

Benemérita Universidad Autónoma de Puebla

Facultad de Ciencias Físico Matemáticas

**Participación laboral de inmigrantes
recién llegados a México:**

**Evidencia basada en la Encuesta
Nacional de Ocupación y Empleo
(ENOE), 2023**

T E S I S

presentada para obtener el grado de:

L I C E N C I A D A E N A C T U A R Í A

P R E S E N T A

Jareth Lizette Mejía Sánchez

DIRECTORA DE TESIS:

Dra. Martha Miranda Muñoz

CO DIRECTOR:

Dr. José Asunción Hernández

Puebla, Pue

Mayo 2024



Agradecimientos

Índice

Agradecimientos	1
Índice.....	2
Introducción	4
Capítulo I. Inmigrantes en México	7
1.1 Migración en tránsito.....	7
1.2 Proporción de migrantes que arriban y eligen permanecer en territorio mexicano	9
1.3 Los desafíos y riesgos del camino hacia la migración a Estados Unidos	15
1.4 Explorando los flujos migratorios: una mirada a la EMIF.....	18
1.5 Participación laboral de los inmigrantes.....	22
Capítulo II. Metodología.....	25
2.1 Motivación de los modelos de regresión empleados en la presente tesis.....	25
2.2 Regresión lineal estimada por mínimos cuadrados ordinarios	25
2.2.1 Estimación usando el método de mínimos cuadrados ordinarios.....	26
2.2.2 Supuestos	29
2.2.3 Explicación de los coeficientes de regresión	29
2.2.4 Problema de estimación.....	30
2.3 Modelo de probabilidad no lineal: probit	31
2.3.1 Estimación de los parámetros vía máxima verosimilitud.....	32
2.3.2 Efectos marginales del modelo probit	33
Capítulo III. Datos ENOE.....	35
3.2 Encuesta nacional de ocupación y empleo (ENOE 2023)	35
3.2.1 Diseño conceptual.....	36
3.2.2 Diseño de los instrumentos de captación de la ENOE	37
3.2.3 Cobertura geográfica	38

3.2.4 Marco de la encuesta	38
3.2.5 Formación de las Unidades Primarias de Muestreo (UPM)	39
3.2.6 Estratificación	40
3.2.7 Tamaño de la muestra	40
3.2.8 Estructura de la base de datos	45
3.3 <i>Procesamiento de los datos de la base de la ENOE</i>	46
3.4 Elaboración de variables de interés	47
Capítulo IV. Resultados	52
5.3 Variables utilizadas en los modelos de regresión	54
5.4 <i>Regresión Lineal de Mínimos Cuadrados Ordinarios</i>	54
5.5 <i>Regresión Probit</i>	55
5.5.1 Efectos marginales del modelo de regresión probit.....	56
5.6 <i>Análisis de los salarios de inmigrantes de corto plazo en México</i>	57
Conclusiones	59
Apéndice	61
Referencias.....	67

Introducción

La movilidad humana es un proceso complejo y diverso que ha sido una constante a lo largo del tiempo y sigue siendo una realidad global en nuestros días. La migración abarca una amplia gama de aspectos que van más allá del simple desplazamiento geográfico de personas, involucrando elementos culturales, económicos, sociales y políticos que influyen en su dinámica y significado.

La migración es el desplazamiento de personas de un lugar a otro, ya sea dentro de un país (migración interna) o entre países (migración internacional), motivado por una variedad de factores que incluyen aspectos económicos, políticos, sociales y ambientales (Puerta et al., 2019).

Una de las causas más significativas de la migración es la búsqueda de mejores oportunidades económicas. Muchas personas abandonan sus lugares de origen en busca de empleo, educación o mejores condiciones de vida en otros lugares. Esta migración económica puede ser tanto temporal como permanente y puede tener un impacto considerable en las economías tanto de los países de origen como de destino.

La participación laboral de los inmigrantes recién llegados a México es un aspecto crucial que refleja las dinámicas migratorias y su impacto en la economía y sociedad del país. En el contexto actual de creciente migración internacional, entender cómo estos nuevos migrantes se integran al mercado laboral mexicano es fundamental para analizar sus efectos en diferentes aspectos socioeconómicos (OIM, 2021).

Uno de los puntos centrales es la dinámica que estos inmigrantes generan en el mercado laboral local. La llegada de nuevos trabajadores puede influir en la oferta y demanda de empleo, así como en las condiciones laborales y salariales. Este fenómeno no solo afecta a los propios migrantes, sino también a los trabajadores nacionales, las empresas y el sistema de seguridad social (INEGI, 2020).

Además, la participación laboral de los inmigrantes recién llegados también está vinculada con temas de inclusión social y derechos laborales. Es importante garantizar que estos trabajadores tengan acceso a empleos dignos y justamente remunerados, así como a condiciones laborales adecuadas y protección social. Esto no solo contribuye al bienestar de los migrantes, sino que también fortalece la cohesión social y el desarrollo económico del país (SER., 2021).

Para comprender mejor esta temática, es necesario analizar datos estadísticos y estudios empíricos que permitan identificar patrones, tendencias y desafíos específicos que enfrentan los inmigrantes en su inserción laboral en México.

Un factor importante que impulsa la migración es la búsqueda de seguridad y protección. Las personas que huyen de conflictos armados, persecución política, violencia o violaciones de derechos humanos pueden buscar refugio y asilo en otros países, lo que genera flujos significativos de migración forzada.

El estudio sobre cómo se integran los inmigrantes en el mercado laboral de México se ha vuelto crucial en el contexto actual, marcado por un aumento significativo de la migración internacional y sus repercusiones en la economía nacional. Comprender este fenómeno no solo es importante para fomentar la inclusión y el desarrollo social de esta población, sino también para identificar los desafíos y las oportunidades que surgen en el ámbito laboral y en la contribución al crecimiento económico del país.

En este sentido, es fundamental reconocer que los inmigrantes pueden aportar de manera significativa al desarrollo económico, al aportar nuevas habilidades y perspectivas al mercado laboral. Además, su integración enriquece la diversidad cultural de la sociedad mexicana, generando beneficios tanto en términos económicos como sociales (OIM, 2020).

Sin embargo, esta inserción laboral puede tener efectos diversos. Por un lado, puede significar la pérdida de talento y mano de obra para los países de origen de los migrantes, lo que plantea desafíos en términos de desarrollo en esas regiones. Por otro lado, la entrada de remesas provenientes de los trabajadores migrantes beneficia a las comunidades de origen, contribuyendo a su desarrollo económico y social.

Es fundamental abordar la migración desde una perspectiva integral y respetuosa de los derechos humanos. Esto incluye garantizar la protección de los migrantes, facilitar vías legales y seguras de migración, promover la inclusión y la cohesión social en las comunidades de acogida, y abordar las causas subyacentes de la migración, como la desigualdad económica, la inseguridad, el conflicto y el cambio climático.

El propósito fundamental de esta tesis es examinar la actividad laboral de los inmigrantes a corto plazo en México, comprendiendo sus particularidades en el ámbito laboral, los campos en los que encuentran empleo y los elementos que afectan su inclusión en el mercado de trabajo, teniendo en cuenta variables como el género, nivel educativo y edad. En este estudio, se busca profundizar en cómo estos factores influyen en su integración laboral en el país.

Mediante la inclusión de la base de datos que nos otorgan en la Encuesta Nacional de Ocupación y Empleo (ENOE), el objetivo de este estudio es esclarecer la situación laboral de los inmigrantes, a quienes caracterizamos como aquellos individuos que i) no nacieron en México y ii) no residían en el país hace 12 meses.

Se pretende analizar y proporcionar datos relevantes sobre la participación laboral y los niveles de remuneración de este grupo de individuos que han migrado a México. El propósito de la investigación es comprender las dinámicas laborales de los inmigrantes en el contexto nacional y su interacción con el mercado de trabajo, contribuyendo así a una comprensión más completa de la situación ocupacional de esta parte de la población.

Los interrogantes fundamentales que guían esta investigación son los siguientes: ¿qué proporción de los inmigrantes que han llegado recientemente a México se encuentran empleados?, ¿en qué áreas económicas se integran con mayor frecuencia y qué tipo de trabajos desempeñan?, ¿cuáles

son los elementos que favorecen u obstaculizan la inclusión laboral de los inmigrantes en el mercado laboral mexicano?

Para abordar estas preguntas, se empleará un enfoque cuantitativo que implica un análisis detallado de los datos proporcionados por la Encuesta Nacional de Ocupación y Empleo (ENOE), centrandó la atención en los registros del año 2023. Utilizaremos técnicas estadísticas para examinar tendencias, relaciones y patrones en la participación laboral de los inmigrantes recién llegados a México captados durante ese periodo.

Este estudio se desarrollará utilizando los siguientes enfoques analíticos:

- Modelo de Regresión Lineal de Mínimos Cuadrados Ordinarios: Este modelo busca establecer una relación lineal entre la variable dependiente y una o más variables independientes, permitiendo predecir el valor de la variable dependiente basándose en las variables independientes.
- Modelo de Regresión Probit: En este enfoque, se emplea la función probit para modelar la probabilidad de que ocurra un evento o resultado, usualmente binario, en función de las variables explicativas. La función probit es una transformación de la distribución normal estándar y se utiliza particularmente en análisis de variables dicotómicas.

El presente estudio se divide en diversas secciones que abarcan aspectos clave relacionados con la migración en transición y la participación laboral de aquellos que deciden quedarse. Estas secciones comprenden una revisión exhaustiva de la literatura sobre este tema, el análisis de datos estadísticos que detallan la llegada de migrantes al país y su decisión de establecerse aquí, así como los desafíos específicos que enfrentan los migrantes en su proceso de tránsito. También se abordan aspectos generales sobre la Encuesta sobre Migración en la Frontera Norte (EMIF) y se examina detalladamente la participación laboral de los inmigrantes en México.

En cuanto al marco metodológico que sustenta nuestro análisis, se emplean herramientas como el análisis de regresión lineal de mínimos cuadrados ordinarios y el análisis de regresión probit. Además, se profundiza en la metodología utilizada, detallando la manipulación de la base de datos y la creación de las variables de interés. Finalmente, se presenta una descripción minuciosa de los resultados obtenidos, así como las conclusiones y recomendaciones derivadas de esta investigación.

Capítulo I. Inmigrantes en México

En este capítulo, discutiremos sobre la migración y aquellos migrantes que eligen establecerse como inmigrantes en México. Además, analizaremos los desafíos y riesgos asociados con la búsqueda del sueño americano, exploraremos la Encuesta sobre Migración en la Frontera (EMIF) y finalmente, examinaremos la participación laboral de los inmigrantes.

La migración se describe como el desplazamiento físico de individuos a través de límites geográficos definidos, con la intención de establecer una residencia de manera permanente o temporal prolongada. Este movimiento puede ser de carácter internacional, es decir, entre distintos países, o interno, cuando ocurre dentro de las fronteras de una nación (Puerta et al., 2019).

La migración de centroamericanos, caribeños, sudamericanos y africanos, que viajan a través de México con la intención de llegar a Estados Unidos de América, ha sido un fenómeno persistente que ha generado importantes desafíos y dinámicas sociales en la región.

1.1 Migración en tránsito

México es un país que enfrenta considerables movimientos migratorios debido a su situación económica y su papel como receptor, emisor y territorio de paso, crear una estrategia migratoria y demográfica representa un desafío significativo. Esta labor requiere armonizar múltiples aspectos de interés público, como salvaguardar la soberanía y gestionar las fronteras, al mismo tiempo que se fomentan los derechos humanos y se atienden las legítimas necesidades de los grupos vulnerables, sin importar su nacionalidad.

La vulnerabilidad de los migrantes indocumentados a lo largo de las fronteras internacionales se ve agravada por su necesidad de ocultarse, lo que facilita que sean víctimas de múltiples abusos en un contexto de impunidad generalizada. Estos abusos suelen quedar fuera de las estadísticas oficiales de criminalidad debido a la falta de denuncias, sanciones y difusión, lo que también limita su impacto en la conciencia pública (Valenzuela M. et al., 1998).

A principios de los años ochenta, la inmigración no era un asunto significativo para México, ya que los flujos eran reducidos y tenían pocas repercusiones políticas. No obstante, desde entonces, la inmigración ha aumentado tanto en diversidad como en cantidad (Castillo et al., 1998).

Según lo señalado por Castillo (2010), durante el periodo del siglo XX, aunque de manera más pronunciada a partir de los años 1980, Centroamérica ha sido identificada como una región de la cual provienen migrantes, principalmente de los países del norte de la región: Guatemala, Honduras y El Salvador. Estos tres países, junto con México y Estados Unidos, conforman un sistema migratorio regional.

Cada año, un gran grupo de migrantes atraviesa México de manera "no autorizada" en su camino hacia Estados Unidos. Durante este viaje, enfrentan serias agresiones y violaciones de sus derechos humanos (Aikin, O. et al. 2013).

México destaca como el principal país de tránsito para los flujos migratorios irregulares que se dirigen hacia el norte. Sin embargo, a partir de la década de 1990, Mesoamérica ha emergido como un destino cada vez más atractivo no solo para los centroamericanos, sino también para migrantes de Sudamérica, África, Asia y Europa Oriental (Kron et al., 2011). Si bien muchos migrantes tienen la esperanza de llegar a su destino final en Estados Unidos, una parte significativa se detiene en diversas regiones de México, y algunos toman la decisión de quedarse en el país.

Según información de Naciones Unidas, entre los años 1985 y 1990, Centroamérica experimentó la mayor emigración de su población en comparación con otras regiones del mundo. Durante este periodo, un promedio de 4.2 personas por cada mil habitantes de la población total emigraron para establecerse en otro país (Puerta et al., 2019).

El comienzo de la era actual de restricciones en las fronteras de la región data de 1986, cuando los legisladores estadounidenses Simpson y Rodino lograron que el congreso de su país aprobara la Ley de Reforma y Control de la Inmigración (IRCA, por sus siglas en inglés). Esta ley constaba de cuatro componentes principales: a) el cierre de las fronteras, b) la reducción del programa para trabajadores agrícolas especiales, c) otorgar amnistía a inmigrantes que llevaban más de cinco años en el país, y d) sanciones para los empleadores que contrataban a migrantes en situación irregular (Durand et al., 2003).

En el contexto de las políticas migratorias en México, se observa un panorama dinámico y cambiante a lo largo de distintas administraciones presidenciales. Desde el gobierno de Vicente Fox hasta la gestión de Enrique Peña Nieto, se han delineado diferentes enfoques y estrategias en relación con la migración, reflejando tanto cambios internos como influencias externas, como los eventos del 11 de septiembre de 2001. Estos enfoques abarcan desde la búsqueda de acuerdos con Estados Unidos hasta iniciativas de "desmigratización" y la integración de políticas migratorias en el desarrollo nacional. Durante el gobierno de Vicente Fox (2000-2006), se estableció una política proactiva a favor de alcanzar un acuerdo migratorio con Estados Unidos, pero esto cambió tras los eventos del 11 de septiembre de 2001. En contraste, la administración de Felipe Calderón se caracterizó por una política opuesta, buscando "desmigratizar" la relación bilateral y enfocándose en la lucha contra el narcotráfico, especialmente a través de la Iniciativa Mérida. Aunque al final de su mandato se promulgó la Ley de Migración de 2011, que representó un avance parcial en materia de extranjería. Por otro lado, la administración de Enrique Peña Nieto continuó con la Ley de Migración y su reglamento, reintegrando la propuesta de incluir un apartado sobre migración en el Plan Nacional de Desarrollo (PND) bajo el nombre de Programa Especial de Migración (PEM). Este programa, desarrollado en consenso con diversas organizaciones de la sociedad civil como el Colectivo Migraciones para las Américas (COMPA), representa una política de estado gestionada por la Secretaría de Gobernación en colaboración con otras entidades gubernamentales (Arp-Nisen, et al., 2019)

Es inaceptable que, sin importar cómo se analice la situación, los migrantes están en una posición extremadamente vulnerable. El notable aumento de la migración desde Centroamérica hacia Estados Unidos provocó una respuesta severa por parte del gobierno mexicano. Se implementó lo que se conoce como la "frontera vertical", que consiste en una serie de estaciones migratorias distribuidas a lo largo del territorio mexicano con el propósito de "controlar" a la población

migrante. Esta medida de control penaliza a los migrantes, a pesar de que su falta de documentos es un asunto del ámbito administrativo.

Algunas medidas que han dificultado el flujo migratorio temporal incluyen (Pérez et al., 2021):

- Implementación de controles fronterizos más estrictos.
- Mayor vigilancia y patrullaje en zonas de cruce frecuente.
- Aplicación de políticas migratorias más restrictivas y sanciones más severas para quienes infringen las leyes de inmigración.

En este escenario, dado que todavía existen migrantes mexicanos y centroamericanos con la intención o necesidad de llegar a Estados Unidos, y las rutas convencionales, que solían ser relativamente seguras, fueron bloqueadas por el control fronterizo, la migración se vio obligada a utilizar senderos más remotos y con condiciones climáticas más extremas. Esto llevó a que los traficantes de personas, conocidos coloquialmente como coyotes o polleros, aumentaran sus tarifas para facilitar el cruce de fronteras (Castillo et al., 2022).

La migración de mexicanos a Estados Unidos de América experimentó un cambio significativo de temporal a permanente debido a varios factores, entre los que destacan el reforzamiento de la frontera México-Estados Unidos y el endurecimiento de las políticas migratorias en ambos países. Estas medidas han dificultado el flujo migratorio temporal y han llevado a que muchos mexicanos opten por establecerse de manera permanente en Estados Unidos, buscando oportunidades económicas, seguridad y estabilidad para sus familias.

En el caso de la migración centroamericana en tránsito por México, se observa un fenómeno similar. La ruta migratoria desde Centroamérica hacia Estados Unidos atraviesa territorio mexicano, y las políticas migratorias restrictivas en la frontera norte de México también han influido en la dinámica migratoria de los centroamericanos. Muchos de ellos, al encontrar mayores obstáculos para llegar a Estados Unidos, han decidido quedarse en México de manera temporal o permanente.

1.2 Proporción de migrantes que arriban y eligen permanecer en territorio mexicano

En el año 2020, México se posicionó como uno de los principales países de origen de migrantes internacionales, ocupando el segundo lugar a nivel mundial después de la India. Esta posición resalta la relevancia de México en el contexto global de la migración, evidenciando la magnitud de los movimientos migratorios que tienen origen en el país (Estadísticas Migratorias Para México. Boletín Anual 2023, s. f.).

La migración hacia México desde Centroamérica y otros continentes ha generado un fenómeno significativo en términos de la cantidad de migrantes que deciden establecerse en el país. Según datos del Instituto Nacional de Migración (INM) de México, de los aproximadamente 140,000 eventos migratorios anuales registrados en el país, alrededor del 78% de los migrantes centroamericanos que llegan a México con la intención de llegar a Estados Unidos optan por

quedarse en territorio mexicano (INM, 2020). Esta cifra resalta la importancia de México como destino final para una parte sustancial de los migrantes que ingresan al país.

Este fenómeno puede explicarse por varios factores. Uno de ellos es la creciente dificultad para cruzar la frontera norte hacia Estados Unidos debido a medidas más estrictas y a la presencia de barreras físicas. Además, en México existen redes de apoyo y oportunidades laborales que resultan atractivas para muchos migrantes. La presencia de comunidades migrantes establecidas y la posibilidad de regularizar su situación migratoria también influyen en la decisión de quedarse en México.

Además, algunos migrantes centroamericanos que llegan a México eventualmente deciden establecerse en el país de manera permanente. Según datos del Instituto Nacional de Estadística y Geografía (INEGI) de México, alrededor del 15% de los migrantes que llegan al país con la intención de llegar a Estados Unidos terminan por establecerse en diversas regiones de México, ya sea por razones laborales, familiares o de seguridad.

Entre los años 2000 y 2020, la cantidad de personas inmigrantes creció un 123% en México. De acuerdo con los datos del censo, en el año 2020, las entidades federativas con las mayores poblaciones de personas migrantes internacionales fueron: Baja California (13%), Ciudad de México (9%), Chihuahua (8%), Jalisco (8%) y Tamaulipas (6%). En 2023, se observaron movimientos migratorios complejos y de rápida evolución en México y la región. Se registró un récord histórico de aproximadamente 44 millones de ingresos regulares al país, lo que representa un aumento del 132% desde el año 2020 (Estadísticas Migratorias Para México. Boletín Anual 2023, s. f.).

Las siguientes figuras se extrajeron del Boletín Anual de la OIM de 2023.

La figura 1, nos muestra los eventos relacionados con personas en situación migratoria irregular reportados por la autoridad migratoria mexicana entre 2017 y 2023.

