

La insoportable levedad del dato

Fernando Cortés*

Este trabajo muestra, por una parte, que la teoría de la modernización y el análisis histórico estructural pueden combinarse en un solo esquema explicativo para dar cuenta de las determinantes de las migraciones en el origen del flujo.

Por otra parte, discute la pertinencia de los datos respecto a las proposiciones contenidas en la teoría e ilustra un proceso de aproximación entre el tipo de preguntas que mueven a la investigación y el tipo de respuestas que entregan los distintos instrumentos estadísticos utilizados.

También se discuten algunos problemas metodológicos de actualidad, tales como construcción del dato, niveles de análisis y diacronía-sincronía pero no en términos generales sino dentro del contexto de esta investigación específica.

A lo largo de la exposición queda claro que la aplicación mecánica de cualquier técnica estadística, sin entrar a considerar la validez de los supuestos en que se basa ni el tipo de estructura que la constituye, produce conocimiento no válido, y que la estadística entregará respuestas empíricas sin significado teórico si no se cuida la relación necesaria entre dicha estructura y el conjunto de las relaciones teóricas.

1. Introducción

La investigación social empírica que recurre al análisis estadístico no sólo debe satisfacer el requisito de precisión conceptual (planteamiento claro de las proposiciones teóricas, especificación de la connotación, denotación y operacionalización de los conceptos, confiabilidad y validez de las mediciones) sino que además debe seleccionar, entre los modelos estadísticos disponibles, aquel o aquellos que ayuden a responder las preguntas y permitan contrastar las hipótesis de la investigación. Se enlazan así dificultades que surgen de los planteamientos teóricos con las que nacen de los supuestos y las características de la técnica.

La relación entre la teoría y el instrumento estadístico presenta problemas. No plantearse como una dificultad la relación armoniosa que debe existir entre teoría y estadística en una investigación particular, puede conducir a imponer a la teoría la camisa de fuerza del modelo estadístico o a extraer conclusiones abusivas de los resultados estadísticos. Cualquier modelo estadístico es una estructura matemática que se impone a los datos introduciendo una serie de relaciones que no necesariamente coincidirán con las teóricamente planteadas a menos que el investigador seleccione la técnica cui-

* Facultad Latinoamericana de Ciencias Sociales, México.

Agradezco los comentarios y sugerencias de Brígida García y Rosa María Rubalcava.

dando la armonía entre la estructura de las relaciones teóricas y la del modelo estadístico.

No cuidar la relación entre teoría y técnica estadística lleva a dos tipos de situaciones polares: i) o la estructura de la técnica se impone a la de la teoría, en cuyo caso adquiere pleno sentido la advertencia de que "hay que cuidarse de la técnica porque no es neutral", ii) o, en caso contrario, se interpretan los datos sin respetar las limitaciones que impone el instrumento; la advertencia simétrica, opuesta a la anterior, debería ser "hay que cuidarse de la teoría porque no es neutral".

En lo que sigue, presentaremos un análisis del isomorfismo entre la estructura de una teoría o, si se quiere, de algunas hipótesis teóricas y de varias técnicas estadísticas de análisis de datos. Para ello se recurrirá a un viejo trabajo sobre migraciones rurales.¹

2. El problema y la aproximación teórica

Para identificar las determinantes estructurales de las migraciones rurales en Chile, en los primeros años de la década de los setenta, se establece la conexión entre el proceso histórico-político y la estructura productiva agraria individualizándose dos puntos de quiebre en el esquema de dominación que entrañaron cambios sustanciales en la estructura productiva agrícola.

La organización productiva tradicional del agro chileno denominada fundo fue modificada entre los años 1965-1970 como parte del proyecto político de la Democracia Cristiana, a la sazón en el poder, dando lugar a un nuevo tipo de organización productiva agraria denominada asentamiento. Un fenómeno similar ocurre en el lapso 1971-septiembre de 1973, en que el proyecto agrario de la Unidad Popular creó "centros de reforma agraria" (Ceras).

