

Problem set 1

Max Brando Serna Leyva
Econometría aplicada
Agosto - Diciembre 2024

Problema 1

Requisito cumplido. (Omito escribir las instrucciones por simplicidad)

Problema 2

1. Visita el sitio de INEGI en microdatos. Describe dos bases de datos que te sirvan para alguna (o algunas) pregunta de investigación de interés. No se puede utilizar ENIGH, ENE/ENOE, Censo.

- **Programa de la Industria Manufacturera, Maquiladora y de Servicios de Exportación (IMMEX).**

Esta base contiene registros administrativos de las unidades económicas beneficiarias del programa de fomento a la Industria Manufacturera, Maquiladora y de Servicios de Exportación (IMMEX), desde 2007 en adelante. El objetivo general de este proyecto es, de acuerdo con el INEGI:

Disponer de estadísticas de corto plazo que muestren las características y evolución de las actividades económicas en el contexto del Programa IMMEX, para un mejor análisis y comprensión de la actividad de las empresas incorporadas en el mismo.

La periodicidad es mensual, y la información para el segmento manufacturero inicia desde julio de 2007, mientras que para el no manufacturero inicia en enero de 2009. La disponibilidad de la información es de 60 días después del mes de referencia. Asimismo, la cobertura es nacional, desglosando por entidad federativa y por municipio.

Me interesa conectar la reciente evolución del salario mínimo con la dinámica del sector de exportación en México. En este sentido, las **preguntas de investigación** a las que considero que este proyecto podría contribuir a responder, son las siguientes:

- ¿Cómo afecta el aumento del salario mínimo a la competitividad de las empresas bajo el programa IMMEX en términos de costos laborales y empleo?
- ¿De qué manera el salario mínimo contribuye a reducir o ampliar la brecha salarial dentro de las empresas IMMEX?
- ¿Cómo varía la estructura salarial en empresas IMMEX tras ajustes en el salario mínimo?
- ¿Existen diferencias regionales en el impacto del salario mínimo sobre el empleo y los salarios dentro de las empresas IMMEX?

- **Balanza Comercial de Mercancías de México (BCMM).**

Esta base se compone también de registros administrativos que reflejan información referente al volumen y valor de las exportaciones e importaciones de mercancías en México. El objetivo general de este proyecto del INEGI es, de acuerdo con su sitio:

Proporcionar información que permita conocer el intercambio comercial de mercancías que realiza México con el resto del mundo y que sirva como insumo para el Sistema de Cuentas Nacionales, los cálculos de la Balanza de Pagos y programas relacionados con la vinculación estadística.

La periodicidad de la información es mensual y anual. Rescato también el diseño conceptual presentado por el INEGI en su sitio. De acuerdo con éste, la información está agrupada de la siguiente manera:

- País origen y de último destino conocido.
- Principales productos del Sistema Armonizado.
- Capítulo, partida, subpartida, fracción arancelaria, fracción arancelaria y Número de Identificación Comercial (NICO) conforme a la Tarifa de la Ley de los Impuestos Generales de Importación y de Exportación (TIGIE).
- Balanzas comerciales por grupo de productos.
- Tipo de bien.
- Modo de transporte.
- Principales aduanas.

Los datos revisados a nivel mensual se publican entre 40 y 42 días después del mes de referencia. Por otro lado, el INEGI publica las cifras definitivas de los datos mensuales cinco meses después del año de referencia, mientras que las anuales, con mayor desagregación, siete meses después. La cobertura es nacional.

Ahora bien, en el contexto de la relocalización de las cadenas de suministro en el que el comercio global se encuentra (el famoso *nearshoring*, encuentro interesante la idea que algunos sostienen acerca de la siguiente dinámica: China ha comenzado a invertir en México para exportar a Estados Unidos desde nuestro país, sirviendo como una suerte de caballo de troya. Consecuentemente, algunas de las **preguntas de investigación** que me interesaría intentar responder, con ayuda de los datos de la Balanza Comercial y alguna fuente de datos de inversión extranjera (probablemente de la Secretaría de Economía), son las siguientes:

- ¿Qué sectores industriales en México están recibiendo mayor inversión directa de China, y cuál es su relación con las exportaciones a Estados Unidos?
- ¿Las empresas de capital chino en México muestran un desempeño exportador superior al de las empresas locales en términos de acceso al mercado estadounidense?
- ¿Ha habido un cambio en la composición de las exportaciones mexicanas a Estados Unidos que podría estar asociado a la inversión china en México?
- ¿Existe evidencia de que estas inversiones están orientadas principalmente a la exportación hacia Estados Unidos?

2. Visita el sitio de datos.gob.mx. ¿Qué datos serían interesantes y qué pregunta de investigación podrías contestar con ellos?

Consistentemente con el segundo punto de la pregunta anterior, me resultan interesantes los datos concernientes a la **”Información estadística de la Inversión Extranjera Directa”**. Al conectar los recursos disponibles en esta categoría con la información del INEGI, podría responder las **preguntas de investigación del punto anterior**. El **Registro Nacional de Inversiones Extranjeras (RNIE)** también podría ser de utilidad para ello. Me interesa también construir una red de transmisión de variaciones salariales por empresas o sectores (hasta un cierto dígito de acuerdo con el SCIAN) para entender **si los choques salariales se originan en algún sector y después se contagian al resto de la economía o no**. Considero que la información de este sitio web y del INEGI podrían ser de utilidad para ello

3. Visitar sitio: <https://ourworldindata.org/>, y <https://www.xaber.org.mx/>, y analiza los datos disponibles.

Describe como respuesta los datos que te parecen interesantes de cada punto.

Del **primer sitio web**, me interesan los datos energéticos a nivel global. Creo que podrían ser de utilidad para intentar entender mejor la distribución mundial de la energía y su consumo, las desigualdades existentes y su relación con otras variables de desarrollo humano. Del **segundo sitio web**, encuentro que la sección **”Estadística del Sistema Educativo Mexicano”** tiene datos interesantes correspondientes a los distintos niveles educativos en el país. Sería útil analizar las tendencias en dichos niveles, y su relación con otras variables de relevancia para el desarrollo económico del país.

Problema 3

- (a) Describe el problema de error de medición. Cuando tienes error de medición únicamente en variables independientes, y luego cuando tienes únicamente el error en variables dependientes. Explica tu respuesta con un modelo econométrico.

