CAPÍTULO 2

METODOLOGÍA

Según la teoría del capital humano, la decisión de migrar puede ser vista por los individuos como una inversión, que será rentable siempre y cuando los beneficios netos sobrepasen los costos. La probabilidad de migrar será mayor en la medida en que los diferenciales salariales entre el lugar de origen y el de destino sean mayores, así como también los costos tanto monetarios como psicológicos sean menores.

Para realizar este análisis se utilizará un modelo desarrollado por Nakosteen y Zimmer (1980), el cual fue aplicado por Norbert M. Fiess y Dorte Verner en el paper "Migration and Human Capital in Brazil" (2003) para explicar la dinámica migratoria que se da entre las regiones Norte y Sur de Brasil.

En el presente estudio se aplicará este modelo para México para probar que los determinantes de la decisión de migrar internamente son en primer lugar, las variables de capital humano, el género, el tamaño de la familia, además de que también influyen de manera significativa los diferenciales salariales y los costos en los que se incurre.

2.1 Hipótesis

Con base en lo anterior se proponen la siguiente hipótesis:

Hipótesis: "Los principales determinantes de la migración interna en México son: la edad, la escolaridad, el género y los diferenciales salariales entre las regiones de origen y de destino."

a) Los determinantes de la migración cambian dependiendo del criterio de identificación del migrante (reciente o de largo plazo).

b) El efecto de los determinantes de la migración en la probabilidad de migrar varía según el género.

2.2 Modelo econométrico

El modelo que se utilizará consta de dos partes, pero sólo se estimará la primera, donde se usa un modelo probit[1] para estimar el efecto de los determinantes de la migración interna propuestos en la hipótesis.

El modelo para estimar los determinantes de la migración involucra dos aspectos: la heterogeneidad y la auto-selección. El primero debido a que habrá una ecuación para cada grupo de individuos (migrantes y no migrantes), ya que cada grupo tiene un salario distinto. El segundo, porque los individuos que migran pueden no ser una muestra aleatoria[2], esto bien porque su nivel de educación es mayor, son más jóvenes, o poseen

[1] Modelo de respuesta binaria en el que la probabilidad de respuesta es la función de densidad acumulada normal estándar, evaluada en una función lineal de las variables explicativas.

[2] Esquema de muestreo en el que cada observación se elige en forma aleatoria de la población. En particular, ninguna unidad tiene mayor probabilidad de ser elegida que cualquier otra y cada elección es independiente de las demás.

alguna otra característica que los hace tener mayor probabilidad de migrar, y esto daría como consecuencia un sesgo de selección[3], esto implicaría un error en la estimación que haría poco precisos los resultados. Para resolver este problema, se usa la corrección de Heckman.

En esta investigación no será necesaria, ya que como se mencionó anteriormente se pretende analizar únicamente los determinantes de la migración.

La ecuación del modelo probit es la siguiente:

$$M^* = \beta' \overline{X} + \varepsilon \qquad (4)$$

Donde:

- M es una variable que indica la probabilidad de ser migrante.
- $\overline{X}$ es un vector compuesto por variables de características individuales, familiares y de redes migratorias para cada individuo.

Se asume que ε tiene una distribución normal con media 0 y varianza 1. Sólo se observa el resultado de la variable latente M^* si el individuo migra o no, así:

$$M = 1 \text{ si } M^* > 0 \qquad (5)$$

$$M = 0 \text{ si } M^* <= 0 \qquad (6)$$

La probabilidad de que $M = 1$ es:

$$\text{Prob}\,(M^* > 0) = \text{Prob}\,(\beta' \overline{X} + \varepsilon > 0)$$

$$= \text{Prob}\,(\varepsilon > \beta' \overline{X}) \qquad (7)$$

Donde:

$$\text{Prob}\,(M^* > 0) = \text{Prob}\,(\varepsilon < \beta' \overline{X})$$

[3] Sesgo en el estimador de MCO que se induce al tomar datos que proceden de una selección muestral endógena.

$$= F(\beta' \overline{X}) \qquad (8)$$

Donde F es la función de distribución acumulada de la normal estándar, así:

$$F(\beta' \overline{X}) = \Phi(\beta' \overline{X}) = \int_{-\infty}^{\beta' \overline{X}} \phi(t)dt \qquad (9)$$

Debido a que el objetivo principal es estimar los efectos del vector $\overline{X}$ en la probabilidad de migrar la P ($M^* = 1 | \overline{X}$) se debe hallar el efecto marginal de las variables que se obtiene calculando la derivada parcial. Por su naturaleza no lineal esta derivada está dada por:

$$\frac{\partial E[M \mid \overline{X}]}{\partial \overline{X}} = \left(\frac{\partial F(\beta' \overline{X})}{\partial (\beta' \overline{X})} \right) \beta$$

$$= f(\beta' \overline{X}) \beta \qquad (10)$$

Donde f (·) es la función de densidad de probabilidad que corresponde a la distribución acumulada de la distribución normal dada x:

$$\frac{\partial E[y \mid \overline{X}]}{\partial \overline{X}} = \Phi(\beta' \overline{X}) \beta \qquad (11)$$

La variable dependiente que se utilizará tiene dos definiciones para poder distinguir entre la migración reciente y de largo plazo:

1) Migración reciente: para evaluarla se tomó la migración según el lugar de residencia de hace 5 años, usando como lugar de origen el estado en el que vivía el individuo en 1995.
2) Migración de largo plazo: para evaluarla se usó la migración según el lugar de nacimiento, es decir, se toma como lugar de origen el lugar de nacimiento del individuo.

Dentro de las variables independientes se incluyen los estados de la República para poder evaluar su condición migratoria y poder saber cuáles estados son los más receptores y expulsores de migrantes en el país en comparación con el Distrito Federal.

Debido a que también se pretende analizar si existe algún cambio en los determinantes de la migración en ambas definiciones, tanto reciente como de largo plazo, se utilizó la interacción de todas las variables del modelo con la variable género. De esta manera se puede estimar el efecto de las variables dependiendo de si el individuo es hombre o mujer. Además se pretende realizar pruebas de significancia conjunta (Test de Wald)[4] para comprobar si es significativo el efecto de los determinantes para los hombres, ya que el grupo base son las mujeres.

Así la hipótesis nula está definida como, Ho:

$$\beta_{ij} + \beta_{ik} = 0 \qquad (12)$$

Donde:

- β_{ij} es el coeficiente de la variable j
- β_{ik} es el coeficiente de la interacción entre la variable k y la variable género.

2.3 Variables

Las variables utilizadas en el modelo anterior se describen en la siguiente tabla.

Características individuales	
Variable	Descripción
Género	Variable dicotómica que toma el valor de 1 si es hombre y el valor de 0 si es mujer.
Edad	Grupo de variables que indican los años

[4] Prueba de significancia de hipótesis conjuntas: prueba que comprende más de una restricción en los parámetros de un modelo.

	cumplidos del individuo.
Estado Civil	Grupo de variables que indican si el individuo es soltero, casado o divorciado.
Escolaridad	Grupo de variables que indican los años aprobados acumulados.
Características familiares	
Tamaño de la familia	Número de personas que componen la familia
Ingresos	Ingreso mensual en pesos
Características redes migratorias	
Lugar de origen	Estado del que proviene el individuo, ya sea su lugar de nacimiento o el lugar en el que vivía hace 5 años.
Lugar de destino	Estado al que migró el individuo.