Los fundos se caracterizaban por la nula participación de los trabajadores en la toma de decisiones en la gestión de la empresa, de propiedad del terrateniente, y una baja participación relativa en el producto generado, que se reducía a un salario que era pagado en especies y dinero, bajo en comparación con los salarios obreros urbanos.²

En los asentamientos, los trabajadores recibían la posesión transitoria de los medios de producción durante un periodo que oscilaba entre tres y cinco años. Una vez cumplidos, los mismos trabajado-

¹ Argüello, Omar, "Estructura agraria, participación y migraciones internas", en *Migración y desarrollo*, núm. 3, CLACSO, Buenos Aires, 1974.

² *Idem, ibidem*, p. 12.

res decidían si procedían a la subdivisión de la tierra o continuaban trabajándola en forma cooperativa. En el periodo de transición las relaciones sociales de producción eran cooperativas, comunitarias y de autogestión, lo que se tradujo en participación plena en las decisiones de producir, en la gestión de la empresa, y en una participación igualitaria en la distribución del producto.³

El proyecto político agrario de la Unidad Popular superaba la idea de empresa de autogestión, intentando construir comunas agrarias cuyos excedentes se destinarían al desarrollo global. Sin embargo, el corto periodo que media entre la puesta en práctica de esta política y el momento en que se levantó la información, combinado con los problemas que surgieron ante el intento de ejecutar el proyecto político global, condujeron a que no se registrasen diferencias esenciales entre los asentamientos y los Ceras.⁴

La descripción de los cambios ocurridos en el proceso político global permiten *entender* las modificaciones estructurales experimentadas por las unidades agrícolas de producción. Los argumentos expuestos muestran cómo la descripción de un proceso macrosocial se utiliza para explicar los cambios de un microproceso. Es una de las maneras posibles para abordar la relación entre distintos niveles de análisis,⁵ aun cuando pareciera ser aplicable sólo en el caso en que la relación sea macro-micro y no a la inversa.⁶

Hasta este punto el argumento sostiene que los cambios en el esquema de dominación permiten identificar tres periodos a los cuales corresponden tres formas de organizar socialmente la producción agraria. El esquema teórico se completa al agregar que los cambios en las relaciones sociales de producción inhibirán las migraciones. Es esta hipótesis la que examinaremos a continuación.

³ *Idem, ibidem*, p. 13.

⁴ *Idem, ibidem*, pp. 14 y 15.

⁵ Véase Rolando García, *Food system and society, a conceptual and methodological change*, United Nations Research Institut for Social Development, Ginebra, Suiza, 1984, p. 26.

⁶ Cuando se trata de pasar de un nivel bajo a uno superior hay que resolver el problema de explicar las "emergencias", es decir, dar cuenta de donde surgen un conjunto de características presentes en el nivel superior, pero que no se encuentran presentes en el inferior. Un tratamiento sistemático de este tema se encuentra en la compilación de Lancelot Whyte, Alber G. Wilson y Donna Wilson, *Las estructuras jerárquicas*, Alianza Editorial, Madrid, 1973. En particular, los artículos de Lancelot Whyte, "Las jerarquías estructurales, o una retadora clase de problemas físicos y biológicos", pp. 19-32, y el de Mario Bunge, "La metafísica, epistemología y metodología de los niveles", pp. 33-46. También se recomienda revisar el trabajo de Rolando García, *Food system and society... op. cit*, pp. 23-40.

En primer lugar, la relación planteada exige, desde el punto de vista lógico, que la unidad de análisis sea tal que permita observar o construir variables que reflejen tanto las diferentes estructuras productivas agrarias como las migraciones. A partir de los datos referidos a cada agricultor y analizados por Argüello, se procedió, en este estudio, a construir medidas asociadas a las unidades de producción.

Esto se hizo con el propósito de rehacer el análisis buscando la correspondencia entre las unidades de análisis y el dominio de los conceptos teóricos. Ya hemos visto en el plano conceptual, que los tres tipos de organización de la producción (fundos, asentamientos y Ceras) corresponden a tres estructuras agrarias distintas, tal como se puede apreciar empíricamente en el cuadro 1 a través de los indicadores estabilidad laboral, participación en las decisiones productivas y participación gremial. El concepto de migración debe referirse a una propiedad de estas unidades de análisis y no a individuos. Por ello hemos incluido en la última columna del cuadro 1, el indicador de porcentaje de agricultores con proyecto de migrar.

