



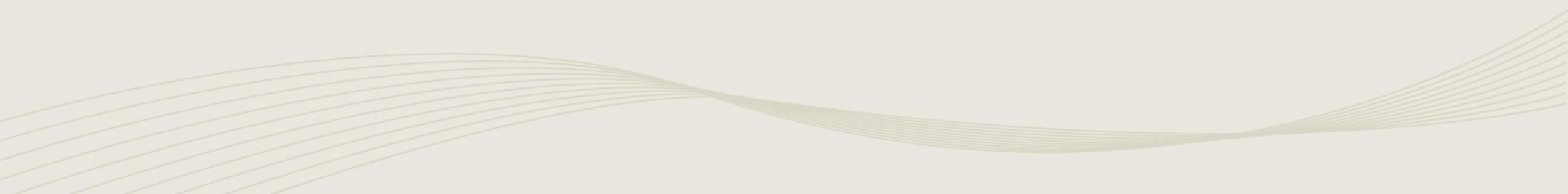
Programa
Universitario
de Estudios
del **Desarrollo**
UNAM

Comparabilidad en medición de pobreza

Hector Najera



Contenidos

- La comparación como razón para medir
 - Dos perspectivas sobre comparabilidad
 - Retos en la comparabilidad en el tiempo y comparación entre grupos
 - Soluciones y métodos para la comparabilidad
 - Límites para el análisis de la comparabilidad
- 



Comparabilidad

Análisis de grupos: Poblaciones, unidades territoriales y entre sociedades

¿Es la pobreza más alta en la *unidad* A que en la B?

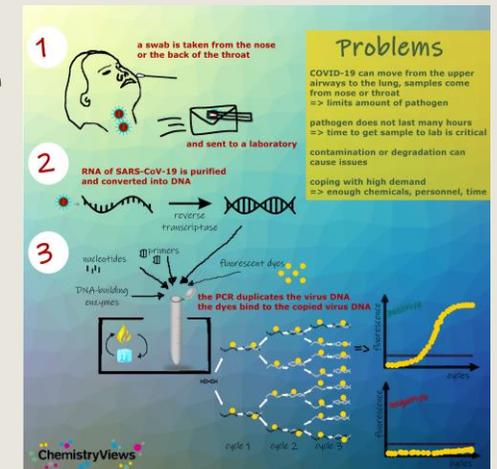
Análisis en el tiempo: Tendencias y cambios

¿Creció, disminuyó o permaneció igual la pobreza?

Combinaciones: Cambios en el tiempo por grupo.

Comparabilidad vs desagregación

- Comparabilidad: Ejercicio que consiste en **contrastar scores equivalentes** entre distintas unidades
- Desagregación: Acción u operación de separación de scores **-equivalentes o no equivalentes-** dada otra variable.
 - Poder desagregar un índice no lo hace comparable (Marco axiomático?)



¿Qué permite tener scores equivalentes?



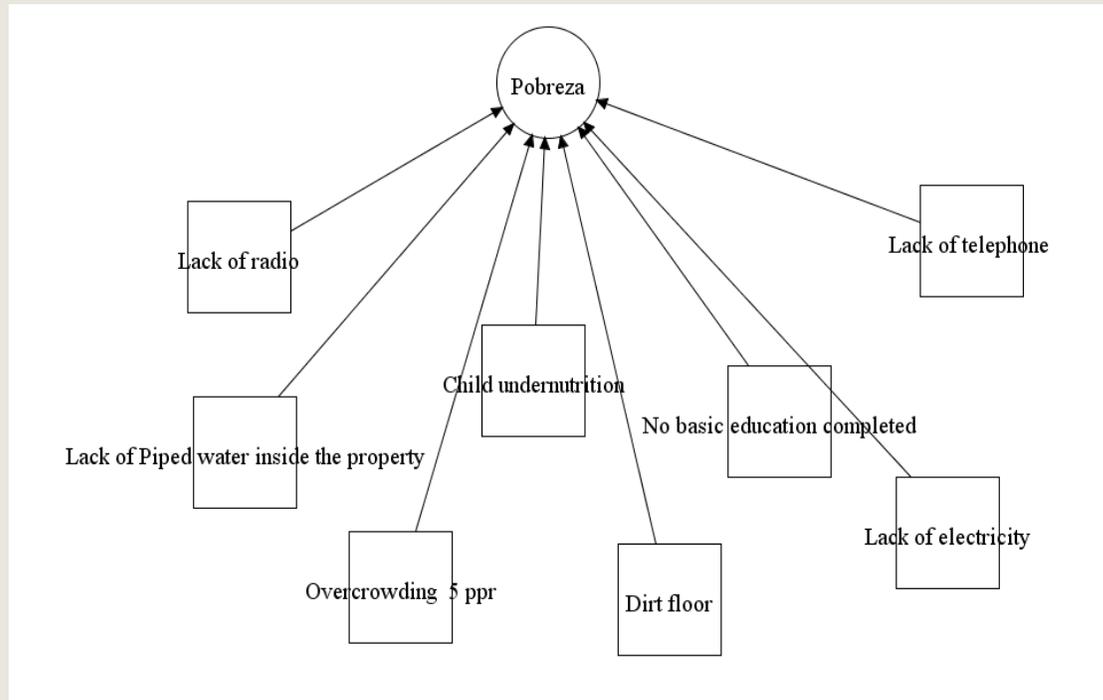


¿Scores equivalentes?

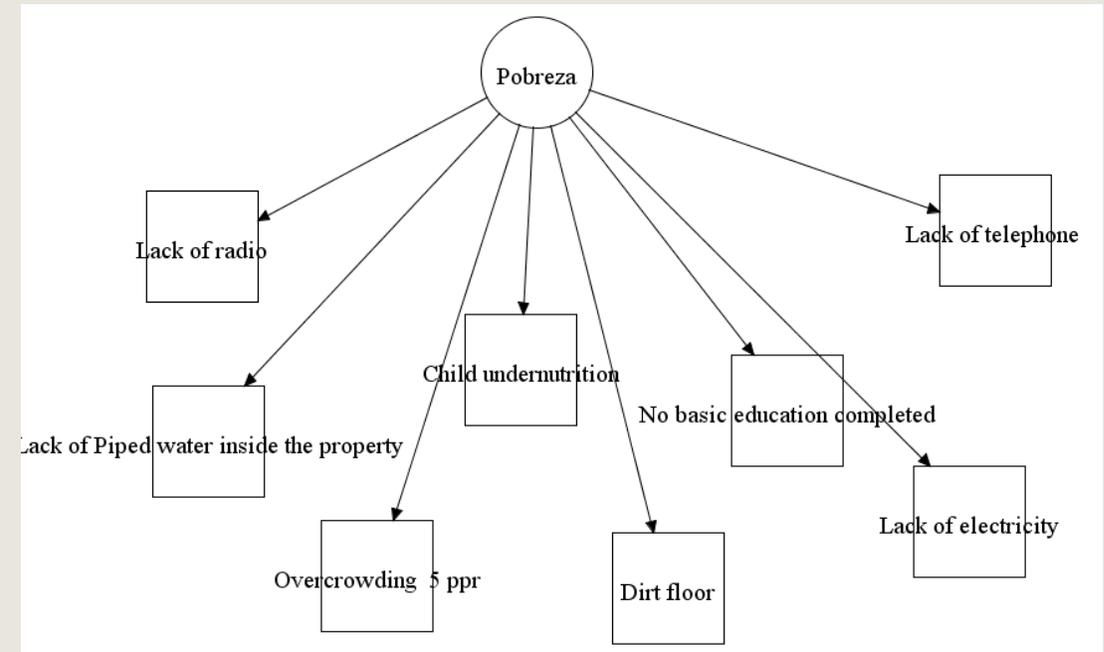
- Son aquellos scores que **independientemente** del grupo de pertenencia reflejan diferencias en pobreza
 - Los scores de un examen de admisión deben reflejar el aprendizaje no si la persona viene de cierto estrato social
 - Una cosa es que el estrato social pueda ser una causa de ciertas ventajas educativas y otra que el examen esté sesgado y favorezca los temarios de ciertas escuelas privilegiadas.
 - Diferencias efectivas en pobreza, no diferencias por lugar de residencia: El techo de palma de las casas Mayas en México podría penalizar áreas rurales a costa de criterios urbanos

