

A satellite image of the Earth, showing the Americas, Europe, and Africa. The image is darkened to serve as a background for text. The title is centered over the North American continent.

# Geografía de la pobreza en México

Dr. Héctor Nájera

# Antecedentes

- Los primeros pasos a la caracterización territorial de las carencias y niveles de vida de la población en México

- Índice de marginación



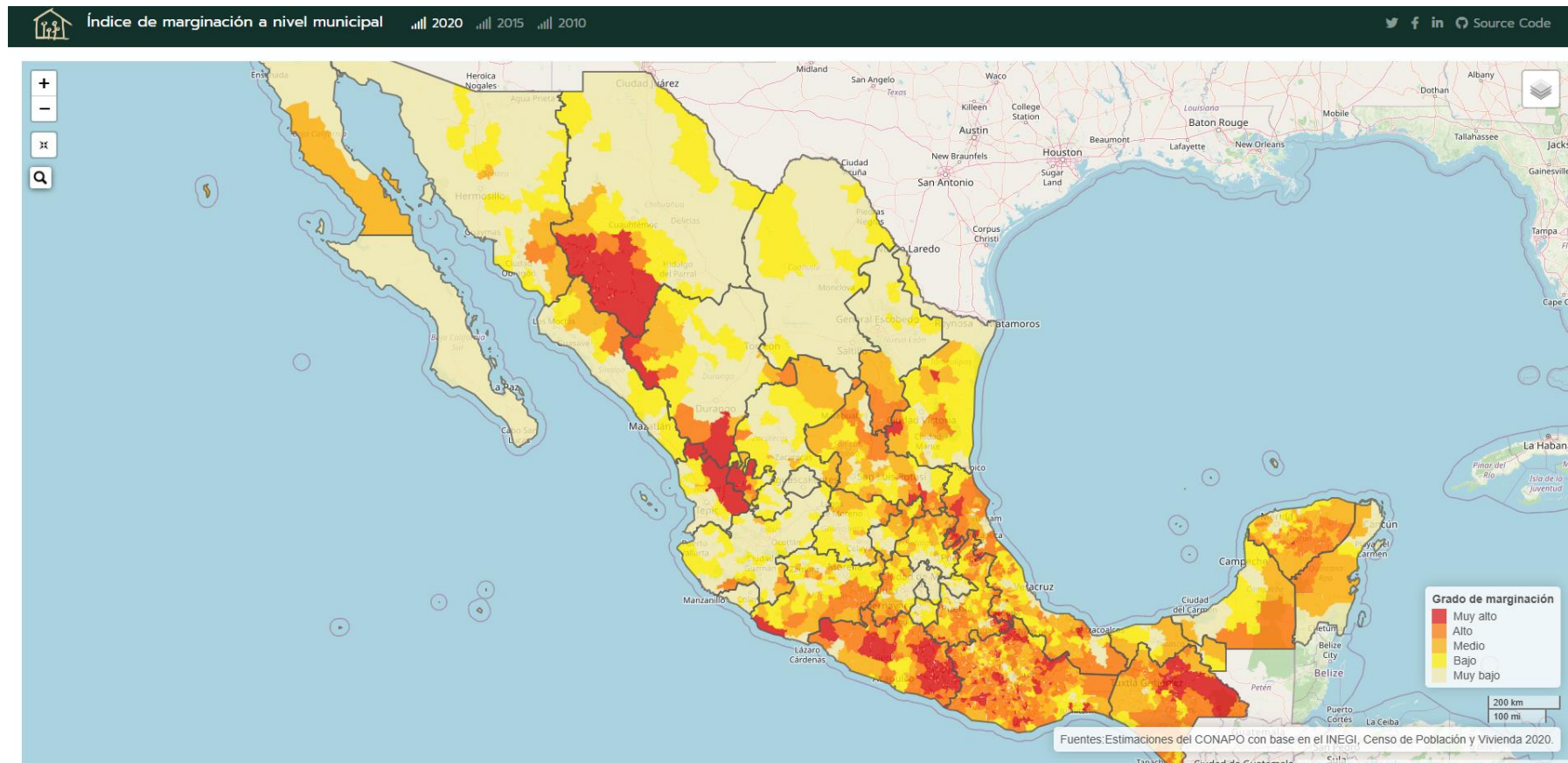
- Índice de rezago social





# Antecedentes

- Índice de marginación: CONAPO
- 1990-2010-2020
- Estado, municipio, localidad, AGEB y colonia



# Índice de marginación

- *“Identificar las disparidades territoriales existentes en el país en un momento determinado, una cualidad que le ha conferido un valor relevante como herramienta analítica y operativa para la definición y focalización de políticas públicas, enfocadas al abatimiento de las carencias socioeconómicas de la población mexicana.”*
- Índice de marginación  $\neq$  Pobreza
  - Aunque pudieran estar correlacionados. Son fenómenos distintos.

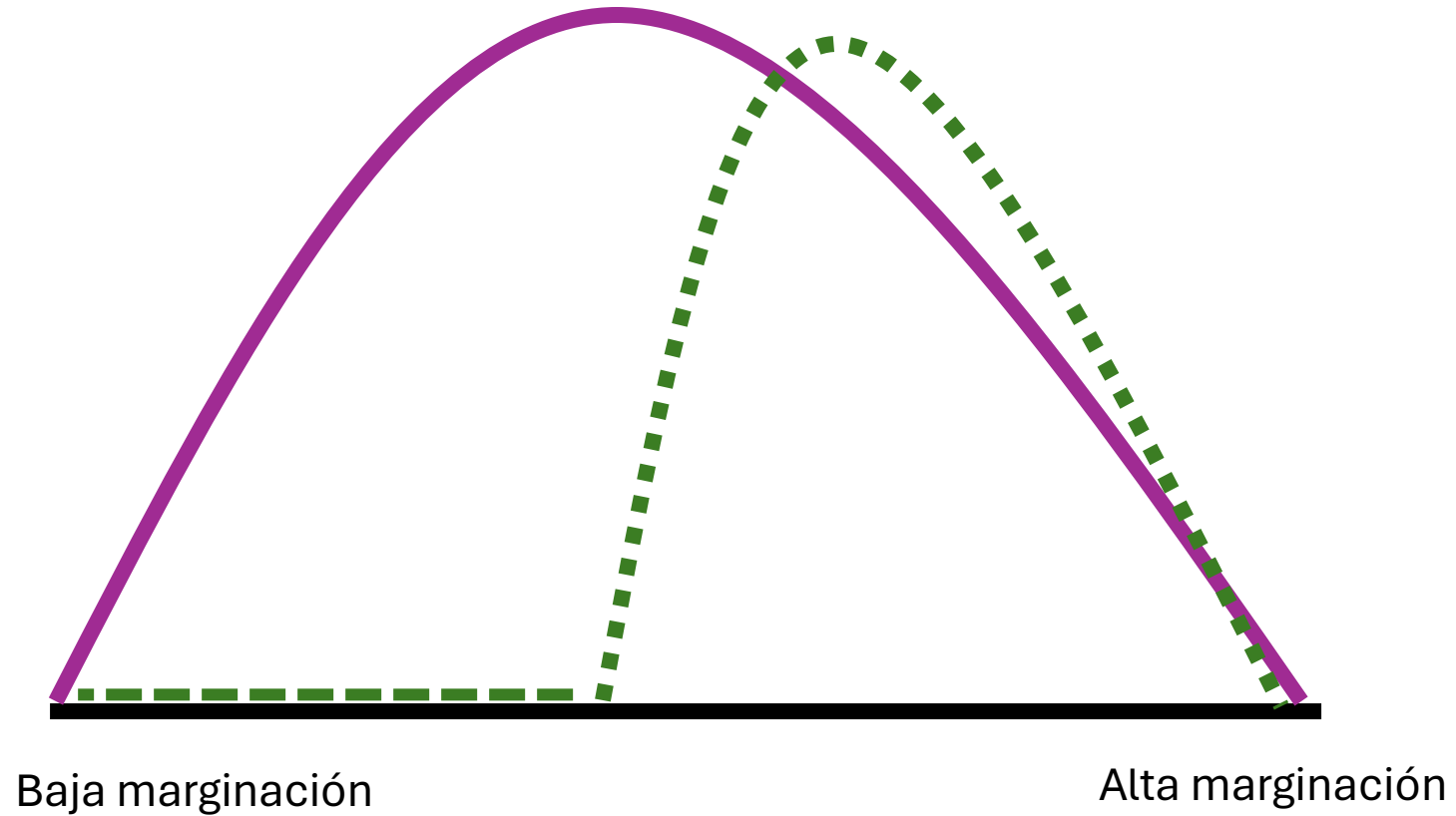
# Problemas: índice de marginación

- El índice es producto de una iniciativa que buscaba aprovechar los datos existentes en los censos de población y las técnicas de reducción de datos (**componentes principales**) para brindar cifras sobre la desigualdad territorial en México
- No tiene un modelo de medición que lo sustente
- La comparabilidad entre distintos puntos en el tiempo está altamente comprometida
- El índice original y el modificado tienen un alto error de medición en el centro y la cola izquierda de la distribución