Del cuadro se desprende que los indicadores de estructura productiva diferencian con toda nitidez los fundos de las unidades del área reformada.

La relación entre estructuras productivas y migraciones es difícil de comprender si no se hace referencia a los agricultores. Tal como ha sido planteada se ha reducido a una relación matemática o, si se quiere, estadística, pero no ha sido fundamentada. Se plantea así, por segunda vez, un problema de niveles de análisis. En este caso la descripción del nivel superior ayudaría a la comprensión del acto individual de migrar sólo si se argumentara que el agricultor internaliza la operación de la estructura que, combinada con características sicosociales, llevaría a que algunos migren y otros no. Esta forma de entender la actuación de la estructura sobre la migración debería primero, explicar por qué las organizaciones productivas reformadas tienden a retener a la población a diferencia de la explotaciones tradicionales y, segundo, explicar la forma como se materializa el proceso de internalización.

Sin embargo, esta no es la única forma de establecer la relación entre los procesos que afectan a las unidades de producción y la migración. En efecto, se puede utilizar la teoría de juegos para mostrar cómo la estructura impone límites a la decisión de migrar,⁷ o bien, la teoría estadística establece que la estructura condiciona (pero no

⁷ Przeworski, Adam, "Teoría sociológica y el estudio de la población: reflexiones sobre el trabajo de la Comisión de Población y Desarrollo de CLACSO", en *Reflexiones teórico-metodológicas sobre investigaciones en población*, El Colegio de México, México, D.F., 1982. pp. 66 a 82.

CUADRO 1

Migraciones potenciales, relaciones sociales y tipos de organización productiva

Tipos de organización de la producción	Porcentaje			
	Empleo estable	Participación en decisiones productivas	Participación gremial	Agricultores con proyecto de migrar
<i>Fundos*</i>				
F ₁	56.2	0.0	12.5	12.5
F ₂	31.8	0.0	0.0	13.6
F ₃	71.4	0.0	7.1	57.1
F ₄	66.7	4.2	43.7	16.7
F ₅	50.0	0.0	7.1	28.6
F ₆	52.3	2.3	25.0	13.6
F ₇	58.8	0.0	5.9	5.9
F ₈	31.2	0.0	31.2	31.2
F ₉	50.0	0.0	10.0	10.0
<i>Asentamientos</i>				
A ₁	97.2	97.2	83.3	0.0
A ₂	91.7	70.8	63.9	0.0
A ₃	100.0	94.1	41.2	0.0
A ₄	100.0	84.9	71.7	1.9
A ₅	77.3	63.6	40.9	9.1
A ₆	100.0	75.0	50.0	8.3
A ₇	100.0	83.3	75.0	0.0
<i>Ceras</i>				
C	100.0	87.0	73.9	4.3
C ₂	100.0	73.4	86.7	6.7
C ₃	100.0	93.6	83.9	0.0
C ₄	100.0	71.9	65.6	3.1
C ₅	100.0	83.9	69.6	1.8
C ₆	100.0	69.6	47.8	4.3
C ₇	100.0	100.0	85.0	0.0

* Debido a que algunos fundos incluían un escaso número de personas, fue necesario juntarlos para así aumentar el tamaño de cada unidad.

determina) la decisión de migrar.⁸ Cualquiera de estas dos opciones permite argumentar que la estabilidad laboral, unida a la mayor participación en las decisiones productivas, combinadas con la inserción en un proyecto de transformación de la sociedad global, cambian el horizonte de los trabajadores agrícolas modificando el campo de opciones, entre ellas, la de permanecer o desplazarse de las actividades agrarias. Esta forma de conceptualizar la relación entre or-

⁸ Cortés, Fernando, "A propósito de una explicación social de las migraciones rurales", FLACSO, Santiago, 1975, p. 8.

ganización de la producción agraria y migraciones implica que las proporciones de migrantes deben explicarse tanto por la operación de la estructura como por las dimensiones sicosocial y "penetración urbana" asociadas al proceso de modernización.⁹ La investigación plantea que el proceso de modernización de los trabajadores agrarios induce las migraciones siempre que su efecto no sea neutralizado o aun superado por el cambio en la organización de la unidad productiva.