Dos visiones: Scores equivalentes

Medición formativa



Medición reflexiva

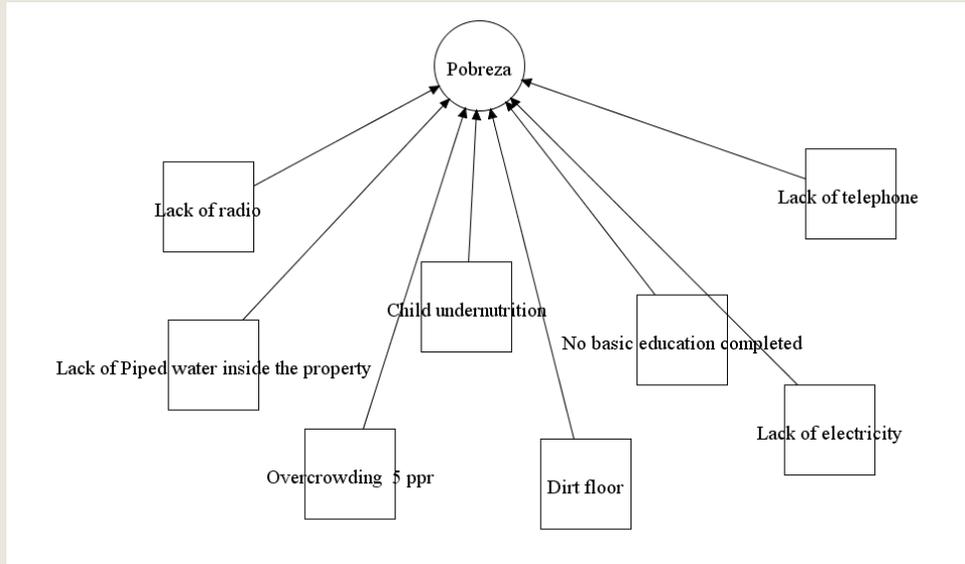


Los indicadores hacen al concepto, la pobreza es lo que está ahí en esos indicadores. ¿Qué pasa si tiro uno? ¿La definición cambia?

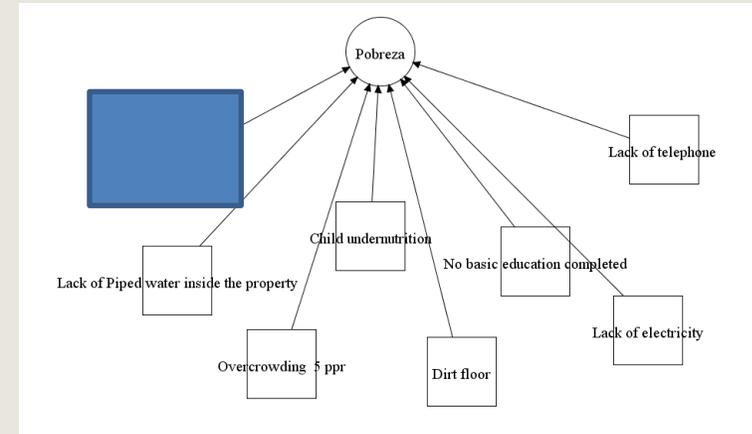
El fenómeno latente es la causa de las manifestaciones de pobreza

Comparabilidad medición formativa

No se asocia con pobreza en el grupo/tiempo 2



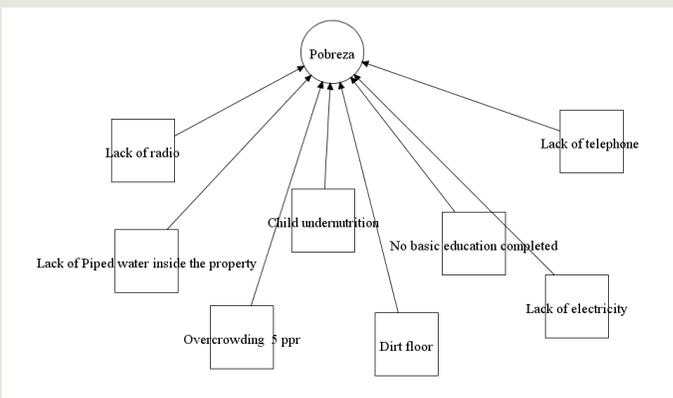
Conjunto I_f
Scores S_{I_f} {0-8}



Conjunto I_{f2}
Scores $S_{I_{f2}}$ {0-7}

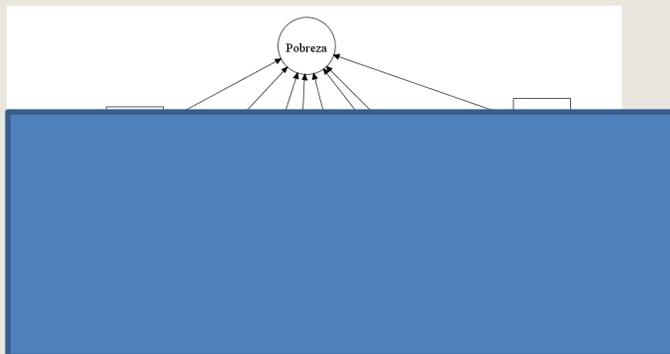
Bajo este marco formativo, la comparabilidad se rompe. El Score ya no es equivalente. Si agregamos un indicador, la definición cambia.

Comparabilidad medición formativa



Conjunto I_f
Scores S_{I_f} {0-8}

Sabemos que $S_{I_f} \geq S_{I_f2}$



Los indicadores muy probablemente no sirven para identificar quién es pobre

Conjunto I_{f2}
Scores $S_{I_{f2}}$ {0}

Esto funcionaría si tuviéramos un censo de todas las dimensiones y todos los indicadores de pobreza pero no lo tenemos y nunca lo vamos a tener!

Esto significa que esas necesidades mejoraron.

¿La pobreza es cero?

Típico caso donde el político piensa que la pobreza se resuelve pensando en que hay que resolver los indicadores!

Eventualmente necesito nuevos indicadores.

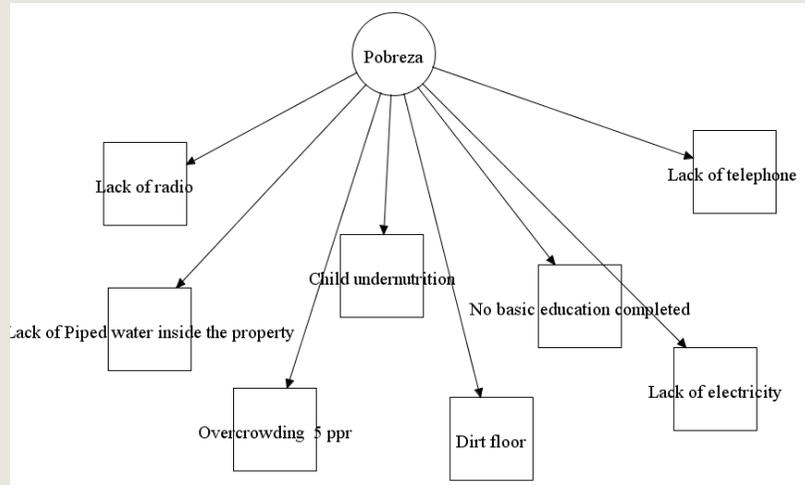
El enfoque formativo es una trampa sin salida

La medición formativa es un caso especial -no muy notable de la reflexiva-.

Buscar comparabilidad a costa de los mismos indicadores es un sinsentido.