# Problemas: índice de marginación

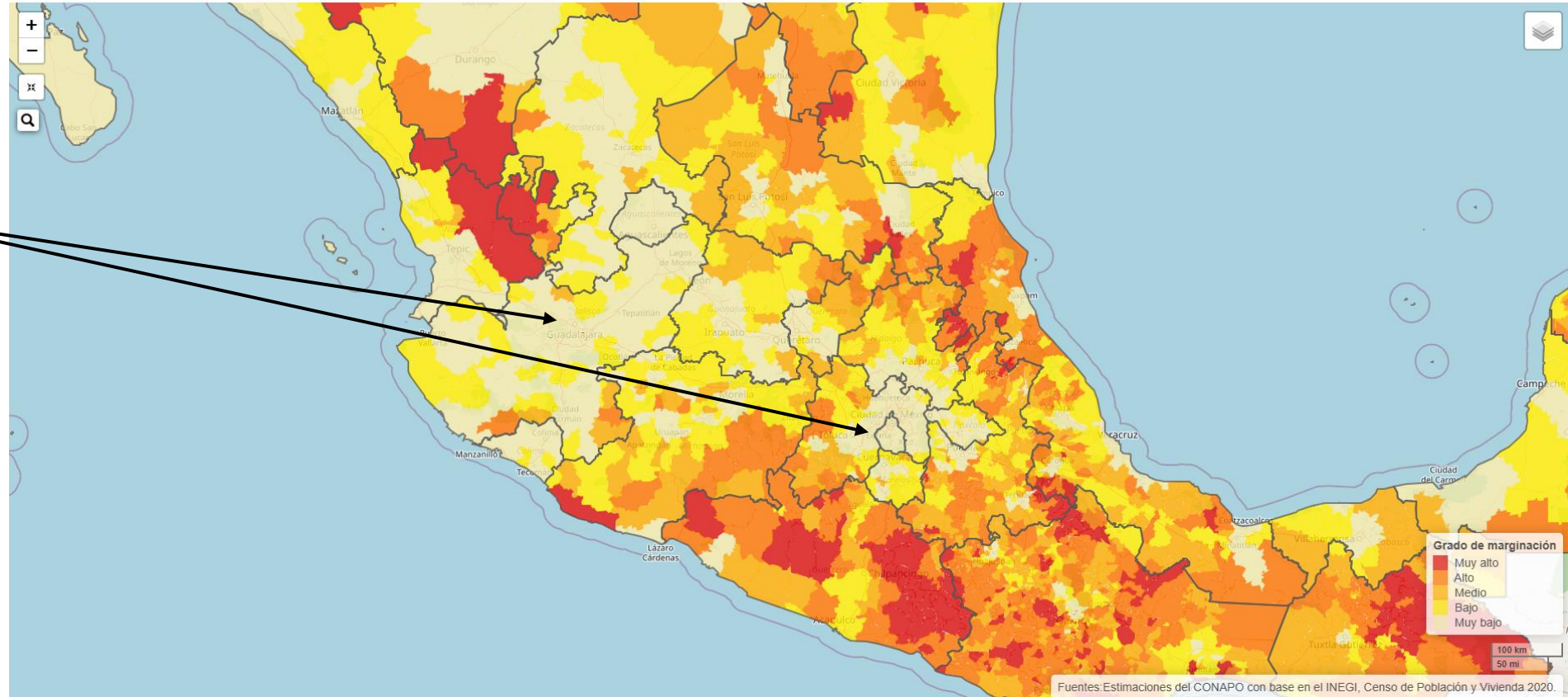
Verdadera

Observada



# Problemas: índice de marginación

Muy Bajo: ¿Es igualmente muy bajo?



# Índice de rezago social

- CONEVAL: De acuerdo a la LGDS debe producir cifras de pobreza a nivel municipal, pero en un inicio no estaba claro cómo podía lograrse
- Al igual que el índice de marginación surge de la necesidad de contar con datos a nivel territorial que reflejen los diferenciales en carencia de ciertos bienes y servicios
- Busca hacer el mejor uso posible de la información disponible para generar información de uso público: **Zonas de Atención Prioritaria (ZAP)**
- **No tiene un modelo de medición.** Explora datos disponibles mediante técnicas de reducción de varianza.

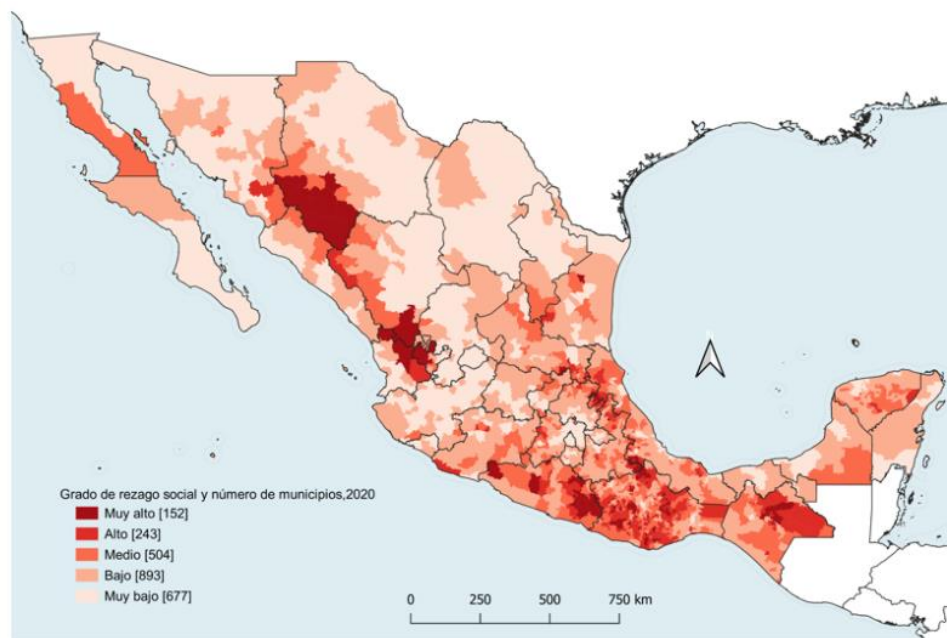


# Índice de rezago social

**coneval**

Lo que se mide se puede mejorar

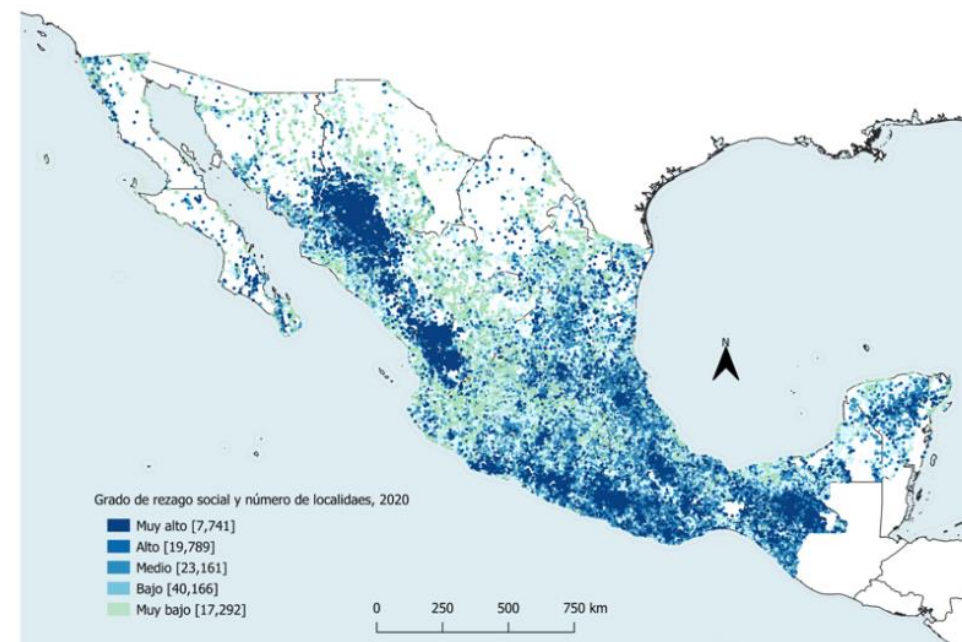
Grado de rezago social a nivel municipal  
2020



**coneval**

Lo que se mide se puede mejorar

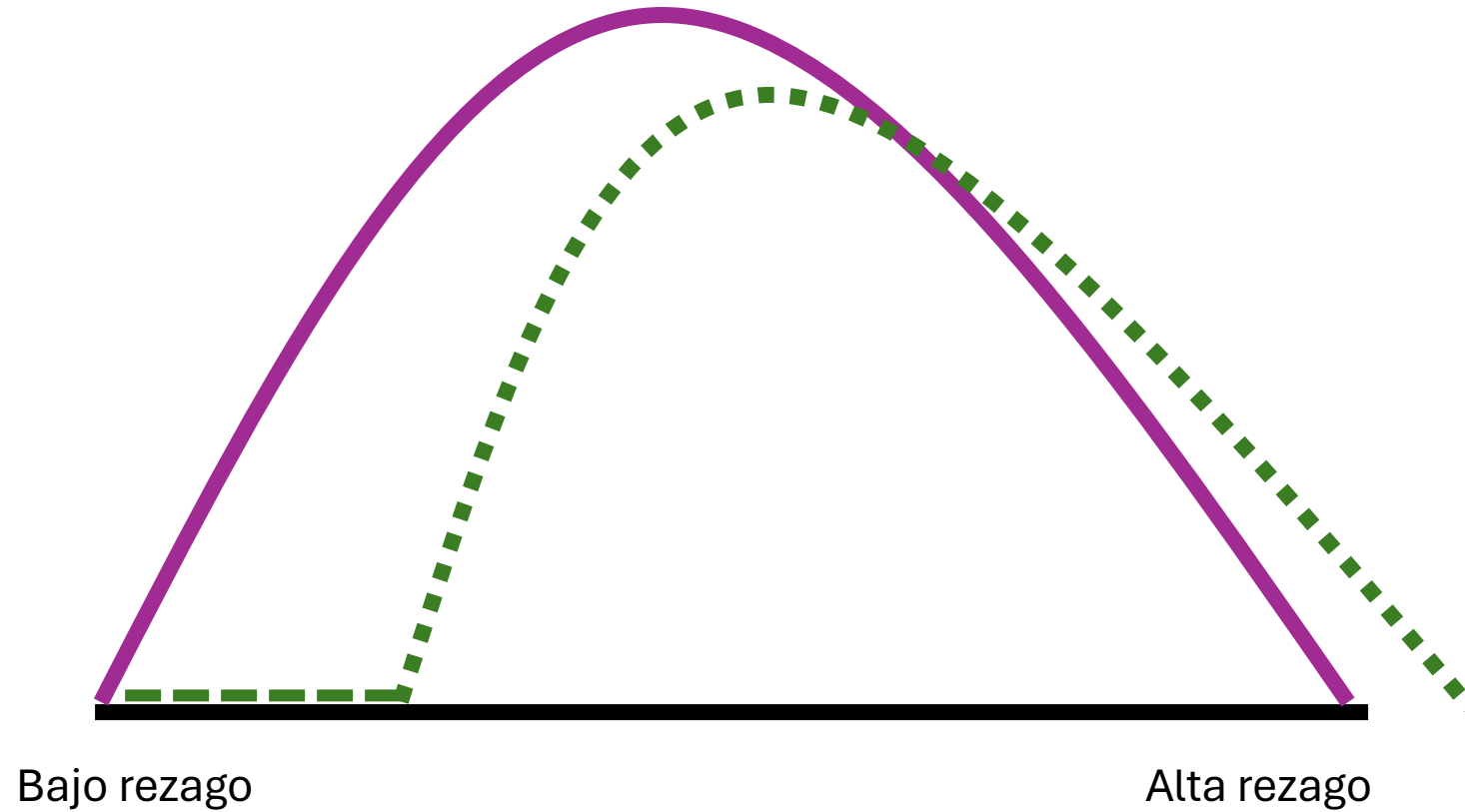
Grado de rezago social a nivel localidad  
2020