La hipótesis que someteremos a contraste estadístico alude al proceso de transformación de las relaciones sociales de producción y a su impacto sobre las migraciones rurales. Por lo tanto, la información se debería haber levantado en, por lo menos, dos puntos del tiempo: uno antes y otro después de haber realizado la transformación de la estructura productiva para cada predio agrícola. Sin embargo, como proviene de un estudio sincrónico hay la posibilidad de que se incurra en falacia temporal. No abordaremos, en el presente trabajo, este tema.

3. La información y el instrumento de captación

Debido a razones de diversos órdenes sólo se cuenta con información obtenida a través de cuestionarios aplicados a trabajadores agrícolas en un periodo relativamente breve.

La información proviene de una pequeña comuna agrícola que fue seleccionada tomando en cuenta que, *i*) en el pasado se hubiese constatado una acentuada migración, *ii*) hubiese variabilidad en los tipos de organización de la producción, *iii*) resultara accesible para la realización del estudio¹⁰ y *iv*) que las tierras tuviesen calidad homogénea y que estuvieran comunicadas con los centros urbanos.¹¹ Con base en estos criterios se decidió realizar el estudio en la comuna de Chépica.

En el área reformada las entrevistas cubren a la totalidad de los trabajadores, mientras que en los fundos se tomó una muestra aleatoria.

En cuanto a la construcción de la variable "migración" hay que considerar que al intentar explicar la migración rural en el origen del flujo, se enfrenta al problema de localizar informantes que ya

⁹ En el proceso de modernización, la penetración de las normas sociales urbanas sobre las zonas rurales se puede diferenciar de las alteraciones sicosociales experimentadas por los agentes sociales.

¹⁰ Argüello, Omar, *op. cit.*, p. 18.

¹¹ Este último criterio intenta controlar el efecto de la productividad de la tierra y de las comunicaciones sobre las migraciones.

no se encuentran en el agro. Dado que la unidad de análisis es el predio, se podría obtener información, para calcular el porcentaje de migrantes, a partir de la memoria de los que aún se encontraban en el agro al momento de la realización del estudio. Sin embargo, no fue ésta la estrategia que se siguió.

La variable a explicar que se eligió para realizar este trabajo es la migración potencial y corresponde a la variable "proyecto de migrar" definida en el trabajo de Argüello.¹² El interés del estudio radica en identificar las determinantes de los diferenciales migratorios entre los predios agrícolas y no en el volumen de migración. Por lo tanto, si el perfil de la migración potencial correspondiera con el de la migración real, los resultados que se obtengan del análisis de la primera serían aplicables a la segunda.

4. El análisis de las correlaciones: alcances y limitaciones

La relación planteada entre estructura agraria y migraciones potenciales no sería rechazada por los datos, si fuesen inversas las correlaciones entre las correspondientes variables, esto es: sería menor el porcentaje de migrantes, si *a*) fuera mayor la proporción de agricultores que gozan de estabilidad laboral, *b*) fuera mayor el porcentaje que participa en las decisiones productivas y *c*) fuera mayor el porcentaje que participa en las organizaciones gremiales.

CUADRO 2

Correlaciones entre los indicadores de organización productiva y el porcentaje de migrantes potenciales

	Porcentaje de agricultores
	con proyecto migratorio
Con empleo estable	-0.6125
que participan en las decisiones de producción	-0.7073
con participación gremial	-0.6420

A pesar de que estos resultados son coherentes con las proposiciones teóricas, deben tomarse con cautela debido a que las correlaciones entre los indicadores de estructura son fuertes, lo que impide asignar satisfactoriamente los pesos relativos. En el cuadro 3

¹² Podríamos haber tomado otra variable "migración" a explicar, por ejemplo, el "proyecto migratorio no rutinizado". Sin embargo, hemos aplicado el principio de intercambiabilidad de los índices (ver Paul Lazarsfeld, "De los conceptos a los índices empíricos", en R. Boudon y Paul Lazarsfeld, *Metodología de las Ciencias Sociales*, vol. I Editorial Laia, Barcelona, 1973).