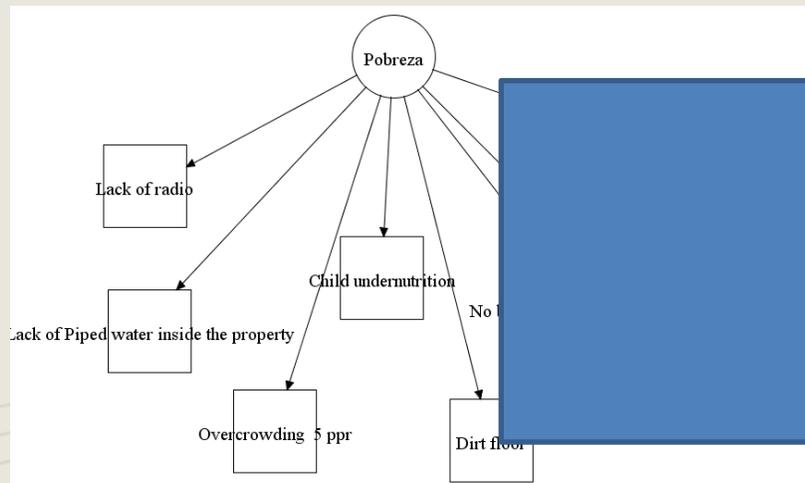


Comparabilidad medición reflexiva. Caso I: Pérdida de indicadores



Conjunto I_r

Scores: S_r



Conjunto I_{r2}

Scores: S_{r2}

Recuerden que lo que tenemos son muestras de D, I, U, P y L

Lo que queremos es buenas muestras **y que sean comparables**

El enfoque reflexivo sí tiene un marco de inferencia para tratar con el problema de la comparabilidad

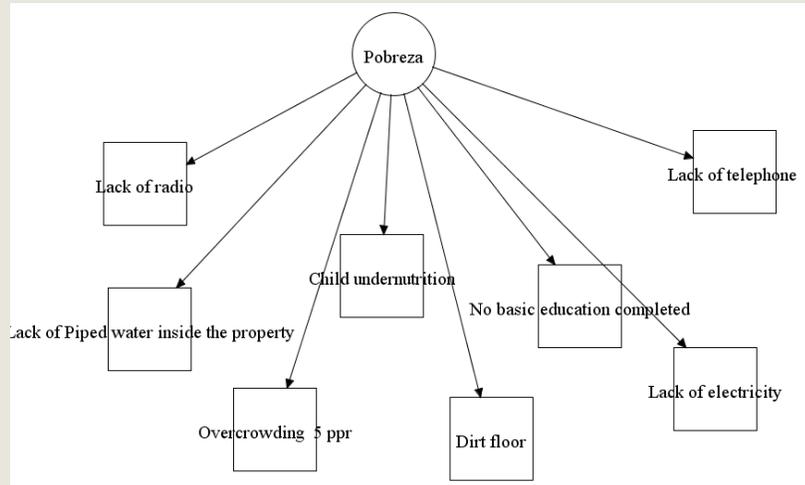
¿Es $S_r = S_{r2}$? -Son equivalentes-

¿Es el modelo equivalente y sus scores también?

¿Si no es equivalente? ¿Qué podría hacer?
-Re-muestreo-

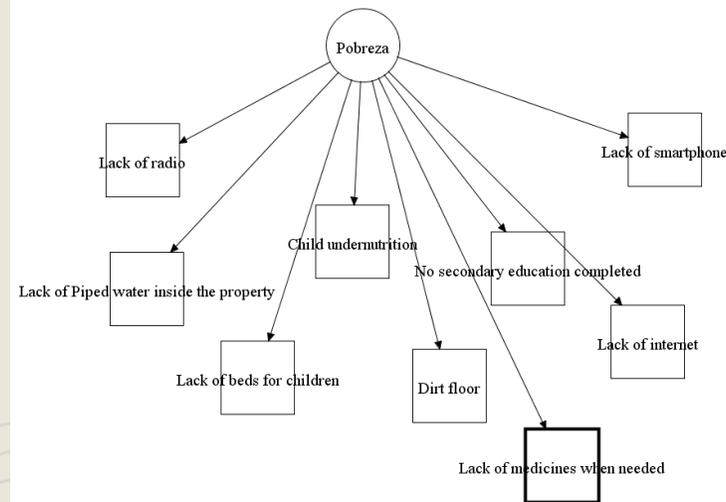


Comparabilidad medición reflexiva. Caso 2: Nuevos y más indicadores



Conjunto I_r

Scores: S_r
{0,8}



Conjunto I_{r2}

Scores: S_{r2}
{0,9}

Recuerden que lo que tenemos son muestras de D, I, U, P y L

El enfoque reflexivo sí tiene un marco de inferencia para tratar con el problema de la comparabilidad

¿Es $S_r = S_{r2}$? -Son equivalentes-

¿Qué significa $S_r = 3$ en términos de $S_{r2} = 4$?

¿Es el modelo equivalente y sus scores también?

¿Si no es equivalente? ¿Qué podría hacer?



Comparabilidad como medición invariante -measurement invariance-

¿Qué significa equivalencia?

- Measurement invariance/equivalence
 - Es condicional en que confiabilidad y validez se sostengan (De hecho esa fue la razón para cambiar los indicadores)
 - Es una condición necesaria para poder realizar comparaciones
 - Ocurre cuando se cumple de

[Published: December 1993](#)

Measurement invariance, factor analysis and factorial invariance

[William Meredith](#)

[Psychometrika](#) 58, 525–543 (1993) | [Cite this article](#)

5764 Accesses | 2606 Citations | 6 Altmetric | [Metrics](#)

PSYCHOMETRIKA—VOL. 57, NO. 2, 289–311
JUNE 1992

ON THE MISUSE OF MANIFEST VARIABLES IN THE DETECTION OF MEASUREMENT BIAS

WILLIAM MEREDITH

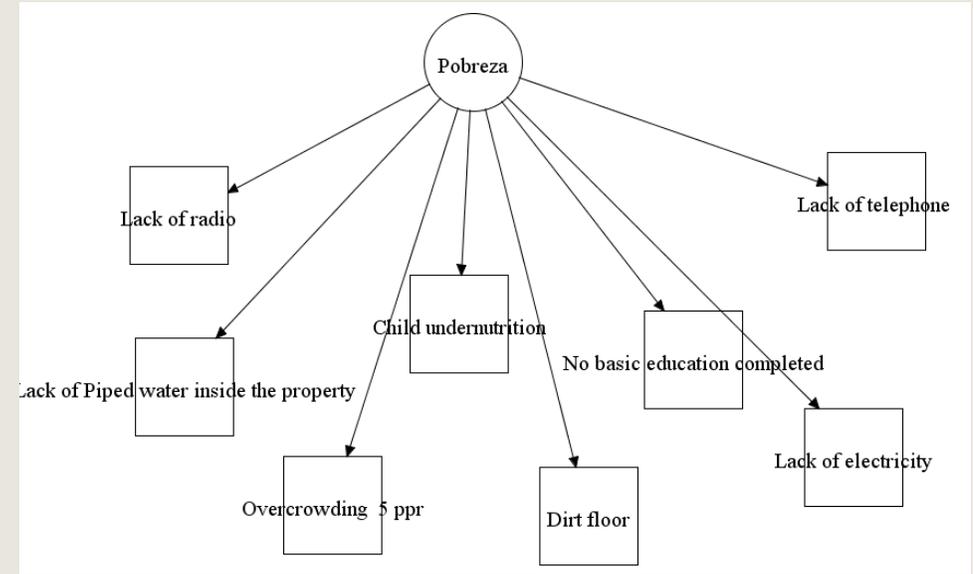
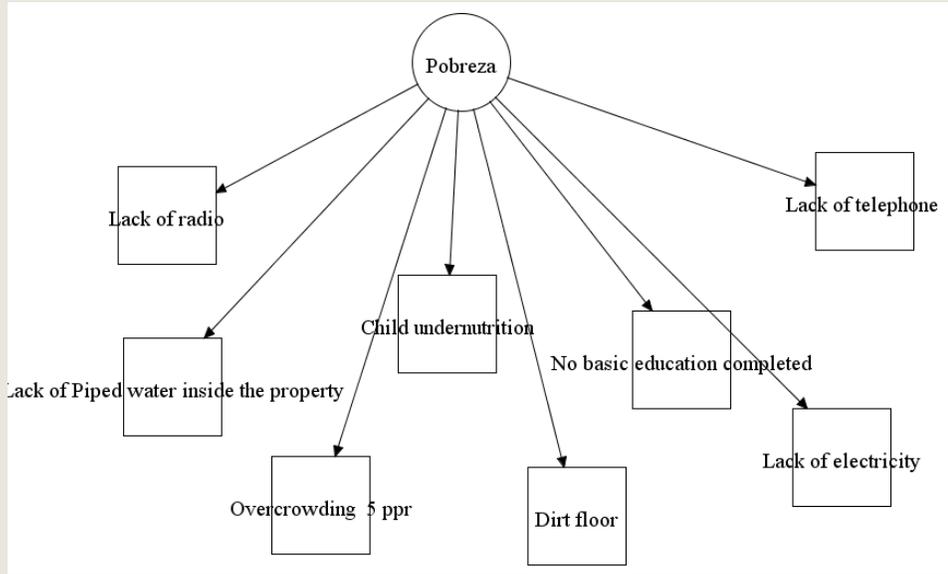
UNIVERSITY OF CALIFORNIA, BERKELEY