Respuesta. Comencemos definiendo nuestro modelo econométrico como sigue:

$$Y = \mathbf{X}\beta + u$$

Los errores de medición pueden surgir tanto en las variables independientes como en la variable dependiente. La consecuencia será un sesgo de atenuación en los parámetros de MCO, es decir, una subestimación de los estimadores. A continuación, una descripción de dos casos de errores de medición:

- **Error de medición en variable dependiente.**

Supongamos que el modelo de regresión verdadero es el mencionado anteriormente, pero nosotros observamos solamente

$$\tilde{Y} = \mathbf{X}\beta + u$$

Donde $\tilde{Y} = Y + v$; y es el término v el que representa el error de medición. Supongamos además que $v \sim N(0, \sigma_v^2)$, $cov(v, \mathbf{X}) = 0$ y $cov(v, u) = 0$.

De este modo, los estimadores de MCO son:

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\tilde{Y}) = (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'(Y + v))$$

Si calculamos la esperanza de $\hat{\beta}$ veremos que los estimadores siguen siendo tanto

condicional como incondicionalmente insesgados:

$$\begin{aligned}
\mathbb{E}[\hat{\beta}|\mathbf{X}] &= \mathbb{E}[(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'(Y + v))|\mathbf{X}] \\
&= \mathbb{E}[(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'(\mathbf{X}\beta + u + v))|\mathbf{X}] \\
&= \mathbb{E}[(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X}\beta + \mathbf{X}'u + \mathbf{X}'v)|\mathbf{X}] \\
&= \mathbb{E}[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'u + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'v|\mathbf{X}] \\
&= \mathbb{E}[\beta|\mathbf{X}] + \mathbb{E}[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'u|\mathbf{X}] + \mathbb{E}[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'v|\mathbf{X}] \\
&= \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbb{E}[u|\mathbf{X}] + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbb{E}[v|\mathbf{X}] \\
&= \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' * 0 + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' * 0 \\
&= \beta
\end{aligned}$$

Lo que demuestra que el estimador es condicionalmente insesgado. Este hecho implica, además, que $\mathbb{E}[\hat{\beta}] = \mathbb{E}[\mathbb{E}[\hat{\beta}|\mathbf{X}]] = \mathbb{E}[\beta] = \beta$, lo que demuestra la insesgadez del estimador.

No obstante, el problema surge con los errores estandar; éstos serán más grandes, dado que la varianza de los residuales es mayor (recordemos que $\beta = \Sigma_X^{-1}\sigma'_{YX}$):

$$\begin{aligned}
\tilde{Y} = \mathbf{X}\beta + u &\iff Y + v = \mathbf{X}\beta + u \\
&\iff u = Y - \mathbf{X}\beta + v \\
&\iff \text{var}(u) = \text{var}(Y - \mathbf{X}\beta + v) \\
&\iff \text{var}(u) = \text{var}(Y - \mathbf{X}\beta) + \text{var}(v) + 2\text{cov}(Y - \mathbf{X}\beta, v) \\
&\iff \text{var}(u) = \text{var}(Y) + \beta'\text{var}(\mathbf{X})\beta - 2\text{cov}(Y, \mathbf{X})\beta + \sigma_v^2 + \\
&\quad 2[\text{cov}(Y, v) - \text{cov}(\mathbf{X}, v)\beta] \\
&\iff \text{var}(u) = \sigma_Y^2 + \beta'\Sigma_x\beta - 2\sigma_{YX}\beta + \sigma_v^2 + \\
&\quad 2[\text{cov}(\mathbf{X}\beta + u, v) - 0 * \beta] \\
&\iff \text{var}(u) = \sigma_Y^2 + (\beta'\Sigma_X - 2\sigma_{YX})\beta + \sigma_v^2 + \\
&\quad 2[\text{cov}(\mathbf{X}, v)\beta + \text{cov}(u, v)] \\
&\iff \text{var}(u) = \sigma_Y^2 + (\sigma_{YX}\Sigma_X^{-1}\Sigma_X - 2\sigma_{YX})\Sigma_X^{-1}\sigma'_{YX} + \sigma_v^2 + \\
&\quad 2[0 * \beta + 0] \\
&\iff \text{var}(u) = \sigma_Y^2 - \sigma_{YX}\Sigma_X^{-1}\sigma'_{YX} + \sigma_v^2 \\
&\iff \text{var}(u) = \sigma_u^2 + \sigma_v^2
\end{aligned}$$

Entonces, $\text{var}(u) = \sigma_u^2 + \sigma_v^2 > \sigma_u^2$; es decir, la varianza de los residuales observados

es mayor a la varianza de los residuales verdaderos. En conclusión, los estimadores siguen siendo insesgados, pero se pierde precisión dada su mayor varianza.

- **Error de medición en variables independientes.**

Al igual que en el caso anterior, vamos a suponer que el modelo de regresión verdadero es $Y = \mathbf{X}\beta + u$, pero nosotros observamos

$$Y = \tilde{\mathbf{X}}\beta + u$$

Donde $\tilde{\mathbf{X}} = \mathbf{X} + w$; y el término w representa el error de medición. Supongamos además que $w \sim N(0, \sigma_w^2)$, $cov(w, Y) = 0$ y $cov(w, u) = 0$.

De este modo, los estimadores de MCO son:

$$\begin{aligned}\hat{\beta} &= (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'Y) \\ &= (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'(\tilde{\mathbf{X}}\beta + u)) \\ &= (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\tilde{\mathbf{X}}\beta + \mathbf{X}'u) \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\tilde{\mathbf{X}}\beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'u \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X} + w)\beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'u \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'w\beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'u \\ &= \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(w\beta + u)\end{aligned}$$

¿Son los estimadores de MCO insesgados? Observemos lo que ocurre:

$$\begin{aligned}\mathbb{E}[\hat{\beta}|\mathbf{X}] &= \mathbb{E}[\beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(w\beta + u)|\mathbf{X}] \\ &= \beta + \mathbb{E}[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(w\beta + u)|\mathbf{X}] \\ &= \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbb{E}[(w\beta + u)|\mathbf{X}] \\ &= \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbb{E}[w\beta|\mathbf{X}] + \mathbb{E}[u|\mathbf{X}]) \\ &= \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbb{E}[w\beta|\mathbf{X}] + 0) \\ &= \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbb{E}[w\beta|\mathbf{X}] \\ &\neq \beta\end{aligned}$$

Por tanto, los estimadores de MCO no son insesgados.

Ahora bien, recordemos que sí observamos el Y verdadero: $Y = \mathbf{X}\beta + u$, y que, además, $\tilde{\mathbf{X}} = \mathbf{X} + w$.