Distribución de algunas variables individuales por tipos de organización productiva

Tipos de organización productiva	Porcentaje de agricultores						
	jóvenes	escolarizados	con alta exposición a medios de comunicación de masas	con alto nivel de aspiración	con experiencia urbana	con contactos urbanos	
Fundos							
F ₁	25.0	50.0	37.5	75.0	12.5	31.3	
F ₂	54.5	50.0	40.9	54.5	13.6	31.8	
F ₃	50.0	85.7	51.1	42.9	14.3	50.0	
F ₄	45.8	70.8	29.2	29.2	4.2	50.0	
F ₅	57.1	71.4	50.0	57.2	0.0	42.8	
F ₆	20.5	61.4	25.0	47.7	6.8	38.6	
F ₇	29.4	47.1	17.6	41.2	11.8	11.8	
F ₈	62.5	81.2	37.5	18.8	12.5	43.7	
F ₉	18.8	62.5	31.3	31.3	0.0	31.3	
F ₁₀	40.0	50.0	40.0	50.0	0.0	30.0	
Asentamientos							
A ₁	19.4	58.3	22.2	22.2	2.8	22.2	
A ₂	29.2	41.7	27.1	37.5	10.4	37.5	
A ₃	11.8	82.4	11.8	23.5	5.9	29.4	
A ₄	34.0	67.9	17.0	15.1	11.3	24.5	
A ₅	18.2	40.9	27.3	81.8	13.6	36.4	
A ₆	16.7	50.0	33.3	0.0	8.3	50.0	
A ₇	41.7	66.7	50.0	16.7	8.3	25.0	
Ceras							
C ₁	34.8	52.2	43.5	47.8	13.0	4.3	
C ₂	53.3	73.3	60.0	33.3	33.3	6.7	
C ₃	45.2	77.4	48.4	22.6	9.7	19.4	
C ₄	40.6	81.2	62.5	25.0	31.2	18.8	
C ₅	39.3	67.8	48.2	19.6	14.8	12.5	
C ₆	30.4	34.8	21.7	21.7	8.7	17.7	

se ha incluido una serie de variables que, según la teoría de la modernización, deberían tener una correlación positiva con las migraciones.

El cuadro 4 muestra que los indicadores de modernización se correlacionan positivamente con las migraciones potenciales.

CUADRO 4

Correlaciones entre algunos indicadores de modernización y el proyecto de migrar

	<i>Porcentaje de agricultores</i>
	con proyecto migratorio
jóvenes	0.2294
escolarizados	0.2807
altamente expuestos a los medios de comunicación de masas	0.2365
con alto nivel de aspiraciones	0.3136
con experiencia urbana	-0.0667
que han tenido contactos urbanos frecuentes	0.5583

El análisis superficial de estos resultados llevaría a concluir que las variables relacionadas con la dimensión sicosocial tendrían menor valor explicativo que la variable que da cuenta de la penetración de los sectores modernos sobre los tradicionales.

Al comparar el vector del cuadro 4 con el del cuadro 2, que contiene las correlaciones entre las variables estructurales y las migraciones, se debería concluir que el principal peso explicativo corresponde a los factores estructurales ya que son los que presentan las mayores correlaciones.

Sin embargo, los elementos de la matriz de correlaciones (cuadro 5) que muestran los encadenamientos de primer orden entre las variables explicativas, hacen necesario realizar análisis intermedios antes de arribar a una conclusión fundamentada.

