sobre los mecanismos de inclusión, generando hipótesis que podrán ser contrastadas con los datos cuantitativos en fases posteriores.

- **Fase II (Integración y triangulación)**: Los hallazgos de ambas metodologías se integrarán al final del proceso. Esta etapa incluirá el ajuste de instrumentos cualitativos basados en los resultados preliminares cuantitativos (ver sección 4.4.1.4 sobre las dos etapas de recolección de información cualitativa). Por ejemplo, si los análisis preliminares cuantitativos con registros administrativos y la encuesta GEIH revelan patrones inesperados en la empleabilidad de ciertos subgrupos, los instrumentos cualitativos también podrán ser ajustados para explorar estos temas en mayor detalle. Además, se elaborarán conclusiones y recomendaciones basadas en la complementariedad de los resultados. Las diferencias encontradas entre mujeres y jóvenes, por ejemplo, podrán ser explicadas no solo mediante diferencias en las características observadas cuantitativamente, sino también a través de las narrativas que arrojen luz sobre barreras o facilitadores contextuales.

La integración de estas metodologías se ajusta a las necesidades particulares de la evaluación del PEI al permitir una comprensión integral y matizada de los resultados. La simultaneidad de la recolección de datos y la posterior triangulación aseguran que los hallazgos tanto cuantitativos como cualitativos sean considerados en las recomendaciones, ofreciendo una visión robusta que combina la precisión numérica con la profundidad contextual.

La metodología para la evaluación del PEI será un enfoque de triangulación concurrente y explicativa secuencial. Implementaremos ambas metodologías, cuantitativa y cualitativa, de manera simultánea, lo que permitirá una retroalimentación dinámica y continua entre ambas durante el proceso. Este diseño garantizará que los datos sean recolectados y analizados en paralelo, maximizando la riqueza de información de ambos enfoques y promoviendo una integración sólida al final del estudio.

En el componente cuantitativo, recopilaremos información a través de la encuesta y los registros administrativos, como PILA y GEIH, con el objetivo de realizar análisis estadísticos descriptivos, modelos de supervivencia y SEM. Estos análisis cuantificarán indicadores clave como empleabilidad, permanencia en el empleo y características sociodemográficas de los beneficiarios. Simultáneamente, en el componente cualitativo, llevaremos a cabo entrevistas en profundidad, grupos focales y talleres participativos con los beneficiarios y las empresas, explorando sus percepciones, experiencias y los **mecanismos** que operan en los procesos de inclusión laboral.

A partir de los resultados de la primera etapa cualitativa, identificaremos patrones, anomalías o temas emergentes. En los análisis cuantitativos, buscaremos diferencias entre subgrupos poblacionales, como mujeres y jóvenes, utilizando redes neuronales y modelos de supervivencia para detectar tendencias significativas. Por otro lado, los análisis cualitativos explorarán las barreras y facilitadores de inclusión laboral, proporcionando narrativas detalladas que complementen las conclusiones numéricas.

En la fase de integración, realizaremos una **triangulación explicativa** para vincular los hallazgos de ambos enfoques. Compararemos y contrastaremos los resultados cuantitativos y cualitativos, creando una matriz de integración que permita identificar áreas de convergencia y divergencia. Este proceso nos permitirá explicar fenómenos complejos al asociar patrones estadísticos con las narrativas contextuales, enriqueciendo así la interpretación de los resultados. Por ejemplo, si los modelos cuantitativos muestran menores tasas de permanencia en el empleo entre ciertos subgrupos, las entrevistas cualitativas explorarán las posibles causas, como las prácticas de exclusión o discriminación en las empresas.

V. Referencias

Akgul, Y., Uymaz, A. O., & Uymaz, P. (2024). Understanding mobile learning continuance after the COVID-19 pandemic: Deep learning-based dual stage partial least squares-structural equation modeling and artificial neural network analysis. *Environment and Social Psychology,* 9(4), 1–20.

Allison, P. D. (2010). *Survival Analysis Using SAS: A Practical Guide*. SAS Institute.

Arrow, K. J. (1973). *The theory of discrimination*. En *Discrimination in labor markets* (pp. 3-33). Princeton University Press.

Bacchi, C. (2009). *The politics of affirmative action: "Women", equality and category politics*. SAGE Publications.

Barry Schouten, Natalie Shlomo & Chris Skinner (2011) Indicators for Monitoring and Improving Representativeness of Response, Journal of Official Statistics, Vol. 27, No. 2, 2011, pp. 1–24.

Becker, G. S. (1957). *The economics of discrimination*. University of Chicago Press.

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education*. University of Chicago Press.