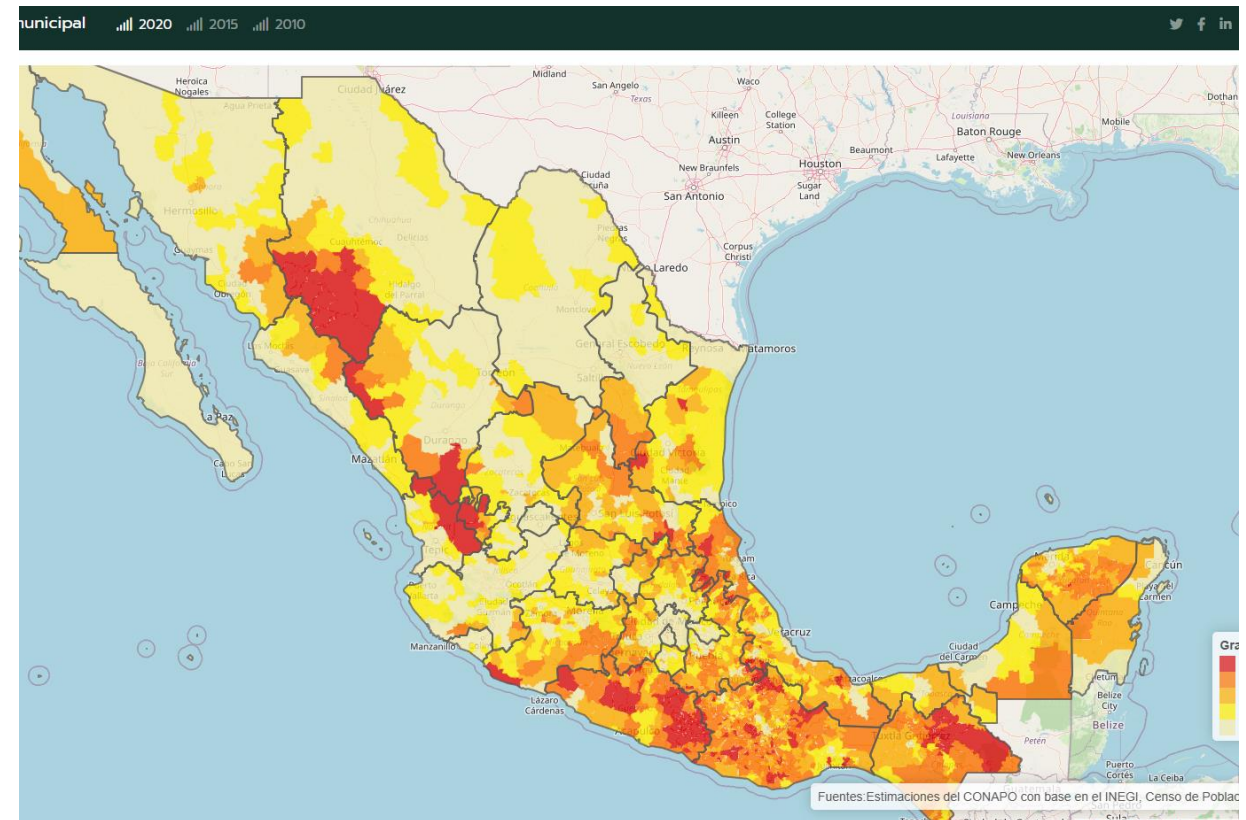
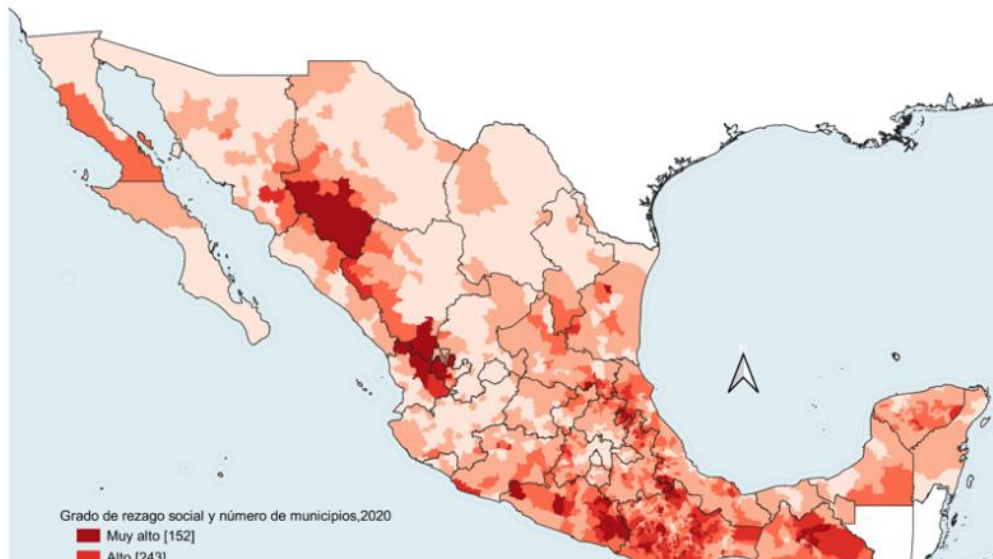
# Problemas: índice de rezago social

Verdadera

Observada

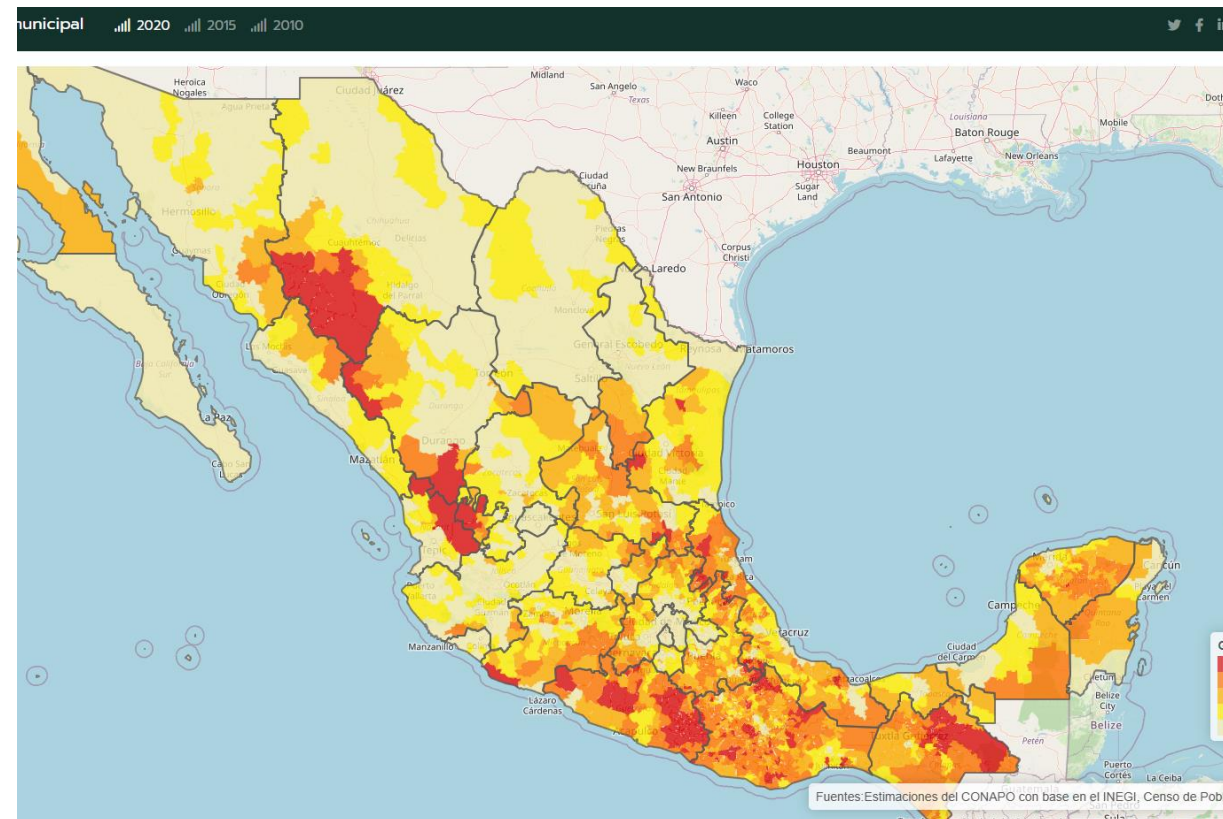
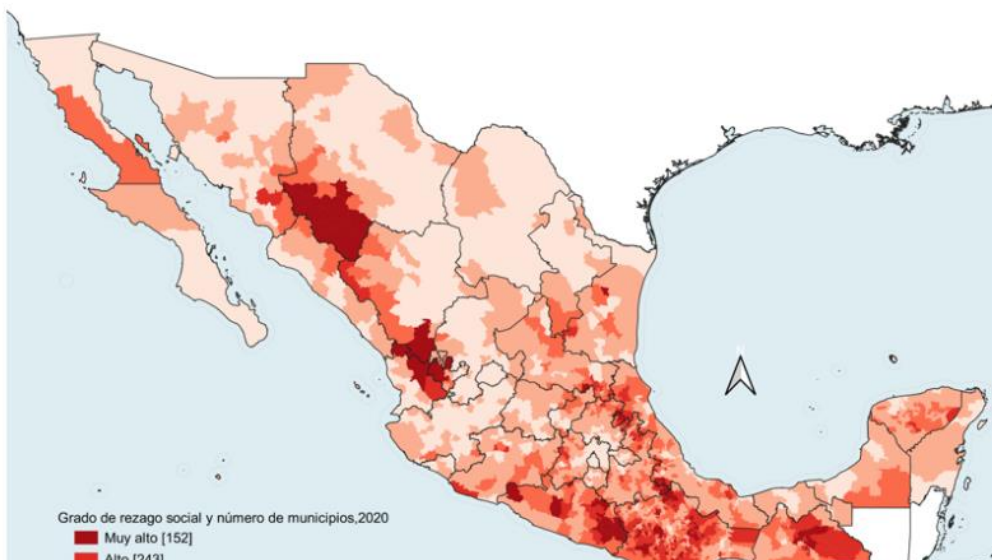


## Grado de rezago social a nivel municipal 2020





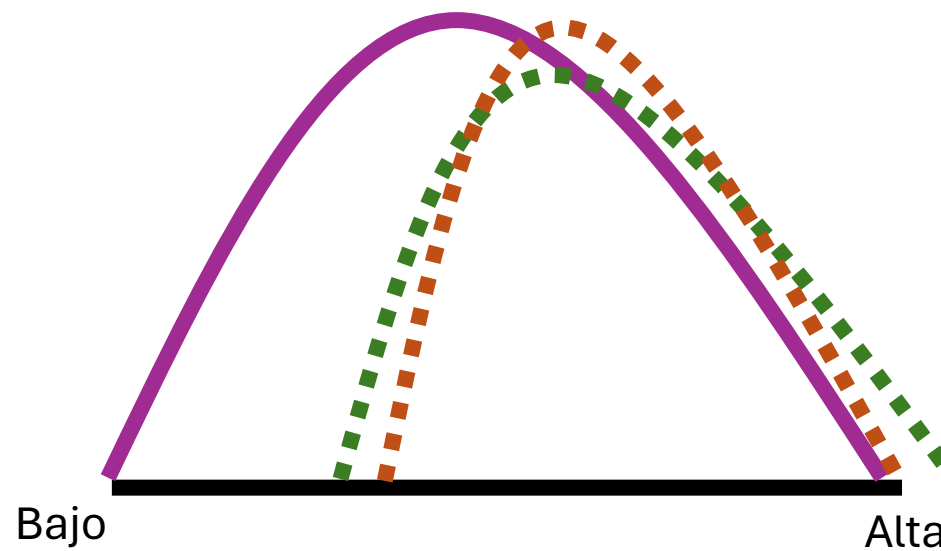
## Grado de rezago social a nivel municipal 2020



Comparten varianza

¿Es la variabilidad que nos interesa?