Para simplificar el argumento se supondrá que las variables contenidas en el cuadro son las únicas que inciden sobre el porcentaje de migrantes y que además son independientes entre sí (este es un supuesto que es contradicho por las correlaciones del cuadro 5). Si estos supuestos se cumplieren, entonces se podría argumentar que el peso explicativo mayor corresponde a los factores estructurales y el menor a las variables de la modernización. Aún más, con estos

CUADRO 5

Matriz de correlaciones entre las variables explicativas

	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉
X ₁	1.00	0.51	-0.14	-0.07	0.14	0.08	0.13	0.67	0.07
X ₂	-	1.00	0.09	0.11	0.20	0.12	0.17	0.42	-0.23
X ₃	-	-	1.00	0.93	0.84	0.23	-0.42	0.09	-0.57
X ₄	-	-	-	1.00	0.89	0.20	-0.43	0.05	-0.62
X ₅	-	-	-	-	1.00	0.29	-0.40	0.21	-0.58
X ₆	-	-	-	-	-	1.00	-0.44	0.41	-0.04
X ₇	-	-	-	-	-	-	1.00	-0.07	0.01
X ₈	-	-	-	-	-	-	-	1.00	0.12
X ₉	-	-	-	-	-	-	-	-	1.00

Porcentaje de trabajadores

X₁: jóvenes.

X₂: escolarizados.

X₃: con estabilidad en el empleo.

X₄: que participan en las decisiones productivas.

X₅: con participación gremial.

X₆: con experiencia urbana.

X₇: de mayores contactos urbanos.

X₈: con alta exposición a los medios de comunicación de masas.

X₉: con alto nivel de aspiraciones.

datos se jerarquizaría el peso explicativo de cada variable según el rango que le corresponda en la ordenación de los coeficientes de correlación.

Al levantar el supuesto de que éstas variables constituyen el universo de factores explicativos, pero se mantiene el de independencia, y si se considera que sólo se han incluido las variables *relevantes*, la jerarquización seguiría siendo válida. La modificación con respecto al caso anterior estaría en que se explicaría menos de 100% de la variabilidad total de la migración.

La información contenida en el cuadro 5 debe llevar a sospechar de estas conclusiones. Se sabe que los coeficientes de correlación lineal (producto-momento de Pearson) sólo dan cuenta de la variación concomitante entre dos variables: dos variables que se mueven conjuntamente, presentarán una correlación alta y positiva y si al crecer una, la otra disminuye sistemáticamente, entonces será negativa y tendrá un alto valor absoluto. Pero la variación conjunta puede deberse a que varían como consecuencia de los cambios experimentados por un conjunto de factores que las afectan en común o porque hay una ligazón directa (no mediada). La ordenación que se ha sugerido sólo sería válida si las correlaciones reflejasen relaciones directas entre las variables.

Para ilustrar lo que ocurre cuando las variables independientes están relacionadas tómese el caso de la variable "contactos urbanos" (X_7). Supongamos que es posible realizar el experimento de hacer variar sólo X_3 (estabilidad en el empleo). El cambio en X_3 tiene un impacto de primer orden sobre un conjunto de variables (efecto transmitido a través de los coeficientes de correlación), por ejemplo, sobre X_7 , la que a su vez "pega" sobre la migración potencial. Al calcular la correlación lineal entre X_7 y Y (las migraciones potenciales), aparecerá una asociación ilusoria.

Este ejemplo ilustra por qué utilizar el modelo de correlación lineal simple, cuando hay relaciones lineales estrechas entre las variables explicativas, conduce a resultados erróneos: no es posible aislar el efecto de cada variable independiente sobre la variable dependiente. El problema se puede resolver utilizando análisis de correlación parcial; sin embargo, se ha demostrado que este tipo de análisis se puede tratar también como parte del modelo de regresión.

5. El modelo de regresión lineal: la colinealidad estrecha entre algunos regresores

Los coeficientes de regresiones de un modelo lineal múltiple miden el efecto de cada variable independiente sobre la variable dependiente, bajo el supuesto de que las otras variables incluidas en el modelo se mantienen constantes. En caso de que haya relaciones lineales estrechas entre las variables explicativas, aparecen dificultades que trataremos más adelante.