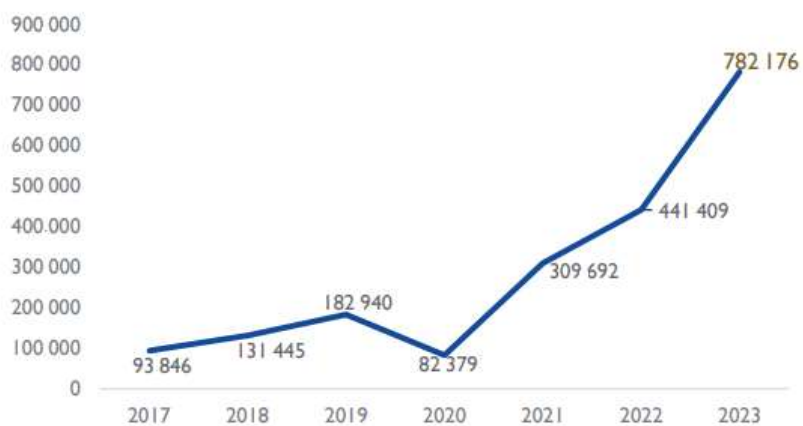


Figura 1. Fuente: Estadísticas Migratorias Para México. Boletín Anual 2023, s. f.

Podemos observar que, en 2023, se registró la mayor cantidad de eventos relacionados con personas en situación migratoria irregular en México, superando el récord anterior de 2022 en un 77%.

La figura 2, nos muestra la cantidad de eventos de personas en situación migratoria irregular registrados por la autoridad migratoria mexicana en cada mes durante 2023.

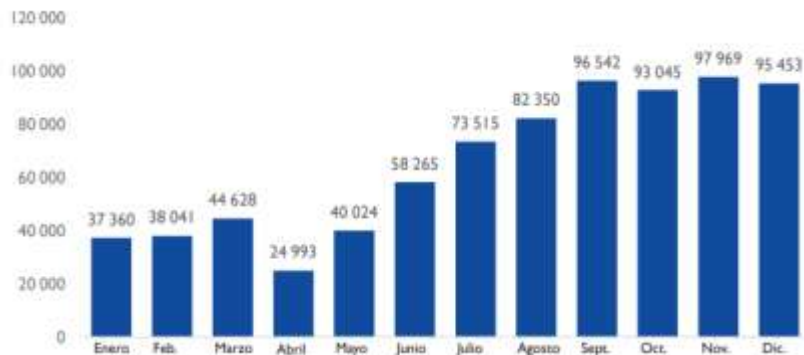


Figura 2. Fuente: Estadísticas Migratorias Para México. Boletín Anual 2023, s. f.

Se puede observar que, durante abril, se notó un número bajo de eventos registrados, seguido por un marcado aumento en los eventos mensuales a partir de mayo. En promedio, se observó un aumento del 14% en los eventos mensuales de junio a diciembre.

La figura 3 muestra la cantidad total de eventos documentados de individuos en condición migratoria irregular, desglosados por género y grupo de edad, desde 2021 hasta 2023.

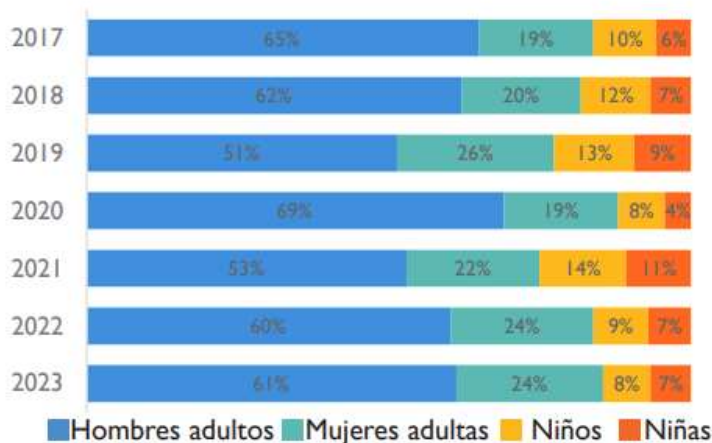


Figura 3. Fuente: Estadísticas Migratorias Para México. Boletín Anual 2023, s. f.

Durante el período comprendido entre 2017 y 2019, se evidencia un aumento significativo del 26% en la proporción de mujeres en situación migratoria irregular, así como un incremento del 22% en la participación de niños y niñas en estas condiciones. Destaca que, en el año 2021, el porcentaje más alto de menores en esta situación alcanzó el 25% del total registrado.

En el año 2023, se documentaron 113,660 casos de niñas, niños y adolescentes (NNAs) en situación migratoria irregular, mostrando un incremento del 60% en comparación con el año anterior, cuando se registraron 71,206 eventos. De este total de 113,660 eventos de NNAs en situación migratoria irregular en 2023, 37,475 (33%) correspondieron a personas de nacionalidad venezolana, 16,596 (15%) de nacionalidad ecuatoriana, 13,925 (12%) de nacionalidad hondureña,

12,370 (11%) de nacionalidad guatemalteca y 7,098 (6%) de nacionalidad colombiana (Estadísticas Migratorias Para México. Boletín Anual 2023, s. f.).

Es relevante resaltar que, en el pasado, la migración estaba principalmente vinculada a hombres, quienes eran vistos como los proveedores del hogar. Sin embargo, en la actualidad, tanto mujeres como menores de edad también participan en este fenómeno migratorio, muchas veces buscando igualar o superar las oportunidades que representa la figura masculina o paterna. Esta transformación ha ampliado nuestra comprensión sobre quiénes integran los flujos migratorios y las diversas motivaciones que impulsan a diferentes grupos a buscar nuevos horizontes fuera de sus lugares de origen.

Entre los migrantes que eligen permanecer en México, algunos deciden buscar asilo en el país. Según datos proporcionados por el Alto Comisionado de las Naciones Unidas para los Refugiados (ACNUR), alrededor del 20% de los migrantes centroamericanos que se quedan en México solicitan oficialmente el estatus de refugiado o protección humanitaria. La cantidad de personas que buscan protección internacional en México está en aumento. En el año 2021, se registró un número récord de más de 131,000 solicitudes de asilo en el país, lo que representa un incremento del 220% en comparación con el año anterior (ACNUR, s.f.). Estas solicitudes son gestionadas por la Comisión Mexicana de Ayuda a Refugiados (COMAR), la cual evalúa cada caso según los criterios establecidos en la legislación mexicana y los tratados internacionales de derechos humanos.

Se requieren la emisión de la Tarjeta de Residente Temporal (TRT), la Tarjeta de Residente Permanente (TRP), la Tarjeta de Visitante por Razones Humanitarias (TVRH), la Tarjeta de Visitante Regional (TVR) y la Tarjeta de Visitante Trabajador Fronterizo (TVFT) para poder permanecer legalmente en México.

- La Tarjeta de Residente Temporal (TRT) se otorga a aquellos que desean ingresar a México como residentes temporales, con la intención de quedarse entre 180 días y 4 años.
- La Tarjeta de Residente Permanente (TRP) se otorga a personas extranjeras que desean residir permanentemente en México.
- La Tarjeta de Visitante por Razones Humanitarias (TVRH) se emite exclusivamente a solicitantes con reconocimiento como refugiados durante el proceso de su solicitud, a migrantes que han sido víctimas de delitos, y a otras personas extranjeras que cumplan con los criterios establecidos en el artículo 52, fracción V de la Ley de Migración. Esta tarjeta se otorga con base en razones humanitarias y especiales circunstancias que justifiquen su emisión.
- La Tarjeta de Visitante Regional (TVR) está destinada a individuos que son residentes u originarios de Guatemala, Belice, El Salvador y Honduras. Esta tarjeta les permite visitar los estados de Chiapas, Campeche, Quintana Roo, Tabasco y Yucatán, con una duración máxima de estancia de 7 días en cada ingreso, sin autorización para trabajar durante ese periodo.
- La Tarjeta de Visitante Trabajador Fronterizo (TVFT) está dirigida a ciudadanos de Guatemala y Belice, quienes pueden llevar a cabo labores remuneradas en la frontera Sur

de México, que incluye los estados de Campeche, Chiapas, Tabasco y Quintana Roo. Esta tarjeta tiene una validez de un año.

En las siguientes gráficas, veremos la solicitud de cada tipo de tarjeta mencionada, así como los países de origen más destacados de quienes solicitan dichas tarjetas.

La figura 4 muestra, en su lado izquierdo, la cantidad de nuevas Tarjetas de Residente Temporal (TRT) emitidas desde 2019 hasta 2023, mientras que en su lado derecho presenta el porcentaje de países que realizan estas solicitudes en el año 2023.

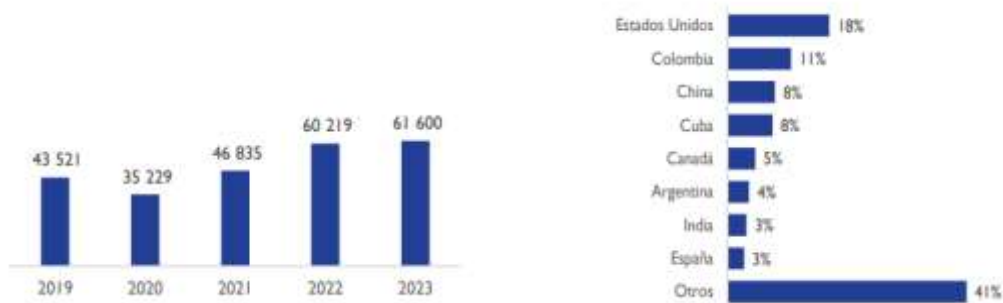


Figura 4. Fuente: Estadísticas Migratorias Para México. Boletín Anual 2023, s. f.

Se puede observar que, en el año 2023, se solicitó la expedición de 61,600 Tarjetas de Residente Temporal (TRT), lo cual representa una disminución de 26,371 tarjetas respecto al año 2020. Entre 2019 y 2020, se observa una reducción de 8,292 TRT emitidas, posiblemente como consecuencia de la pandemia de COVID-19. Los solicitantes de esta tarjeta procedieron de países como Estados Unidos, Colombia, China, Cuba, Canadá, Argentina, India y España.

La figura 5 muestra, en su lado izquierdo, la cantidad de nuevas Tarjetas de Residente Permanente (TRP) emitidas desde 2019 hasta 2023, mientras que en su lado derecho presenta el porcentaje de países que realizan estas solicitudes en el año 2023.

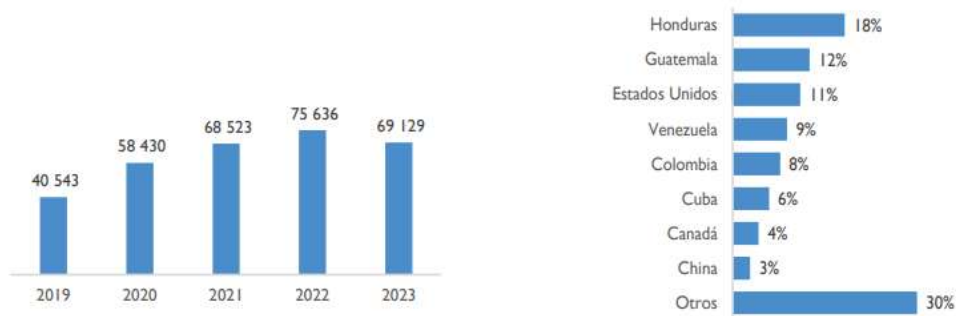


Figura 5. Fuente: Estadísticas Migratorias Para México. Boletín Anual 2023, s. f.

Se puede notar que el año 2022 registró el mayor número de emisiones de esta tarjeta, alcanzando un total de 75,636 tarjetas. Desde el año 2019 hasta el 2022, hubo un aumento constante en las emisiones de la TRP. No obstante, se observa una disminución del 8% entre los años 2022 y 2023. Los solicitantes de esta tarjeta procedieron de países como Honduras, Guatemala, Estados Unidos, Venezuela, Colombia, Cuba, Canadá y China.

La figura 6 muestra, en su lado izquierdo, la cantidad de nuevas Tarjetas de Residente Visitante por Razones Humanitarias (TVRH) emitidas desde 2019 hasta 2023, mientras que en su lado derecho presenta el porcentaje de países que realizan estas solicitudes en el año 2023.

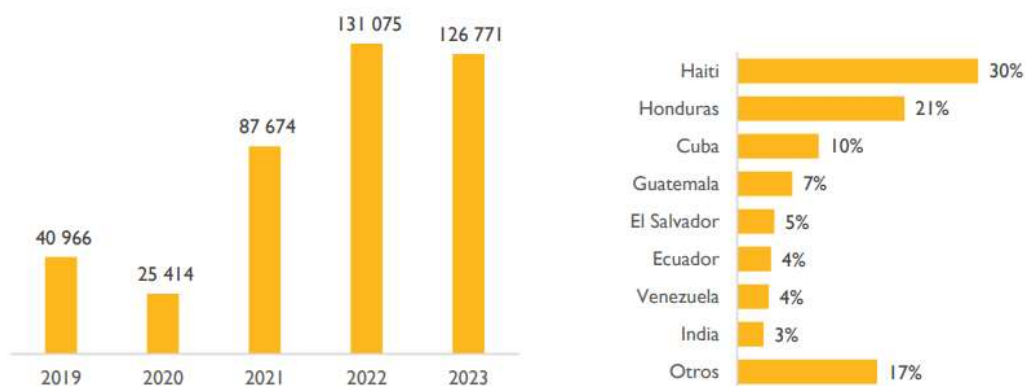


Figura 6. Fuente: Estadísticas Migratorias Para México. Boletín Anual 2023, s. f.

Observamos que en el año 2022 se emitieron 131,075 de estas tarjetas, marcando el punto más alto de emisión entre 2019 y 2023. Además, hubo un aumento del 209% en las TVRH emitidas durante ese periodo. No obstante, este crecimiento no se mantuvo, ya que en 2023 se registró un descenso del 3% con respecto a 2022. Los solicitantes de esta tarjeta procedieron de países como Haití, Honduras, Cuba, Guatemala, El Salvador, Ecuador, Venezuela e India.

La figura 7 muestra la cantidad de nuevas Tarjetas de Visitantes Regionales (TVR) emitidas desde 2019 hasta 2023.

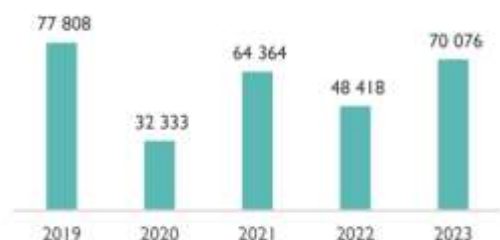


Figura 7. Fuente: Boletín Anual 2023, s. f.

Podemos observar que, en el año 2019, se emitieron un total de 77,808 tarjetas, lo cual representa 7,732 tarjetas más que en el año 2023.

En el 2020, se registraron 32,333 emisiones de esta tarjeta, siendo la cantidad más baja durante el periodo de 2019 a 2023, probablemente debido a la pandemia de COVID-19.

La figura 8 muestra la cantidad de nuevas Tarjetas de Visitantes Trabajadores Fronterizos (TVTF) emitidas desde 2019 hasta 2023.



Figura 8. Fuente: Boletín Anual 2023, s. f.

Se puede observar que en 2019 se emitieron un total de 10,018 tarjetas, lo que supone 3,382 tarjetas más que en 2023. Además, se aprecia una disminución en la emisión de estas tarjetas desde 2020 hasta 2022, manteniendo una media aritmética constante de 3942.67 durante ese periodo.

Es importante resaltar que estas cifras y tendencias pueden variar y están sujetas a cambios que dependen de factores como las políticas migratorias en México y Estados Unidos, las condiciones económicas y sociales en los países de origen y destino, y las dinámicas específicas de los flujos migratorios en la región. Además, estos datos reflejan la complejidad de los movimientos migratorios y la importancia de considerar a México no solo como país de tránsito, sino también como un destino final para muchos migrantes. Esta realidad subraya la necesidad de políticas migratorias integrales que aborden no solo la movilidad transfronteriza, sino también la integración y protección de los migrantes que optan por quedarse en México.

1.3 Los desafíos y riesgos del camino hacia la migración a Estados Unidos

La búsqueda del "sueño americano" ha sido un viaje que ha cautivado a millones de personas en todo el mundo. Sin embargo, detrás de esta aspiración de una vida mejor se esconden una serie de desafíos y riesgos significativos. El camino hacia la migración a Estados Unidos está marcado por una compleja red de dificultades que los migrantes enfrentan desde el momento en que deciden emprender esta travesía.

Los migrantes enfrentan múltiples desafíos que van más allá de los riesgos físicos. La falta de documentos y la entrada irregular a países como México y Estados Unidos los deja expuestos a posibles abusos por parte de traficantes y delincuentes, así como a la vulnerabilidad frente a las autoridades migratorias. Esta incertidumbre sobre su estatus migratorio y el miedo constante a la deportación son aspectos recurrentes en la vida de quienes deciden emprender este viaje.

La movilidad humana, o la situación de los migrantes, está marcada por altos niveles de violencia e impunidad. Su condición de inestabilidad e "ilegalidad" los convierte en objetivos frecuentes del crimen organizado y los deja desamparados frente a la justicia. Esta realidad peligrosa se ha intensificado, convirtiendo el simple acto de migrar en un viaje de alto riesgo. Los migrantes, tanto mexicanos como centroamericanos en tránsito, enfrentan una doble carga: son criminalizados en Estados Unidos por su estatus ilegal y victimizados en México por su condición de migrantes (de la Torre et al., 2019).

Los migrantes experimentan una combinación de peligros físicos, riesgos legales y sociales, así como una situación generalizada de violencia e impunidad. Esta compleja realidad ha convertido el proceso migratorio en un viaje marcado por el temor constante y la incertidumbre, donde la vulnerabilidad y el riesgo son compañeros constantes del camino hacia un futuro mejor.

Este proceso migratorio conlleva un costo tanto en términos financieros como en seguridad personal. Las tarifas exigidas por los traficantes, conocidos peyorativamente como "polleros" o "coyotes", han experimentado un incremento significativo, lo que dificulta aún más el acceso a la migración para aquellos que buscan una vida mejor en Estados Unidos. Esta situación económica adversa se ve agravada por los riesgos asociados con el viaje, como el peligro de ser secuestrado en territorio mexicano, una realidad que muchos migrantes enfrentan en su camino hacia el norte (Palacios et al., 2015).

Uno de los mayores desafíos que enfrentan los migrantes en su camino hacia Estados Unidos es la peligrosa travesía a través de rutas migratorias irregulares. Muchos de ellos recurren a coyotes o

traficantes de personas para cruzar fronteras de manera clandestina, exponiéndose a condiciones extremas y situaciones de riesgo. En el ámbito centroamericano, el término "coyote" engloba a diversos promotores de la migración con fines comerciales, y el "coyotaje" abarca una amplia gama de prácticas relacionadas con facilitar el cruce de fronteras, ya sea de manera no autorizada o semi ilegal. Estas prácticas van desde brindar alojamiento y alimentación a migrantes irregulares, hasta sobornar a autoridades fronterizas y funcionarios de inmigración, y transportar personas a través de puntos no vigilados, además de incluir formas organizadas y transnacionales de tráfico ilícito de migrantes que atraviesan varias fronteras, países e incluso continentes (Kron et al., 2011).

También existen casos excepcionales de migrantes indocumentados que atraviesan la frontera por tierra de manera exitosa sin la asistencia de un coyote. Sin embargo, lo más común y demandado es utilizar los servicios de un coyote, incluso solo para cruzar la frontera entre México y Estados Unidos. Los precios del "coyotaje" difieren según la fama que tenga el "coyote". Los servicios más económicos oscilan entre 1,500 y 2,000 dólares, dirigidos a personas con menos recursos económicos. Para aquellos de nivel económico medio, los precios suelen ser de 2,500 a 3,000 dólares, siendo este el pago más frecuente. Además, el autor menciona un servicio más exclusivo y profesional dirigido a personas de nivel socioeconómico más alto, comenzando en 3,000 dólares por cliente y pudiendo llegar hasta 6,000 dólares, dependiendo de la situación particular del solicitante, el tamaño de la familia que viaja y la urgencia de llegar a su destino una vez en territorio estadounidense (Puerta et al., 2019).

Estos viajes a menudo involucran largas caminatas por terrenos inhóspitos, travesías en balsas precarias a través de ríos peligrosos, e incluso escondites en camiones o trenes de carga. Esta vulnerabilidad física y emocional es solo el comienzo de los desafíos que enfrentan.

Durante los meses de septiembre de 2008 a febrero de 2009, la Comisión Nacional de los Derechos Humanos de México reveló la magnitud del problema del secuestro de migrantes centroamericanos al documentar 198 incidentes de secuestro que afectaron a 9,857 personas (CNDH, 2009). En un informe posterior se indicó que la situación empeoró, con 214 casos de secuestro y 11,333 personas afectadas entre abril y septiembre de 2010, según la Comisión Nacional de los Derechos Humanos (CNDH, 2011). Además, un informe de Amnistía Internacional (2010) mencionaba el aumento generalizado del secuestro de migrantes en México, así como casos de tortura, desaparición y asesinato, junto con el temor de los migrantes a ser entregados a la delincuencia organizada por parte de funcionarios públicos.