ROGER E. MILLSAP

BARUCH COLLEGE, THE CITY UNIVERSITY OF NEW YORK



Caso 1: Mismos indicadores



- ¿Son los modelos equivalentes para dos unidades?
- ¿Son las cargas factoriales equivalentes?
- ¿Son las ordenadas al origen equivalentes?
- ¿Son las varianzas de los residuales equivalentes?

- Invarianza débil
- Invarianza fuerte
- Invarianza estricta

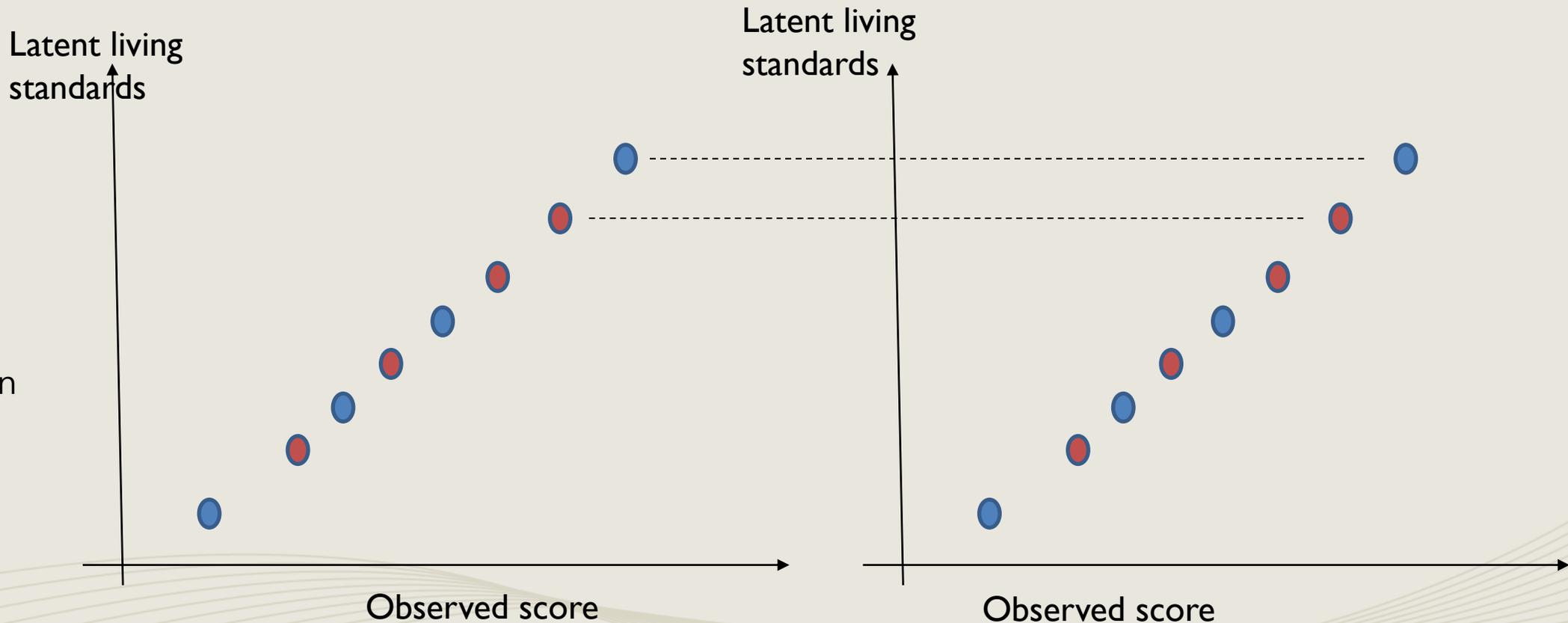


Invarianza estricta

La invarianza estricta implica que tanto los residuales, las ordenadas al origen y las pendientes son las mismas entre grupos.

Al sostenerse,
los scores son
comparables
entre grupos.

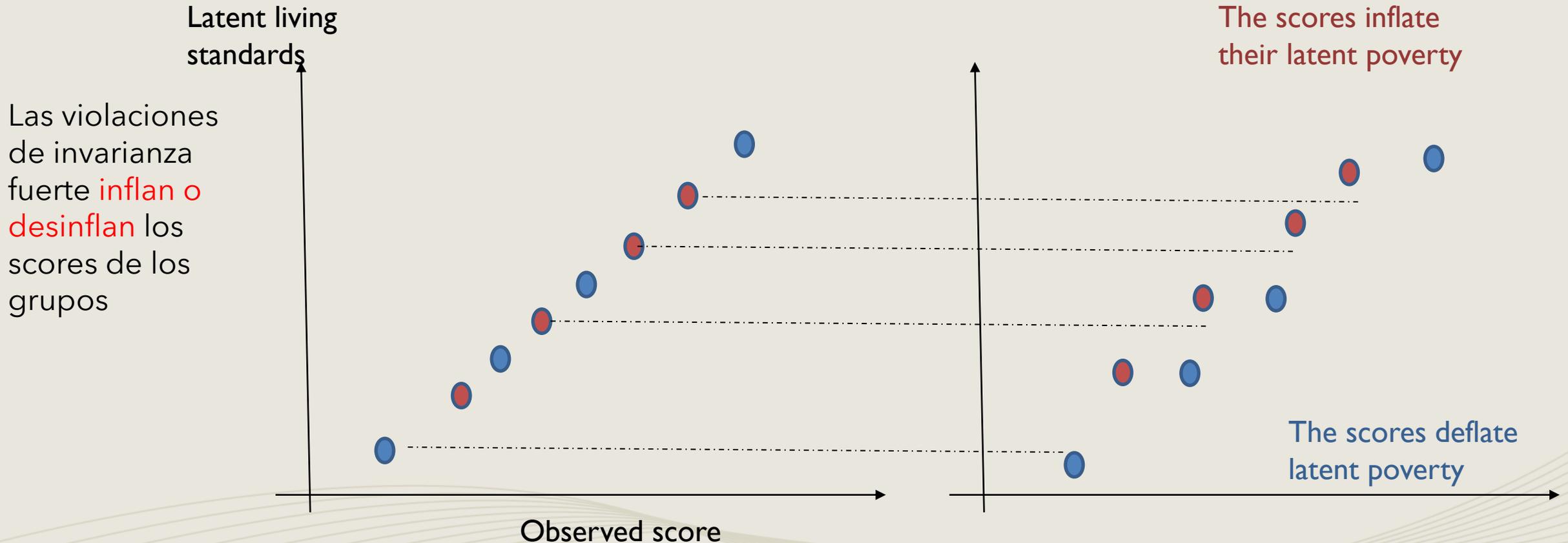
Puedo hacer
desagregación





Violaciones de invarianza fuerte

Las violaciones de invarianza fuerte implican cambio de ordenada al origen de los scores entre grupos. Unos se inflan mientras otros de desinflan.