Dado el modelo:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik}$$

y si se satisface el supuesto de que las X_i son independientes, entonces el impacto lineal de X_i sobre Y_i , que se mide a través de la derivada parcial $\delta Y_i / \delta X_i$, será:

$$\delta Y_i / \delta X_i = \beta_i$$

El ajuste de la ecuación de regresión lineal múltiple entre la variable explicada "migración potencial" y las nueve variables explicativas ya utilizadas en el análisis de correlación, entregó el siguiente resultado:

$$Y_i = -11.141 - 0.004X_1 + 0.190X_2 + 0.170X_3 - 0.203X_4 - 0.187X_5 + \\ + 0.178X_6 + 0.290X_7 + 0.197X_8 - 0.084X_9$$

$F = 4.560$ y $R^2 = 0.750$. Por otra parte, los valores F asociados a cada una de las variables individuales fueron:

Núm. de variable	1	2	3	4	5	6	7	8	9
Valor de F	2.90	1.99	22.17	8.29	1.20	0.13	3.44	0.65	0.27

El valor de F lleva a rechazar la hipótesis de que *todos* los coeficientes de regresión son iguales a cero. El coeficiente de determinación (R^2) indica que el modelo explica 74.56% de la varianza de las migraciones potenciales.¹³ Los valores de F asociados a cada variable permiten concluir que no tienen un efecto significativo las número 1, 2, 5, 6, 7, 8 y 9. En todos los casos se ha utilizado un nivel de significación de 5% y los supuestos clásicos de normalidad sobre los que se construye la teoría de las pruebas de hipótesis aplicables al modelo de regresión.

Los resultados expuestos llevarían a considerar que las variables explicativas de las migraciones potenciales (Y_i) son la estabilidad en el empleo (X_3) y la participación en las decisiones productivas (X_4), por lo tanto las únicas determinantes de las migraciones en el origen del flujo, son las relacionadas con el cambio estructural ya que las asociadas con la modernización no tienen impacto significativo.

Sin embargo, el valor del coeficiente β_3 indica que al controlar los efectos lineales de las otras variables incluidas en el modelo, a mayor estabilidad laboral correspondería mayor migración potencial. Como los asentamientos y Ceras tienen una mayor proporción con trabajadores con empleo estable que los fundos, entonces la proporción de migrantes potenciales tendería a ser más alta en los primeros que en los segundos. Este resultado no sólo contradice la concepción, sino que también atenta contra el conocimiento factual y la intuición. ¿Cómo se ha llegado a él?

Se ha establecido que los coeficientes de regresión del modelo lineal múltiple miden el impacto lineal de una variable manteniendo las otras constantes sólo si los factores explicativos son linealmente independientes, en caso contrario, el cambio en los valores de una variable inducirá modificaciones en los valores de las relacionadas con ella, y todas en conjunto harán variar a Y_i . En este caso el modelo será incapaz de distinguir aquella parte de la varia-

¹³ Esta interpretación sólo es válida en caso de que las variables explicativas sean independientes; luego la afirmación del texto debe entenderse como una licencia de redacción, aunque es de uso habitual. Para detalles ver, Otis Duncan, "Partial, Partitions and Paths", en *Sociological Methodology 1970*, Jossey-Bass Inc., San Francisco, 1970, pp. 38-47.

ción en la variable explicada que se debe a uno u otro de los factores. La información del cuadro 5 muestra que en este problema hay varias correlaciones lineales estrechas entre las variables explicativas. En las aplicaciones del modelo de regresión en que hay fuertes colinealidades (relaciones lineales) se sobreestiman los coeficientes de regresión de las variables más ligadas a la dependiente y se subestiman los de las más débiles.¹⁴

Interesa aclarar el por qué en este caso particular el coeficiente de regresión ha tomado un valor positivo y significativamente distinto de cero para la estabilidad laboral. Para ello es conveniente escribir el modelo de regresión de la siguiente manera:

$$Y_i = \beta_0 + \sum \beta_k X_k + U_i, (k = 1, 2, \dots, 9); (i = 1, 2, \dots, 24)$$