Un gran número de migrantes son secuestrados y sufren abusos sexuales, torturas y asesinatos, una situación que actores internacionales han descrito como una tragedia humanitaria (Audio ante la audiencia ante la Comisión Internacional de Derechos Humanos, CIDH). (Aikin et al., 2013)

La explotación laboral y la trata de personas son otros riesgos graves que enfrentan los migrantes en su búsqueda de una vida mejor en Estados Unidos. Muchos de ellos se ven obligados a trabajar en condiciones precarias y sin derechos laborales básicos, siendo víctimas de la explotación y el abuso por parte de empleadores sin escrúpulos. Además, la vulnerabilidad de los migrantes a ser engañados por redes de trata de personas es una realidad preocupante que contribuye a los riesgos asociados con la migración irregular.

No podemos ignorar tampoco los riesgos para la salud mental y emocional que conlleva este viaje migratorio. La separación de las familias, la exposición a situaciones traumáticas durante el viaje y la incertidumbre sobre el futuro generan niveles significativos de estrés, ansiedad y depresión en los migrantes. Estos impactos psicológicos pueden perdurar incluso después de haber alcanzado su destino, afectando su capacidad para adaptarse y prosperar en su nuevo entorno.

Los emigrados enfrentan costos y riesgos adicionales que afectan diversos aspectos de su vida (Puerta et al., 2019):

- Endeudamiento: Al tener que solicitar préstamos para cubrir los gastos de transporte, alojamiento, alimentación, entre otros, muchos migrantes se ven obligados a vender sus pertenencias o utilizarlas como garantía para obtener el préstamo necesario.
- Travesía: Los costos y peligros aumentan especialmente si el migrante decide viajar por tierra sin documentos en regla. Las mujeres enfrentan el riesgo adicional de la explotación sexual, tanto por parte de compañeros de viaje como de traficantes y criminales que se benefician del tráfico de personas, incluyendo a niños, generando ganancias millonarias en la región centroamericana.
- Pérdida del talento local: Las comunidades de origen pierden a sus líderes cuando estos migran, impactando sectores como iglesias, cooperativas, organizaciones comunitarias y especialmente las escuelas, donde los hijos de los emigrados quedan bajo el cuidado de familiares y amigos, generando una nueva dinámica familiar y social.
- Explotación laboral: Muchos empleadores en Estados Unidos se aprovechan de la condición de "ilegalidad" de los migrantes para contratarlos por salarios inferiores al mínimo y sin otorgarles los beneficios laborales que la ley estadounidense garantiza a los trabajadores, como días de enfermedad, vacaciones y protección en caso de accidentes laborales.
- Materialismo y consumismo: La cultura estadounidense vincula el éxito con la posesión de dinero y bienes materiales, fomentando un estilo de vida basado en el consumo y el bienestar material.
- Ilegalidad: Algunos migrantes se ven tentados por opciones laborales vinculadas al crimen organizado, como las maras, el narcotráfico, la prostitución, buscando obtener riquezas rápidas y sin esfuerzo en Estados Unidos.

En resumen, el camino hacia la migración a Estados Unidos está plagado de desafíos y riesgos que van más allá de la simple búsqueda de oportunidades económicas. La vulnerabilidad física, legal y emocional de los migrantes, junto con la explotación laboral y la trata de personas, son aspectos críticos que requieren una atención urgente y soluciones integrales por parte de los gobiernos y la sociedad en general. La protección de los derechos humanos y la garantía de condiciones dignas y seguras para todos los migrantes deben ser objetivos fundamentales en cualquier política migratoria responsable y humanitaria.

1.4 Explorando los flujos migratorios: una mirada a la EMIF

La EMIF es un estudio migratorio amplio realizado por El Colegio de la Frontera Norte (COLEF) en colaboración con la Secretaría del Trabajo y Previsión Social y el Consejo Nacional de Población, a lo largo de la frontera entre Estados Unidos y México. Estas encuestas incluyen un componente que representa de manera significativa a la población general de inmigrantes mexicanos no autorizados que son detenidos en Estados Unidos y luego deportados a México. Estos estudios se llevan a cabo anualmente en México, específicamente en los puntos donde las autoridades migratorias estadounidenses entregan a los deportados de regreso a México (Amuedo-Dorantes et al., 2014).

La EMIF (Encuesta sobre Migración en las Fronteras de México) es una herramienta esencial para comprender la dinámica migratoria en las zonas limítrofes del país. A través de esta encuesta, se recopila información detallada sobre los flujos migratorios, las características de los migrantes, las razones de su desplazamiento y otros aspectos relevantes relacionados con la migración en las fronteras mexicanas.

En El Colegio de la Frontera Norte, la Encuesta sobre Migración en la Frontera Norte (EMIF) nace con el propósito de medir y describir de manera directa los movimientos laborales migratorios entre México y Estados Unidos, así como los desplazamientos hacia y desde las localidades fronterizas del norte de México. Su objetivo es estimar tanto la magnitud del volumen de estos movimientos como las características que los definen. Para lograr esto, la EMIF adapta y emplea técnicas utilizadas en disciplinas como la oceanografía y la biología, las cuales son efectivas para capturar movimientos periódicos, estacionales o cíclicos de entidades que se desplazan de un lugar a otro, siendo aplicadas en este caso para estudiar los desplazamientos laborales en la frontera norte de México (Téllez et al., 1998).

La metodología de las Encuestas sobre Migración en la Frontera (EMIF) se basa en técnicas similares a las usadas en biología para medir movimientos periódicos, estacionales o cíclicos. Se compara el desplazamiento de especies migratorias con los flujos de migrantes humanos entre regiones y países.

Los movimientos migratorios tienden a agruparse en puntos particulares de la frontera, donde el cruce puede ocurrir con o sin documentos. Estos puntos, que a menudo son estrechos como las puertas de una estación de autobuses, permiten medir la magnitud del flujo y describir a las personas involucradas mediante cuestionarios y análisis estadísticos.

La Encuesta sobre Migración en la Frontera se divide en dos partes: la EMIF Norte y la EMIF Sur (colef.mx/emif/):

- La EMIF Norte de México proporciona información precisa y fiable sobre la dinámica, la magnitud y las características de los movimientos migratorios en la frontera entre México y Estados Unidos, así como en algunos aeropuertos del centro del país. Esta encuesta se inició en 1993 y se llevó a cabo de manera continua hasta 2020. Durante 2021, el levantamiento de datos se interrumpió, pero se reanudó en los períodos de julio a noviembre de 2022 y de julio a octubre de 2023.

- La EMIF Sur ofrece información para analizar y describir los flujos migratorios procedentes de Guatemala, Honduras y El Salvador que se trasladan hacia México y/o Estados Unidos con el objetivo de trabajar en estos países. Esta encuesta se ha realizado desde 2004, y al igual que la EMIF Norte, experimentó una interrupción en 2021, retomándose en los períodos de julio a noviembre de 2022 y de julio a octubre de 2023.

Las encuestas son herramientas fundamentales en la investigación científica, especialmente cuando se busca recopilar información cuantitativa sobre un fenómeno en particular dentro de un grupo específico de la población, conocido como muestra. Su propósito principal radica en inferir conclusiones válidas para toda la población a partir de los datos observados en esta muestra, aunque siempre existe cierta incertidumbre asociada a estas conclusiones debido a la naturaleza del muestreo.

En el contexto de las Encuestas sobre Migración en la Frontera (EMIF), se emplea un método de muestreo probabilístico que garantiza que todos los elementos de interés dentro de la población estudiada tengan una probabilidad conocida y mayor que cero de ser seleccionados para participar en la encuesta. Esta metodología permite no solo obtener datos representativos de la realidad migratoria en la frontera, sino también cuantificar la incertidumbre inherente al proceso de inferencia, lo que brinda mayor fiabilidad a las conclusiones derivadas de la investigación.

El fenómeno estudiado es la migración en las fronteras norte y sur de México en un período determinado, usualmente un trimestre. La muestra está compuesta por los movimientos migratorios captados por los encuestadores siguiendo métodos rigurosos de muestreo y medición. Las EMIF subdividen la población de interés en subgrupos para facilitar el análisis y la comprensión del fenómeno migratorio.

Los movimientos migratorios en las fronteras de México pueden ocurrir en dos direcciones: de sur a norte o de norte a sur. Algunos de estos desplazamientos tienen causas identificables, como cuando son personas devueltas por autoridades migratorias. Debido a que quienes participan en desplazamientos similares tienden a tener otras características en común, se agrupan en flujos migratorios según dirección y, en ocasiones, causa. Estos flujos migratorios son la base para definir las poblaciones objetivo en las Encuestas sobre Migración en la Frontera (EMIF), según su destino final.

En las fronteras sur y norte de México, las encuestas sobre migración detectan movimientos que ocurren en dos direcciones distintas: de sur a norte y de norte a sur. Dependiendo de la dirección del movimiento y de si es voluntario o impuesto por las autoridades migratorias, la EMIF Norte identifica tres principales flujos Migratorios.



Figura 9. Fuente: www.colef.mx/emif

Las dos encuestas son responsabilidad de varias instituciones, que incluyen El Colegio de la Frontera Norte (El COLEF), la Secretaría de Gobernación (SEGOB) a través de la Unidad de Política Migratoria, Registro e Identidad de Personas (UPMRIP), el Consejo Nacional de Población (CONAPO), el Consejo Nacional para Prevenir la Discriminación (CONAPRED), la Secretaría del Trabajo y Previsión Social (STPS), la Secretaría de Relaciones Exteriores (SRE), y la Secretaría de Bienestar (BIENESTAR).

Cobertura Geográfica. La amplitud geográfica de esta encuesta varía según la localización del área que se esté encuestando. Esto significa que la extensión territorial abarcada por la encuesta está determinada por la ubicación geográfica de la región que está siendo investigada.

En la EMIF norte se identifican dos tipos de flujos migratorios en función de su dirección. Por un lado, el flujo proveniente del sur abarca ciudades como Tijuana, Mexicali, San Luis Río Colorado, Nogales, Agua Prieta, Ciudad Juárez, Ciudad Acuña, Piedras Negras, Nuevo Laredo, Reynosa y Matamoros, junto con los aeropuertos ubicados en Tijuana, Mexicali, Ciudad Juárez, Nuevo Laredo, Reynosa y Matamoros. Por otro lado, el flujo procedente del norte incluye ciudades como Tijuana, Mexicali, San Luis Río Colorado, Nogales, Agua Prieta, Ciudad Juárez, Ciudad Acuña, Piedras Negras, Nuevo Laredo, Reynosa y Matamoros, junto con los aeropuertos de la Ciudad de México, Guadalajara, Morelia, León y Monterrey.

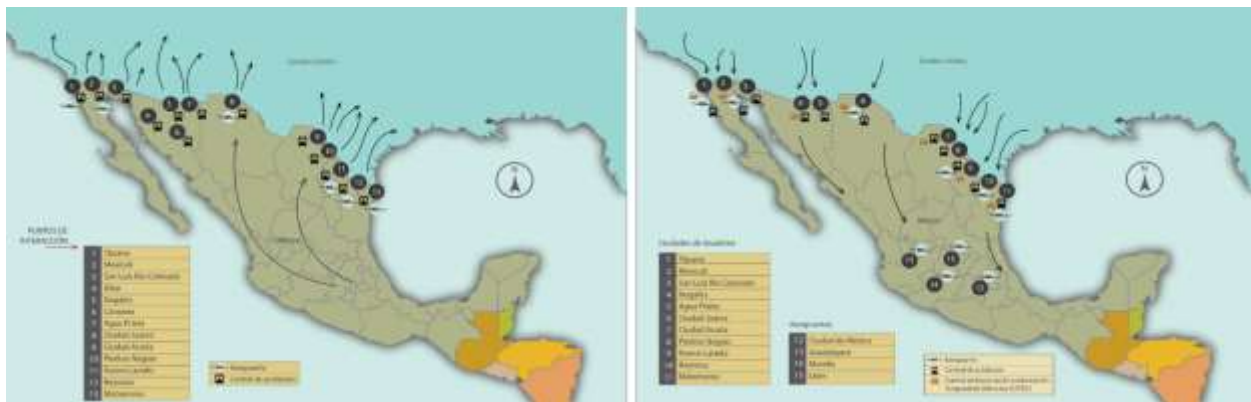


Figura 10. Fuente: www.colef.mx/emif

En la EMIF Sur, se distinguen dos tipos de flujos migratorios según el país de origen de las encuestas realizadas. Por un lado, está el flujo migratorio proveniente de Guatemala, que abarca localidades como Santa Elena, La Mesilla, El Carmen y Tecún Umán. Por otro lado, está el flujo migratorio procedente de México, que incluye las localidades de La Mesilla, El Carmen y Tecún Umán.



Figura. 11 Fuente: www.colef.mx/emif.

La EMIF también abarca las áreas geográficas donde ocurren los procesos de "devolución" de migrantes, tanto en la frontera norte entre Estados Unidos y México:



Figura 12. Fuente: www.colef.mx/emif.

como en la frontera sur entre México y Centroamérica:



Figura 13. Fuente: www.colef.mx/emif.

Así como entre Estados Unidos y Centroamérica:



Figura 14. Fuente: www.colef.mx/emif.

Esto implica que la encuesta incluye no solo los movimientos migratorios en sí, sino también los eventos de devolución que tienen lugar en estas regiones.

Es importante anotar que la EMIF registra movimientos migratorios en lugar de personas individuales, ya que una misma persona puede ser registrada en más de una ocasión. Por ejemplo, si un migrante deportado intenta cruzar a Estados Unidos varias veces y es devuelto por las autoridades, puede ser contado como varias personas distintas, aunque sea la misma persona reincidiendo. Por esta razón, la EMIF se enfoca en los flujos de migrantes en lugar de contar individuos específicos. A diferencia de otras encuestas que recopilan información en hogares o lugares de trabajo fijos, la EMIF recopila datos sobre personas en movimiento que forman parte de los flujos migratorios (Téllez et al., 1998).

La EMIF, bajo la coordinación de diversas instituciones como El COLEF, SEGOB a través de UPMRIP, CONAPO, CONAPRED, STPS, SRE y BIENESTAR, representa un esfuerzo integral para comprender y abordar los fenómenos migratorios en las fronteras norte y sur de México. A través de métodos científicos rigurosos, estas encuestas proporcionan información valiosa sobre los flujos migratorios, contribuyendo así a la formulación de políticas públicas más informadas y efectivas en materia migratoria y poblacional.

1.5 Participación laboral de los inmigrantes

Nuestro enfoque de estudio se centra en la inserción laboral como uno de los principales motivos que conducen a los migrantes a establecerse en un lugar que inicialmente era de tránsito. La inserción laboral no solo proporciona medios de sustento, sino que también marca el inicio de un sentido de pertenencia y arraigo en la comunidad receptora. Además, la posibilidad de una permanencia a largo plazo puede influir en la transformación de los migrantes de "en tránsito" a "temporales", ya que la estabilidad laboral y la perspectiva de futuro en el lugar de destino pueden alterar sus planes migratorios originales.

En México, el ámbito laboral se divide en dos categorías principales: el formal y el informal. En el sector formal, los empleadores ofrecen no solo un salario definido, sino también beneficios como seguridad social y otros privilegios a sus empleados. Esta formalidad se verifica a través del número de trabajadores registrados en el Instituto Mexicano del Seguro Social (IMSS). Sin

embargo, para los migrantes, acceder al mercado laboral formal en el país es difícil debido a la limitada disponibilidad de empleo, incluso para los ciudadanos mexicanos. Además, se enfrentan a obstáculos administrativos relacionados con su estatus migratorio y la documentación requerida por los empleadores. En contraste, el sector informal comprende actividades económicas que operan al margen de las regulaciones y leyes establecidas. Esta falta de regulación conduce a la evasión de impuestos y a que los trabajadores no disfruten de los beneficios y protecciones legales que deberían tener (Mendiburo et al., 2023).

La participación laboral de los migrantes que deciden quedarse en México es un tema complejo que involucra diversos aspectos socioeconómicos y políticos. Según datos del Instituto Nacional de Migración (INM) de México, se observa una creciente tendencia de migrantes que deciden quedarse en el país, ya sea de manera temporal o permanente, debido a diversos factores como la búsqueda de oportunidades laborales, la reunificación familiar o la percepción de seguridad en ciertas áreas de México.

En cuanto al tipo de empleo, muchos migrantes que se quedan en México encuentran trabajo en el sector informal, caracterizado por la falta de seguridad social y condiciones laborales precarias. Esto se debe en parte a la alta incidencia de empleo informal en México, donde más del 50% de la fuerza laboral se encuentra en esta situación, según datos del Instituto Nacional de Estadística y Geografía (INEGI) (INEGI, s.f.). El sector informal incluye actividades económicas no reguladas y no reportadas a las autoridades fiscales y laborales, lo que representa un desafío para garantizar empleos dignos y derechos laborales para todos los trabajadores, incluidos los migrantes.

La mayoría de los empleos que los inmigrantes suelen ocupar se encuentran en el sector de servicios, como trabajos domésticos, limpieza de oficinas y en negocios de comida rápida, supermercados, entre otros. También suelen trabajar en empresas agrícolas y de construcción. Aunque en muchos casos, el salario que reciben es el mínimo establecido por la ley, que es el más bajo posible (Puerta et al., 2019).

Por otro lado, algunos migrantes también logran insertarse en el sector formal de la economía, especialmente aquellos con habilidades específicas o educación que les permite acceder a empleos formales. Sin embargo, esto suele ser más difícil debido a barreras como la falta de reconocimiento de títulos y certificaciones extranjeras, la discriminación laboral y la competencia en ciertos sectores.

El cumplimiento de los requisitos legales, como la inscripción en el Instituto Nacional de Migración (INM), es crucial para que los inmigrantes accedan al empleo formal en México. Esta medida no solo regula la contratación de extranjeros por parte de las empresas, sino que también contribuye a la gestión ordenada de los flujos migratorios, fomentando beneficios sociales, económicos y culturales para el país (INM, s.f.).

En México, la llegada de migrantes se encuentra en interacción directa con los desafíos que el país enfrenta en cuanto a la generación de empleos dignos para todos los ciudadanos en edad activa laboral. A medida que la Población Económicamente Activa (PEA) aumenta cada año, tanto por el crecimiento demográfico como por la llegada de migrantes, se intensifica la competencia por los

puestos de trabajo disponibles. Esto crea tensiones y conflictos entre los diferentes grupos que buscan empleo, lo que puede ejercer presión a la baja en los salarios en ciertos sectores.

Aunque México cuenta con políticas y programas destinados a fomentar la inclusión laboral y proteger los derechos de los trabajadores migrantes, todavía persisten desigualdades y brechas en el mercado laboral. Estas disparidades requieren una atención constante por parte de las autoridades y la sociedad en general. La necesidad de una estructura económica más dinámica capaz de absorber el incremento en la oferta de mano de obra es evidente, especialmente en un contexto donde la migración sigue siendo un factor influyente en la dinámica laboral del país.

La intención de esta investigación es brindar una perspectiva clara sobre la situación laboral de los inmigrantes en México. Estos inmigrantes se definen como personas que no nacieron en el país y que no residían en la región hace doce meses. Este análisis se llevará a cabo mediante la herramienta estadística de la Encuesta Nacional de Ocupación y Empleo (ENOE). Se busca analizar y proporcionar información detallada sobre su participación en el mercado laboral y los niveles salariales que experimentan.

El propósito de esta tesis es comprender las dinámicas laborales de los inmigrantes en México que son identificados a través de una dirección específica. Sin embargo, es importante tener en cuenta que hay una cantidad significativa de inmigrantes en el país que no son captados por esta herramienta, ya que no tienen una dirección particular. Esto sugiere que muchos de ellos enfrentan condiciones precarias y se ven obligados a recurrir a trabajos informales, como la mendicidad, el trabajo sexual, actividades delictivas y otras formas de supervivencia.

La meta de esta investigación es profundizar en las dinámicas laborales de los inmigrantes en México y su conexión con el panorama laboral nacional. Nuestro objetivo va más allá de simplemente analizar sus niveles de participación y salarios, buscamos comprender cómo estas dinámicas se integran en el contexto laboral más amplio del país. Además, nos proponemos ofrecer una visión integral de la situación ocupacional de este grupo de población, lo cual nos permitirá identificar no solo desafíos y oportunidades, sino también las valiosas contribuciones económicas y sociales que los inmigrantes aportan a México.

Capítulo II. Metodología

En este capítulo, se abordarán conceptos teóricos que resultarán fundamentales tanto para la comprensión como para la implementación efectiva de los modelos de regresión.