Bohórquez L. y Merchán C. 2005. Una propuesta de subsidio al salario para Colombia. Documento 279, Archivos de Economía. Departamento Nacional de Planeación.

Byrne, B. M. (2010). *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming*. Routledge.

Card D., Kluve J., Weber A. 2015. What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations. IZA DP No. 9236.

Carroll, A. B. (1991). The pyramid of corporate social responsibility: Toward the moral management of organizational stakeholders. *Business Horizons*, 34(4), 39-48.

Centro de Microdatos (2012). Evaluación de Impacto del Programa Subsidio al Empleo Joven.

Cheng, P. (2022). Employment Data Screening and Destination Prediction of College Students Based on Deep Learning. *Wireless Communications and Mobile Computing*.

Collett, D. (2015). *Modelling survival data in medical research* (3rd ed.). CRC Press.

Comisión Económica para América Latina y el Caribe (CEPAL), Diseño y análisis estadístico de las encuestas de hogares de América Latina, Metodologías de la CEPAL, N° 5 (LC/PUB.2023/14-P), Santiago, 2023.

Cooper, D. (2013). The Effect of Unemployment Duration on Future Earnings and Other Outcomes. *Econometric Modeling: Microeconometric Models of Household Behavior eJournal*.

Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory, and antiracist politics. *University of Chicago Legal Forum*, 1989(1), 139-167.

Dawis, R. V., & Lofquist, L. H. (1984). *A psychological theory of work adjustment: An individual-differences model and its applications*. University of Minnesota Press.

Dema G.; Díaz J.; Chaclatana J. 2015. ¿Qué sabemos sobre los programas y políticas de Primer Empleo en América Latina? Lima: OIT; Oficina Regional para América Latina y el Caribe.

Di Franco, G. (2017). *Machine learning, artificial neural networks and social research*. *Quality & Quantity, 51*(4), 1891–1910.

DNP 2020. Evaluación del Programa 40.000 Primeros Empleos. Producto 4. Informe de Resultados.

DNP 2022. Evaluación de Resultados del Programa de Apoyo al Empleo Formal – PAEF. Evaluación interna.

DNP 2024. Evaluación de Resultados del Programa Nuevo Empleo Evaluación interna.

Ely, R. J., & Thomas, D. A. (2001). Cultural diversity at work: The effects of diversity perspectives on work group processes and outcomes. *Administrative Science Quarterly*, 46(2), 229-273.

Emmel, N. (2013). Sampling and Choosing Cases in Qualitative Research: A Realist Approach. SAGE Publications.

Escudero V., Kluve J., López e., Pignatti C. 2017. Active labour market programmes in Latin America and the Caribbean: Evidence from a meta análisis. International Labour Office. Research Department Working Paper No. 20

Fedesarrollo (2018). *Evaluación de resultados del Mecanismo de Protección al Cesante*. DNP, Sinergia. Bogotá.

Fraser, N. (2008). *Scales of justice: Reimagining political space in a globalizing world*. Columbia University Press.

Guio Daniela 2024. Subsidios laborales en tiempos post pandemia: impacto del programa Nuevo Empleo sobre la creación de empleo formal en Colombia

Hillage, J., & Pollard, E. (1998). *Employability: Developing a framework for policy analysis*. Institute for Employment Studies.

Hoyle, R. H. (Ed.). (2012). *Handbook of Structural Equation Modeling*. Guilford Press.

Hugo Andrés Gutiérrez (2009). Estrategias de muestreo. Diseño de encuestas y estimación de parámetros. Editorial Universidad Santo Tomás, Bogotá, Colombia.

Hugo Andrés Gutiérrez (2020) samplesize4surveys: Sample Size Calculations for Complex Surveys. R package version 4.1.1, <https://CRAN.R-project.org/package=samplesize4surveys>.

Hugo Andrés Gutiérrez, Luisa Fernanda Rojas (2013) Ajuste de estimadores mediante la técnica de propensity score en encuestas complejas. Revista IB, vol 3, no. 1, pg. 105-114.

ILO and World Bank (2022). The impact of active labour market programmes on youth. An updated systematic review and meta-analysis.

International Labour Organization (ILO). (2003). *Working out of poverty*. International Labour Office.

Jenkins, S. P. (2005). *Survival Analysis*. Manuscrito no publicado, Instituto de Investigación Social y Económica, Universidad de Essex.

Judijanto, L., & Sudi, D. (2023). Analysis of Economic, Education, and Skill Factors against High Unemployment: A Case Study in West Java Province. *West Science Journal Economic and Entrepreneurship*.

Jun Shao & Dongsheng Tu (1995) The Jackknife and Bootstrap, Springer.