¿Qué significa que fenómenos que se etiquetan diferente, compartan varianza?





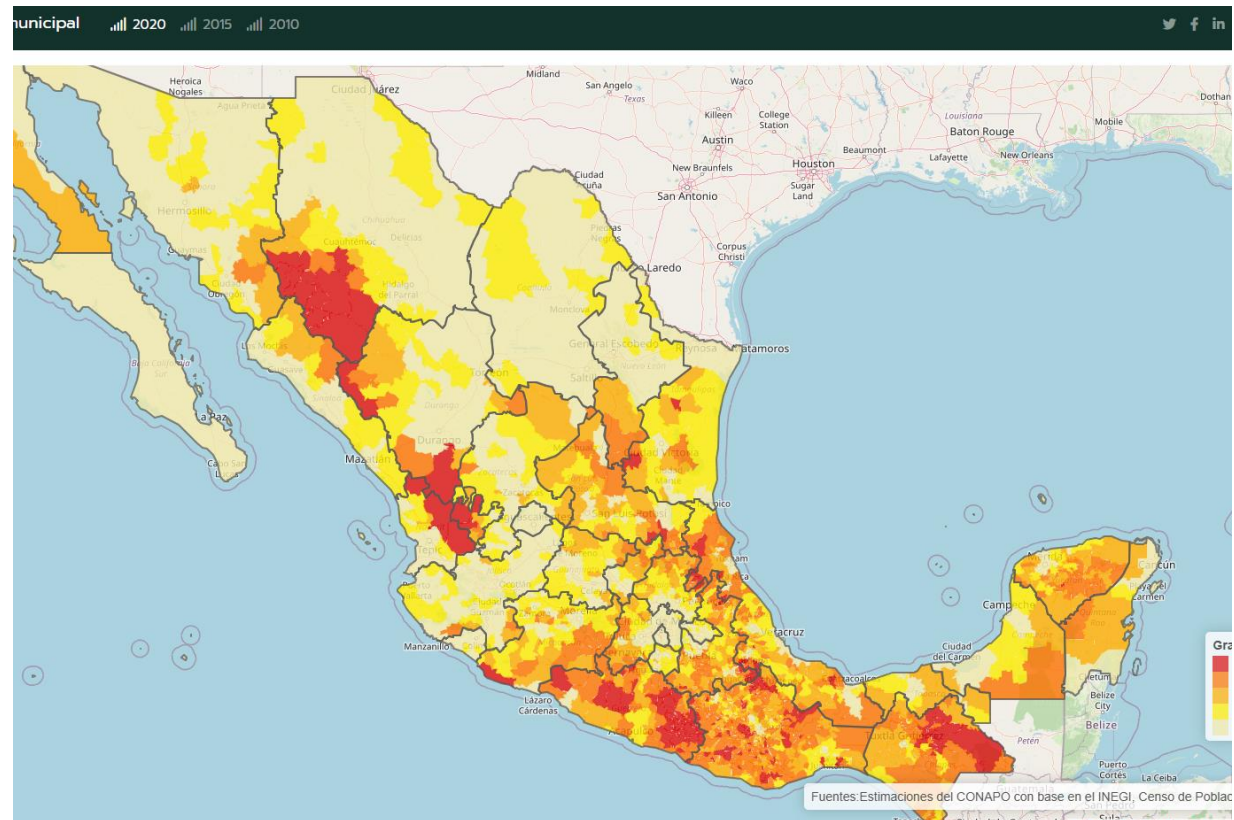
# ¿2 puntos extra en el primer ejercicio?

- Gráfica, en R, de la correlación entre ambos índices. En equipos:
  - 2010
  - 2015
  - 2020

24 horas

Ninguno de estos índices nos permite hacer inferencia sobre la distribución de la pobreza en México

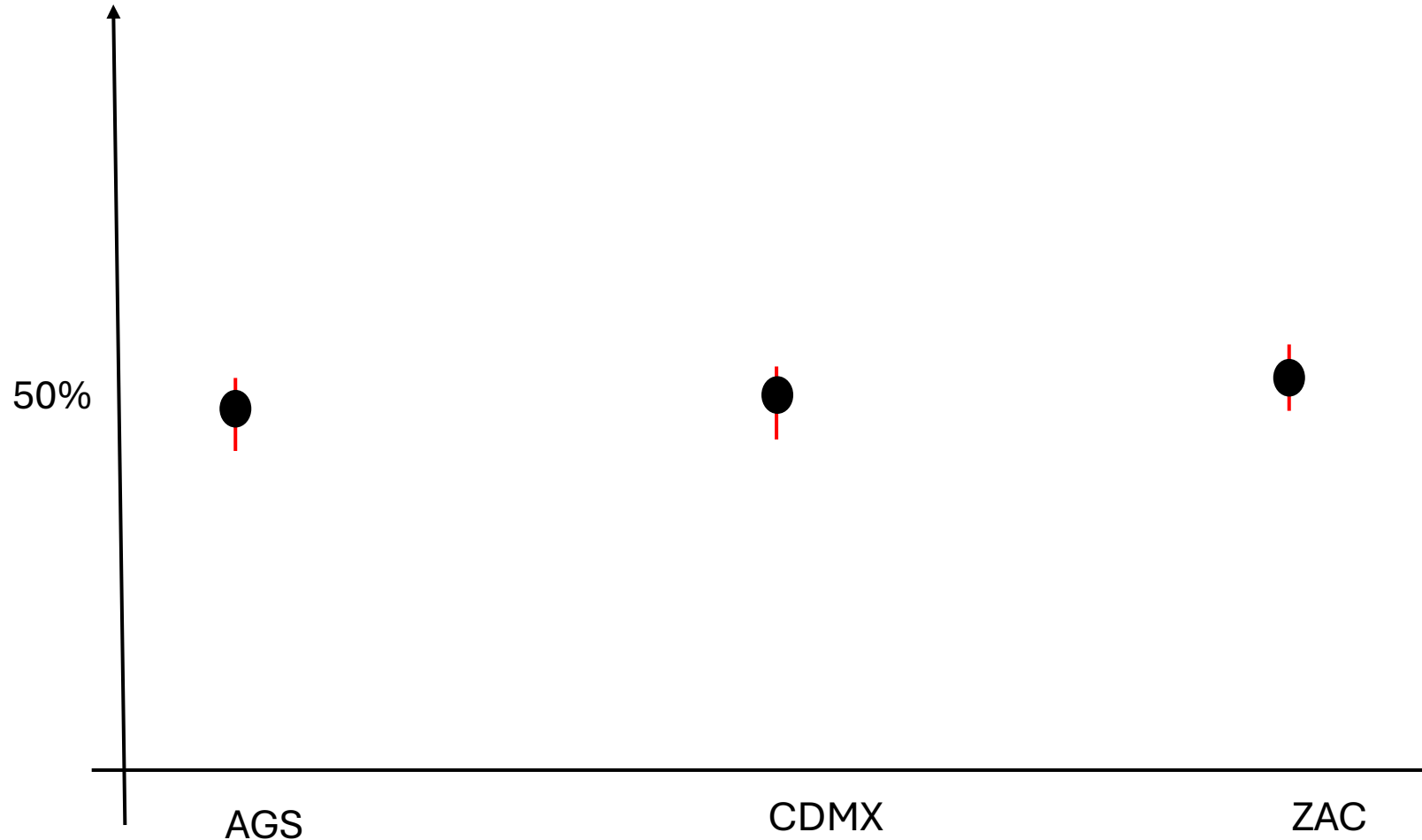
Ambos pueden llegar a ser útiles para hacer caracterizaciones gruesas y generales de la variabilidad territorial del nivel de vida



# Problema

- Las encuestas (ENIGH) con las que medimos pobreza son representativas a nivel nacional, urbano/rural y estatal
- El nivel estatal es insuficiente para hacer buena inferencia sobre la geografía de la pobreza en México
  - Alta variabilidad al interior de cada estado
  - Delimitaciones políticas que son incompatibles con el proceso geográfico