Utilizando lo anterior en la definición de los estimadores de MCO tenemos que:

$$\begin{aligned}
\hat{\beta} &= \frac{\text{cov}(\tilde{\mathbf{X}}, Y)}{\text{var}(\tilde{\mathbf{X}})} \\
&= \frac{\text{cov}(\mathbf{X} + w, \mathbf{X}\beta + u)}{\text{var}(\tilde{\mathbf{X}})} \\
&= \frac{\text{cov}(\mathbf{X}, \mathbf{X}\beta) + \text{cov}(\mathbf{X}, u) + \text{cov}(w, \mathbf{X}\beta) + \text{cov}(w, u)}{\text{var}(\tilde{\mathbf{X}})} \\
&= \frac{\text{var}(\mathbf{X})\beta + 0 + 0 + 0}{\text{var}(\tilde{\mathbf{X}})} \\
&= \frac{\text{var}(\mathbf{X})}{\text{var}(\tilde{\mathbf{X}})}\beta
\end{aligned}$$

Sabiendo que $\text{var}(\mathbf{X}) = \sigma_X^2$ y que $\text{var}(\tilde{\mathbf{X}}) = \text{var}(\mathbf{X} + w) = \sigma_X^2 + \sigma_w^2$; el resultado anterior implica que

$$\mathbb{E}[\hat{\beta}] = \mathbb{E}\left[\frac{\sigma_X^2}{\sigma_X^2 + \sigma_w^2}\beta\right] = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_w^2}\beta$$

Donde el cociente es conocido como la razón "noise-to-signal". A medida que la varianza de w crece, el sesgo de atenuación de los estimadores de MCO incrementa; esto puede provocar que, en casos extremos donde la varianza es demasiado grande, el efecto de \mathbf{X} en Y sea cero. En el caso donde la varianza de w fuera cero (w es constante), los estimadores serían insesgados.

- (b) Describe el problema de sesgo por variables omitidas. Deriva el sesgo con una regresión lineal simple, y explica el sesgo positivo o negativo dependiendo de las correlaciones de la fórmula.

Respuesta. El sesgo por variables omitidas ocurre cuando en un modelo de regresión lineal no se incluye una o más variables relevantes que están correlacionadas tanto con la variable dependiente como con una o más variables independientes incluidas en el modelo. Esto puede llevar a estimaciones sesgadas de los coeficientes de regresión, lo que da lugar a interpretaciones incorrectas.

Para ilustrar este problema y calcular el sesgo en los estimadores de MCO, consideremos el siguiente modelo de regresión lineal:

$$Y = \alpha + \beta_X X + \beta_Z Z + u$$

Si omitimos la variable Z , el modelo a estimar sería

$$Y = \alpha + \beta_X X + v$$

Donde $v = \beta_Z Z + u$.

Ahora bien, nos interesa saber si existe un sesgo en el estimador de MCO del modelo con variable omitida. Sabemos que:

$$\hat{\beta}_X = \frac{\text{cov}(X, Y)}{\text{var}(X)}$$

Sustituyendo el valor de Y y desarrollando tenemos:

$$\begin{aligned} \hat{\beta}_X &= \frac{\text{cov}(X, \alpha + \beta_X X + v)}{\text{var}(X)} \\ &= \frac{\text{cov}(X, \alpha) + \text{cov}(X, \beta_X X) + \text{cov}(X, v)}{\text{var}(X)} \\ &= \frac{0 + \beta_X \text{var}(X) + \text{cov}(X, \beta_Z Z + u)}{\text{var}(X)} \\ &= \beta_X + \frac{\beta_Z \text{cov}(X, Z) + \text{cov}(X, u)}{\text{var}(X)} \\ &= \beta_X + \frac{\beta_Z \text{cov}(X, Z) + 0}{\text{var}(X)} \\ &= \beta_X + \beta_Z \frac{\text{cov}(X, Z)}{\text{var}(X)} \end{aligned}$$

El sesgo en β_X es entonces:

$$\text{Sesgo} = \mathbb{E}[\hat{\beta}_X] - \beta_X = \beta_Z \frac{\text{cov}(X, Z)}{\text{var}(X)}$$

Este sesgo entonces puede ser positivo o negativo según sea el caso:

- $\text{cov}(X, Z) > 0$ y $\beta_Z > 0 \implies \text{Sesgo} > 0$
- $\text{cov}(X, Z) > 0$ y $\beta_Z < 0 \implies \text{Sesgo} < 0$
- $\text{cov}(X, Z) < 0$ y $\beta_Z > 0 \implies \text{Sesgo} < 0$
- $\text{cov}(X, Z) < 0$ y $\beta_Z < 0 \implies \text{Sesgo} > 0$

Es decir, si suponemos que la variable omitida (Z) tiene un efecto positivo sobre la variable dependiente ($\beta_Z > 0$), el sesgo en el estimador de MCO será positivo si la variable independiente (X) está correlacionada positivamente con la variable omitida (i.e. $cov(X, Z) > 0$) o negativo si la correlación es negativa. Dependiendo de las condiciones, el signo del sesgo puede variar.

Problema 4

Con los datos del problema 1 para cada año mencionado en la ENIGH, ENOE, IMSS (cada trimestre, por ejemplo elegir el mes de en medio en trimestre), restringiendo a personas de 20-65 años de edad: (con IMSS lo que se pueda)

1. Realiza una tabla con el número de observaciones: hogares e individuos, el % con grado de educación universitaria, el % que trabaja, el % sector rural. Define claramente estas variables de tal manera que sean comparables en el tiempo. Todo con el comando putexcel (estout, xtable, o cualquier otro), no se permite copy-paste. Dos decimales máximo.

Respuesta.

ENIGH: Características Sociodemográficas 1992 - 2022

Año	Individuos	Hogares	Grado Universitario	Sí trabaja	Vivienda Rural
1992	41.3	17.7	5.4	60.7	23.1
1994	44.1	18.8	5.6	62.1	22.8
1996	46.0	19.7	5.9	63.7	22.0
1998	48.3	21.1	6.4	65.0	21.9
2000	51.4	22.5	8.3	64.9	21.7
2002	53.2	23.3	7.4	66.1	21.8
2004	54.7	24.1	15.0	65.3	21.1
2006	57.4	25.8	15.7	69.4	20.5
2008	59.8	26.2	15.9	65.4	20.4
2010	62.6	27.8	17.5	63.9	20.5
2012	65.2	29.5	17.8	70.5	20.6
2014	67.1	29.5	18.8	68.4	20.8
2016	68.4	30.7	19.8	71.8	21.0
2018	71.1	32.0	21.4	72.1	22.2
2020	74.4	33.0	22.9	68.6	21.0
2022	76.2	34.5	23.8	72.2	22.5

Individuos y hogares en millones; el resto en porcentajes

Elaboración propia con datos de la ENIGH

2. Tabla. ¿Puedes identificar el % de mujeres que trabaja, cómo ha evolucionado en el tiempo? ¿Cómo ha evolucionado para los años mencionados por nivel educativo o por estado civil o por si tienen hijos? Todo con el comando putexcel, no se permite cypaste. Dos decimales máximo.