Presentaremos dos modelos para examinar cómo se relacionan la probabilidad de trabajar de los inmigrantes con sus principales características sociodemográficas, como el sexo, la edad y la educación.

2.1 Motivación de los modelos de regresión empleados en la presente tesis

Los modelos que aplicaremos son: el modelo de regresión lineal con el método de estimación de mínimos cuadrados ordinarios y el modelo de regresión probit.

Consideramos que el modelo probit debido a que tenemos una variable dicotómica (trabajamx), y este modelo implica supuestos que ofrecen tener mayor consistencia en la predicción, a costa de perder eficiencia debido a que la varianza de los coeficientes (betas) aumentan. Por esta razón decidimos contrastarlos con un modelo más simple, como la regresión de mínimos cuadrados ordinarios (MCO). El modelo de regresión de mínimos cuadrados ordinarios es eficiente en términos de tener los mejores estimadores dentro de los modelos lineales, ya que se caracteriza por tener estimadores con mínima varianza y ser insesgados (debido a que no hace supuestos).

Es fundamental entender que cualquier modelo estadístico, no pretende ser una representación absoluta de la realidad; su propósito principal es ser una representación práctica y útil de la realidad. Los modelos nos ayudan a explorar relaciones entre variables y a realizar pronósticos precisos basados en esas relaciones, aun cuando no sean una verdad absoluta. Además, los modelos son temporales y reflejan una versión provisional de nuestras ideas sobre el proceso aleatorio que estamos estudiando. Pueden cambiar basados en análisis con el modelo actual, la selección entre diferentes modelos, la incorporación de nuevos datos, etcétera (Chatterjee et al., 2013).

2.2 Regresión lineal estimada por mínimos cuadrados ordinarios

El método de regresión de mínimos cuadrados ordinarios destaca por su eficiencia al ofrecer estimaciones de alta precisión y libres de sesgo, lo que lo convierte en una elección óptima entre los modelos lineales (BLUE, best linear unbiased estimator).

Este método de MCO es una herramienta fundamental dentro del marco de la regresión lineal ordinaria, considerando que proporciona los coeficientes que describen la relación lineal de manera óptima según el criterio de mínimos cuadrados que definiremos más adelante (Kutner et al., 2004).

Los datos incluyen n conjuntos de observaciones $\{x_{1i}, x_{2i}, \dots, x_{pi}, y_i\}$, tomados de una población más grande de forma aleatoria. Se asume que estas observaciones cumplen una relación lineal (Chatterjee et al., 2013),

$$y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{ip} + \varepsilon_i, \quad (i)$$

siendo los coeficientes β parámetros de naturaleza desconocida, mientras los ε_i son los términos de error aleatorios. Al referirnos a un modelo lineal, estamos indicando que la relación del modelo se establece de manera lineal en relación con estos parámetros.

Cuando tenemos un modelo de regresión simple con una sola variable predictora ($p = 1$), estamos tratando un caso especial que se representa en la figura 15 (Chatterjee et al., 2013):

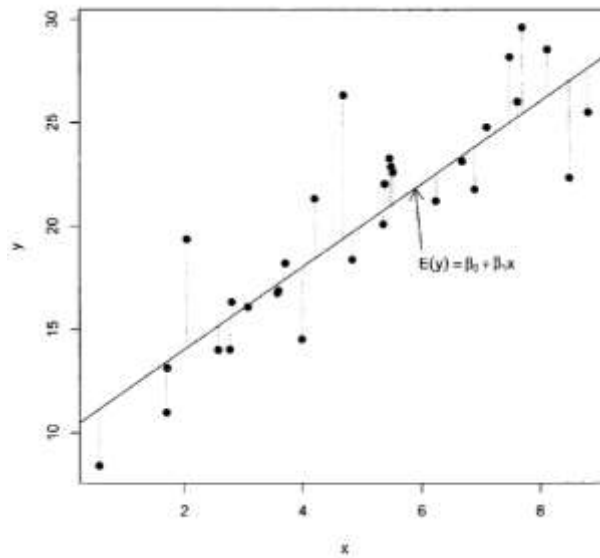


Figura 15. Fuente. Chatterjee et al., 2013

En esta figura, la línea continua representa la verdadera relación de regresión, es decir, el valor esperado de y en función de x . Por otro lado, las líneas de puntos representan los errores aleatorios ε_i , los cuales explican por qué no hay una asociación perfecta entre la variable predictora y las variables objetivo.

2.2.1 Estimación usando el método de mínimos cuadrados ordinarios

La función de regresión verdadera describe la relación que se espera entre las variables predictoras y las variables objetivo, aunque esta relación sea desconocida. Un propósito fundamental del análisis de regresión es calcular esta relación, lo cual implica estimar los parámetros β que no conocemos. Esto se logra utilizando un criterio o regla basada en datos que nos brinde una estimación adecuada. El método estándar para esto es la regresión por mínimos cuadrados, donde las estimaciones se seleccionan de manera que minimicen (Chatterjee et al., 2013):

$$\sum_{i=1}^n [y_i - (\beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi})]^2 . \quad (ii)$$

La figura 16 muestra una representación gráfica de mínimos cuadrados derivada de la figura 15 (Chatterjee et al., 2013):

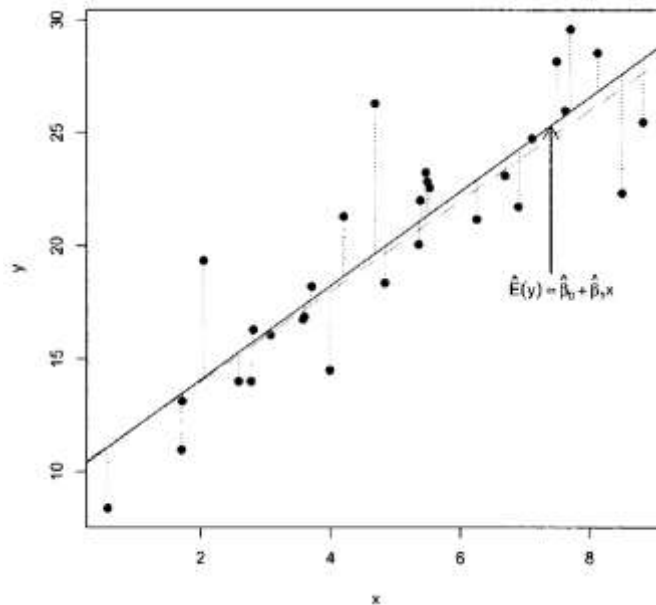


Figura 16. Fuente. Chatterjee et al., 2013

En esta nueva representación, la línea gris representa la verdadera línea de regresión, mientras que la línea continua negra es la línea de regresión estimada. Esta última está diseñada para estimar la línea gris (que es desconocida) de la manera más precisa posible. Para cualquier conjunto de parámetros estimados $\hat{\beta}$, el valor esperado estimado de la respuesta, dado los valores observados de las variables predictoras, se calcula de la siguiente manera (Chatterjee et al., 2013):

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \dots + \hat{\beta}_p x_{pi} ,$$

y se denomina valor ajustado. La diferencia entre este valor ajustado \hat{y}_i y el valor observado y_i se conoce como residuo ($\hat{e}_i = y_i - \hat{y}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{1i}$, Wooldridge et al., 2010). En la figura 16, se representa por las longitudes de las líneas de puntos.

La línea de regresión de mínimos cuadrados se calcula de forma que reduce al mínimo la suma de los cuadrados de las diferencias entre los valores observados y los valores predichos; en otras palabras, las estimaciones de mínimos cuadrados ordinarios (MCO) buscan minimizar la suma de cuadrados de los residuos.

Cuando nos movemos a dimensiones superiores (cuando tenemos más de una variable predictora), las relaciones de regresión verdadera y estimada se representan como planos (cuando $p = 2$) o hiperplanos (cuando $p > 3$), pero los conceptos básicos son los mismos. La figura 17, muestra este caso con dos variables predictoras (Chatterjee et al., 2013):

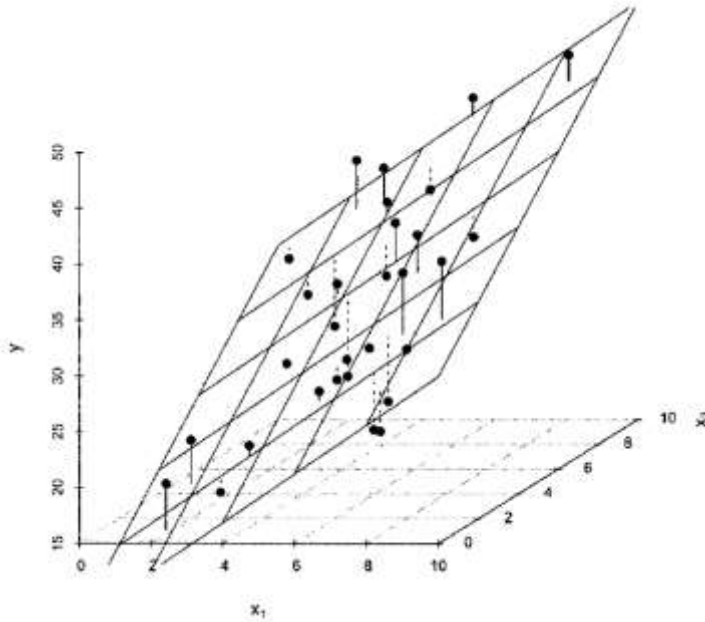


Figura 17. Fuente. Chatterjee et al., 2013

Cada línea vertical representa un residuo, siendo las líneas continuas para residuos positivos y las discontinuas para residuos negativos. El plano de mínimos cuadrados que se ajusta a las observaciones se selecciona para minimizar la suma de los cuadrados de estos residuos.

Podemos expresar el modelo de regresión lineal de manera más concisa utilizando notación matricial. Para ello, definimos la siguiente matriz y vectores de la siguiente manera (Chatterjee et al., 2013):

$$X = \begin{pmatrix} 1 & x_{11} & \dots & x_{p1} \\ \vdots & \vdots & & \vdots \\ 1 & x_{1n} & \dots & x_{pn} \end{pmatrix} \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

El modelo de regresión lineal (i), es entonces (Chatterjee et al., 2013):

$$Y = X\beta + \varepsilon$$

Utilizando cálculo multivariado, se puede demostrar que las ecuaciones normales, las cuales determinan el valor que minimiza (ii), son (Chatterjee et al., 2013):

$$(X'X)\hat{\beta} = X'Y$$

lo que significa que las estimaciones de mínimos cuadrados cumplen con (Chatterjee et al., 2013):

$$\hat{\beta} = (X'X)^{-1}X'Y.$$

Por lo tanto, los valores ajustados son (Chatterjee et al., 2013):

$$\hat{y} = X\hat{\beta} = X(X'X)^{-1}X'Y \equiv HY, \quad (iii)$$

Siendo $H = X(X'X)^{-1}X'$ es definida como la matriz “sombrero”, puesto que toma y a \hat{y} .

Así, los residuos $\varepsilon = y - \hat{y}$ cumplen que (Chatterjee et al., 2013):

$$\varepsilon = y - \hat{y} = y - X(X'X)^{-1}X'y = (I - X(X'X)^{-1}X')Y, \quad (iv)$$

o

$$\varepsilon = (I - H)Y$$

2.2.2 Supuestos

Para que el criterio de mínimos cuadrados produzca resultados sensatos, deben cumplirse los siguientes supuestos (Chatterjee et al., 2013):

- El valor esperado de los errores es nulo ($E(\varepsilon_i) = 0$ para todo i). Esto significa que no puede suceder que el modelo no debe subestimar ni sobreestimar sistemáticamente las observaciones. Si este supuesto se incumple, se presentarán dificultades para estimar β . Además, esto indica que el modelo está omitiendo un componente sistemático esencial, que en su lugar ha sido incorporado en los términos de error.
- La varianza de los errores debe ser constante ($V(\varepsilon_i) = \sigma^2$ para todo i). Esto significa que el modelo no puede ser más preciso para algunas partes de la población (con σ menor) y menos preciso para otras (con σ mayor). Este supuesto de varianza constante se llama homocedasticidad. Si no se cumple se llama heterocedasticidad, puesto que las estimaciones de mínimos cuadrados no serán tan eficientes y los intervalos de confianza y predicción estarán mal calibrados.
- Los errores no deben estar correlacionados entre sí. Esto significa que conocer el error de una observación no debe influir en el error de otra. Esta violación es común en datos de series de tiempo, donde los errores cercanos en el tiempo tienden a ser similares (autocorrelación). Incumplir este supuesto puede resultar en evaluaciones engañosas sobre la precisión de la regresión.
- Los errores deben seguir una distribución normal para realizar pruebas de hipótesis, construir intervalos de confianza o predicción de manera confiable.

Si estos supuestos se incumplen, las pruebas y los intervalos pueden ser poco confiables. Por eso, es crucial verificar estos supuestos usando gráficos, pruebas y diagnósticos en cualquier análisis de regresión.

2.2.3 Explicación de los coeficientes de regresión

Los coeficientes derivados de la regresión de mínimos cuadrados tienen interpretaciones precisas que a veces se malentienden. Por tanto, es crucial comprender claramente lo que representan y lo que no representan (Chatterjee et al., 2013):

- Intercepto $\hat{\beta}_0$: El valor esperado de la variable de interés cuando todas las variables predictoras tienen un valor de cero.
Dado que un valor cero para los predictores podría ser irrealizable o poco probable en los datos observados, carece de sentido asignar una interpretación física a este valor. Ahora bien, si todos los predictores se centran en valores significativos, como una media cero, entonces el intercepto $\hat{\beta}_0$ es igual \bar{Y} en el modelo de regresión lineal, lo que representa la media muestral de los valores objetivo. Observemos que si cada predictor tiene un valor particular significativo y todas las variables están centradas alrededor de esos valores, el intercepto representa una estimación de $E(y)$ cuando todos los predictores tienen esos valores significativos.
- Coeficiente estimado $\hat{\beta}_j$ para el predictor j -ésimo ($j = 1, \dots, p$): El cambio estimado esperado en la variable de interés al variar una unidad en la variable predictora j , manteniendo constantes todos los demás aspectos.
Esta interpretación tiene varios puntos importantes a considerar. Es crucial saber que, aunque la variable objetivo y la variable predictora están asociadas esto no implica que un cambio en la variable objetivo sea directamente causado por un cambio en el predictor, sino que están relacionados entre sí. Esto significa que la correlación no implica necesariamente causalidad.

2.2.4 Problema de estimación

Ahora, considerando n pares de observaciones de y y x , nuestro objetivo es determinar la función de regresión muestral (FRM) de forma que la suma de los residuos (Gujarati et al., 2010)

$$\sum_{i=1}^n \hat{e}_i = \sum_{i=1}^n (y_i - \hat{y}_i) , \quad (vi)$$

sea la más pequeña posible.

Se asigna la misma relevancia o peso a todos los residuos, sin considerar qué tan próximas o dispersas estén las observaciones individuales respecto a la FRM. Como resultado, es muy probable que la suma algebraica de los \hat{e}_i sea bastante pequeña (incluso cero), a pesar de que los \hat{e}_i estén altamente distribuidos (dispersos) en torno de la FRM.

Se puede evitar este problema utilizando el criterio de los mínimos cuadrados, lo cual permite plantear la FRM de la siguiente manera (Gujarati et al., 2010):

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \bar{Y}_i)^2 = \sum_{i=1}^n (\bar{Y}_i - \hat{\beta}_1 - \hat{\beta}_2 x_i)^2 , \quad (vii)$$

puesto que el objetivo es minimizar la suma de los cuadrados de los residuos (e_i^2), es decir, que la suma sea lo más pequeña posible. Como se mencionó anteriormente, al minimizar la $\sum_{i=1}^n \hat{e}_i$, esta

suma puede ser pequeña incluso si los $\hat{\varepsilon}_i$ se encuentran muy dispersos en torno de la FRM. Tomemos en cuenta que cuanto mayor sean los valores absolutos de $\hat{\varepsilon}_i$, será mayor la $\sum_{i=1}^n e_i^2$. Otra razón para utilizar el criterio de método de mínimos cuadrados es que los estimadores obtenidos mediante este método tienen características estadísticas altamente favorables.

2.3 Modelo de probabilidad no lineal: probit

A razón de que el modelo de probabilidad lineal estimado mediante mínimos cuadrados ordinarios proporciona eficiencia (sin supuestos) pero carece de consistencia, fue necesario corregir esta deficiencia mediante la aplicación del modelo de probabilidad no lineal, conocido como probit. Este ajuste nos permite obtener estimaciones más robustas y coherentes al abordar la naturaleza dicotómica de nuestra variable dependiente “trabajamx”.

El modelo probit es un modelo de respuesta binaria en el cual la probabilidad de respuesta se obtiene evaluando la función de distribución acumulativa normal estándar en una función lineal de las variables explicativas (Wooldridge et al., 2010).

Antes que nada, debemos comprender lo que representa una variable latente determinada como (Wooldridge et al., 2010):

$$y_i^* = x_i\beta + \varepsilon_i, \quad (viii)$$

siendo $i = 1, 2, \dots, n$. Y ε_i es independiente de x_i , además de ser $\varepsilon_i \sim Normal(0, 1)$.

De aquí partimos de la premisa que y^* está relacionada de manera lineal con las variables x_i en torno el modelo estructural.

Esta variable latente y_i^* está conectada a la variable y_i a través de la ecuación de medición (Wooldridge et al., 2010):

$$y_i = \begin{cases} 1 & \text{si } y_i^* > 0 \\ 0 & \text{si } y_i^* \leq 0 \end{cases} \quad (ix)$$

Esto implica que, en el modelo probit, se hace una regresión con menos detalles; es decir, el signo de la variable registrada y_i indica si la variable latente y_i^* es positiva o negativa.

De la expresión (ix) podemos afirmar que (Wooldridge et al., 2010):

$$P[y_i = 1|x_i] = P[y_i^* > 0|x_i], \quad (x)$$

puesto que $y_i = 1$ cuando $y_i^* > 0$.

Al reemplazar $y_i^* = x_i\beta + \varepsilon_i$, en la expresión anterior, tenemos que (Wooldridge, JM, 2010):

$$P[y_i^* = 1|x_i] = P[x_i\beta + \varepsilon_i > 0|x_i]. \quad (xi)$$

Restar $x_i\beta$ en cada lado de la desigualdad, tenemos que (Wooldridge et al., 2010):

$$P[y_i^* = 1|x_i] = P[\varepsilon_i > x_i\beta|x_i]. \quad (xii)$$

Cómo las funciones de distribución acumulativas indican la probabilidad de que una variable sea inferior a cierto valor, es necesario invertir la dirección de la desigualdad. Por consiguiente (Wooldridge et al., 2010):

$$P[y_i^* = 1|x_i] = P[\varepsilon_i \leq x_i\beta|x_i]. \quad (xiii)$$

La expresión anterior (xiii), no es más que la función de distribución acumulativa normal estándar (Φ) del error calculada en $x_i\beta$, por último (Wooldridge et al., 2010):

$$P[y_i^* = 1|x_i] = \int_{-\infty}^{x_i\beta} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \Phi(x_i\beta). \quad (xiv)$$

La función de probabilidad acumulativa se escribe de la siguiente manera (Gujarati et al., 2010):

$$Pr(Z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{t^2}{2}} dt$$

2.3.1 Estimación de los parámetros vía máxima verosimilitud

El modelo probit emplea el método de máxima verosimilitud para determinar sus parámetros. Este método busca identificar los valores de los coeficientes que optimizan la probabilidad de que los datos observados se ajusten al modelo propuesto.

Dado que la relación entre la variable dependiente y las variables explicativas no es lineal, se utilizará el método de máxima verosimilitud, que se deriva de la densidad conjunta.

Para especificar la ecuación de verosimilitud, defina k con la siguiente expresión (Long et al., 1997):

$$k_i = \begin{cases} P[y_i^* = 1|x_i] & \text{si } y_i^* = 1 \\ 1 - P[y_i^* = 1|x_i] & \text{si } y_i^* = 0 \end{cases} \quad \text{siendo } i = 1, 2, \dots, n. \quad (xv)$$

El planteamiento de $P[y_i^* = 1|x_i]$ está dada por la forma de (xiv). Cuando las observaciones son independientes, tenemos que la ecuación de verosimilitud es (Long et al., 1997):

$$L(\beta|y, X) = \prod_{i=1}^n k_i \quad (xvi)$$

Uniendo las ecuaciones (xv) y (xvi) tenemos que (Long et al., 1997):

$$L(\beta|y, X) = \prod_{y=1} P[y_i^* = 1|x_i] \prod_{y=0} 1 - P[y_i^* = 1|x_i] \quad , \quad (xvii)$$

siendo $i = 1, 2, \dots, n$.

La multiplicación, nos dice que el producto es tomado en donde los casos son $y = 1$ y $y = 0$, en cada caso.