# Problema: Estimar % Hombres y Mujeres

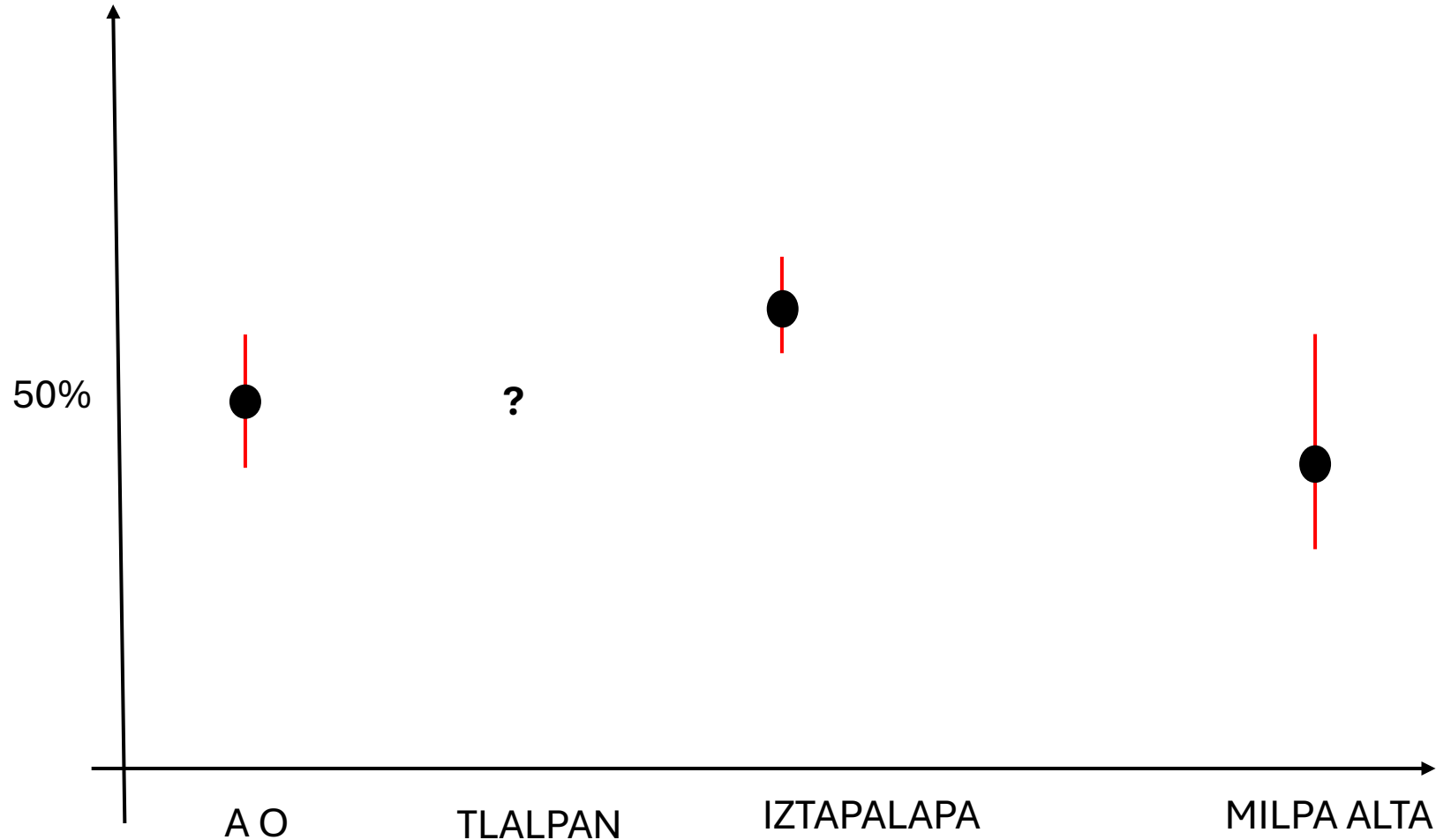


El estimador de la ENIGH es:

Confiable (Tiene poco error aleatorio)



# ¿Qué pasa a nivel municipal?

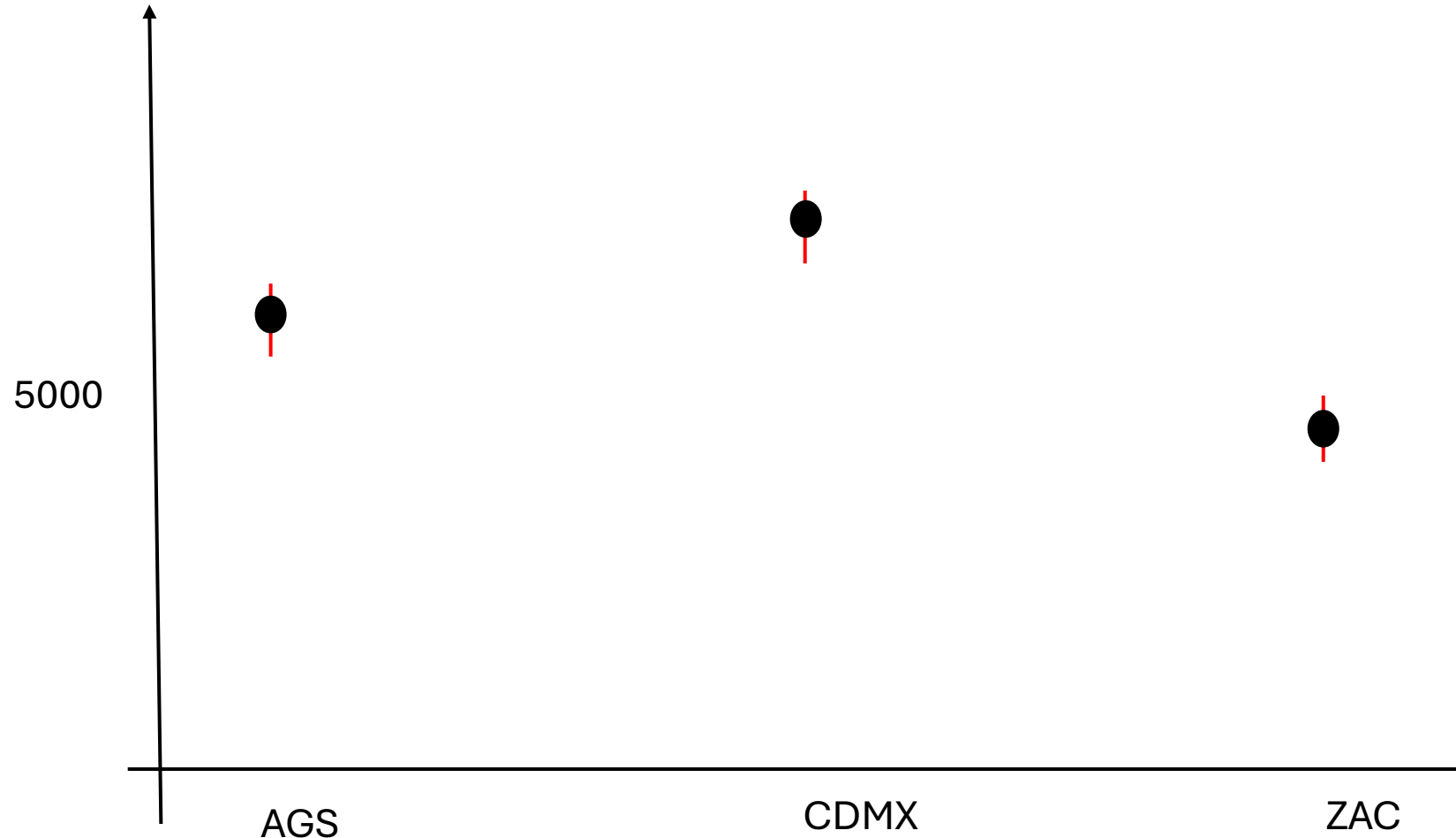


La muestra de la ENIGH tendrá sesgos

Aunque no tenga sesgo, probablemente tendrá poca precisión, i.e. intervalos de confianza muy grandes

Hay municipios sin muestra

# Problema: Estimar ICTPC

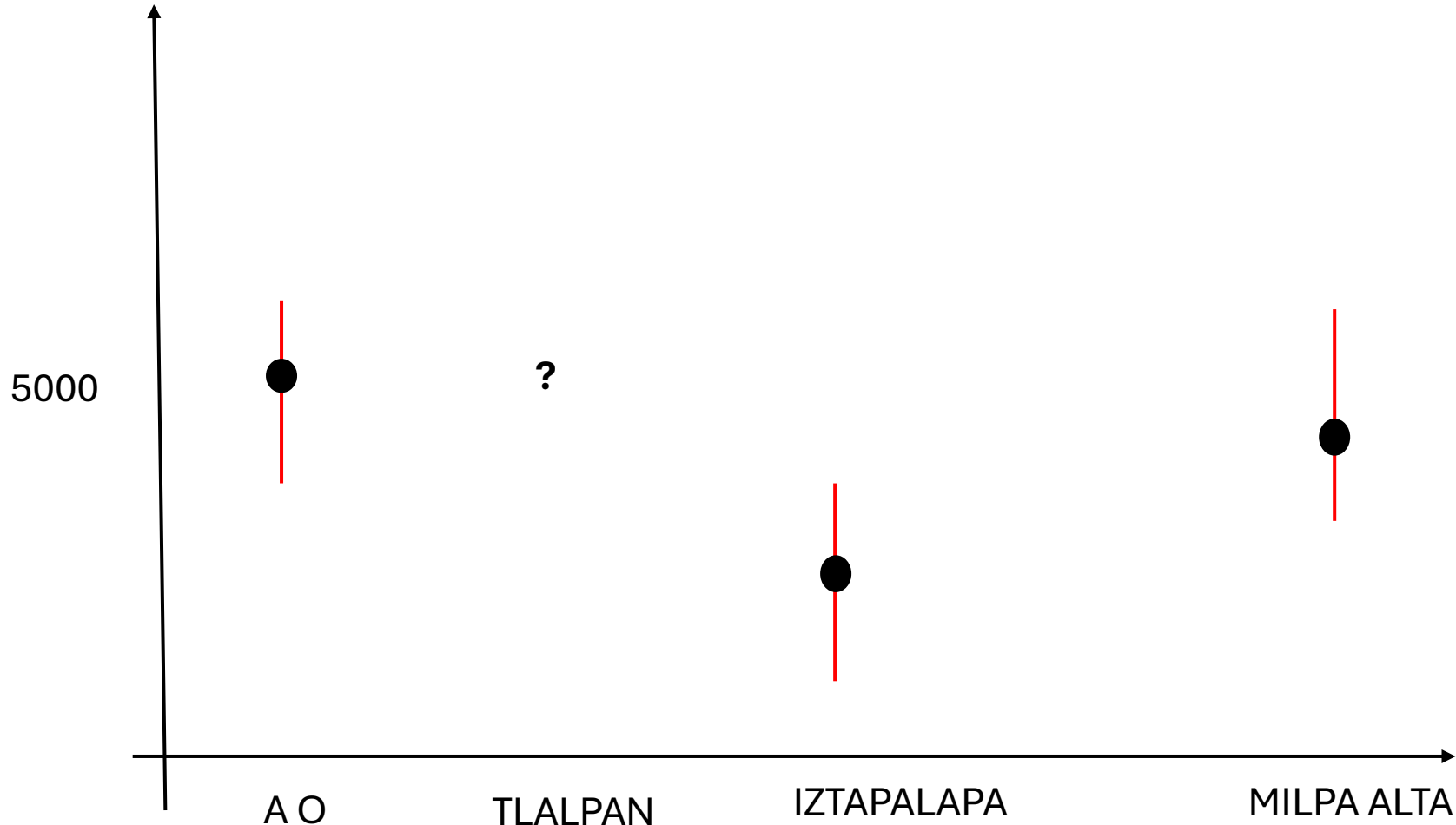


Para indicadores con mayor variabilidad, el problema se agudiza porque se requiera más poder estadístico

El estimador de la ENIGH es:

Confiable (Tiene poco error aleatorio)