Respuesta. Siguiendo página.

ENIGH: Participación laboral femenina 1992 - 2022

Año	Participación laboral femenina (%)
1992	34.9
1994	39.5
1996	43.0
1998	44.8
2000	43.8
2002	47.4
2004	45.9
2006	52.6
2008	47.4
2010	46.1
2012	55.7
2014	52.3
2016	57.3
2018	57.7
2020	54.9
2022	59.0

Elaboración propia con datos de la ENIGH

La tabla anterior muestra la evolución de la participación laboral femenina desde 1992 hasta 2022. En la siguiente página se encuentra la tabla con la evolución por nivel educativo.

ENIGH: Participación laboral femenina por nivel educativo 1992 - 2022

Año	Sin Educación	Primaria	Secundaria	Preparatoria	Universidad	Posgrado
1992	54.1	56.6	63.2	71.0	82.7	90.3
1994	48.6	59.1	64.6	68.8	87.1	90.4
1996	49.1	61.5	65.6	72.5	84.9	91.2
1998	0.0	63.9	66.9	72.2	83.1	94.8
2000	0.0	63.7	68.3	69.9	85.3	90.9
2002	0.0	63.9	69.6	70.2	82.1	70.6
2004	47.5	61.6	69.0	70.5	71.0	86.0
2006	59.3	64.6	72.4	74.5	73.2	87.6
2008	48.2	60.3	68.6	71.3	71.9	85.6
2010	48.3	58.4	66.6	69.8	67.7	87.2
2012	60.4	68.7	71.9	73.3	70.9	92.3
2014	53.2	65.9	70.4	70.4	70.2	85.9
2016	60.0	69.9	72.8	74.2	72.7	86.6
2018	61.8	69.9	72.6	74.4	73.3	86.3
2020	57.5	66.9	69.2	70.8	69.0	83.3
2022	58.9	69.4	72.7	74.8	73.4	86.7

Elaboración propia con datos de la ENIGH

ENIGH: Participación laboral femenina por estado civil 1992 - 2022

Año	Unión libre	Casada	Separada	Divorciada	Viuda	Soltera
1996	66.8	63.4	71.8	76.3	54.4	67.6
1998	69.2	63.9	70.7	76.7	59.2	70.5
2000	66.2	64.5	75.0	78.4	54.0	69.4
2002	69.3	65.9	74.9	76.8	57.0	67.6
2004	63.8	73.6	77.3	55.1	64.0	68.0
2006	70.6	77.1	79.7	60.8	67.5	71.5
2008	65.7	71.6	76.9	56.4	63.7	68.2
2010	64.9	62.2	72.7	75.4	51.2	65.7
2012	72.9	69.2	76.6	84.4	54.7	70.4
2014	69.1	67.4	77.9	76.1	61.5	67.8
2016	73.0	70.2	79.6	80.9	62.6	71.6
2018	73.8	70.5	80.5	80.2	62.9	71.2
2020	71.5	66.8	76.2	74.0	57.1	67.4
2022	74.4	70.2	80.0	79.7	62.5	71.7

Elaboración propia con datos de la ENIGH

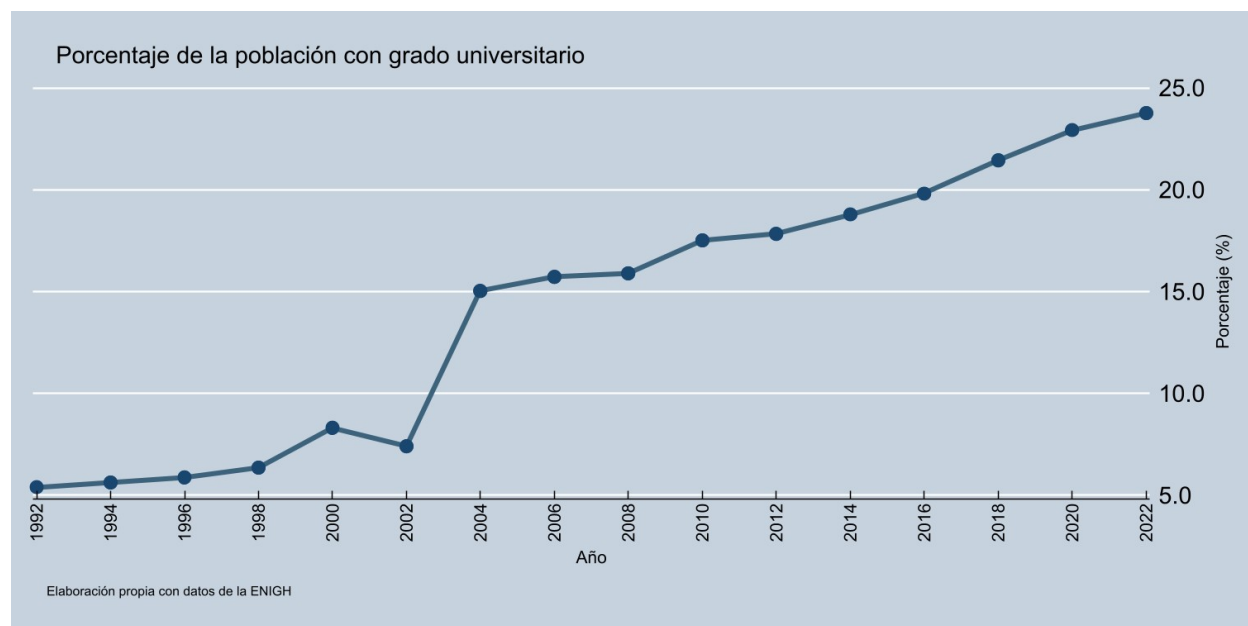
Table 5: ENIGH: Participación laboral femenina con hijos y sin hijos 1992 - 2022