Las β 's se incorporan dentro de la ecuación de verosimilitud al sustituir el lado derecho de la ecuación anterior (xvii) por la ecuación (xiv), tenemos que (Long et al., 1997):

$$L(\beta|y, X) = \prod_{y=1} \Phi(x\beta) \prod_{y=0} [1 - \Phi(x\beta)] \quad (xviii)$$

Por último, al realizar el cálculo de logaritmos, obtenemos la expresión posterior (Long et al., 1997):

$$\ln L(\beta|y, X) = \sum_{y=1} \ln(\Phi(x\beta)) + \sum_{y=0} \ln[1 - \Phi(x\beta)] \quad (xix)$$

Por lo tanto, los estimadores de máxima verosimilitud del modelo exhiben propiedades de consistencia, normalidad y eficiencia asintóticas. Esto significa que, entre todos los estimadores consistentes para un parámetro, asintóticamente normales y eficientes. Es decir, entre todos los estimadores consistentes de un parámetro β , los de máxima verosimilitud son los de varianza mínima.

2.3.2 Efectos marginales del modelo probit

El propósito fundamental del modelo probit es explicar cómo los efectos de las variables x_i , afectan la probabilidad de éxito $y_i^* = x_i\beta + \varepsilon_i$. La representación de la variable latente resalta la relevancia de los efectos individuales de cada x_i en la variable dependiente y_i^* (Wooldridge et al., 2010).

En el probit la influencia de x_i tanto en la $E(y|x) = P[y_i = 1|x_i]$, como en $E(y^*|x) = \beta_0 + x\beta$, siempre sigue la misma dirección. Para este modelo, buscamos estimar el efecto de x_i sobre $P[y_i = 1|x_i]$, que es nuestra probabilidad de éxito, pero esta tarea se vuelve compleja puesto que $\beta_0 + x\beta$ es intrínsecamente no lineal (Wooldridge et al., 2010).

Las estimaciones de $\hat{\beta}$ pueden emplearse para determinar la variación parcial en la probabilidad de un evento, como se indica en la ecuación (Wooldridge et al., 2010):

$$P[y_i = 1|x_i] = \Phi(x\beta). \quad (xx)$$

Para calcular el cambio parcial (efecto marginal) de variables que son aproximadamente continuas en la probabilidad de respuesta requiere realizar cálculos específicos. Cuando x_i es una variable aproximadamente continua, su efecto parcial en $p(x) = P[y_i = 1|x_i]$ se determina a través de la derivada parcial (Wooldridge et al., 2010):

$$\frac{\partial p(x)}{\partial x_j} = \frac{\partial P[y_i = 1|x_i]}{\partial x_j} = \frac{\partial \Phi(x\beta)}{\partial x_j} = \frac{d\Phi(x\beta)}{dx\beta} \frac{\partial(x\beta)}{\partial x_j} = \Phi(x\beta)\beta_j. \quad (xxi)$$

El efecto marginal del modelo probit, indica como cambia la probabilidad de éxito $P[y_i = 1|x_i]$, ante un cambio marginal en la variable x_i . Esta variación se calcula multiplicando la función acumulativa de distribución normal estándar $\Phi(x\beta)$, por el coeficiente β_j , correspondiente a x_j . Lo cual esta dado por la siguiente expresión (Wooldridge et al., 2010):

$$\frac{\partial P[y_i = 1|x_i]}{\partial x_j} = \Phi(x\beta)\beta_j . \quad (xxii)$$

Capítulo III. Datos ENOE

En este capítulo se abordará de manera integral la base de datos empleada en este estudio. Se explicará el origen de la Encuesta Nacional de Ocupación y Empleo (ENOE), su método de muestreo y la selección de la muestra, así como sus objetivos principales. Además, se detallará la metodología utilizada, incluyendo las preguntas de interés para nuestra investigación.

3.1 Información de interés nacional

La ENOE, considerada de importancia nacional, experimentó una interrupción debido a la pandemia en abril de 2020, reiniciando su recolección de datos regularmente a partir de 2023. Durante el período comprendido entre 2018 y 2022, el INEGI llevó a cabo consultas públicas y evaluaciones de impacto para implementar cambios metodológicos en la encuesta, que incluyeron:

- Agregar preguntas relacionadas con migración y lugar de trabajo.
- Actualizar las categorías de parentesco, carreras, actividades económicas, ocupaciones y ubicaciones geográficas.
- Utilizar entrevistas telefónicas para recopilar información.

En consecuencia, el programa estadístico emergente ENOEn se considera finalizado, y sus resultados se difundieron en diciembre de 2022.

3.2 Encuesta nacional de ocupación y empleo (ENOE 2023)

En 1972, la Dirección General de Estadística comenzó a llevar a cabo encuestas en hogares con la Encuesta Nacional en Hogares (ENH). Posteriormente, se separó el segmento de empleo y se creó la Encuesta Continua de Mano de Obra (ECMO 1973-1974), la cual evolucionó a la Encuesta Continua Sobre Ocupación (ECSO 1974-1984). Para el estudio del mercado laboral, en 1983-1984 se estableció la Encuesta Nacional de Empleo Urbano (ENEU). Durante un tiempo, tanto la ENEU como la ECSO se llevaron a cabo de manera simultánea, hasta que en 1985 se desarrolló un nuevo cuestionario específico para la ENEU, centrado en aspectos como las condiciones laborales y la búsqueda de empleos secundarios.

Durante dos décadas, la Encuesta Nacional de Empleo Urbano (ENEU) estuvo en funcionamiento, realizando ajustes en sus métodos y expandiendo su alcance. En 1994, se llevaron a cabo actualizaciones en los cuestionarios y en los procedimientos de procesamiento de datos, lo que resultó en una mejora en la rapidez con la que se proporcionaban los indicadores clave.

Se estableció la Encuesta Nacional de Empleo (ENE) en 1988 con el propósito de analizar la ocupación en zonas rurales. Esta encuesta proporcionó datos de manera bianual desde 1991 hasta 1995, luego de forma anual con representación por entidad federativa en años pares, y posteriormente de manera trimestral de 2000 a 2004.

Este capítulo se basa en el informe de la ENOE: “INEGI. *Cómo se hace la ENOE. Métodos y procedimientos. 2023*”. Para más detalles, consulte el enlace: <https://www.inegi.org.mx/app/biblioteca/ficha.html?upc=889463909743>

En contraste con varios países de la Organización para la Cooperación y el Desarrollo Económicos (OCDE), en México se iniciaron las encuestas de empleo con alcance urbano para después expandirse a nivel nacional.

El INEGI realizó una revisión exhaustiva del método de recolección de datos de las encuestas de empleo (ENE-ENEU) y desarrolló un nuevo modelo de encuesta denominado Encuesta Nacional de Ocupación y Empleo (ENOE), aprovechando la base sólida y las ventajas de las encuestas previas. Este nuevo enfoque está diseñado según las necesidades actuales del país, teniendo en cuenta las transformaciones ocurridas en el mercado laboral mexicano en los últimos años.

A partir del inicio de 2023, la Encuesta Nacional de Ocupación y Empleo (ENOE) recobra su importancia como un recurso clave de datos a nivel nacional. En esta nueva fase, se incorporan preguntas adicionales sobre migración y empleo, y se emplean tanto entrevistas presenciales como telefónicas como parte de su metodología de recolección de información.

3.2.1 Diseño conceptual

Los objetivos de la ENOE se dividen en generales y particulares para asegurar una representación completa de la información estadística sobre el trabajo en México (ENOE, Cómo Se Hace la ENOE, 2023):

Objetivos Generales:

- Asegurar información estadística básica a nivel nacional, estatal y de las principales ciudades sobre las características laborales de la población.
- Complementar el análisis de las características laborales con información sociodemográfica.
- Ampliar la gama de indicadores estratégicos para comprender la realidad nacional y orientar decisiones políticas laborales.

Objetivos Particulares:

- Describir con precisión la actividad laboral de la población, identificando quiénes están en la fuerza laboral y quiénes no, y los perfiles sociodemográficos correspondientes.
- Identificar a las personas que buscan trabajo, pero no lo encuentran (desempleados).
- Conocer las características de quienes no participan en el mercado laboral (inactivos) y las razones por las que no lo hacen.
- Analizar la naturaleza y calidad de los empleos, incluyendo la diferenciación entre formal e informal.
- Estudiar las características de los lugares de trabajo, como el tamaño, sector de propiedad y productos/servicios.
- Registrar cambios en los empleos y las circunstancias que los rodean.
- Explorar cómo las personas toman decisiones laborales.

La ENOE también tiene como objetivo aumentar la comparabilidad con otros estudios al actualizar sus sistemas de clasificación y al introducir sistemas automatizados para disminuir la incidencia de errores.

3.2.2 Diseño de los instrumentos de captación de la ENOE

La ENOE utiliza dos cuestionarios principales para recopilar información: el Cuestionario Sociodemográfico (CS) y el Cuestionario de Ocupación y Empleo (COE). El CS se enfoca en obtener datos sociodemográficos de los residentes de las viviendas seleccionadas, mientras que el COE se centra en recabar información sobre la actividad laboral de la población de 12 años en adelante y las características de las unidades económicas donde trabajan.

Cada vivienda en la muestra es visitada en cinco ocasiones cada tres meses, según el diseño estadístico establecido. En cada visita, se aplican ambos cuestionarios, comenzando con el CS seguido del COE. Es importante destacar que en la primera visita se recopila toda la información sociodemográfica inicial, y en las visitas posteriores se actualizan solo algunos rubros específicos de información. Esto asegura una recopilación exhaustiva y actualizada de los datos relevantes a lo largo del tiempo.

El Cuestionario Sociodemográfico aborda varios temas relacionados con la ubicación de las viviendas seleccionadas, el número de hogares y residentes en estas, la disposición del trabajo doméstico y la información sociodemográfica de los residentes por grupos de edad y pertenencia a un hogar específico.

El concepto de hogar se fundamenta en la convivencia y en el presupuesto compartido para la alimentación. Un hogar, también conocido como unidad doméstica, es una estructura social formada por personas que viven juntas, ya sea con o sin lazos de parentesco, y comparten una misma vivienda (incluso una sola persona puede constituir un hogar).

La información sobre los residentes de un hogar se organiza según su relación con la persona que encabeza el hogar, determinada por el reconocimiento de los demás miembros. Esto permite clasificar los hogares en tipos (familiares y no familiares) y clases (nucleares, ampliados o compuestos en el caso de familiares, y unipersonales o de corresidentes en el caso de no familiares).

El concepto de hogar se refleja en la variable de gasto común, que incluye dos preguntas complementarias: gasto común y número de hogares. Esta información es fundamental para estudios sobre distribución de ingresos, pobreza y desigualdad, especialmente en hogares con jefatura femenina, que suelen ser más vulnerables económicamente. La importancia de estos análisis ha llevado a una mayor presencia del tema en el Sistema Nacional de Información Estadística.

La estructura temática del Cuestionario de Ocupación y Empleo (COE) difiere del Cuestionario Sociodemográfico (CS) en varios aspectos. Mientras que el CS se aplica una vez que se ha recopilado toda la información sociodemográfica del hogar, el COE se dirige a cada miembro individualmente, a partir de los 12 años, aunque los indicadores relacionados con la Población Económicamente Activa (PEA) y la Población No Económicamente Activa (PNEA) se basen en los datos de personas de 15 años en adelante, una distinción importante para usuarios que utilizan la base de datos basándose en la edad mínima previa.

3.2.3 Cobertura geográfica

La ENOE está estructurada en tres áreas principales de estudio. El primer sector engloba las 39 ciudades auto-representadas, originadas a partir de la Encuesta Nacional de Empleo Urbano (ENEU). El segundo sector abarca las áreas urbanas de alta densidad, complementando el primero, mientras que el tercer sector se enfoca en las zonas rurales. A continuación, se describen detalladamente cada uno de estos sectores de análisis (ENOE, 2023):

- Ciudades auto representadas: Se inicia con la referencia al Sistema Urbano Nacional 2018 (SUN) establecido por el Consejo Nacional de Población (CONAPO) y la Secretaría de Desarrollo Agrario, Territorial y Urbano (SEDATU), donde se clasifican las ciudades del país en tres categorías: Zonas Metropolitanas (tipo 1), conurbaciones (tipo 2) y centros urbanos (tipo 3). Esta clasificación se basa en criterios como el tamaño, la estructura funcional y la continuidad física entre las localidades.

En el contexto de la Encuesta Nacional de Ocupación y Empleo (ENOE), se seleccionan una o más ciudades de cada una de las 32 entidades federativas de acuerdo con su relevancia, siendo consideradas auto representadas y con un tamaño de muestra mínimo para generar resultados independientes. Los criterios para esta elección incluyen aspectos políticos, de desarrollo económico, crecimiento anual y desarrollo urbano, en consonancia con la clasificación del SUN.

El grupo de ciudades auto representadas incluye 39 localidades, principalmente formando áreas metropolitanas, mayormente conformando zonas metropolitanas, salvo Ciudad del Carmen que se identifica como un centro urbano. Este conjunto representa alrededor del 45.7% de la población nacional

- Complemento urbano de alta densidad: El área urbana densamente poblada en la ENOE se subdivide en tres categorías: la primera comprende ciudades con una población de 100,000 habitantes o más que no son auto representadas; la segunda incluye localidades con entre 15,000 y 99,999 habitantes; y la tercera abarca localidades con entre 2,500 y 14,999 habitantes. Esta área adicional representa aproximadamente el 31.2% de la población nacional.
- Rural: El área rural abarca poblaciones con menos de 2,500 habitantes y constituye alrededor del 23.1% de la población nacional.

La ENOE proporciona resultados desagregados por nivel nacional, entidad federativa, ciudades de interés y diferentes tamaños de localidades para cada trimestre del año.

3.2.4 Marco de la encuesta

La estructura de muestreo de la ENOE se define por ser probabilística, lo que implica que los hallazgos de la encuesta se extrapolan a toda la población. Además, se trata de un diseño bietápico, estratificado y basado en conglomerados, enfocado en individuos de 15 años o más en el momento de la entrevista.

Un diseño bietápico implica llevar a cabo el muestreo en dos fases separadas. En la primera fase, se eligen conjuntos o conglomerados de elementos, mientras que, en la segunda fase, se seleccionan elementos individuales dentro de esos conjuntos o conglomerados para conformar la muestra.

La selección de la submuestra de la encuesta se realizó empleando la Muestra Maestra 2012, la cual se fundamentó en el Marco Maestro de Muestreo 2012 del INEGI. Este marco se elaboró utilizando información cartográfica y demográfica del Censo de Población y Vivienda 2010.

La Muestra Maestra se adhiere a un diseño de muestreo probabilístico, estratificado, unietápico y por conglomerados, en el cual los conglomerados son las unidades principales de muestreo.

Un diseño de muestreo unietápico implica que la muestra se selecciona en una sola fase. Esto significa que los elementos individuales que integrarán la muestra se eligen directamente, sin dividir el proceso en múltiples etapas de selección.

3.2.5 Formación de las Unidades Primarias de Muestreo (UPM)

Las unidades primarias de muestreo son conjuntos de viviendas que comparten características específicas basadas en su ubicación geográfica, como se describe a continuación (ENOE, 2023):

a. En urbano alto (áreas urbanas densamente pobladas):

La cantidad mínima de viviendas ocupadas en una unidad primaria de muestreo es de 80, mientras que el máximo es de 160. Estas unidades pueden incluir:

- Una cuadra.
- La combinación de dos o más cuadras contiguas dentro de la misma área geográfica específica (AGEB, Área Geoestadística Básica).
- La combinación de dos o más cuadras contiguas de diferentes AGEB específicas dentro de la misma localidad.
- La combinación de dos o más cuadras contiguas de diferentes localidades, siempre y cuando tengan el mismo tamaño de localidad.

b. En áreas de complemento urbano:

La cantidad mínima de viviendas ocupadas en una unidad primaria de muestreo es de 160, mientras que el límite máximo es de 300. Estas unidades pueden incluir:

- Una cuadra.
- La combinación de dos o más cuadras contiguas dentro de la misma área geográfica específica (AGEB, Área Geoestadística Básica).
- La combinación de dos o más cuadras contiguas de diferentes áreas geográficas básicas específicas dentro de la misma localidad.

c. En rural:

La cantidad mínima de viviendas ocupadas en una unidad primaria de muestreo es de 160, en contraste con el número máximo de 300 viviendas. Estas unidades pueden incluir:

- Una cuadra.
- Fragmento de un área geográfica básica (AGEB, Area Geoestadística Básica).
- La combinación de dos o más cuadras colindantes del mismo municipio.
- La fusión de un área geográfica específica (AGEB) con una parte de otra AGEB colindante del mismo municipio.

3.2.6 Estratificación

Después de establecer las UPM, se clasifican aquellas con atributos similares como vemos a continuación.

La estructura administrativa nacional y la variedad de localidades en función de su tamaño contribuyen a una estratificación geográfica natural. Dentro de cada estado, la muestra se focaliza específicamente en áreas urbanas con alta densidad poblacional, las cuales se dividen en distintas zonas, como se explica en la tabla que sigue (ENOE, 2023):

Cuadro 1. Estratificación. Fuente: ENOE, 2023

Ámbito	Zona	Tamaño de localidad
Urbano alto	01	Ciudades con 100 000 o más habitantes.
	... 09	
Complemento urbano	25	De 50 000 a 99 999 habitantes.
	35	De 15 000 a 49 999 habitantes.
	45	De 5 000 a 14 999 habitantes.
	55	De 2 500 a 4 999 habitantes.
Rural	60	Localidades menores de 2 500 habitantes.

Se establecieron cuatro categorías sociodemográficas que engloban todas las Unidades Primarias de Muestreo (UPM) del país. Esta clasificación se fundamenta en aspectos como la composición demográfica de los habitantes y las condiciones físicas y de equipamiento de las viviendas, representadas por 34 indicadores. Estos indicadores se elaboraron a partir de datos del Censo de Población y Vivienda 2010 y técnicas estadísticas avanzadas.

Se asignó cada Unidad Primaria de Muestreo (UPM) a una única categoría que considera aspectos geográficos y sociodemográficos, resultando en un total de 683 grupos en todo el país.

3.2.7 Tamaño de la muestra

Se estableció el tamaño de la muestra para cada área de estudio a partir de la Tasa de Desempleo Abierto (TDA), la cual es la variable principal de la encuesta y requiere la muestra más amplia. Esto garantiza una representación adecuada de las estimaciones de otras variables de interés con este tamaño de muestra. La fórmula utilizada para calcularlo es la siguiente (ENOE, 2023):

$$n = \frac{z^2 \cdot q \cdot \text{DEFF}}{r^2 \cdot p \cdot (1 - \text{tnr}) \cdot \text{TNP} \cdot \text{PHV}} ,$$

siendo:

- n = el tamaño de la muestra,
- p = el cálculo de la proporción deseada: la tasa de empleo activo,
- q = 1-p
- r = el máximo error relativo permitido,
- z = la cantidad registrada en las tablas estadísticas de la dist. normal estándar que asegura obtener las estimaciones con una confianza previamente establecida,
- DEFF = el factor de diseño determinado como la relación entre variabilidad en la estimación del diseño empleado y la variabilidad obtenida al considerar un muestreo aleatorio simple con el mismo tamaño de muestra,
- tnr = la máxima tasa de personas que se espera no responda,
- TNP = el porcentaje de participación efectiva y,
- PHV = la media de personas de 15 años o más por hogar.

Con base en un nivel de confianza del 90 %, un efecto de diseño de 4.00 observado en investigaciones previas, una tasa de participación neta del 60.4 %, un promedio de 2.70 personas por hogar, un margen de error relativo máximo esperado del 5.05 % para una tasa de desempleo abierto (TDA) del 2 % y una tasa de no respuesta del 15 %, se optó por seleccionar una muestra nacional de 150,033 viviendas, la cual fue ajustada a 150,449 hogares.

La Encuesta Nacional de Ocupación y Empleo (ENOE) se divide en dos submuestras: una que utiliza entrevistas cara a cara con 132,383 viviendas y otra que emplea entrevistas telefónicas a partir de la segunda entrevista con 18,066 viviendas. A nivel estatal, los tamaños de muestra varían según la distribución geográfica y el número de ciudades representativas en la ENOE; se establece un mínimo de 3,460 viviendas por entidad para obtener resultados estadísticamente confiables. Este tamaño de muestra garantiza una cobertura adecuada para tasas de desempleo iguales o superiores al 5 % con un nivel de confianza del 90 %, un error máximo esperado del 13.67 %, un efecto de diseño de 1.68, una tasa de no respuesta máxima del 15 %, un promedio de 2.6 habitantes por hogar mayores de 15 años y una tasa de participación del 60 %.