# ¿Qué pasa a nivel municipal?

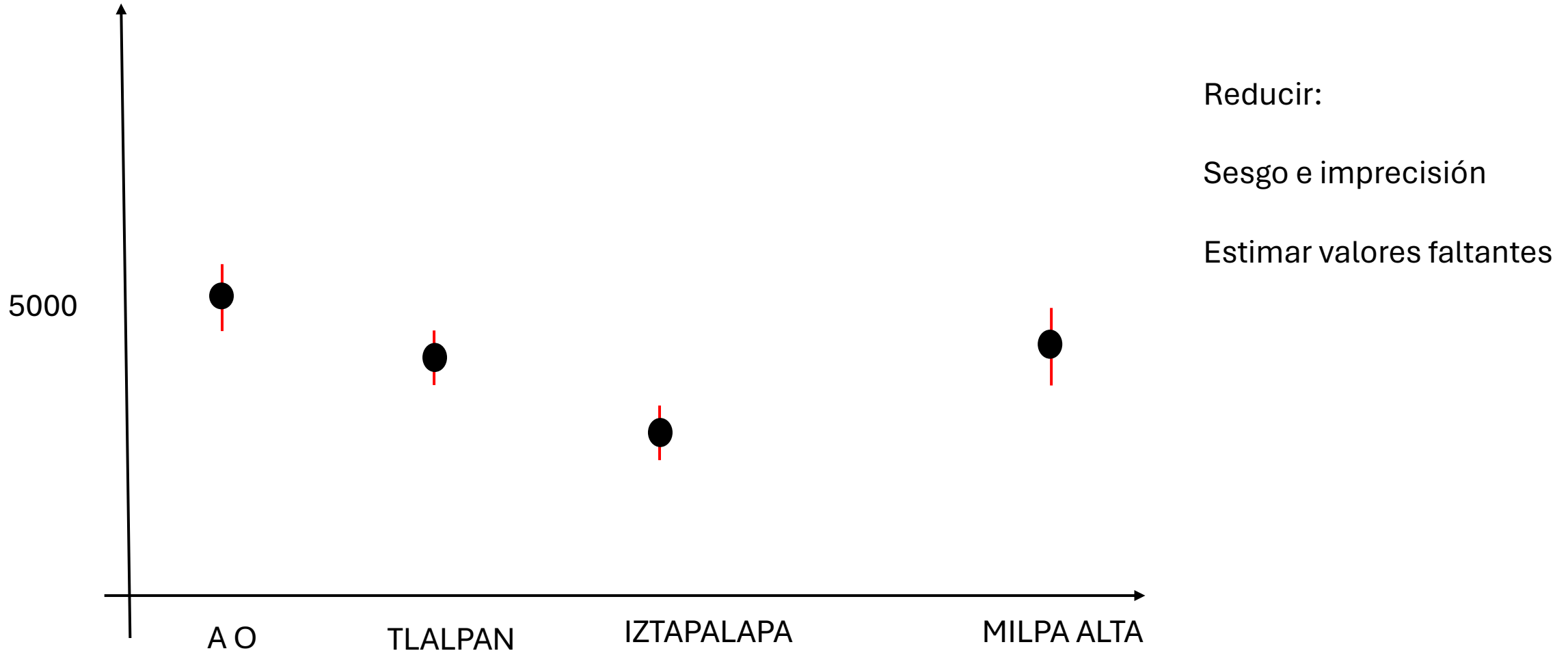


La muestra tendrá sesgos

Aunque no tenga sesgo, probablemente tendrá poca precisión, i.e. intervalos de confianza muy grandes

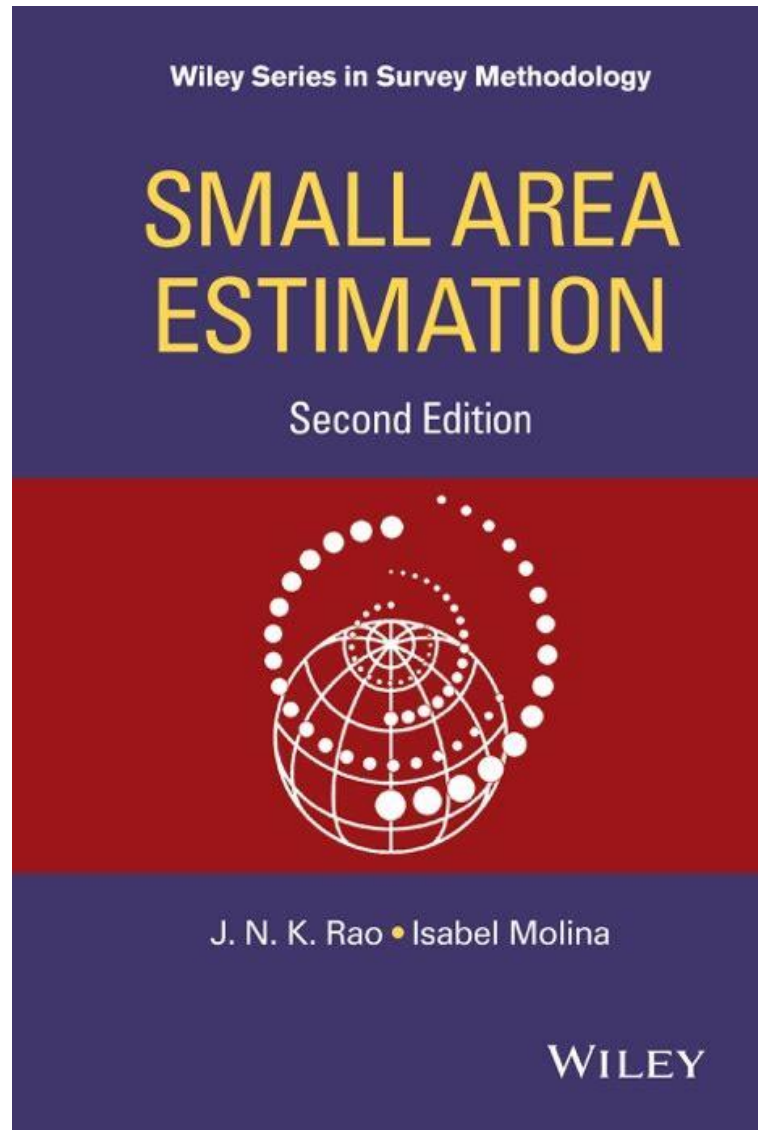
Para fenómenos más complejos, se requiere mayor muestra (poder estadístico). Sin ello, la inferencia es imposible.

# ¿Cómo lograr alguna corrección?

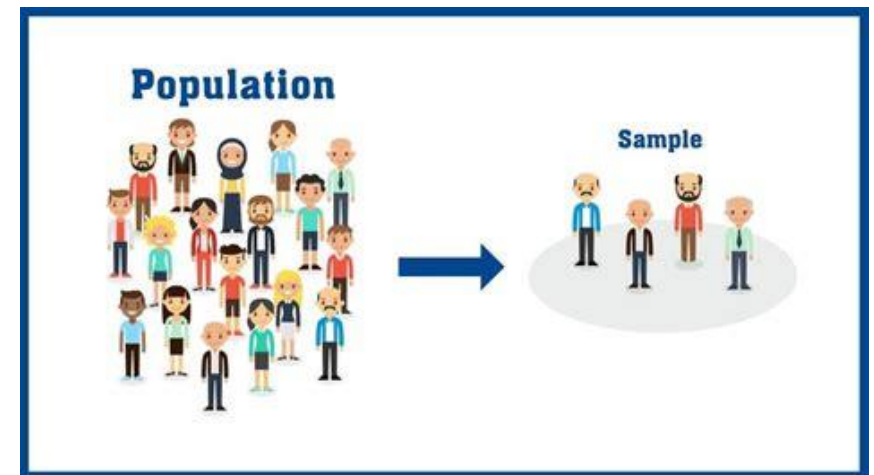
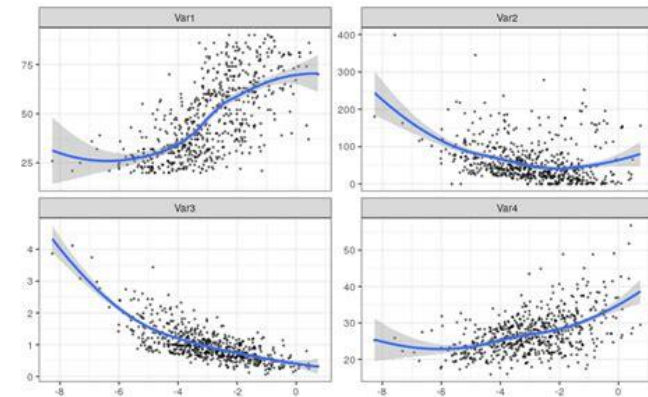
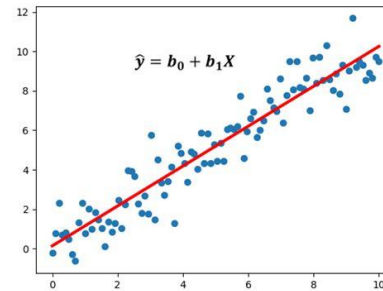