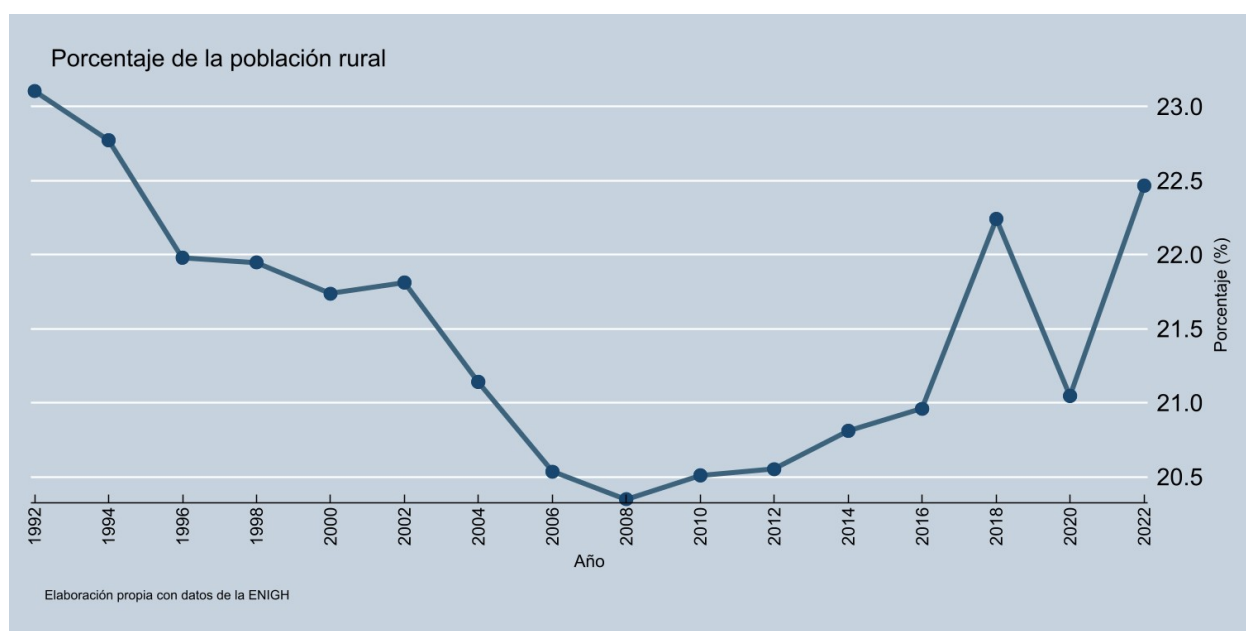
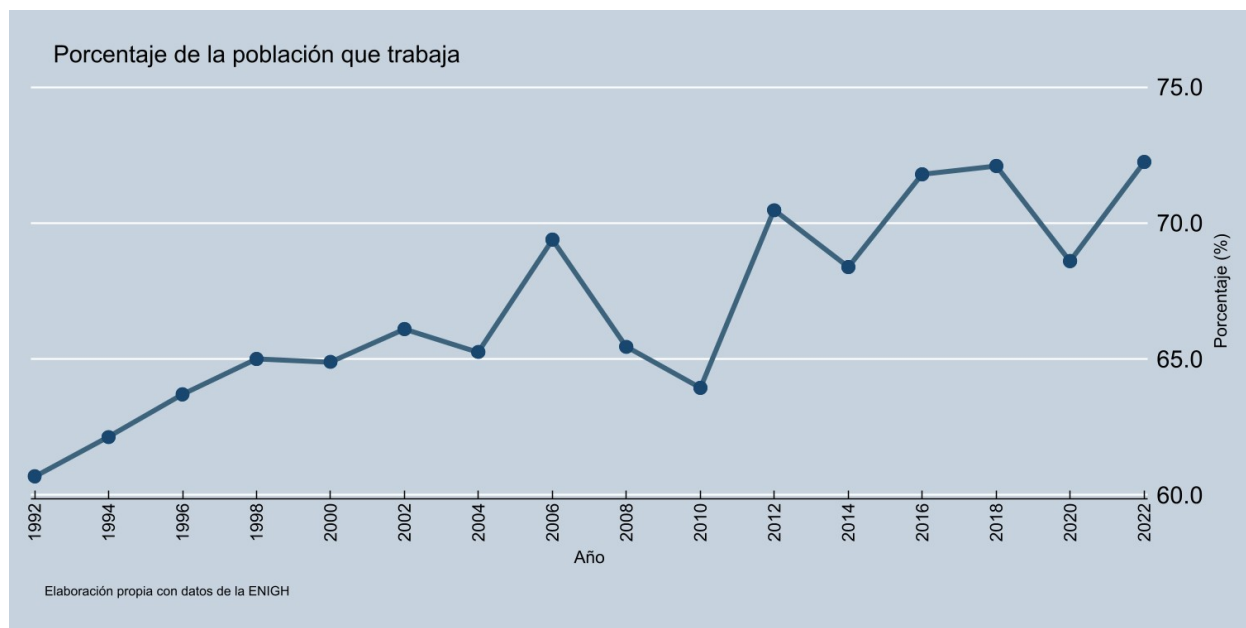
Año	Sin hijos	Con hijos
2004	81.8	42.9
2006	84.3	49.7
2008	34.9	44.5
2010	56.4	43.5
2012	46.4	53.9
2014	40.9	50.4
2016	62.2	55.6
2018	59.2	56.2
2020	65.8	53.4
2022	58.0	57.4

Elaboración propia con datos de la ENIGH

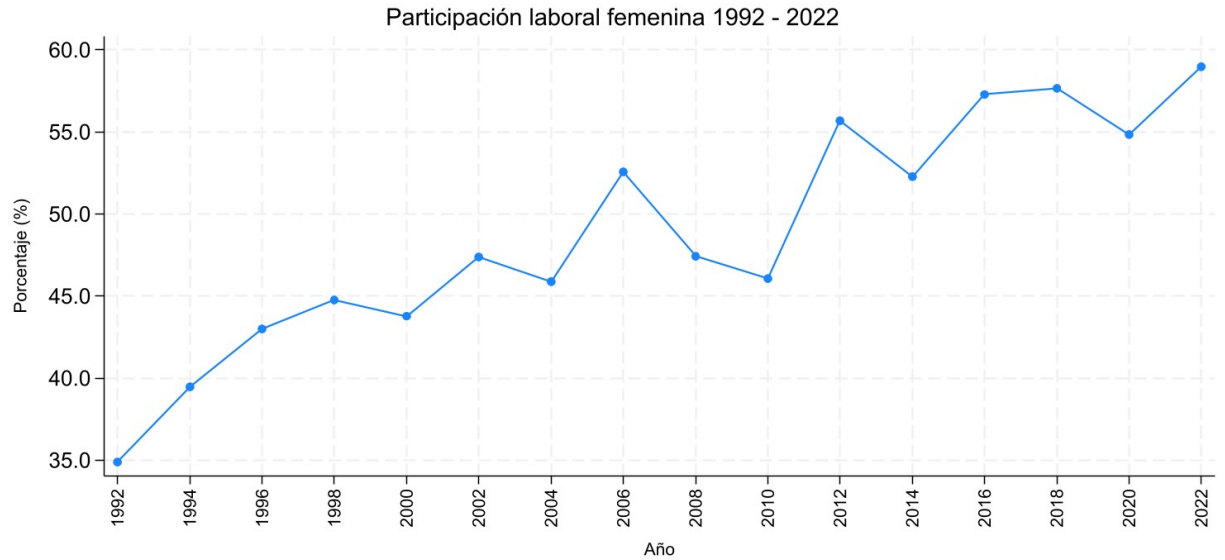
3. La parte 1 de este problema realiza figuras similares al The Economist. Busca el comando `grstyle` y `palette` en Stata para realizarlas. Para la parte 2 usa los colores standard de Stata pero con background blanco y leyenda bonita entendible. El eje x debe tener el formato de fecha en stata. Verlos comandos `ym` o `yq` para definir fechas.