Kabeer, N. (1999). Resources, agency, achievements: Reflections on the measurement of women's empowerment. *Development and Change*, 30(3), 435-464.

Kirk Wolter (1985). Introduction to Variance Estimation, Springer

Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling*. Guilford Press.

Lazar, M. M. (2005). Feminist Critical Discourse Analysis: Gender, Power and Ideology in Discourse. Palgrave Macmillan.

Leslie Kish (1965). Survey Sampling. J. Wiley

Lin, Y., Huang, I., & Wang, Y. (2014). Outcomes of home-based employment service programs for people with disabilities and their related factors—a preliminary study in Taiwan. *Disability and Rehabilitation*, 36(17), 1457–1463.

Lumley. T., Gao, P., Schneider, B. (2024) Survey: analysis of complex survey samples. R package, version 4.4.

Meng, X., Ren, G., & Huang, W. (2021). A Quantitative Enhancement Mechanism of University Students. Employability and Entrepreneurship Based on Deep Learning in the Context of the Digital Era. *Sci. Program.*, 2021, 7245465:1-7245465:12.

Ministerio del Trabajo 2015. Manual Operativo Programa 40 mil primeros empleos. Anexo II.

Mor Barak, M. E. (2015). *Inclusion is the key to diversity management, but what is inclusion? Human Service Organizations: Management, Leadership & Governance*, 39(2), 83-88.

Núñez, J. (2005). *Éxitos y fracasos de la reforma laboral en Colombia*. Universidad de los Andes, Facultad de Economía, CEDE. No. 2005-43.

OIT 2016. Soluciones eficaces: Políticas Activas del Mercado de Trabajo en América Latina y el Caribe.

Olivera M. 2006. Evaluación del Programa de Apoyo Directo al Empleo – PADE. Informe final Departamento Nacional de Planeación.

Patton, M. Q. (2015). Qualitative Research and Evaluation Methods (4th ed.). SAGE Publications.

Pawson, R., & Tilley, N. (1997). *Realistic Evaluation*. Sage Publications.

Rao, JNK. 2021. On making valid inferences by integrating data from surveys and other sources. Sankhya B 83: 242-72.

Riessman, C. K. (2008). Narrative Methods for the Human Sciences. SAGE Publications.

Rogers, E. M. (2007). *Diffusion of innovations* (5th ed.). Free Press.

Salvia, A.; Ravallion, M.; Galasso, E. 2004. "Assisting the transition from workfare to work: A randomized experiment", en Industrial and Labor Relations Review, Vol. 57, No. 5, pp. 1-32.

Särndal C., Swensson, B., Wretman, J. (2003) Model Assisted Survey Sampling. Springer

Schumacker, R. E., & Lomax, R. G. (2016). *A Beginner's Guide to Structural Equation Modeling*. Routledge.

Servicio Nacional de Capacitación y Empleo. (2024). *Evaluación ex ante - Proceso formulación presupuestaria 2025: Subsidio al empleo joven (SEJ) (Versión 3)*. Ministerio de Trabajo y Previsión Social.Sharon L. Lohr (2010). Sampling: Design and Analysis. Brooks/Cole Cengage Learning

Singer, J. D., & Willett, J. B. (2003). *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. Oxford University Press.

Valliant, R., Dever, J., and Kreuter, F. (2018). Practical Tools for Designing and Weighting Survey Samples, 2nd edition, New York: Springer

William G.. Cochran (1977). Sampling Techniques. Wiley.

Zhang, Y., & Cheng, T. (2020). A Deep Learning Approach to Infer Employment Status of Passengers by Using Smart Card Data. *IEEE Transactions on Intelligent Transportation Systems*, 21, 617-629.

CERTIFICACIÓN DE CALIDAD

Ipsos Napoleón Franco cumple con los códigos éticos establecidos por ESOMAR/CCI, y se encuentra certificada en dos normas ISO.

- ISO 9001:2015: Focalizada en la satisfacción del cliente. Todos sus requerimientos son genéricos y aplican a todo tipo de negocio.

- ISO 20252:2019: Norma es específica para empresas y requerimientos de servicio de compañías y profesionales que trabajen en Investigación de Mercados, Social, y de la Opinión. Es el *Benchmark* global de mejores prácticas para Investigación de Mercados. Fue una respuesta a las iniciativas de ESOMAR y EFAMRO y fue aprobada por ISO en 2006 como el resultado del trabajo del Comité Técnico ISO/TC 225.

icontec ISO 20252 EO-CER894739

icontec ISO 9001 CO-SC-CER894736

IQNET CERTIFIED MANAGEMENT SYSTEM