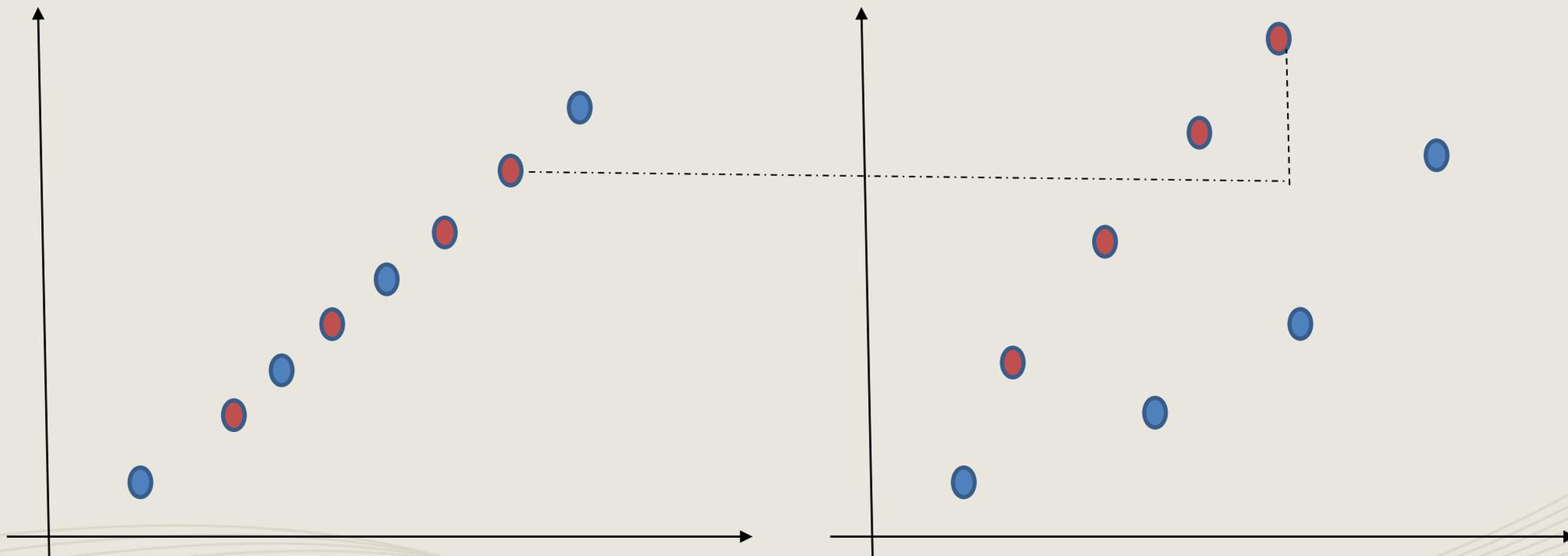


Violaciones de Invarianza débil

La invarianza débil se viola cuando las cargas factoriales son diferentes entre grupos

The bias
gets
worse

Las violaciones de invarianza débil implican desviaciones en los scores entre grupos



Puede haber violaciones de pendientes y no de ordenadas al origen!



Caso I

- Bajo esquemas de modelos equivalentes con mismas estructuras y variables, el análisis de invarianza me permite conocer si las comparaciones son justas entre unidades
- Violaciones de invarianza me harán más difícil las comparaciones porque estarán sesgadas
- Si las violaciones de invarianza son graves -i.e. invarianza débil no se sostiene, ni siquiera parcialmente-, tengo que modificar la escala para hacer comparaciones en el tiempo o entre grupos
- Esto es similar al caso de medición formativa, sin embargo, en el marco reflexivo existen alternativas

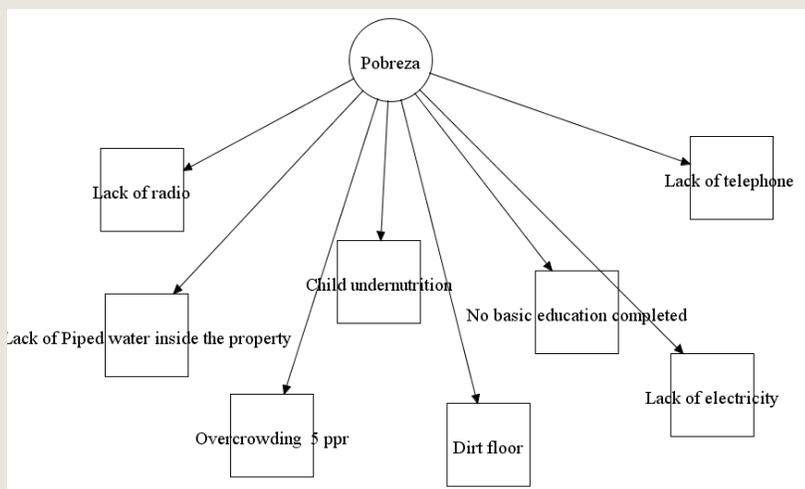
Caso I

Si la invarianza estructa se cumple –pero incluso la fuerte de manera parcial- puedo hacer comparaciones

Bajo invarianza estricta el subset $S_r =$ subset S_r

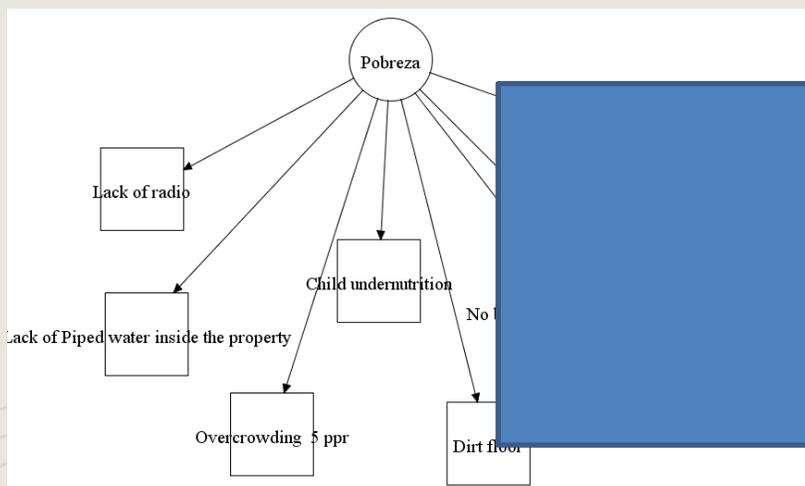
Si quiero preservar la escala original La pregunta es qué significa $S_r = 2$ respecto a $S_r = 3$.

“Test-equating and linking”