En las ciudades representativas, los tamaños de muestra varían según la población y la actividad económica, estableciendo un mínimo de 1,965 viviendas para obtener resultados confiables. Este tamaño de muestra permite abarcar tasas de desempleo del 5 % o más con un nivel de confianza del 90 %, un margen de error máximo esperado del 14.96 %, un efecto de diseño de 1.15, una tasa de no respuesta máxima del 15 %, un promedio de 2.47 habitantes por hogar mayores de 15 años y una tasa de participación del 64 %. La mayoría de las ciudades cuentan con muestras que superan las 2,100 viviendas, siendo solo algunas, como Guadalajara, León, Monterrey, Puebla y Torreón, las que tienen más de 3,000 viviendas en la muestra. Por otro lado, el Área Metropolitana de la Ciudad de México cuenta con una muestra de 6,149 viviendas.

La distribución de la muestra en cada estado se realizó de manera equitativa entre los diferentes estratos en función de su tamaño, siguiendo la fórmula siguiente (ENOE, 2023):

$$n_{eh} = \frac{N_{eh}}{N_e} n_e .$$

La fórmula para calcular la cantidad de unidades primarias de muestreo (UPM) a seleccionar es la siguiente (ENOE, 2023):

$$k_{eh} = \frac{n_{eh}}{b} ,$$

siendo:

- n_{eh} = la cantidad de hogares en muestra en el h-ésimo estrato, en la e-ésima entidad,
- n_e = la cantidad de hogares en muestra en la e-ésima entidad, según su distribución proporcional,
- N_{eh} = la cantidad neta de hogares en el h-ésimo estrato, en la e-ésima entidad,
- N_e = la cantidad neta de hogares en la e-ésima entidad,
- k_{eh} = la cantidad de unidades primarias de muestreo en muestra en el h-ésimo estrato, en la e-ésima entidad y,
- b = la cantidad de hogares a seleccionar por unidades primarias de muestreo.

Las UPM elegidas para la ENOE conforman la muestra principal que se utilizará para crear submuestras en todas las encuestas de viviendas del INEGI. Estas UPM fueron seleccionadas utilizando un método que considera su tamaño relativo en relación con el total de UPM disponibles, lo que se conoce como probabilidad proporcional al tamaño (ENOE, 2023):

$$P\{u_{ehi} \in S_m\} = \frac{k_{eh}m_{ehi}}{m_{eh}} ,$$

siendo:

- u_{ehi} = la i-ésima unidad primaria de muestreo, del h-ésimo estrato, de la e-ésima entidad,
- S_m = la muestra principal,
- k_{eh} = la cantidad de unidades primarias de muestreo en la muestra principal, en el h-ésimo estrato, en la e-ésima entidad,
- m_{ehi} = el número de viviendas en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad en el Censo de Población y Vivienda 2010 y,
- m_{eh} = el número de hogares en el h-ésimo estrato, en la e-ésima entidad en el CPV 2010.

En la fase siguiente, la elección de viviendas se ajustó según el entorno al que pertenecen, como se detalla a continuación.

En urbano alto

En áreas urbanas densamente pobladas, la selección de la muestra se llevó a cabo de manera individual para cada entidad y estrato, utilizando el siguiente método:

- De las k_{eh} Unidades Primarias de Muestreo (UPM) incluidas en la muestra principal en el estrato h-ésimo, se seleccionaron k_{eh} UPM para la Encuesta Nacional de Ocupación y Empleo (ENOE) en la entidad e-ésima.
- Se seleccionaron cinco hogares dentro de cada UPM con la misma probabilidad para participar en la ENOE.

La fórmula para calcular la probabilidad de seleccionar la i-ésima UPM en el estrato h-ésimo y en la entidad e-ésima, utilizando la probabilidad proporcional al tamaño en la muestra maestra es (ENOE, 2023):

$$P_{1ehi} = \frac{k_{eh}m_{ehi}}{m_{eh}} .$$

La fórmula para calcular la posibilidad de elegir la j-ésima vivienda de la i-ésima UPM en el h-ésimo estrato de la e-ésima entidad es (ENOE, 2023):

$$P_{2ehij} = \frac{5}{m_{eh}^*} .$$

De esta manera, la probabilidad general de seleccionar la j-ésima vivienda en la i-ésima UPM, en el h-ésimo estrato y en la e-ésima entidad se calcula multiplicando las probabilidades individuales de selección (ENOE, 2023):

$$P_{ehij} = P_{1ehi} * P_{2ehij} = \frac{k_{eh}m_{ehi}}{m_{eh}} \frac{5}{m_{eh}} = \frac{5k_{eh}m_{ehi}}{m_{eh}m_{ehi}^*} .$$

El factor de expansión se calcula de la siguiente manera (ENOE, 2023):

$$F_{ehij}^{UA} = \frac{m_{eh}m_{ehi}^*}{5k_{eh}m_{ehi}} ,$$

siendo:

- k_{eh} = la cantidad de UPM seleccionadas en el h-ésimo estrato y la e-ésima entidad para el marco de la muestra principal,
- m_{ehi} = la cantidad de hogares en la i-ésima UPM, en el h-ésimo estrato y la e-ésima entidad, según los datos del Censo de Población y Vivienda 2010,
- m_{eh} = la cantidad de hogares en el h-ésimo estrato y la e-ésima entidad y,
- m_{ehi}^* = la cantidad d de hogares en la i-ésima UPM, en el h-ésimo estrato y la e-ésima entidad, según la lista de hogares actualizada.

El factor de expansión se describe como el valor obtenido al invertir la probabilidad de selección. Según la Norma Técnica del Proceso de Producción de Información Estadística y Geográfica del Instituto Nacional de Estadística y Geografía, este factor se denomina "Ponderador".

En complemento urbano

- Se seleccionaron k_{eh} unidades primarias de muestreo (UPM) de la muestra principal en el h-ésimo estrato y la e-ésima entidad, se eligieron k_{eh} UPM para la ENOE.
- Dentro de cada UPM seleccionada, se eligieron 20 viviendas con la misma probabilidad.

La probabilidad de seleccionar la i-ésima UPM en el h-ésimo estrato y la e-ésima entidad, utilizando una probabilidad proporcional al tamaño para la muestra principal, es (ENOE, 2023):

$$P_{1ehi} = \frac{k_{eh}m_{ehi}}{m_{eh}} .$$

La probabilidad de elegir la j-ésima vivienda de la i-ésima UPM $_{ehi}$, en el h-ésimo estrato y la e-ésima entidad es la siguiente (ENOE, 2023):

$$P_{2ehij} = \frac{20}{m_{ehi}^*} .$$

Entonces, la probabilidad total de elegir la j-ésima vivienda, en la i-ésima Unidad Primaria de Muestreo (UPM), en el h-ésimo estrato y en la e-ésima entidad es igual al resultado de multiplicar las probabilidades individuales de selección (ENOE, 2023):

$$P_{ehij} = P_{1ehi} * P_{2ehij} = \frac{k_{eh}m_{ehi}}{m_{eh}} \frac{20}{m_{eh}m_{ehi}^*} .$$

El factor de expansión se determina como (ENOE, 2023):

$$F_{ehij}^{CU} = \frac{m_{eh}m_{ehi}^*}{20 k_{eh} m_{ehi}} ,$$

siendo:

- k_{eh} = la cantidad de unidades primarias de muestreo (UPM) seleccionadas en el estrato h-ésimo y en la entidad e-ésima para el marco de la muestra principal.
- m_{eh} = la cantidad de hogares en el estrato h-ésimo y en la entidad e-ésima.
- m_{ehi} = la cantidad de hogares en la UPM i-ésima, en el estrato h-ésimo y en la entidad e-ésima, según los datos del Censo de Población y Vivienda 2010.
- m_{ehi}^* = la cantidad de hogares en la UPM i-ésima, en el estrato h-ésimo y en la entidad e-ésima, según el listado de viviendas.

En rural

- De las k_{eh} unidades primarias de muestreo que forman la muestra principal en el estrato h-ésimo y en la entidad e-ésima, se escogieron k_{eh} UPM para la ENOE

- En cada una de las k_{eh} unidades primarias de muestreo seleccionadas, se eligieron cuatro grupos de viviendas, cada uno con aproximadamente cinco viviendas, de manera equitativa.

La posibilidad de seleccionar la i -ésima UPM en el h -ésimo estrato y la e -ésima entidad mediante la aplicación de una probabilidad proporcional al tamaño en la muestra principal es (ENOE, 2023):

$$P_{1ehi} = \frac{k_{eh}m_{ehi}}{m_{eh}} .$$

La posibilidad de elegir la vivienda número j -ésima de la UPM $_{eih}$, en el estrato h -ésimo y la entidad e -ésima es (ENOE, 2023):

$$P_{2ehij} = \frac{4 * 5}{m_{ehi}^*} \frac{20}{m_{ehi}^*} .$$

Entonces, la probabilidad total de elegir la vivienda j en la UPM i , en el estrato h y en la entidad e , se obtiene multiplicando las probabilidades de selección individuales (ENOE, 2023):

$$P_{ehij} = P_{1ehi} * P_{2ehij} = \frac{k_{eh}m_{ehi}}{m_{eh}} \frac{20}{m_{ehi}^*} = \frac{20 k_{eh}m_{ehi}}{m_{eh}m_{ehi}^*} .$$

El factor de expansión se determina como (ENOE, 2023):

$$F_{ehij}^R = \frac{m_{eh}m_{ehi}^*}{20 k_{eh} m_{ehi}} ,$$

siendo:

- k_{eh} = la cantidad de Unidades Primarias de Muestreo (UPM) seleccionadas en el estrato h -ésimo y en la entidad e -ésima para el marco de la muestra principal.
- m_{eh} = la cantidad de hogares en el estrato h -ésimo y en la entidad e -ésima.
- m_{ehi} = la cantidad de hogares en la UPM i -ésima, en el estrato h -ésimo y en la entidad e -ésima, según los datos del Censo de Población y Vivienda 2010.
- m_{ehi}^* = la cantidad de hogares en la UPM i -ésima, en el estrato h -ésimo y en la entidad e -ésima, según el listado de viviendas.

3.2.8 Estructura de la base de datos

Definimos la investigación presente, gracias a las bases de datos que nos proporciona la ENOE, la cual es la principal fuente de datos sobre el mercado laboral en México, ofreciendo información detallada mensual y trimestral sobre la fuerza laboral, ocupación, informalidad laboral, subocupación y desempleo. Es un proyecto estadístico continuo de gran alcance, brindando cifras a nivel nacional y para diversas categorías geográficas.

La Encuesta Nacional de Ocupación y Empleo (ENOE) tiene como objetivo principal generar indicadores fundamentales sobre la situación laboral y ocupacional de la población mexicana. Para lograr esto, utiliza dos herramientas de recolección de datos: el Cuestionario Sociodemográfico (CS) y el Cuestionario de Ocupación y Empleo (COE), este último disponible en versiones básica y ampliada.

El espacio tiempo de interés es el año 2023, debido a que la ENOE ofrece los datos de manera trimestral, se descargaron del apartado de microdatos: I trimestre, II trimestre, III trimestre y IV trimestre. Estos datos se descargan en zip, cada zip contiene cinco bases de datos: COE1, COE2, HOG, SDEM y VIV. También se descargaron los cuestionarios que corresponden a las bases de datos: cuestionario_amp_v6a, cuestionario_bas_v7 y cuestionario_sdem_v5a.

3.3 Procesamiento de los datos de la base de la ENOE

Ahora bien, se realizó una tabla con las preguntas que fueron esenciales para la realización de esta investigación. La tabla contiene la ubicación de las preguntas en la base de datos, es decir, el nombre de la variable que se le asigna a la pregunta, en qué base de datos se encuentra dicha variable y si los trimestres del 2023 cuentan con esta variable en sus respectivas bases de datos.

Preguntas de la ENOE	Pregunta	Base de datos	1er trimestre-2023	2do trimestre-2023	3er trimestre-2023	4to trimestre-2023
<i>Datos que nos interesan</i>						
Lugar de nacimiento	<i>l_nac_e</i>	SDEM	Si	Si	Si	Si
Sexo	<i>sex</i>	SDEM	Si	Si	Si	Si
Edad	<i>edo</i>	SDEM	Si	Si	Si	Si
Hasta qué grado aprobó... en la escuela	<i>es_pl3_1</i>	SDEM	Si	Si	Si	Si
Hace un año, en (mes) de (año), ¿en qué estado de la República o en qué país vivía?	<i>es_p20r_1</i>	SDEM	Si	Si	Si	Si
¿La semana pasada trabajó por lo menos una hora?	<i>pl</i>	COET1	Si	Si	Si	Si
Opción 1 ¿... ha tratado de buscar trabajo en otro país o hacer preparativos para cruzar la frontera?	<i>p2_1</i>	COET1	Si	Si	Si	Si
Opción 2 ¿... ha tratado de buscar trabajo aquí en el país?	<i>p2_2</i>	COET1	Si	Si	Si	Si
¿Cuánto ganó o en cuánto calcula sus ingresos?	<i>p052</i>	COET2	Si	Si	Si	Si
Opción 1 Durante los últimos tres meses, ¿ha tratado de buscar trabajo en otro país o hacer preparativos para cruzar la frontera?	<i>p0_1</i>	COET2	Si	Si	Si	Si

Cuadro 2. Tabla de ubicación de variables. FUENTE: Elaboración propia con datos de ENOE 2023

Investigamos la llave que nos ayudó a juntar las bases de datos que tenían las variables de nuestro interés. “La llave o clave única, se construye a través de ciertos campos, que vincula las tablas con el fin de darle una identidad digital irreplicable al microdato” de INEGI. Encuesta Nacional de Ocupación y Empleo. ENOE. Conociendo la base de datos. 2023. Las variables que conforman la llave son: *cd_a*, *ent*, *con*, *v_sel*, *n_hog*, *h_mud*, *n_ren*, *tipo* y *mes_cal*.

Gracias a esta llave, se realizó la unión de las bases de datos en el software STATA con el comando “merge”. Las bases de datos que necesite unir fueron: SDEM, COE1 y COE2. Para hacer el merge de estas tres bases de datos, lo primero que se realizó fue abrir la base de datos de SDEM y hacer un sort con las nueve variables que conforma la llave, acto seguido, se realizó la unión utilizando:

```
merge 1:1 cd_a ent con v_sel n_hog h_mud n_ren tipo mes_cal using "C: path"
```

Para observar que no se había perdido ningún dato, se ejecutó describe. Una vez guardada esta base de datos como PART1_SdemCoe1_Tn (sea n= 1, 2, 3 y 4, respectivamente al número del trimestre del año 2023). Ahora bien, para unir la base de datos con el merge realizado a la base de datos COE2, se utilizó:

```
merge m:1 cd_a ent con v_sel n_hog h_mud n_ren tipo mes_cal using "C: path", no gen
```

La m en la línea anterior especifica que es una unión many-to-one, es decir, en COE2 pueden existir varias observaciones para una misma variable (cd_a, ent, con, etcétera), sin embargo, en el merge de SDEM y COE1, cada variable aparece una sola vez. El comando no gen, evita la generación de variables fusión.

Para corroborar que no se perdieron datos en el merge many-to-one, se eliminó la variable “merge_” de la base de datos de PART1_SdemCoe1_Tn para así, hacer un merge one-to-one con la base de datos COE2. Al unir las bases de esta manera, comprobamos que no se pierde ningún dato utilizando el merge many-to-one. Sin embargo, por cuestiones de seguridad para el análisis de la investigación utilizaremos la base de datos que realizamos con el merge one-to-one.

3.4 Elaboración de variables de interés

Para llevar a cabo nuestro análisis, es fundamental generar nuestras variables de interés utilizando las preguntas disponibles en la base de datos ENOE. Esto se debe a que la información específica que necesitábamos no estaba directamente disponible en la base de datos en su forma actual. Por lo tanto, fue necesario construir nuestras propias variables utilizando las preguntas existentes en la ENOE para poder profundizar en los aspectos que deseábamos estudiar. Este proceso de creación de variables nos permitió obtener datos más relevantes y ajustados a nuestros objetivos de investigación, garantizando así la calidad y la precisión de nuestro análisis.

3.4.1 Variable para captar a los inmigrantes: “migrante”

En la base de datos tenemos las dos siguientes variables que nos permiten saber si son migrantes:

- l_nac_c = Lugar de nacimiento y,
- cs_p20a_1 = Hace un año, en (mes) de (año), ¿en qué estado de la República o en qué país vivía?

Utilizamos el comando label list para observar las etiquetas de estas variables l_nac_c y cs_p20a_1 (Véase en Apéndice A y B, respectivamente).

Observemos que en la tabla de l_nac_c nos da los códigos de los estados y países que se tomaron en cuenta al ejercer las entrevistas.

En el formato clasificaciones_enoe_123.pdf, nos dice que los códigos dados por 100, 200, 300, 400 y 500 corresponden a África, América, Asia, Europa y Oceanía respectivamente.

Una vez dicho lo anterior, definimos que son migrantes si pasan dos cosas:

- no nació aquí en el país de México y,

- no vivía el año pasado aquí en México.

Para crear esta variable utilizamos la función generate y replace en STATA de la siguiente manera:

```
gen migrante =.
```

```
replace migrante=1 if (l_nac_c>=100) & (cs_p20a_1==3|cs_p20a_1==4)
```

```
replace migrante=0 if migrante==.
```

De la siguiente manera se definió la variable migrante

- 1: si no nació aquí y si no residía aquí hace un año
- 0: nació aquí y si residía aquí hace un año

Dejándonos así, una población de 304 personas que no vivían aquí hace un año, y tampoco nacieron en el país de México. (Véase Apéndice C)

3.4.2 Variable “trabajamx”

Para definir la variable “trabaja”, utilizamos las siguientes variables:

- p1 = ...¿La semana pasada trabajó por lo menos una hora?
- p2_2 = ¿... ha tratado de buscar trabajo aquí en el país?

Utilizamos el comando label list, para observar las etiquetas de estas variables en la base de datos. (Véase Apéndice D)

Por tanto, se definió la variable trabaja si sí trabajó la semana pasada y también, si no trabajó la semana pasada, pero busca un trabajo aquí en el país.

Generamos la variable trabaja de la siguiente manera:

```
gen trabajamx =.
```

```
replace trabajamx = 1 if (p1==1|p2_2==2)
```

```
replace trabajamx = 0 if trabajamx==.
```

Finalmente tenemos la variable trabajamx definida de la siguiente forma:

- 1: Si sí trabaja y si no trabaja, pero está en búsqueda de un trabajo.
- 0: No trabaja y no está en búsqueda de un trabajo.

3.4.3 Variable “salariomensual”

Para definir esta variable, utilizamos el cuestionario ampliado de la ENOE para localizar la pregunta siguiente: 6b. ¿Cada cuando obtiene... sus ingresos o le pagan? ¿Cuánto ganó o en cuánto calcula sus ingresos?

En la base de datos buscamos esta pregunta y estaba dada por las siguientes variables:

- p6b1 = ¿Cada cuándo obtiene ... sus ingresos o le pagan?
- p6b2 = ¿Cuánto ganó o en cuánto calcula sus ingresos?

Ejecutamos label list en la variable p6b1 (véase Apéndice F). También se ejecutó el mismo comando para la variable p6b2, la cual tiene diferentes montos salariales, por tanto, se decidió definir la variable salariomensual de la siguiente manera:

```
generate salariomensual = .
replace salariomensual = p6b2 if (p6b1 == 1)
replace salariomensual = p6b2*2 if (p6b1 == 2)
replace salariomensual = p6b2*4 if (p6b1 == 3)
replace salariomensual = p6b2*30 if (p6b1 == 4)
```

3.4.4 Variable “educ”

La variable educación de creo a partir de la siguiente pregunta:

- cs_p13_1 = *Hasta que grado aprobó ... en la escuela*

Ejecutamos label list en la variable anterior para observar cómo se encuentra definida (véase Apéndice G).

Ahora bien, para definir la variable educ se realizó lo siguiente:

```
generate educ = .
replace educ=1 if (cs_p13_1==0|cs_p13_1==1|cs_p13_1==2)
replace educ=2 if (cs_p13_1==3|cs_p13_1==4)
replace educ=3 if (cs_p13_1==5|cs_p13_1==6|cs_p13_1==7|cs_p13_1==8|cs_p13_1==9)
```

Por tanto, la variable educ quedo definida de la siguiente manera:

- 1 = Baja (ninguna, preescolar y primaria)
- 2 = Media (secundaria y preparatoria o bachillerato)
- 3 = Alta (normal, carrera técnica, profesional, maestría y doctorado)

3.4.5 Variable “emigrar”

También creamos la variable “emigrar” definidas por las siguientes variables:

- p2_1 = ¿... ha tratado de buscar trabajo en otro país o hacer preparativos para cruzar la frontera?
- p8_1 = Durante los últimos tres meses, ¿... ha tratado de buscar otro trabajo en otro país o hacer preparativos para cruzar la frontera?