Respuesta. Parte 1.

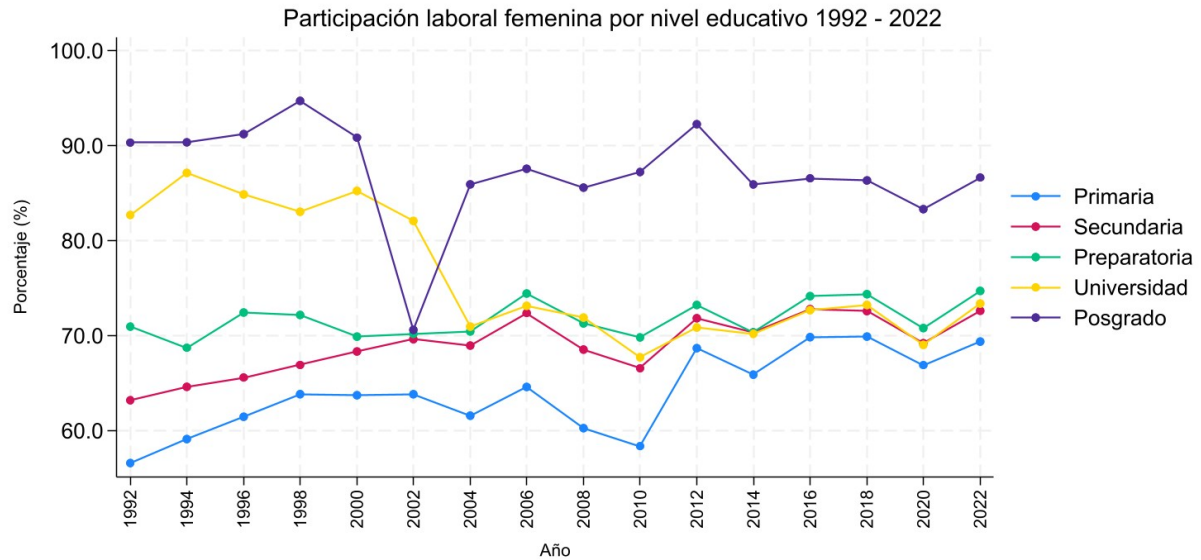




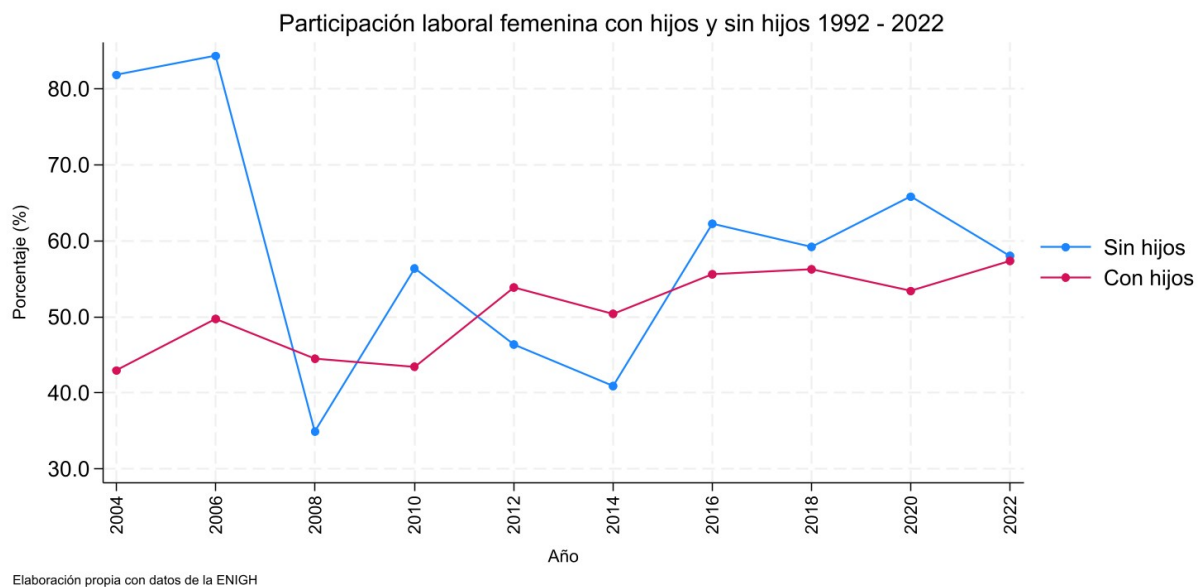
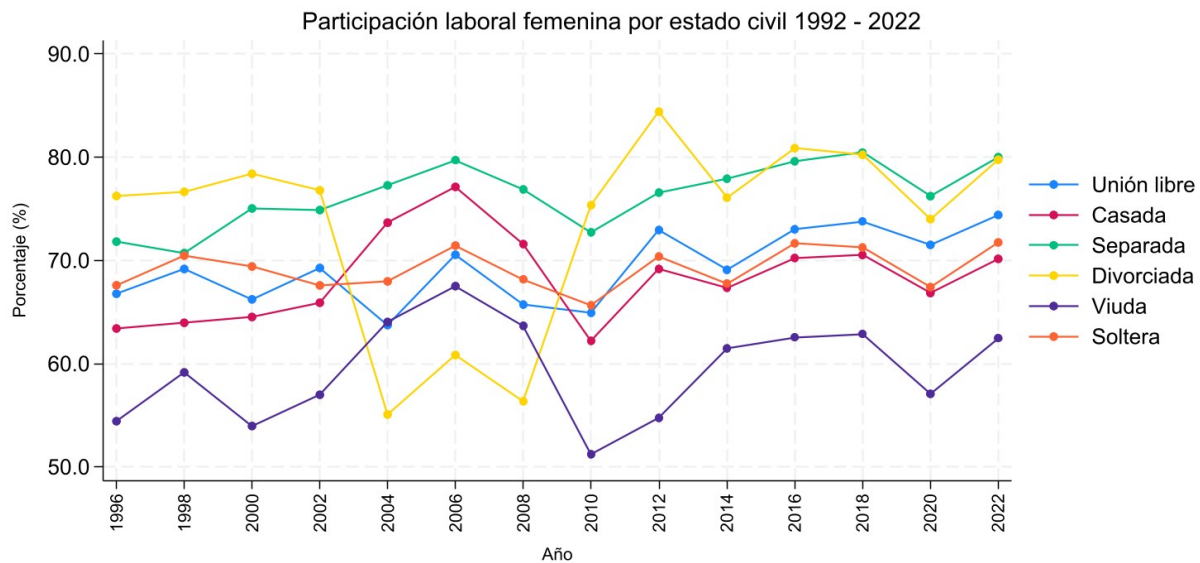
Parte 2.