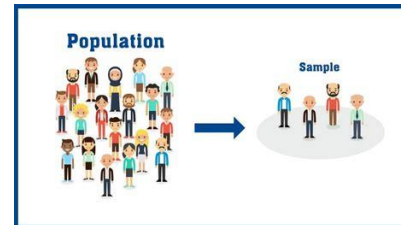
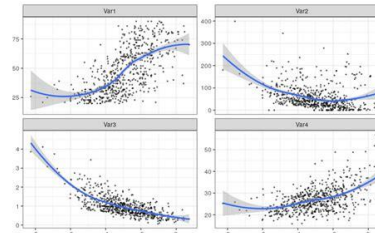
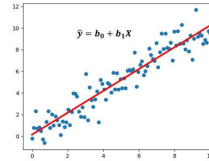
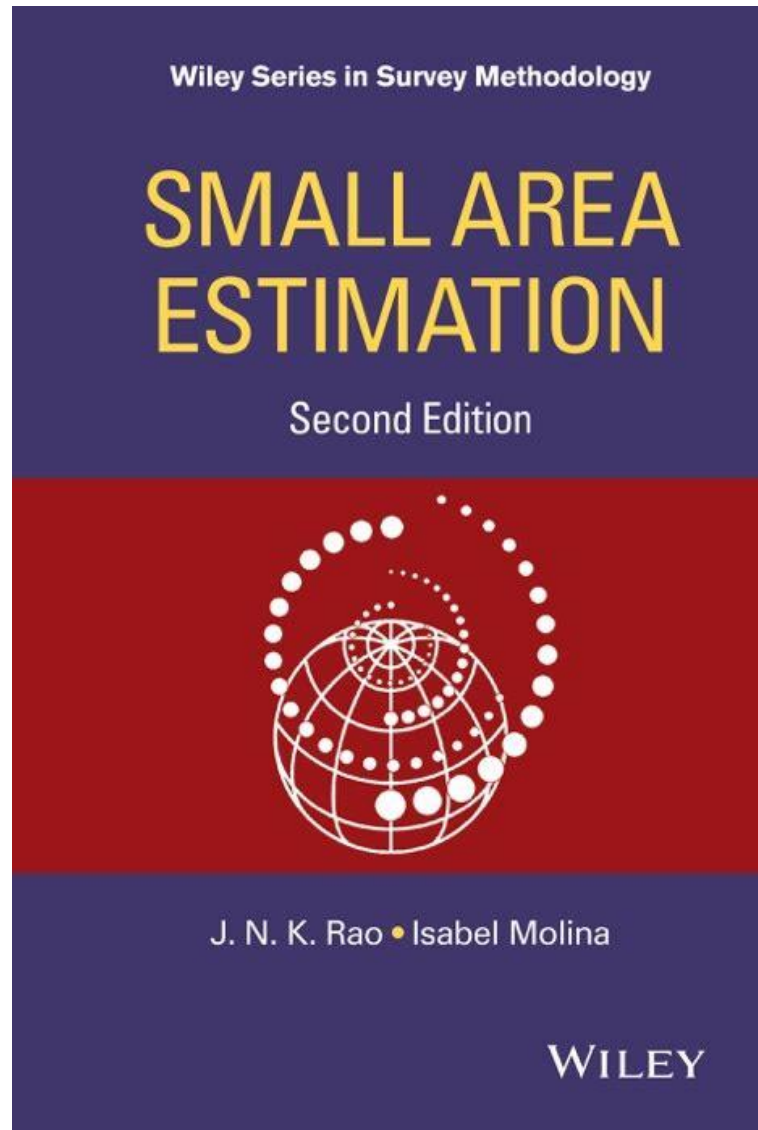
# Estadística teórica y aplicada



Objetivo: Producir estimaciones a nivel de área con bajo **sesgo y alta precisión**.



# Estadística teórica y aplicada



Modelo estadístico

- Auxilio de modelos de aprendizaje de máquina

Uso de información adicional:

- Registros administrativos
- Información satelital

# Estadística teórica y aplicada

## Combining satellite imagery and machine learning to predict poverty

NEAL JEAN, MARSHALL BURKE, MICHAEL XIE, W. MATTHEW DAVIS, DAVID B. LOBELL, AND STEFANO ERMON [Authors Info & Affiliations](#)

SCIENCE • 19 Aug 2016 • Vol 353, Issue 6301 • pp. 790-794 • DOI: 10.1126/science.aaf7894

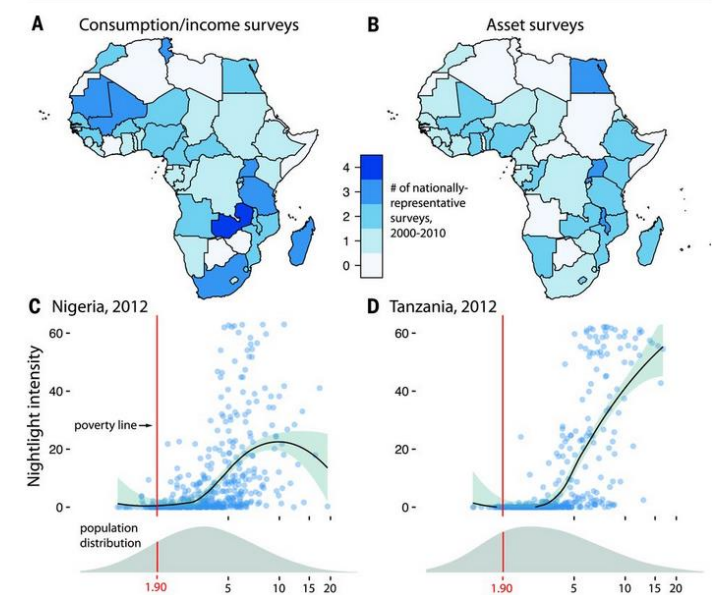
↓ 12,696    715



### Measuring consumption and wealth remotely

Nighttime lighting is a rough proxy for economic wealth, and nighttime maps of the world show that many developing countries are sparsely illuminated. Jean *et al.* combined nighttime maps with high-resolution daytime satellite images (see the Perspective by Blumenstock). With a bit of machine-learning wizardry, the combined images can be converted into accurate estimates of household consumption and assets, both of which are hard to measure in poorer countries. Furthermore, the night- and day-time data are publicly available and nonproprietary.

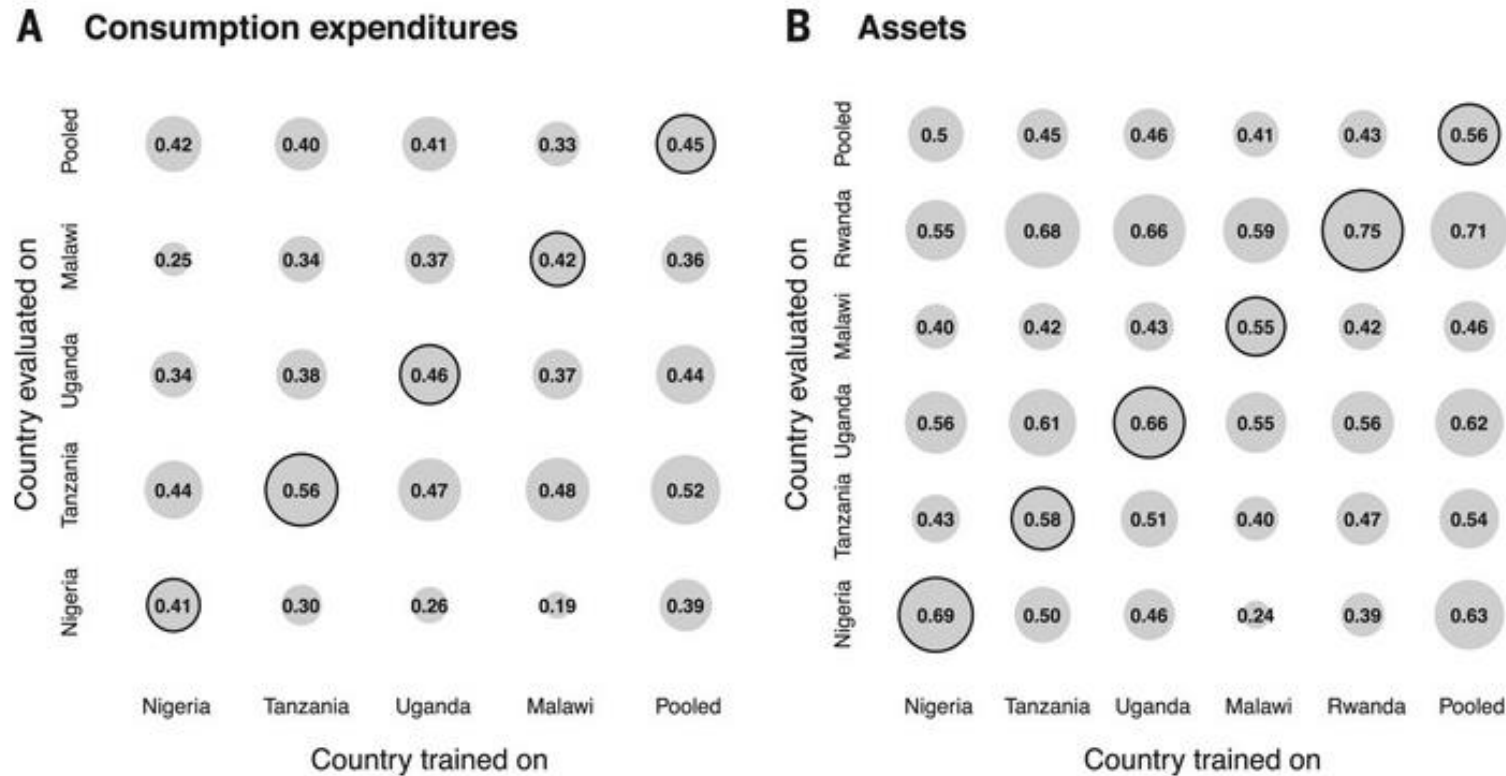
*Science*, this issue p. [790](#); see also p. [753](#)



Luz de noche:

¿Puede funcionar en México?

# Luz de noche: ¿Puede funcionar en México?



¿Qué piensan de esos R2?

**Fig. 5 Cross-border model generalization.**

(A) Cross-validated  $r^2$  values for consumption predictions for models trained in one country and applied in other countries. Countries on x axis indicate where model was trained, countries on y axis where model was evaluated. Reported  $r^2$  values are averaged over 100 folds (10 trials, 10 folds each). (B) Same as in (A), but for assets.



# 2010



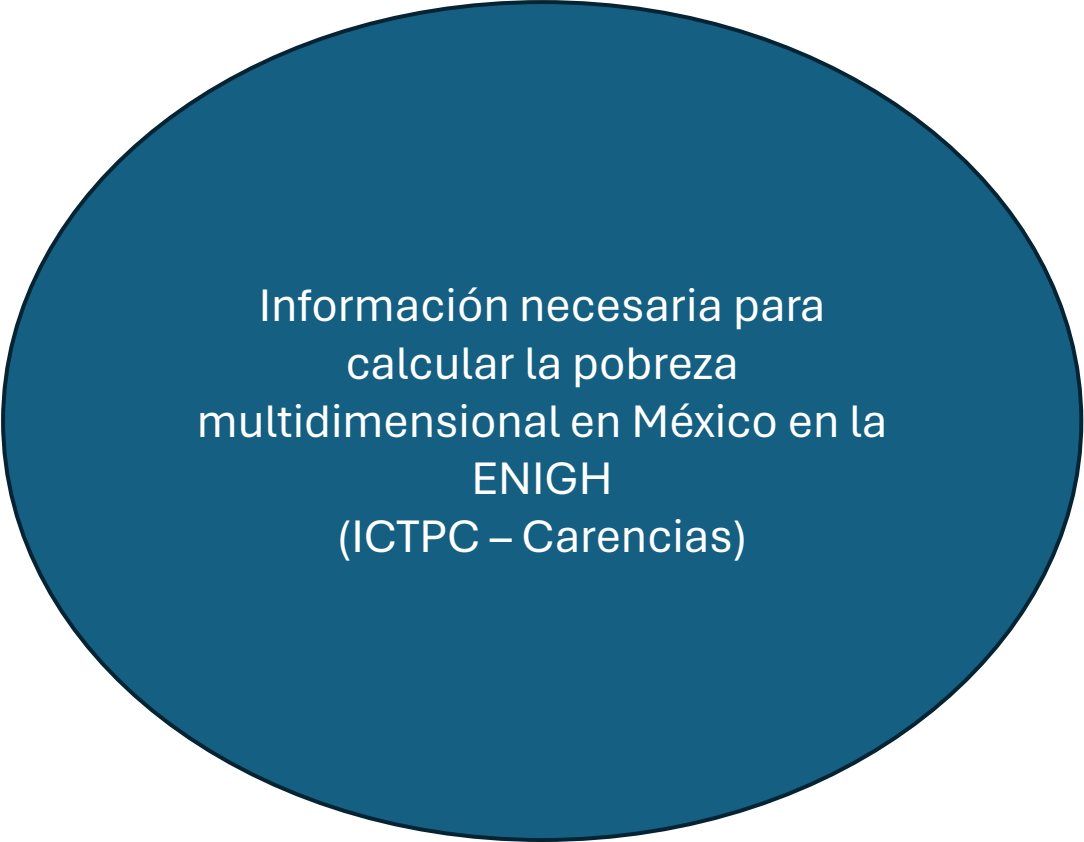
## Metodología para la medición de la pobreza en los municipios de México **2010**