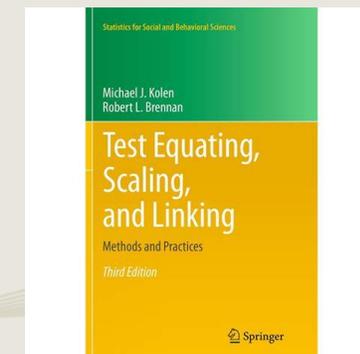
Conjunto I_r

Scores: S_r

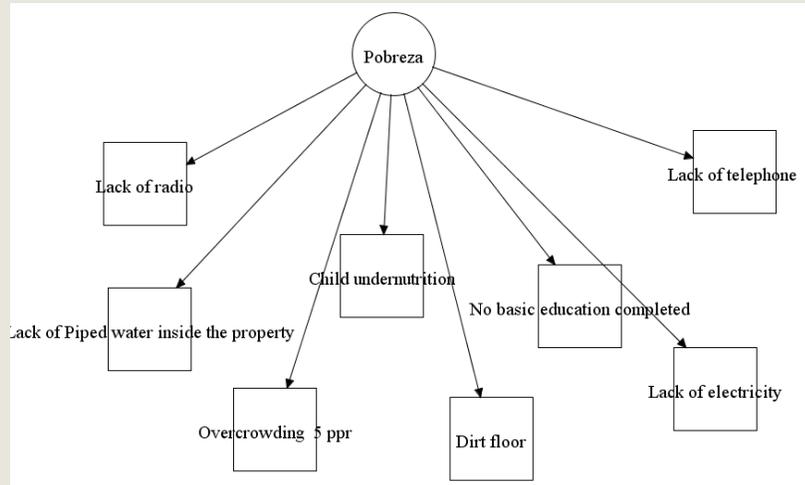


Conjunto I_{r2}

Scores: S_{r2}

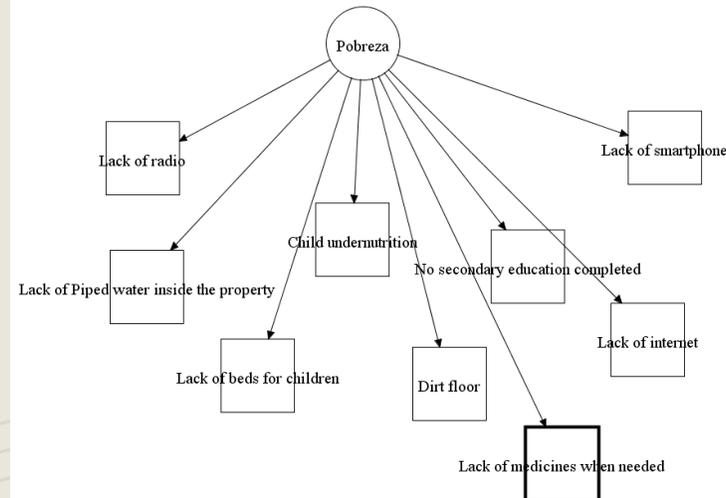


Caso 2: Diferentes indicadores y diferentes totales



Conjunto I_r

Scores: S_r
{0,8}



Conjunto I_{r2}

Scores: S_{r2}
{0,9}

Primera opción:

1. Identificar ítems invariantes -débil y fuerte- y *ajustar los scores totales a partir de estos ítems (anclas) -scale equating-*

Segunda opción:

1. Hipótesis: Ítems reemplazan previos ítems -Esto es útil para indicadores de una misma dimensión o para actualizaciones de umbrales-
2. Uso análisis de invarianza
3. Los scores totales se autoajustan - aunque seguimos con 0-8 vs 0-9

Primera opción

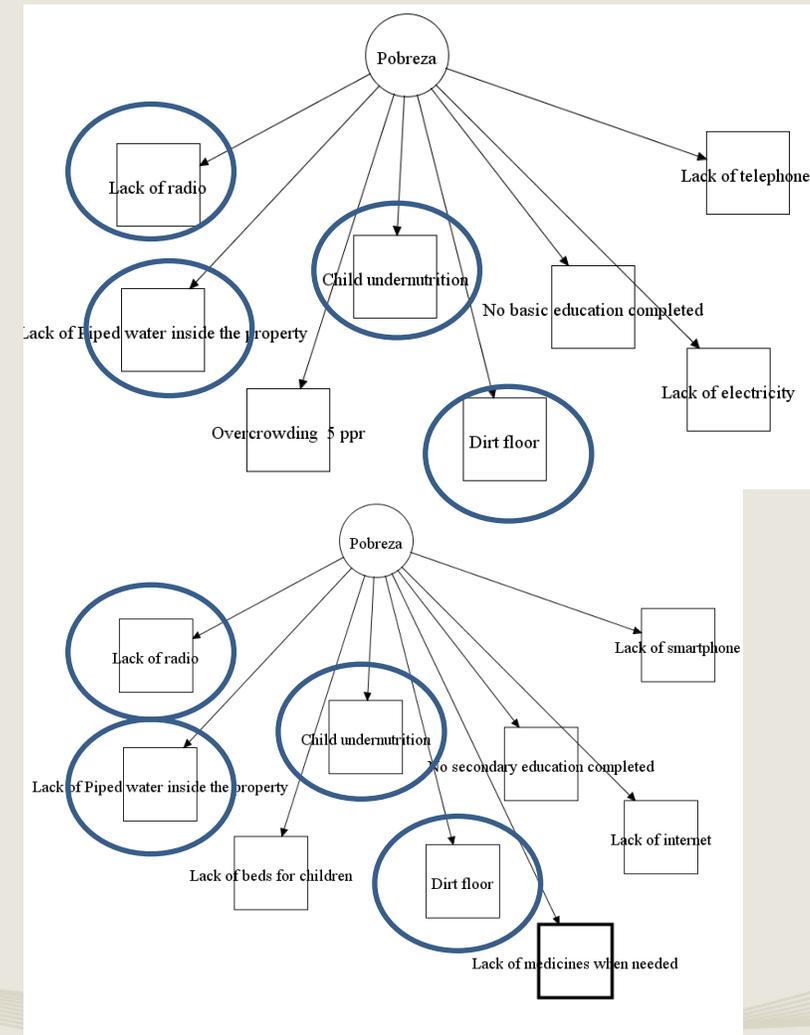
Primera opción:

1. Identificar ítems invariantes - débil y fuerte- y ajustar los scores totales a partir de estos ítems (anclas)

Puedo evaluar el segundo modelo a partir de la información que tengo del primero.

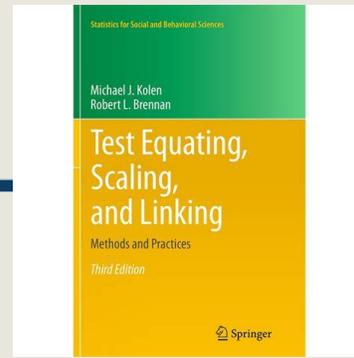
Esta evaluación es en términos de severidad, lo cuál me sirve para ajustar los scores totales

Condiciones: Se necesita un número grande de anclas. 60% de la escala, más de 10?





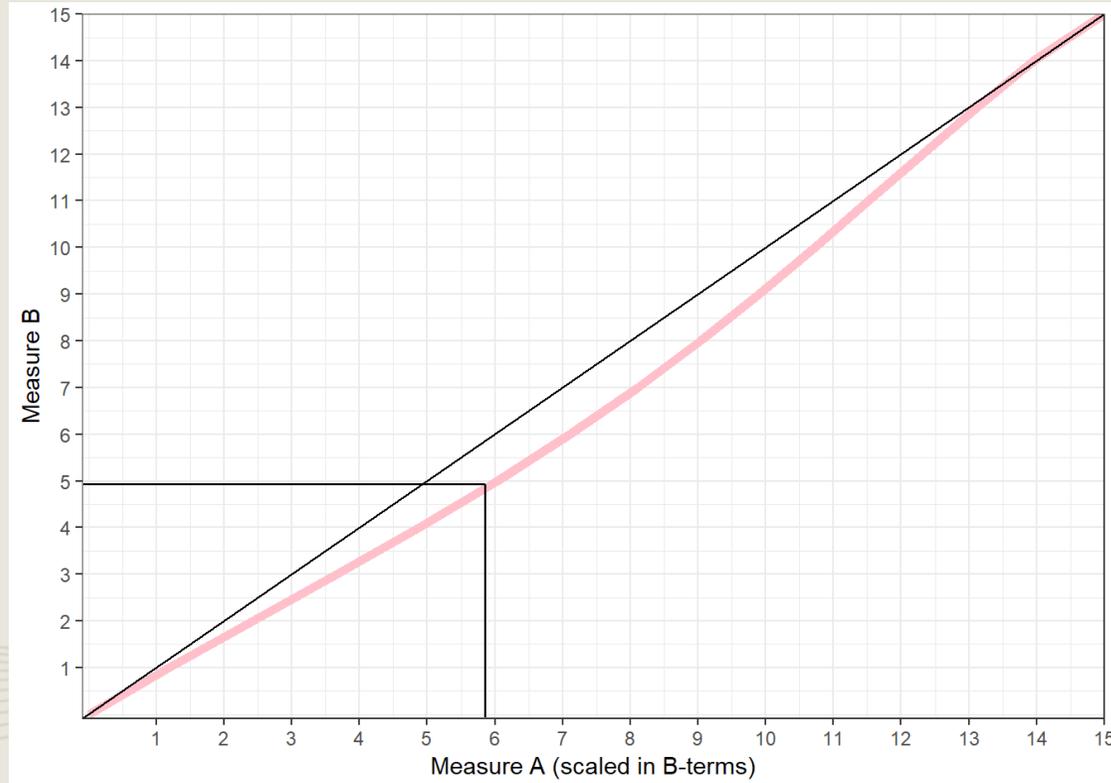
Primera opción: segundo paso



Test equating and scaling es un procedimiento que permite transformar scores y ponerlos en los mismos términos.