Elaboración propia con datos de la ENIGH



Elaboración propia con datos de la ENIGH



Problema 5

Existe una discusión sobre si los experimentos controlados nos sirven para combatir pobreza, si son éticos, si es la forma que debemos hacer economía. Leer los siguientes artículos y escribir un ensayo al respecto de 2 páginas a doble espacio.

<https://voxeu.org/article/alleviating-poverty-experimental-research-2019-nobel-laureates>

<https://voxeu.org/article/what-randomisation-can-and-cannot-do-2019-nobel-prize>

<https://voxeu.org/article/limitations-randomised-controlled-trials>

<https://theconversation.com/2019-nobel-prize-in-economics-the-limits-of-the-clinical-trial-method-125888>

<https://www.economist.com/1843/2024/03/01/how-poor-kenyans-became-economists-guinea-pigs>

Respuesta. EL ENSAYO COMIENZA EN LA SIGUIENTE PÁGINA.

Claroscuros del enfoque experimental en la Economía del Desarrollo.

El enfoque experimental en economía, particularmente el de las Pruebas Controladas Aleatorizadas (RCTs por sus siglas en inglés), ha ganado prominencia en los últimos años, marcando una notable transición desde una economía predominantemente teórica hacia una más empírica. El reconocimiento de Abhijit Banerjee, Esther Duflo y Michael Kremer con el Premio Nobel de Economía en 2019 subraya el impacto significativo de este método en la investigación sobre la pobreza. No obstante, esta metodología ha suscitado intensas críticas que van más allá de cuestiones metodológicas, abordando aspectos epistemológicos, políticos y éticos.

El artículo “Alleviating poverty with experimental research: The 2019 Nobel laureates” refleja cómo este cambio hacia un enfoque experimental ha llevado a algunos economistas a desapegarse del enfoque teórico tradicional para adentrarse al trabajo de campo. Aunque este giro ha permitido avances significativos, el artículo sostiene que el trabajo teórico no debe ser relegado. La verdadera fortaleza de la investigación económica reside en la integración sinérgica de los enfoques teórico y experimental. El Nobel a Banerjee, Duflo y Kremer no solo celebra sus logros con los RCTs, sino que también subraya la necesidad de combinar teoría robusta con evidencia empírica para explorar nuevas preguntas y responder de manera más completa a los fenómenos sociales. A pesar de estos avances, el artículo “What randomisation can and cannot do: The 2019 Nobel Prize” destaca las limitaciones inherentes a los RCTs. Aunque estos estudios ofrecen una valiosa evidencia empírica y han avanzado significativamente en la capacidad de diseñar y controlar experimentos, su validez externa, o la capacidad para generalizar resultados a contextos más amplios, sigue siendo una preocupación. Los RCTs tienden a simplificar problemas complejos y no siempre capturan las dinámicas subyacentes del desarrollo económico. Esta crítica es fundamental para entender que, aunque los RCTs representan un avance, es crucial complementar estos estudios con otros enfoques que puedan ofrecer una visión más completa. El artículo “The limitations of randomised controlled trials” profundiza en los inconvenientes asociados con la metodología

de los RCTs cuando no se ejecutan con la rigurosidad adecuada. Los problemas incluyen los sesgos en los tratamientos; outliers en las muestras; problemas de inferencia; la influencia de factores previos en los resultados; entre otros. Además, se critica la falta de conexión entre el "conocimiento previo" y la metodología experimental, destacando la necesidad de un marco teórico sólido para interpretar correctamente los resultados y evaluar su aplicabilidad en contextos diversos (el problema de la "transportación"). Por su parte, "2019 Nobel Prize in Economics: The limits of the clinical trial method" cuestiona la capacidad de los RCTs para abordar problemas estructurales y sistémicos de la pobreza y la desigualdad. La crítica sugiere que el enfoque experimental, en lugar de analizar las causas estructurales y sistémicas, tiende a centrarse en intervenciones aisladas, lo que podría desviar la atención de problemas más amplios que requieren soluciones integrales. El artículo también plantea preocupaciones sobre el predominio de los RCTs y su efecto de "crowding-out" sobre otras metodologías, cuestionando si este enfoque es científicamente válido y socialmente deseable. Finalmente, el artículo "How poor Kenyans became economists' guinea pigs" aborda los dilemas éticos asociados con los RCTs, especialmente en contextos de alta vulnerabilidad como en África subsahariana, reflejando también la controversia detrás de este enfoque. Rescato la posibilidad de que los sujetos de los experimentos, y sus familias, se conviertan en conejillos de indias sin suficiente consideración ética, como el artículo parece denunciar sutilmente en algunos párrafos. La combinación de académicos, filántropos y gobiernos débiles parece influir en la proliferación de estos estudios en regiones marginadas (no creo que sea coincidencia), lo que plantea interrogantes sobre el impacto y la ética de estos experimentos. En conclusión, aunque los RCTs han sido fundamentales para avanzar en la investigación económica y mejorar las políticas de desarrollo, es esencial reconocer y abordar sus limitaciones. La integración de enfoques teóricos con metodologías experimentales, junto con una consideración ética y crítica, podría permitir una comprensión más completa y efectiva de los problemas económicos y sociales complejos. Solo mediante un enfoque integrado se podrán formular políticas efectivas y escalables para abordar los desafíos del desarrollo económico.

Problema 6

Usando los microdatos de las ENIGH de 2016 a 2022 calcula pobreza por ingresos y pobreza multidimensional. Replica el ejercicio hecho por CONEVAL (los programas están disponibles en su página de internet) para obtener el % de personas en situación de pobreza. Incluir figuras similares y cuadros similares.