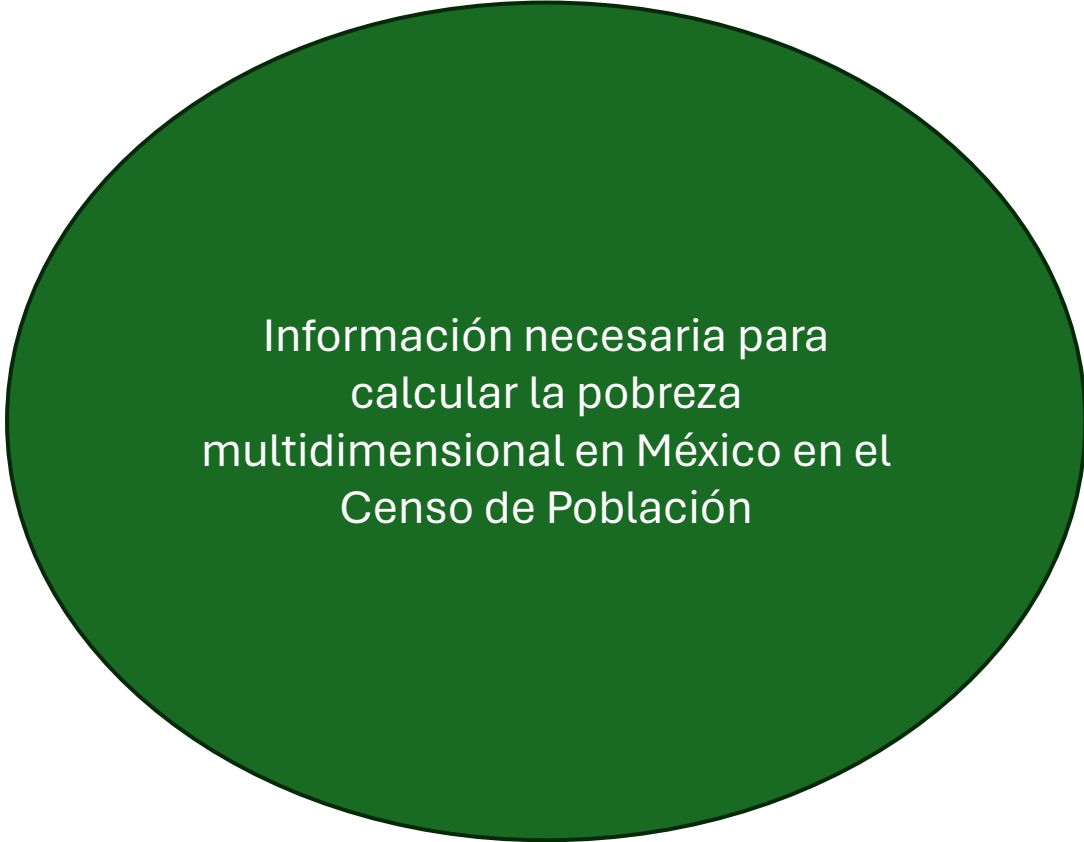
CONEVAL realiza por primera vez en América Latina un proyecto de Estimación de Pobreza a Nivel de Áreas pequeñas (EAP).

Este proyecto tuvo el reto de innovar en un contexto en el que había más dudas que certezas sobre cómo hacer EAP en un contexto como el mexicano

# El reto de hacer EAP en México: ¿Cómo se planteó el problema?



Información necesaria para  
calcular la pobreza  
multidimensional en México en la  
ENIGH  
(ICTPC – Carencias)



Información necesaria para  
calcular la pobreza  
multidimensional en México en el  
Censo de Población



# El reto de hacer EAP en México



El CPV tiene  
información sobre 4 de  
las 6 carencias

Le falta:

ICTPC  
Carencia Alimentaria  
Carencia de Seguridad  
Social

# ¿Cómo se abordó el problema de EAP?

i	CA	SS	CS	SBV	CV	CE	Ic	I	P
1	1	1	0	1	0	1	4	3000	1
2	0	1	0	0	0	0	1	2500	1
...	...	...	...	...	...	...	...	...	...
n	0	0	0	0	0	0	0	5000	0

Si estimo los vectores que faltan  
ya está ¿cierto?

$$\hat{P}_i | \hat{I}_i \cap \hat{C}_i$$

$$\hat{C}_i = \widehat{CA}_i + \widehat{CASS}_i + SA_i + ED_i + CV_i + SB_i$$



# Todos los modelos están mal, algunos son útiles

- ¿Cómo sumar los errores individuales para producir estimaciones a nivel de área?

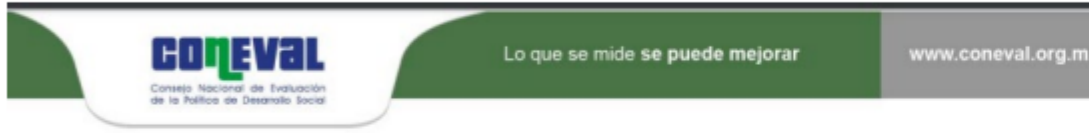
i	CA	SS	CS	SBV	CV	CE	Ic	I	P
1	1	1	0	1	0	1	4	3000	1
2	0	1	0	0	0	0	1	2500	1
...	...	...	...	...	...	...	...	...	...
n	0	0	0	0	0	0	0	5000	0

$$\hat{P}_i | \hat{I}_i \cap \hat{C}_i$$

$$\hat{C}_i = \widehat{CA}_i + \widehat{CASS}_i + SA_i + ED_i + CV_i + SB_i$$

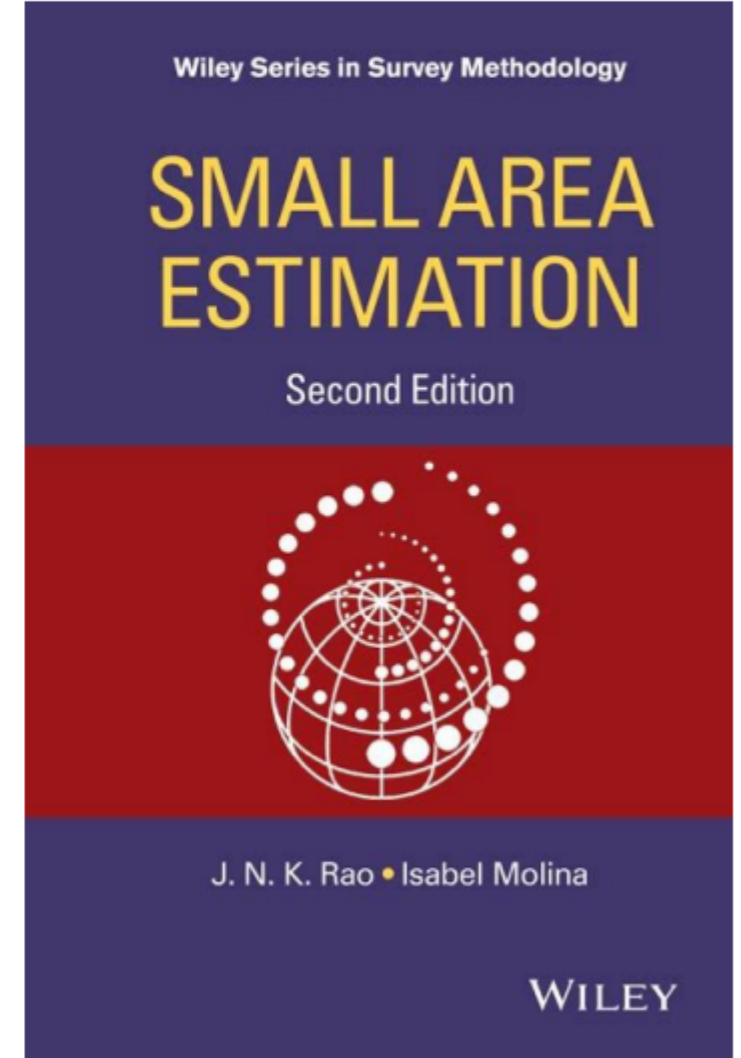
$\widehat{P}_{ijk}$  utilizando  $\widehat{P}_i \mid \widehat{I}_i \cap \widehat{C}_i$

$$\widehat{P}_j^{EB} = \frac{1}{N} (n_j \bar{y}_j + \sum p_{ij})$$



## Metodología para la medición de la pobreza en los municipios de México, 2020

Vs



# Conclusiones

- La estrategia metodológica del CONEVAL produce estimaciones por carencia y de la pobreza multidimensional con un error, en promedio, **casi 70% veces mayor** a la estrategia convencional del campo de EAP.
- La magnitud del error estimado es optimista puesto que hay que sumarlos errores de medición, los errores de clasificación en P, errores de medición de las carencias en el Censo y los errores de un marco muestral más complejo
- Al no reportar el error total de la estimación, es imposible hacer inferencias sobre cambios en pobreza en el tiempo entre distintas unidades
- Esto no se puede hacer ni siquiera cuando hay cambios que **oscilen por los 5 pp**
- Se recomienda optar por la estrategia convencional de EAP para reportar cifras de pobreza multidimensional

# Entonces...

- El CONEVAL y el PUED-UNAM están trabajando para mejorar las estimaciones oficiales y crear una nueva serie 2015-2025
- Por ahora, las cifras del CONEVAL hay que usarlas con cautela
- Hay otras alternativas:

<https://pued-unam.shinyapps.io/EAP-Pobreza-Desigualdad/>