Primera opción:

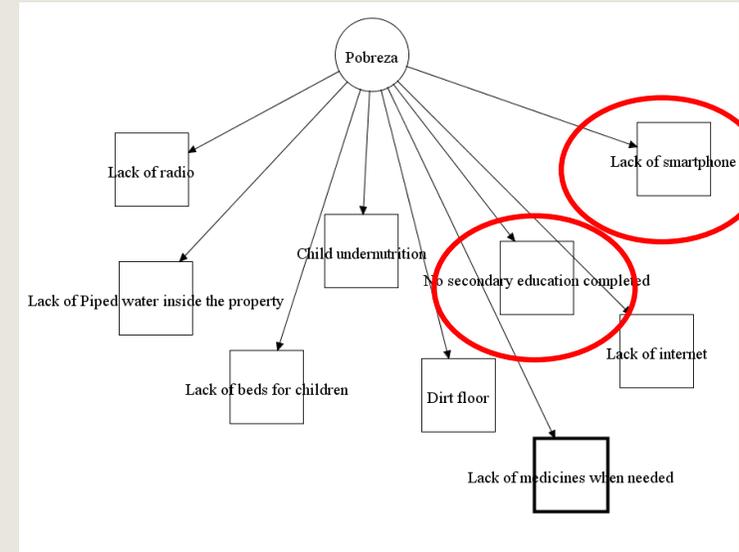
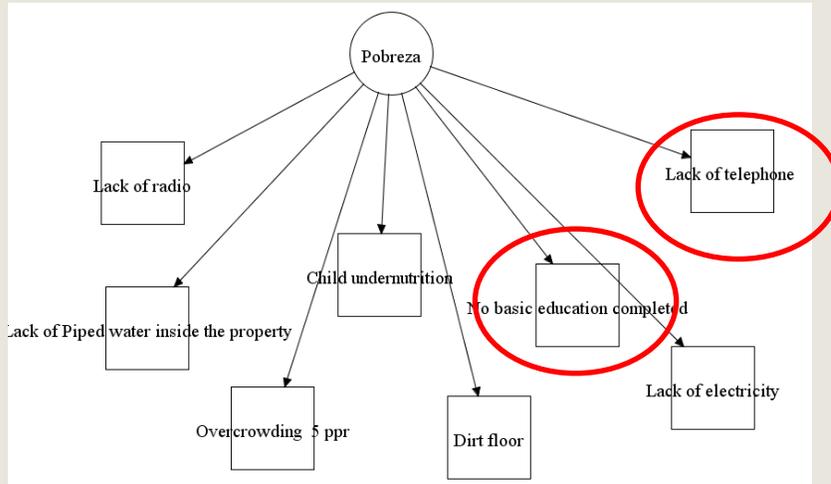
1. Identificar ítems invariantes -débil y fuerte- y ajustar los scores totales a partir de estos ítems (anclas)
2. En principio se necesita un buen número de anclas - más del 60% de la escala-



What's a deprivation score of 6 in terms of the previous measure.

It can be done the other way around!

Segunda opción: Items reemplazables



¿Son equivalentes? –Mismos valores de los parámetros: severidad y discriminación–

Si lo son, entonces puedo hacer reemplazos simples –Como en los tests Educación–

La pregunta de diferencias en los scores totales se abordaría con



¿Si quiero comparar sociedades con diferentes indicadores?

- Si la pobreza es relativa a cada sociedad, las prevalencias simples nos dirían cuál es la pobreza en dicha sociedad en sus propios términos. Esto a muchas veces es suficiente
- Si quiero hacer comparaciones con un referente, la situación cambia. Se trata de otra pregunta.
 - ¿Cuál es la severidad de la pobreza de Uganda en términos de la severidad de la pobreza en México?
 - ¿Cuál es la severidad de la pobreza en México en términos de la severidad de la pobreza en Uganda?
 - ¿Cuál es la severidad de la pobreza en determinado país condicional en un referente mundial de pobreza?

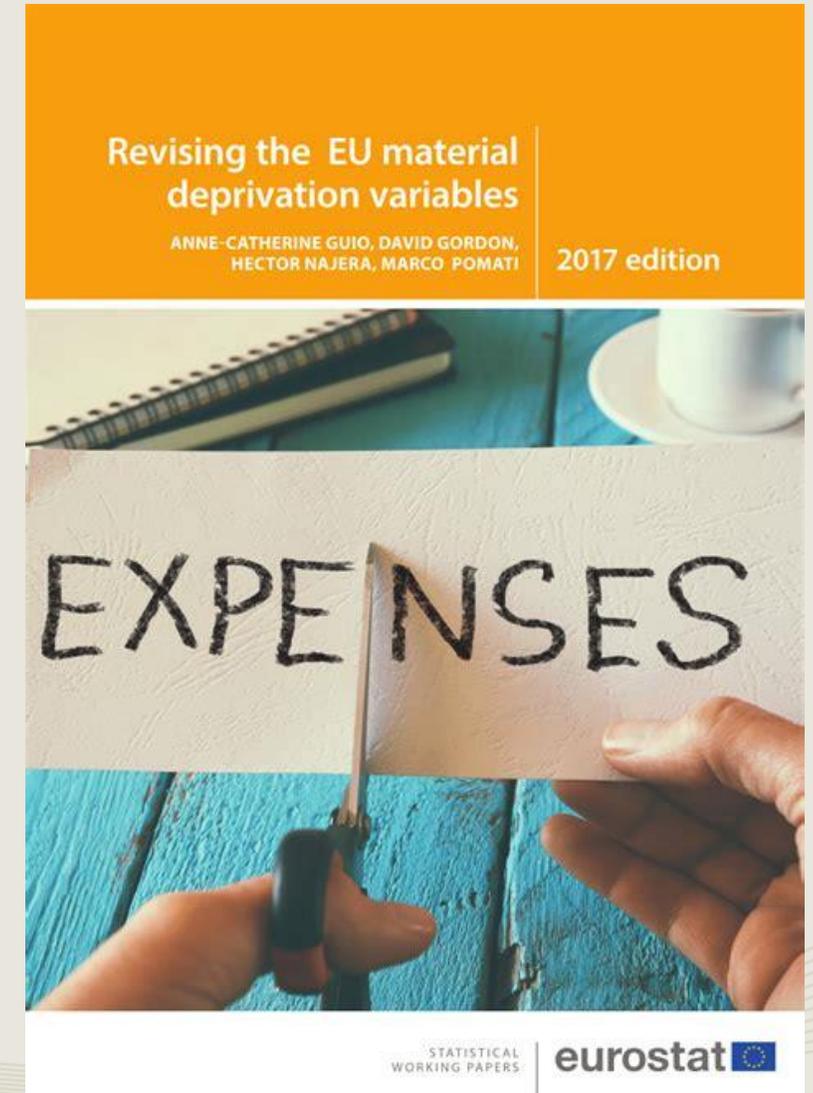
El tratamiento es similar al anterior.

Invarianza y scale-equating



¿Aplicaciones?

- No muchas!
- Comparación de pobreza en Europa bajo un estándar común europeo de necesidades
- Dado un set de indicadores, quisieramos comparar la pobreza de Bulgaria con la de Noruega bajo un mismo estándar
- No es el uso de mismos indicadores sino el uso de modelos equivalentes de medición!