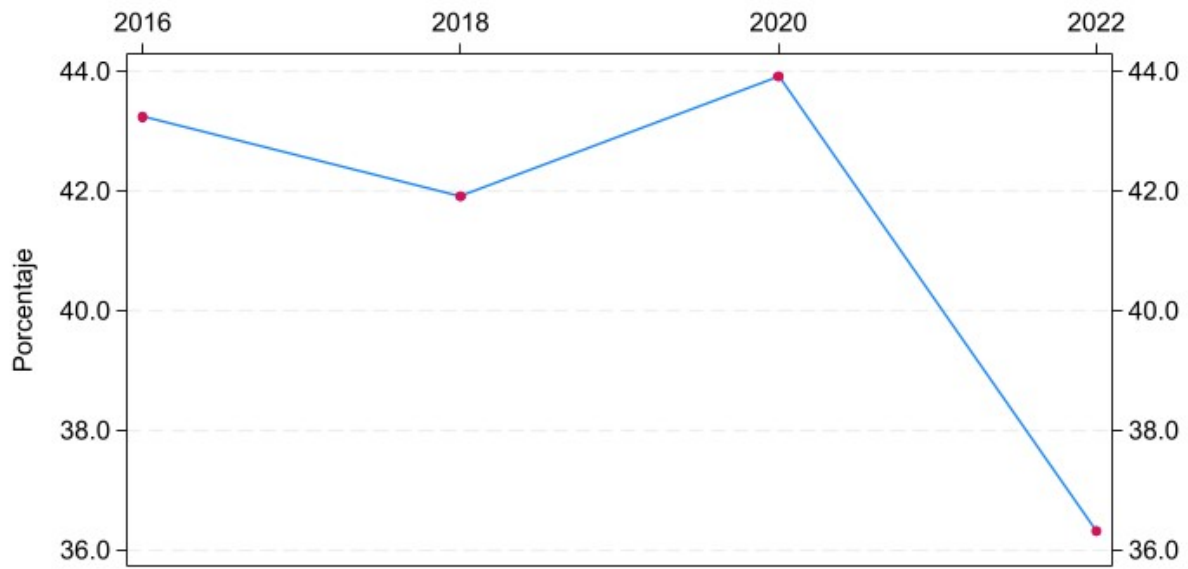
Respuesta. Replicando los programas del CONEVAL, se trabajaron las bases de datos generadas por los mismos. Mediante un programa, se construyó la siguiente tabla:

Indicadores de pobreza multidimensional

Indicador / Año	2016	2018	2020	2022
Población en situación de pobreza	43.2	41.9	43.9	36.3
Población en situación de pobreza moderada	36.0	34.9	35.4	29.3
Población en situación de pobreza extrema	7.2	7.0	8.5	7.1
Población vulnerable por carencias sociales	25.3	26.4	23.7	29.4
Población vulnerable por ingresos	7.6	8.0	8.9	7.2
Población no pobre y no vulnerable	24.0	23.7	23.5	27.1
Población con al menos una carencia social	68.5	68.3	67.6	65.7
Población con al menos tres carencias sociales	20.0	20.2	23.0	24.9
Rezago educativo	18.5	19.0	19.3	19.4
Carencia por acceso a los servicios de salud	15.6	16.2	28.2	39.1
Carencia por acceso a la seguridad social	54.1	53.5	52.0	50.2
Carencia por calidad y espacios de la vivienda	12.0	11.0	9.3	9.1
Carencia por acceso a los servicios básicos en la vivienda	19.2	19.6	17.9	17.8
Carencia por acceso a la alimentación nutritiva y de calidad	21.9	22.2	22.5	18.2
Población con ingreso inferior a la línea de pobreza extrema por ingresos	14.9	14.0	17.2	12.1
Población con ingreso inferior a la línea de pobreza por ingresos	50.8	49.9	52.8	43.5

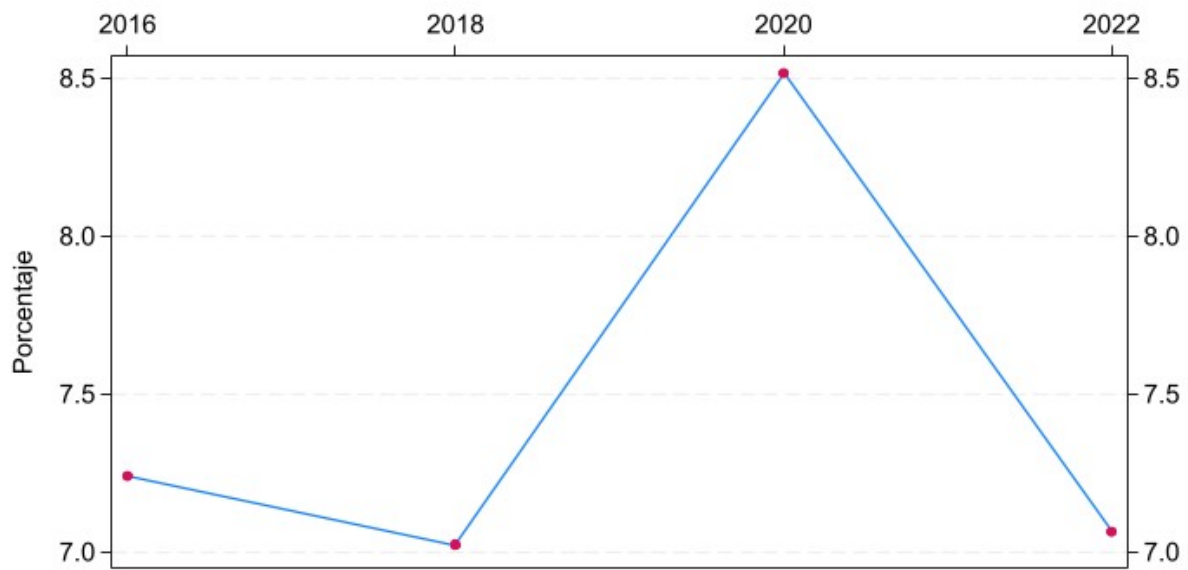
El archivo .do utilizado para calcular los porcentajes y construir la tabla en Latex se compartirá en el Teams. Asimismo, se generaron las gráficas de las series de tiempo del porcentaje de la población en pobreza y pobreza extrema. Las figuras se encuentran en la siguiente página.

Población en situación de pobreza



Fuente: Elaboración propia con datos del INEGI

Población en situación de pobreza extrema



Fuente: Elaboración propia con datos del INEGI