Ejecutamos label list en las variables mencionadas para ver cómo están definidas en la base de datos (véase Apéndice H).

Por tanto, se definió la variable emigrar si la persona no se encuentra ocupada, pero ha tratado de buscar trabajo en otro país o hacer preparativos para cruzar la frontera.

Generamos la variable emigrar de la siguiente manera:

```
gen emigrar = 0
```

```
replace emigrar=1 if (p2_1==1|p8_1==1)
```

Finalmente, la variable emigrar se encuentra definida en las siguientes líneas:

- 1: sí buscó y busca un trabajo en otro país o pensó y piensa en cruzar la frontera y,
- 0: no busca un trabajo en otro país ni piensa en cruzar la frontera.

3.4.6 Modificación de la variable sexo

Al revisar las variables que utilizaremos para la regresión, nos percatamos que la variable sex se encontraba definida con la numeración 1 y 2 (véase Apéndice J).

Así que, generamos la variable sexo:

```
gen sexo = .
```

```
replace sexo = 0 if sex==2
```

```
replace sexo = 1 if sex==1
```

Finalmente, la variable sexo se definió de la siguiente manera: 0: mujer y, 1: hombre.

Después de haber definido cuidadosamente las variables que formarán parte de nuestros modelos de mínimos cuadrados ordinarios y probit, estamos en condiciones de dar inicio al análisis estadístico. Este proceso implica una serie de pasos fundamentales que nos permitirán comprender mejor la relación entre nuestras variables independientes y la variable dependiente de interés.

En primer lugar, procederemos a ejecutar el modelo de mínimos cuadrados ordinarios (MCO). Este método nos proporcionará una visión detallada de cómo las variables independientes influyen en la variable dependiente, teniendo en cuenta la tendencia lineal entre ellas. Analizaremos los coeficientes estimados, sus errores estándar, los valores t asociados, así como la significancia de cada variable en el modelo.

Posteriormente, nos adentraremos en el modelo probit, una herramienta poderosa para analizar variables binarias o dicotómicas. Este modelo nos permitirá evaluar la probabilidad de ocurrencia de un evento específico como trabajar en México en el caso de inmigrantes de corto plazo, en función de las variables explicativas seleccionadas (sexo, edad y educación). Analizaremos los coeficientes estimados, los valores z, los errores estándar y la significancia de cada variable en el modelo probit.

A lo largo de este análisis estadístico, nos enfocaremos en interpretar los resultados de manera profunda y significativa. Esto implica no solo observar los valores numéricos de los coeficientes y los estadísticos asociados, sino también comprender la relevancia práctica de estos resultados en el contexto de nuestro estudio.

Nuestro objetivo es obtener una comprensión sólida de cómo las variables independientes afectan a la variable dependiente, para así poder generar conclusiones significativas y fundamentadas en evidencia investigativa.

Capítulo IV. Resultados

En este capítulo, una vez que hemos detallado la base de datos en nuestro trabajo, el siguiente paso es realizar la estimación del modelo probit utilizando el programa STATA.

Se optó por utilizar STATA como software debido a su practicidad y su capacidad para calcular los estimadores utilizando el método de máxima verosimilitud de manera sencilla y eficiente. Esta elección se hizo considerando la conveniencia que ofrece STATA en términos de realizar cálculos complejos de manera automatizada y confiable. Utilizar este software nos permite realizar este proceso de manera eficiente y precisa, asegurando resultados confiables en nuestro análisis.

5.1 Análisis Estadístico con STATA: Comandos Utilizados

STATA es un software estadístico utilizado para el análisis de datos, la manipulación de datos, la visualización de datos y la generación de informes. Es ampliamente utilizado en campos como la economía, la sociología, la política, la salud y otras disciplinas que requieren análisis estadístico avanzado. STATA permite realizar análisis estadísticos complejos, como regresiones, análisis de series temporales, análisis de supervivencia, análisis multivariante y más. Es conocido por su capacidad para manejar grandes conjuntos de datos y su flexibilidad en la generación de gráficos e informes.

En el contexto de analizar esta base de datos, empleamos las siguientes funciones y herramientas disponibles en el software seleccionado (Gould et al., 2010):

- Comando: En STATA, un comando es una instrucción que se utiliza para realizar una acción específica, como cargar datos, realizar cálculos estadísticos, generar gráficos, entre otros. Los comandos en STATA suelen seguir una sintaxis específica y se utilizan para interactuar con los datos y llevar a cabo análisis estadísticos y manipulaciones de datos.
- Llave: En STATA, el término "llave" se refiere a la variable o conjunto de variables que se utilizan para combinar dos conjuntos de datos en una operación de fusión (merge). Estas variables deben estar presentes en ambos conjuntos de datos y se utilizan para identificar las observaciones que se corresponden entre sí en los dos conjuntos de datos que se están fusionando.
- merge: La fusión (merge) en STATA es una operación que combina dos conjuntos de datos en función de una o varias variables clave (llaves). La fusión puede ser de diferentes tipos, como la fusión interna (inner join), la fusión externa (outer join), la fusión izquierda (left join) y la fusión derecha (right join), dependiendo de cómo se deseen combinar las observaciones de los conjuntos de datos.
- label list: es una lista de etiquetas que se asignan a los valores de una variable para proporcionar una descripción más comprensible de esos valores. Por ejemplo, si una variable tiene valores numéricos que representan categorías, un label list puede asociar etiquetas descriptivas a esos valores numéricos para facilitar la interpretación de los datos.
- generate: se utiliza para crear nuevas variables en un conjunto de datos. Se puede utilizar para realizar cálculos, asignar valores predeterminados, aplicar transformaciones a

variables existentes y más. Por ejemplo, se puede generar una nueva variable que calcule el índice de masa corporal a partir de variables de peso y altura.

- **replace:** se utiliza para reemplazar los valores de una variable existente en un conjunto de datos. Puede ser útil cuando se necesita corregir o actualizar los valores de una variable sin tener que crear una nueva variable. Por ejemplo, se puede usar el comando "replace" para corregir un error de codificación en una variable específica.

El enfoque de investigación se dirige hacia el análisis estadístico de la ocupación laboral de los migrantes, con el objetivo de proporcionar información relevante sobre su participación en el mercado laboral y sus niveles salariales. El estudio se propone ofrecer insights significativos acerca de a qué se dedican estas personas (empleo formal o informal) y cómo son sus ingresos.

El software estadístico STATA nos ofrece la posibilidad de llevar a cabo análisis de regresión siempre y cuando tengamos una base de datos en formato ".dta".

5.2 Los modelos de regresión utilizados en la investigación: Un análisis detallado

El planteamiento de dos modelos de regresión se hace con el propósito de comparar y evaluar diferentes enfoques en el análisis de datos, especialmente cuando se trabaja con una variable dicotómica. Al considerar tanto la regresión probit como la regresión de mínimos cuadrados ordinarios (MCO), se busca entender cómo varían las predicciones y la eficiencia de los estimadores bajo diferentes supuestos y enfoques metodológicos.

A continuación, se presentan dos distintos modelos de regresión para el estudio de las variables explicativas sobre la variable *trabajamx*, considerando cada una de las variables binarias desarrolladas en el capítulo anterior.

Cuadro 3. Descripción de los modelos utilizados en la investigación.

Modelo	Descripción
<p>Método de Mínimos Cuadrados Ordinarios (regress)</p> <p>Syntax</p> <pre>regress depvar [indepvars] [if] [in] [weight] [, options]</pre>	<p>Se utiliza para ajustar un modelo lineal a los datos. En este tipo de regresión, se supone que la variable dependiente es continua y que la relación entre la variable dependiente y las variables independientes es lineal.</p> <p>El resultado de una regresión lineal ordinaria se interpreta en términos de la magnitud y dirección de la relación lineal entre las variables.</p>
<p>Regresión Probit</p>	<p>Se utiliza para ajustar un modelo probit a los datos. En la regresión probit, la variable dependiente es binaria y se asume que está</p>

<p>(probit)</p> <p>Syntax</p> <pre>probit <i>depvar</i> [<i>indepvars</i>] [<i>if</i>] [<i>in</i>] [<i>weight</i>] [, <i>options</i>]</pre>	<p>relacionada con las variables independientes a través de una función de distribución normal acumulativa (probit).</p> <p>El resultado de una regresión probit se interpreta en términos de la probabilidad de ocurrencia del evento binario en función de las variables explicativas.</p>
--	--

En resumen, "regress" se usa para variables continuas y ajusta un modelo lineal, mientras que "probit" se usa para variables binarias y ajusta un modelo probit para estimar probabilidades.

5.3 Variables utilizadas en los modelos de regresión

Esta investigación se enfoca exclusivamente en la población inmigrante que no residía en este lugar hace un año y cuyo país de origen es distinto al de México. Se limita el análisis a este grupo específico de inmigrantes para obtener una comprensión más precisa y detallada de su situación y características en el contexto estudiado.

Ahora bien, tenemos la siguiente expresión matemática:

$$trabajamx = \beta_0 + \beta_1 \text{sexo} + \beta_2 \text{eda} + \beta_3 \text{educ}.$$

En la siguiente tabla, se da una descripción de cada una de las variables de interés involucradas en los modelos de regresión que anteriormente fueron presentados:

<i>Variable</i>	<i>Descripción</i>
trabajamx	Representa a la población que se encuentra trabajando o está en búsqueda de un trabajo en el país mexicano.
sexo	Especifica el género del individuo.
eda	Representa la edad del individuo en cuestión.
educ	Indica el grado educativo alcanzado por el individuo.

Cuadro 4. Descripción de variables.

Las variables en color lila son nuestras variables explicativas o independientes y la variable en color azul será la variable dependiente.

5.4 Regresión Lineal de Mínimos Cuadrados Ordinarios

Al utilizar el comando "regress" seguido de la variable dependiente y las variables independientes de interés, obtenemos en la ventana un resumen detallado que incluye los elementos esenciales de una regresión lineal de mínimos cuadrados ordinarios.

Esta función nos permite estimar la relación entre una variable dependiente continua y una o más variables independientes. Esto nos da a entender cómo varía la variable dependiente en función de

los cambios en las variables independientes, evaluando así el impacto de estas últimas en la variable objetivo.

Al examinar la regresión lineal de mínimos cuadrados ordinarios (véase el Apéndice K), observamos lo siguiente:

- Coeficientes y Significancia:
 - La variable "sexo" tiene un coeficiente estimado de 0.1706075, con un error estándar de 0.058456. El valor t asociado es 2.92, con un valor p significativo de 0.004 (menor que 0.05), lo que indica que el género (sexo) tiene un efecto significativo en la variable dependiente "trabajamx".
 - La variable "eda" tiene un coeficiente estimado de 0.0034541, con un error estándar de 0.0019886. El valor t asociado es 1.74, con un valor p de 0.084 (mayor que 0.05), lo que indica que la edad (eda) podría no tener un efecto significativo en "trabajamx" a nivel de significancia típico (0.05).
 - La variable "educ" tiene un coeficiente estimado de 0.0887728, con un error estándar de 0.0435927. El valor t asociado es 2.04, con un valor p significativo de 0.043 (menor que 0.05), lo que indica que el nivel educativo (educ) también tiene un efecto significativo en "trabajamx".
 - El coeficiente de la constante (_cons) es 0.0801543, con un error estándar de 0.0969295. El valor t asociado es 0.83, con un valor p de 0.409 (mayor que 0.05), lo que sugiere que la constante no es significativa en el modelo.
- Bondad de Ajuste: el R-cuadrado (R^2) es 0.0647, lo que significa que aproximadamente el 6.47% de la variabilidad en "trabajamx" se explica por las variables independientes incluidas en el modelo. Esto sugiere que el modelo explica una cantidad limitada de la variabilidad en la variable dependiente.
El R-cuadrado ajustado (Adj R^2) es 0.0544, que tiene en cuenta el número de variables en el modelo y es ligeramente más bajo que el R-cuadrado.
- Análisis de Significancia Global: el resultado de la prueba F indica que el modelo en su conjunto es significativo, con un valor de Prob > F de 0.0004 (menor que 0.05). Esto sugiere que al menos una de las variables independientes tiene un efecto significativo en la variable dependiente "trabajamx".
- Root MSE: la raíz del error cuadrático medio (Root MSE) es 0.48441, lo que indica la dispersión promedio de los residuos alrededor de la línea de regresión.

En resumen, este análisis proporciona información sobre la significancia individual de las variables independientes, la capacidad del modelo para explicar la variabilidad en la variable dependiente y la significancia global del modelo en su conjunto.

5.5 Regresión Probit

Al utilizar el comando "probit" seguido de la variable dependiente y las variables independientes de interés, obtenemos en la ventana un resumen detallado que incluye los elementos esenciales de una regresión binaria.

Esta función nos permite explorar la relación entre las variables predictoras y la variable objetivo de manera efectiva, proporcionándonos información valiosa sobre la probabilidad de ocurrencia de un evento específico en función de las variables explicativas.

Al examinar la regresión probit observamos lo siguiente (véase el Apéndice L):

- Log-likelihood y Pseudo R2: el logaritmo de la verosimilitud (log likelihood) es -181.22295, lo que indica qué tan bien se ajusta el modelo a los datos. Un log likelihood más alto indica un mejor ajuste del modelo.
El pseudo R2 es 0.0486, que proporciona una medida de la variabilidad explicada por el modelo. Un pseudo R2 más alto indica un mejor ajuste, similar al R-cuadrado en modelos de regresión lineal.
- Estadístico de Prueba (LR chi2): el estadístico de prueba LR chi2 es 18.52 con 3 grados de libertad, lo que indica que el modelo en su conjunto es significativo. El valor p asociado es 0.0003 (Prob > chi2), lo que indica que al menos una de las variables independientes tiene un efecto significativo en la variable dependiente "trabajamx".
- Coeficientes y Significancia:
 - La variable "sexo" tiene un coeficiente estimado de 0.4476639, con un error estándar de 0.1551256 y un valor z de 2.89. El valor p asociado es 0.004 (menor que 0.05), lo que indica que el género (sexo) tiene un efecto significativo en la probabilidad de "trabajamx" en el modelo probit.
 - La variable "eda" tiene un coeficiente estimado de 0.0094473, con un error estándar de 0.0053185 y un valor z de 1.78. El valor p asociado es 0.076 (mayor que 0.05), lo que sugiere que la edad (eda) podría no ser significativa en el modelo a nivel típico de significancia (0.05).
 - La variable "educ" tiene un coeficiente estimado de 0.2363906, con un error estándar de 0.1150195 y un valor z de 2.06. El valor p asociado es 0.040 (menor que 0.05), lo que indica que el nivel educativo (educ) también tiene un efecto significativo en la probabilidad de "trabajamx" en el modelo probit.
 - El coeficiente de la constante (_cons) es -1.122469, con un error estándar de 0.2681158 y un valor z de -4.19. El valor p asociado es 0.000 (menor que 0.05), lo que indica que la constante es significativa en el modelo.

En resumen, el modelo probit analizado tiene un buen ajuste a los datos, es significativo en su conjunto y muestra la significancia de algunas variables independientes en la probabilidad de "trabajamx". Sin embargo, la edad no parece ser significativa a nivel típico de significancia.

5.5.1 Efectos marginales del modelo de regresión probit

Este análisis muestra los efectos marginales condicionales en términos de la probabilidad de trabajar en México ($\Pr(\text{trabajamx})$) en relación con las variables independientes sexo, edad (eda) y educación (educ) en los valores medios de estas variables. Así que estimamos su derivada en la media de las variables.

Los efectos marginales condicionales miden el cambio en la probabilidad de "trabajamx" debido a un cambio de una unidad en las variables independientes (sexo, edad y educación), manteniendo todas las demás variables constantes y evaluadas en sus valores medios (atmeans).

El análisis detallado de los efectos marginales condicionales del modelo probit es el siguiente (véase en el Apéndice M):

- Efectos Marginales por variable:
 - Sexo (sexo): el efecto marginal de la variable "sexo" es 0.1768578, con un error estándar de 0.0612799 y un valor z de 2.89. El valor p asociado es 0.004 (menor que 0.05), lo que indica que un cambio en el género (de mujer a hombre, por ejemplo) aumenta la probabilidad de "trabajamx" en aproximadamente 0.177 unidades, siendo significativo.
 - Edad (eda): el efecto marginal de la variable "edad" es 0.0037323, con un error estándar de 0.0021005 y un valor z de 1.78. El valor p asociado es 0.076 (mayor que 0.05), lo que sugiere que el cambio en la edad podría no ser significativo en la probabilidad de "trabajamx" a nivel típico de significancia (0.05).
 - Educación (educ): el efecto marginal de la variable "educación" es 0.0933904, con un error estándar de 0.0454326 y un valor z de 2.06. El valor p asociado es 0.040 (menor que 0.05), lo que indica que un aumento en el nivel educativo aumenta la probabilidad de "trabajamx" en aproximadamente 0.093 unidades, siendo significativo.

En resumen, los efectos marginales condicionales muestran cómo cambia la probabilidad de "trabajamx" en respuesta a cambios específicos en las variables independientes. La significancia de estos efectos marginales se evalúa mediante los valores z y los valores p asociados.

5.6 Análisis de los salarios de inmigrantes de corto plazo en México

Hemos examinado la probabilidad de empleo entre los inmigrantes considerando sus características sociodemográficas fundamentales. En nuestra continua exploración de los inmigrantes de corto plazo, otro aspecto relevante a considerar es su salario, específicamente su salario mensual.

Primero, ejecutamos el comando summarize para la variable salariomensual (véase Apéndice N), lo cual nos muestra información sobre salarios mensuales en una población específica, con un total de 59 observaciones, es decir de 59 individuos migrantes.

Aquí está el análisis de los estadísticos descriptivos proporcionados:

- Media (Mean): el salario mensual promedio en esta población es de aproximadamente 77,673.05 unidades monetarias.
- Desviación estándar (Std. Dev.): la desviación estándar de los salarios mensuales es de alrededor de 98,592.03 unidades monetarias. Esto indica la dispersión de los salarios alrededor de la media. Una desviación estándar alta sugiere que los salarios están más

dispersos, mientras que una desviación estándar baja indica que los salarios están más concentrados alrededor de la media.

- Valor Mínimo (Min): el salario mensual más bajo registrado en la población inmigrante es de 5,140 unidades monetarias. Este valor representa el salario más bajo entre todas las observaciones.
- Valor Máximo (Max): el salario mensual más alto registrado en la población inmigrante es de 451,500 unidades monetarias. Este valor representa el salario más alto entre todas las observaciones.

En resumen, estos estadísticos descriptivos proporcionan una idea general de la distribución de los salarios mensuales en la población estudiada, incluyendo el rango de salarios (desde el mínimo al máximo), la dispersión alrededor de la media (desviación estándar), y el salario promedio en la población (media).

Nos interesa comprender la distribución de estos salarios mensuales. Para aproximarnos a esta distribución, optaremos por un análisis descriptivo utilizando la técnica de densidad de kernel.

En la descripción de nuestra gráfica de la densidad de kernel, tenemos lo siguiente (véase apéndice Ñ):

- En el eje Y - Densidad: la densidad representa qué tan concentrados están los datos en diferentes rangos de salario mensual. La escala de 0 a 0.0015 indica la densidad de datos en esa área específica.
- En el eje X - Salario Mensual: el eje X muestra los valores de salario mensual, que van desde 6000 hasta 10000. La curva muestra cómo se distribuyen los salarios en este rango.
- Forma de la Curva: la curva comienza alrededor de 0.001 en el eje Y aumenta gradualmente hasta aproximadamente 0.0013, luego se mantiene relativamente constante hasta alrededor de 6900 en el eje X, donde se vuelve una línea horizontal hasta 8900. Después de 8900, hay un pequeño ascenso muy gradual hasta cerca de 0.000002. Esta forma indica una distribución de datos sesgada a la derecha, lo que significa que hay más observaciones con salarios más bajos y una cantidad menor de observaciones con salarios más altos, pero algunos salarios muy altos están presentes.

En resumen, la representación gráfica revela que la mayoría de los inmigrantes tienen salarios mensuales en el extremo inferior del espectro, con una minoría que recibe salarios más altos. Esta tendencia indica una distribución de salarios sesgada hacia la derecha.

Conclusiones

Esta investigación de tesis revela que los migrantes provenientes de Centroamérica, al cruzar la frontera sur, tienen como objetivo principal asegurar estabilidad económica para ellos y sus familias. Aunque la mayoría tiene la meta de llegar a Estados Unidos, las estrictas leyes migratorias los llevan a quedarse en México, convirtiéndolos en inmigrantes de corto plazo. A través de la Encuesta Nacional de Ocupación y Empleo, aunque no es diseñada específicamente para captar datos de inmigrantes en México, se obtuvo información valiosa.