Invarianza aproximada

Measurement Invariance

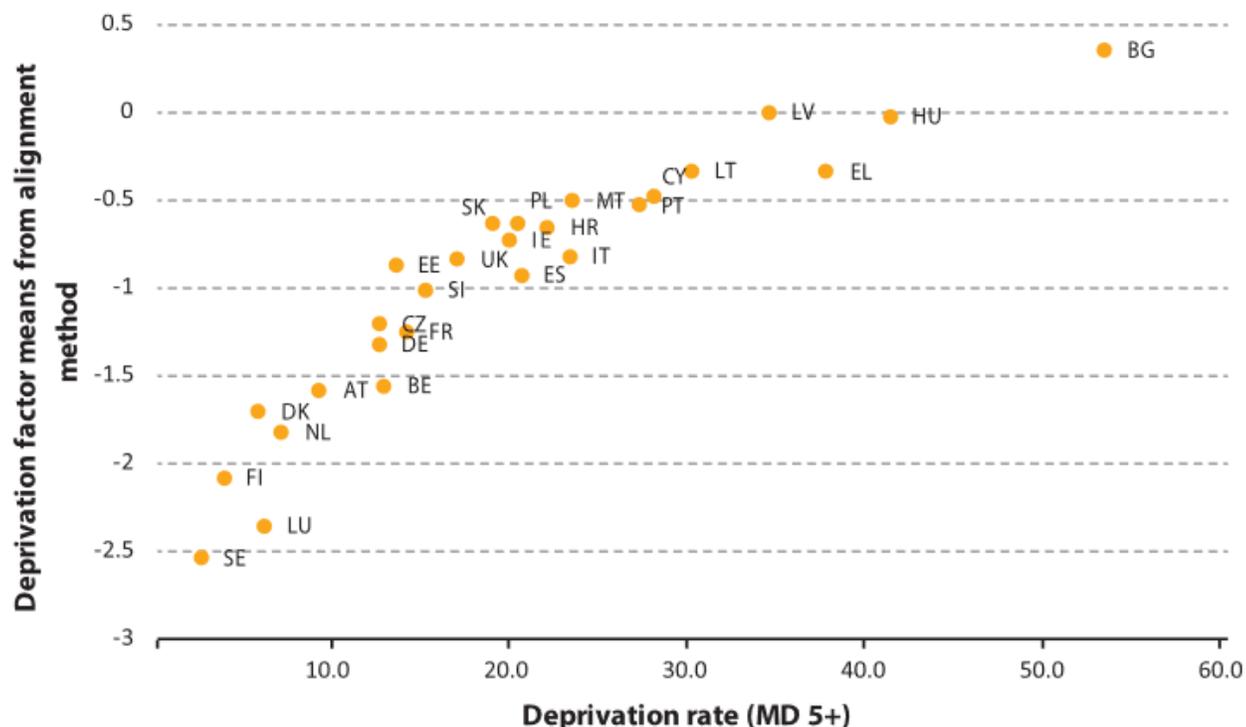
Alignment

- Billet, J., Meeusen, C. & Abts, K. (2021). [The relationship between \(sub\)national identity, citizenship conceptions, and perceived ethnic threat in Flanders and Wallonia for the period 1995-2020: A measurement invariance testing strategy](#). Forthcoming in *Frontiers in Political Science*. DOI: 10.3389/fpos.2021.676551
- DeMars, C.E. (2019). [Alignment as an alternative to anchor purification in DIF analyses](#). *Structural Equation Modeling: A Multidisciplinary Journal*, DOI: 10.1080/10705511.2019.1617151

Special Issue: Measurement Invariance - Sociological Methods & Research

- Ciecuch, J., Davidov, E., Algesheimer, R. & Schmidt, P. (2018). [Testing for approximate measurement invariance of human values in the European Social Survey](#). *Sociological Methods & Research*, 47:4 665-686. DOI: 10.1177/0049124117701478
- Davidov, E., Dülmer, H., Ciecuch, J., Kuntz, A., Seddig, D. & Schmidt, P. (2018). [Explaining measurement nonequivalence using multilevel structure equation modeling: The case of attitudes toward citizenship rights](#). *Sociological Methods & Research*, 47:4 729-760. DOI: 10.1177/0049124116672678
- Munck, I., Barber, C., & Torney-Purta, J. (2018). [Measurement invariance in comparing attitudes toward immigrants among youth across Europe in 1999 and 2009](#). *Sociological Methods & Research*, 47:4 687-728. DOI: 10.1177/0049124117729691
- Muthén, B. & Asparouhov, T. (2018). [Recent methods for the study of measurement invariance with many groups: Alignment and random effects](#). *Sociological Methods & Research*, 47:4 637-664. DOI: 10.1177/0049124117701488 [Mplus scripts](#).

Figure 9: Relationship between adjusted deprivation scores and the deprivation rate (proportion of people lacking at least 5 items out of 13), 2014



Note: see Annex 1 for country abbreviations.

Source: EU-SILC 2014 cross-sectional data, authors' computation.



¿Aplicaciones? Actualización de umbrales



QUALITY & QUANTITY

springer.com

Qual Quant. 2017; 51(3): 1217–1241.

Published online 2016 Mar 17. doi: [10.1007/s11135-016-0327-0](https://doi.org/10.1007/s11135-016-0327-0)

PMCID: PMC5406581

PMID: [28502997](https://pubmed.ncbi.nlm.nih.gov/28502997/)

Does measurement invariance hold for the official Mexican multidimensional poverty measure? A state-level analysis 2012

Hector Ernesto Najera[✉]

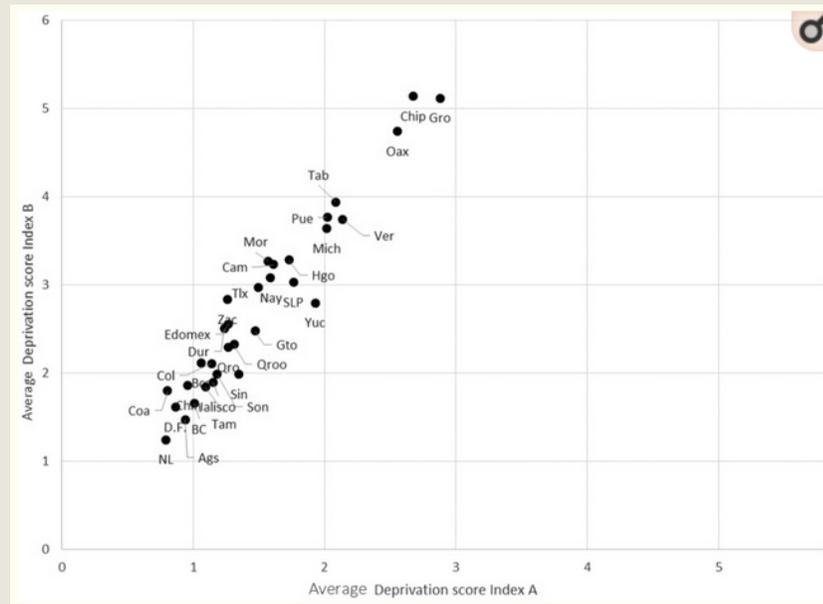
Does changing the thresholds lead to violations of MI?

Table 2

Percentage of the population deprived in relation to the given item. Mexico, 2012

	Index A (official thresholds)	Index B (adjusted)
Education	19	19
Food deprivation	23	44
Minimum social protection floor	62	62
Flooring material	4	4
Roofing material	2	25
Walling material	2	14
Overcrowding	10	10
Access to water	9	47
Sanitation	9	40
Fuel	13	13

Source Estimations based on INEGI-CONEVAL (2012)



Yes, updating the thresholds improves the mexican scale.

It affects invariance but midly

Could it be posible to use scale equating?

Yes but it is posible to find partial strong invariance. I guess this Will be stronger for the most recent data.



Conclusiones

- La comparabilidad es uno de los principales objetivos en medición
- Hay dos formas de pensar comparabilidad pero una es un callejón sin salida
- Hay que hacer cambios/actualizaciones para preservar la comparabilidad no para romperla!
- Estadísticamente hablando hay un marco de referencia claro para hacerlo pero poca investigación al respecto
- Los métodos son muy nuevos y algunas de las propiedades aún se desconoce.