Con respecto a la conclusión estadística, con la ayuda de los comandos "regress" y "probit" en STATA nos brindaron un análisis de: la regresión lineal de mínimos cuadrados ordinarios (MCO) y la regresión probit. Ambos modelos se emplearon para estimar la relación entre nuestra variable dependiente (trabajamx), y las variables independientes de nuestro interés (sexo, eda y educ). Es decir, estimar cómo varía la probabilidad de que los inmigrantes trabajen en México en función de sus características sociodemográficas básicas.

El análisis de ambas regresiones nos dice que las variables sociodemográficas básicas (sexo, edad y educación) son estadísticamente significativas (al 99%, 92% y 96% respectivamente), por tanto, los resultados del modelo probit son consistentes y muy parecidos con el modelo por MCO.

Los efectos marginales condicionales en el modelo probit muestran cómo cambia la probabilidad de "trabajamx" en respuesta a cambios específicos en las variables independientes. En esta estimación tenemos que la probabilidad de que el inmigrante trabaje es mayor cuando el sexo del inmigrante es masculino y aumenta en una unidad cuando aumenta: la educación del inmigrante y la edad.

Otro hallazgo de esta tesis fue que, del total de 304 inmigrantes captados, 125 inmigrantes trabajan, mientras 179 no tienen trabajo. Del total de los 125 que trabajan, solo 59 inmigrantes reportaron tener un salario mensual, mientras que 66 inmigrantes no reportan salarios mensuales.

Estos 59 salarios mensuales reportados nos dicen que el 5% de los salarios es más bajo que el salario mensual mínimo (\$7,468 pesos, en 2024), mientras el 20% tiene un salario mensual arriba de \$100,00.

Se observa una brecha significativa entre los salarios de los inmigrantes y el salario mínimo nacional (\$266.71 diarios), lo que confirma que la condición de inmigrante sin documentación legal afecta su acceso a empleos dignos. En contraste, aquellos inmigrantes con documentación legal tienen salarios mucho más altos, llegando hasta \$451,500 mensuales, reflejando la importancia de la regularización migratoria en el acceso al mercado laboral.

Con respecto al gráfico de densidad kernel, debido a que es una manera no paramétrica de estimar la función de densidad de probabilidad de una variable aleatoria, nos ayuda a visualizar esta brecha salarial de los inmigrantes de manera visual.

En términos de género y educación, se destaca que los hombres con educación tienen una mayor probabilidad de empleo en México, mientras que la falta de educación reduce esta probabilidad. En el caso de las mujeres, tener educación no aumenta significativamente la probabilidad de empleo. La edad, por otro lado, no parece ser un factor relevante en la probabilidad de empleo para ambas categorías de género.

En conclusión, los inmigrantes de corto plazo en México enfrentan desafíos significativos en términos de acceso al empleo digno y salarios justos, especialmente aquellos sin documentación legal. La regularización migratoria y el acceso a la educación son aspectos clave para mejorar su integración en el mercado laboral mexicano.

Apéndice

Apéndice A. Lista de etiquetas de la variable l_nac_c, que corresponde al lugar de nacimiento.

```
. label list l_nac_c
l_nac_c:
  1 Aguascalientes
  2 Baja California
  3 Baja California Sur
  4 Campeche
  5 Coahuila
  6 Colima
  7 Chiapas
  8 Chihuahua
  9 Distrito Federal
 10 Durango
 11 Guanajuato
 12 Guerrero
 13 Hidalgo
 14 Jalisco
 15 México
 16 Michoacán
 17 Morelos
 18 Nayarit
 19 Nuevo León
 20 Oaxaca
 21 Puebla
 22 Querétaro
 23 Quintana Roo
 24 San Luis Potosí
 25 Sinaloa
 26 Sonora
 27 Tabasco
 28 Tamaulipas
 29 Tlaxcala
 30 Veracruz
 31 Yucatán
 32 Zacatecas
 33 Entidad federativa no especificada
100 Cualquier país de Africa
200 Otro país de América
221 Estados Unidos de América
225 Guatemala
300 Cualquier país de Asia
400 Otro país de Europa
415 España
500 Cualquier país de Oceanía
600 País no especificado
```

Apéndice B. Lista de etiquetas de la variable cs_p20a_1

```
. label list cs_p20a_1
cs_p20a_1
  1 Aquí, en otro estado
  2 En otro estado
  3 En los EUA
  4 En otro país
```

Apéndice C. Tabulador de la variable definida migrante.

```
. tab migrante
migrante   Freq.   Percent   Cum.
      1      304      100      100
  Total      304      100
```

Apéndice D. Lista de etiquetas de la variable p1 y de la variable p2_2

```
. label list p1
p1
  1 Sí trabajó
  2 No trabajó

. label list p2_2
p2_2
  1 Sí buscó trabajo aquí en el país
  2 No buscó trabajo aquí en el país
```

Apéndice E. Tabulador de la variable definida trabajamx si el individuo es migrante.

```
. tab trabajamx if migrante ==1
```

trabajamx	Freq.	Percent	Cum.
0	179	58.88	58.88
1	125	41.12	100
Total	304	100	

Apéndice F. Lista de etiquetas de la variable p6b1.

```
. label list p6b1
```

```
p6b1
```

- 0 No aplica
- 1 Cada mes
- 2 Cada 15 días
- 3 Cada semana
- 4 Diario
- 5 Otro periodo de pago
- 6 Le pagan por pieza producida o vendida, servicio u obra realizada
- 7 No supo estimar
- 8 Se negó a contestar esta pregunta

Apéndice G. Lista de etiquetas de la variable cs_p13_1.

```
. label list cs_p13_1
```

```
cs_p13_1
```

- 0 Ninguno
- 1 Preescolar
- 2 Primaria
- 3 Secundaria
- 4 Preparatoria o bachillerato
- 5 Normal
- 6 Carrera técnica
- 7 Profesional
- 8 Maestría
- 9 Doctorado
- 99 No sabe

Apéndice H. Lista de etiquetas de las variables p2_1 y p8_1.

```
. label list p2_1
```

```
p2_1
```

- 1 Sí buscó trabajo en otro país o piensa cruzar la frontera
- 0 No aplica

```
. label list p8_1
```

```
p8_1
```

- 1 Buscar otro trabajo en otro país o piensa cruzar la frontera
- 0 No aplica

Apéndice I. Tabulador de la variable emigrar si el individuo es migrante.

`. tab emigrar if migrante ==1`

trabajamx	Freq.	Percent	Cum.
0	300	98.68	98.68
1	4	1.32	100
Total	304	100	

Apéndice J. Lista de etiquetas de la variable sex y sexo.

`. label list sex`

sex

1 Hombre
2 Mujer

`. label list sexo`

sexo

1 Hombre
0 Mujer

Apéndice K. Modelo de regresión lineal de mínimos cuadrados ordinarios.

`regress trabajamx sexo eda educ if migrante==1`

Source	SS	df	MS	Number of obs	277
Model	4.43002033	3	1.47667344	F(3, 273)	6.29
Residual	64.0609544	273	0.23465551	Prob> F	0.0004
Total	68.4909747	276	0.24815571	R-squared	0.0647
				Adj R-squared	0.0544
				Root MSE	0.48441

trabajamx	Coef.	Std. Err.	t	P > t	[95% Conf. Interval]	
sexo	0.1706075	0.58456	2.92	0.004	0.0555256	0.2856894
eda	0.0034541	0.0019886	1.74	0.084	-0.0004609	0.0073691
educ	0.0887728	0.0435927	2.04	0.043	0.0029522	0.1745933
_cons	0.0801543	0.0969295	0.83	0.409	-0.11067	0.2709786

Apéndice L. Modelo de regresión Probit.

`probit trabajamx sexo eda educ if migrante==1`

Iteration 0: log likelihood = -190.48093
 Iteration 1: log likelihood = -181.24641
 Iteration 2: log likelihood = -181.22296
 Iteration 3: log likelihood = -181.22295

Probit regression	Number of ob	=	277
	LR chi2 (3)	=	18.52
	Prob > chi2	=	0.0003
Log likelihood = -181.22295	Pseudo R2	=	0.0486

trabajamx	Coef.	Std. Err.	z	P > z	[95% Conf. Interval]	
sexo	0.4476639	0.1551256	2.89	0.004	0.1436234	0.7517044
eda	0.0094473	0.0053185	1.78	0.076	-0.0009767	0.0198713
educ	0.2363906	0.1150195	2.06	0.04	0.0109564	0.4618247
_cons	-1.122469	0.2681158	-4.19	0	-1.647966	-0.5969715

Apéndice Ñ. Tabulador de la variable salariomensual de los individuos inmigrantes.

tab salariomensual if migrante==1

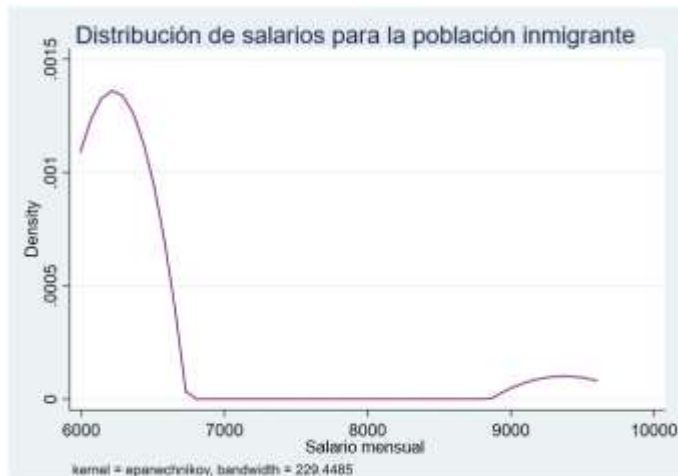
salariomensual	Freq.	Percent	Cum.	salariomensual	Freq.	Percent	Cum.
5140	1	1.69	1.69	43000	2	3.39	62.71
6000	1	1.69	3.39	51600	1	1.69	64.41
8000	1	1.69	5.08	61920	1	1.69	66.1
10320	1	1.69	6.78	77400	1	1.69	67.8
12000	1	1.69	8.47	92880	1	1.69	69.49
12040	2	3.39	11.86	96000	1	1.69	71.19
12900	1	1.69	13.56	103200	1	1.69	72.88
13000	1	1.69	15.25	116100	4	6.78	79.66
13760	1	1.69	16.95	129000	1	1.69	81.36
17200	5	8.47	25.42	154800	3	5.08	86.44
18920	1	1.69	27.12	161250	1	1.69	88.14
20640	6	10.17	37.29	180000	1	1.69	89.83
22020	1	1.69	38.98	232200	1	1.69	91.53
22360	2	3.39	42.37	270000	1	1.69	93.22
24000	1	1.69	44.07	288100	1	1.69	94.92
24080	2	3.39	47.46	309600	1	1.69	96.61
25800	5	8.47	55.93	387000	1	1.69	98.31
34400	1	1.69	57.63	451500	1	1.69	100
40000	1	1.69	59.32	Total	59	100	

Apéndice O. Grafica de la distribución de salarios para la población migrante (densidad Kernel).

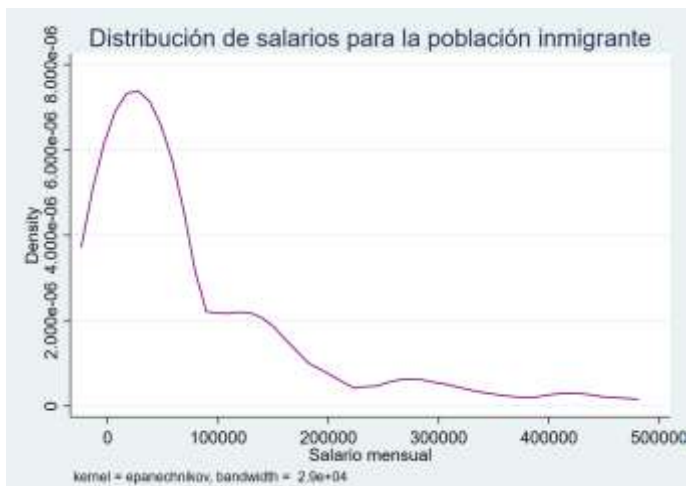
Para estimar la densidad del núcleo de los salarios, tenemos que:

- Tener un conjunto de datos que contenga información de los salarios (variable salariomensual)
- Escoger una función de suavizado como la gaussiana, Epanechnikov o uniforme.
- Seleccionar el ancho de banda: Este parámetro determina la suavidad de la estimación. Un ancho de banda menor produce más detalle, mientras que uno mayor genera una estimación más suave. Métodos comunes para seleccionar el ancho de banda incluyen la validación cruzada y las reglas de Scott o Silverman.
- Calcular la densidad del núcleo, esto se hace aplicando la función de núcleo centrada en cada punto de datos y suma las contribuciones de todos los puntos para obtener la estimación de densidad.
- Ajustar (normalizar) la estimación para que el área bajo la curva sume 1, permitiendo su interpretación como una función de densidad de probabilidad.
- Aplicar el comando del gráfico y visualizar la estimación de densidad para entender la distribución de los salarios.

Herramientas como SciPy en Python, la función `density` en R y comandos en STATA facilitan estos cálculos manejando la selección del núcleo, la estimación del ancho de banda y la normalización automáticamente.



Se agregó esta gráfica debido a que podemos observar con mayor precisión la distribución de salarios para la población inmigrante a comparación de la gráfica anterior que se corta hasta el monto de nuestros de \$10,000.



Referencias

1. Aikin, O., & Muñoz, A. A. (2013). Crisis de derechos humanos de las personas migrantes en tránsito por México: redes y presión transnacional. *Foro Internacional*, 53(1(211)), 143–181. Recuperado el 15 de diciembre del 2022 de <http://www.jstor.org/stable/23608361>
2. Alto Comisionado de las Naciones Unidas para los Refugiados (ACNUR): www.acnur.org
3. Amnistía Internacional. 2010. *Víctimas invisibles. Migrantes en movimiento en México*. Madrid: Editorial Amnistía Internacional. Recuperado el 15 de diciembre del 2022 de <http://amnistiainternacional.org/publicaciones/108-victimas-invisibles-migrantes-en-movimiento-en-mexico.html>
4. Amuedo-Dorantes, C., & Pozo, S. (2014). On the Intended and Unintended Consequences of Enhanced U.S. Border and Interior Immigration Enforcement: Evidence from Mexican Deportees. *Demography*, 51(6), 2255–2279. Recuperado el 15 de diciembre del 2022 de <http://www.jstor.org/stable/43697504>
5. Arp-Nisen, J. D., Morayta, G. C., & Mabire, B. (2019). Política migratoria - immigration policy: entre el discurso, la práctica y la coyuntura. *Foro Internacional*, 59(3-4 (237-238)), 1021–1048. Recuperado el 15 de diciembre del 2022 de <https://www.jstor.org/stable/26663297>
6. Castillo, G. M. (2022). Desaparición de migrantes en tránsito por Centroamérica-México-Estados Unidos.: Quiénes son, cómo y por qué son desaparecidos. In C. Pedone & A. Hinojosa (Eds.), *Vidas en movimiento: migración en América Latina* (pp. 135–190). CLACSO. Recuperado el 15 de diciembre del 2022 de <https://doi.org/10.2307/j.ctv2v88chc.6>
7. Castillo, M. Á. (1998). La política de inmigración en México: un breve recuento. In M. Á. Castillo, A. Lattes, & J. Santibáñez (Eds.), *Migración y fronteras* (1st ed., pp. 425–452). El Colegio de México. Recuperado el 15 de diciembre del 2022 de <https://doi.org/10.2307/j.ctv6jmxmm.21>
8. Chatterjee, S. y Simonoff, JS (2013). Manual de análisis de regresión. John Wiley & sons.
9. de la Torre, R., Béliveau, V. G., Steil, C. A., Toniol, R., Algranti, J., Bordes, M., Pinheiro, L. F., Lozano, F., Catoggio, M. S., Pérez, R., Gumucio, C. P., Robledo, G., Ameigeiras, A., Mallimaci, F., Setton, D., & Mosqueira, M. (2019). La transmigración en México: el cruce de los caminos de la fe y los derechos humanos. In V. G. Béliveau (Ed.), *La religión ante los*

- problemas sociales: Espiritualidad, poder y sociabilidad en América Latina* (pp. 229–258). CLACSO. Recuperado el 15 de diciembre del 2022 de <https://doi.org/10.2307/j.ctvnp0jnq.12>
10. Durand, J. y Massey D. (2003), *Clandestinos. Migración México-Estados Unidos en los albores del siglo XXI*, México, Universidad Autónoma de Zacatecas / Miguel Ángel Porrúa.
 11. EMIF: *Encuestas sobre Migración en las Fronteras Norte y Sur de México*. (s. f.). Recuperado el 15 de diciembre del 2022 de www.colef.mx/emif
 12. ENOE, *Cómo se hace la ENOE: Métodos y procedimientos* (3ra. ed.). (2023). [INEGI]. Instituto Nacional de Estadística y Geografía. Recuperado el 15 de diciembre del 2022 de <https://www.inegi.org.mx/programas/enoe/15ymas/#publicaciones>
 13. *Estadísticas migratorias para México. Boletín anual 2023*. (s. f.). México. Recuperado el 15 de diciembre del 2022 de <https://mexico.un.org/es/264927-estad%C3%ADsticas-migratorias-para-m%C3%A9xico-bolet%C3%ADn-anual-2023>
 14. García, B. Á., & Econometría, I. I. (2008). Modelos de elección binaria. *Econometría II*. Recuperado el 15 de diciembre del 2022 de https://alvarez.webs.uvigo.es/teaching_archivos/ectria2_0708/binary.pdf
 15. Gould, W. W. (2010). *Introduction to STATA*. College Station, TX: Stata Press. Recuperado el 15 de diciembre del 2022 de <https://www.stata.com/manuals/p.pdf>
 16. Gujarati, D., & Porter, D. (2010). *Econometría* (quinta edición). México: Editorial Mc. Graw Hill.
 17. INEGI *Cómo se hace la ENOE. Métodos y procedimientos*. 2023. (2023). En *ENOE*. Instituto Nacional de Estadística y Geografía. Recuperado el 15 de diciembre del 2022 de <https://www.inegi.org.mx/contenidos/productos>
 18. Instituto Nacional de Migración (INM). (2020). *Estadísticas migratorias anuales*. Recuperado de www.gob.mx/inm
 19. Kron, S. (2011). Gestión migratoria en norte y centroamérica: manifestaciones y contestaciones. *Anuario de Estudios Centroamericanos*, 37, 53–85. Recuperado el 15 de diciembre del 2022 de <http://www.jstor.org/stable/41306433>
 20. Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2004). *Applied Linear Statistical Models*. McGraw-Hill Education.
 21. Long, J. S. (1997). *Regression models for categorical and limited dependent variables* (Advanced Quantitative Techniques in the Social Sciences). NY: Sage Publications.

22. Mendiburo, A. D., Moro, R. J. D., Castillo, P. G. M., Govea, D. M., & Muñoz, C. E. R. (2023). Los retos de la inserción laboral del migrante en México: una aproximación desde la integración a través de las organizaciones de la sociedad civil. *Inter Disciplina*, 11(29), 79-104. Recuperado el 15 de diciembre del 2022 de <https://doi.org/10.22201/ceiich.24485705e.2023.29.84481>
23. Organización Internacional para las Migraciones (OIM): www.iom.int
24. Palacios, S. P. I. (2015). Los transmigrantes centroamericanos en México. *Latin American Research Review*, 50(4), 49–68. Recuperado el 15 de diciembre del 2022 de <http://www.jstor.org/stable/44000319>
25. Puerta, R. (2019). Entendiendo y explicando la migración hondureña a Estados Unidos. In R. Romero (Ed.), *Antología del pensamiento hondureño contemporáneo* (pp. 309–352). CLACSO. Recuperado el 15 de diciembre del 2022 de <https://doi.org/10.2307/j.ctvnp0kc9.18>
26. Téllez, M. E. A. (1998). Migración laboral interna e internacional captada en la frontera norte mexicana.: Diferencias por sexo y sector de ocupación. In A. Lattes, J. Santibáñez, & M. Á. Castillo (Eds.), *Migración y fronteras* (1st ed., pp. 262–278). El Colegio de México. Recuperado el 15 de diciembre del 2022 de <https://doi.org/10.2307/j.ctv6jmxmm.13>
27. Valenzuela M., J. (1998). El programa “beta”.: La protección de los derechos humanos de los migrantes indocumentados desde una perspectiva policiaca no convencional. In A. Lattes, J. Santibáñez, & M. Á. Castillo (Eds.), *Migración y fronteras* (1st ed., pp. 463–474). El Colegio de México. Recuperado el 15 de diciembre del 2022 de <https://doi.org/10.2307/j.ctv6jmxmm.23>
28. Wooldridge, J. M. (2006). *Introducción a la Econometría*. 4 edición. Un enfoque moderno. México: Editorial Thomson-Learning. 849 p.
29. Wooldridge, JM (2010). *Análisis econométrico de datos de sección transversal y panel*. Prensa del MIT.