Si las X_k son independientes entre sí, entonces:

$$\delta Y_i / \delta X_k = \beta_k, (k = 1, 2, \dots, 9)$$

Pero si hay una relación funcional entre dos o más de las X , entonces los coeficientes estimados miden el impacto de la *combinación lineal de las variables relacionadas*. Para mostrar este resultado supongamos:

$$X_7 = f(X_3),$$

entonces:

$$\delta Y_i / \delta X_3 = \beta_3 + \beta_7 f'(X_3),$$

en que $f'(X_3)$ simboliza la derivada de $f(X_3)$ respecto a X_3 . El impacto lineal de la estabilidad laboral (X_3) sobre las migraciones potenciales se compone del efecto propio (β_3), al cual se agrega el efecto de los contactos urbanos X_7 (β_7) inducido por la correlación entre éstos y la estabilidad laboral (a menores contactos urbanos mayor estabilidad laboral). Por lo tanto, aunque el efecto propio de la estabilidad laboral pudiese ser negativo ($\beta_3 < 0$), el de la combinación lineal puede ser positivo. Para ello bastaría que:

$$0 < \beta_7 f'(X_3) > |\beta_3|$$

¹⁴ Ver Johnston, J., *Econometric models*, McGraw Hill, N. York, 1972, pp. 161 y 162. Para una presentación no matemática se puede consultar: H. Blalock, *Conceptualization and measurement in the social sciences*, Sage Publications, California, 1982, cap 1.

El coeficiente de regresión estimado, asociado a la estabilidad laboral, tendrá una relación positiva con las migraciones, siempre que el efecto neto de los contactos urbanos (β_7) combinado con el cambio inducido por la variación en la estabilidad laboral, $f'(X_3)$ (manteniendo constantes los restantes factores) fuese positivo y menor que el valor absoluto del efecto neto de la estabilidad laboral.

Las correlaciones lineales estrechas entre las variables independientes (cuadro 5) señalan dos hechos de importancia: i) el análisis de regresión tendrá que enfrentar problemas de multicolinealidad y ii) que la medición de las relaciones sociales y la modernización son confiables según el criterio de consistencia interna.¹⁵

Es claro que el análisis presentado se ha limitado a analizar relaciones entre variables olvidando que la discusión teórica remite a incorporar como determinantes de las migraciones el cambio estructural y la modernización (en sus dos dimensiones). El análisis estadístico ha establecido las relaciones entre los indicadores de los conceptos y se ha perdido el isomorfismo entre la estructura de la teoría y el instrumental estadístico.

A continuación se tomarán las variables como indicadores de los conceptos estructura y modernización y se procederá a construir índices sumatorios ponderados aplicando análisis de componentes principales.

6. Construcción de índices: componentes principales

Hasta ahora el análisis se ha limitado a los elementos de la matriz de intercorrelaciones, es decir, a las correlaciones entre las variables tomadas en pares. El análisis de componentes principales sirve para detectar la malla de relaciones y agrupar en un solo índice las variables que tengan un perfil similar.

CUADRO 6

Factores y sus contribuciones a la varianza total

Factores	1	2	3	4	5	6	7	8	9
Contribuciones	0.40	0.24	0.16	0.07	0.04	0.04	0.03	0.01	0.01
Contribuciones Ac.	0.40	0.64	0.80	0.87	0.91	0.95	0.98	0.99	1.00

¹⁵ Ver, por ejemplo, George W. Bohrnstedt, "Evaluación de la confiabilidad y validez en la medición de actitudes", en Gene Summers, *Medición de actitudes*, Trillas, México, 1978.

La aplicación de esta técnica a los datos de la matriz de intercorrelaciones (cuadro 5) arrojó los siguientes resultados:

El análisis de componentes principales es una técnica que permite reducir la matriz de datos en el sentido de las columnas (variables). Con los resultados que arroja se puede reducir el número de variables. Siempre entrega tantas combinaciones lineales (factores) como variables. El investigador debe decidir *i)* el criterio de resumen de la información, es decir, con cuántos factores trabajará y *ii)* cómo se interpretan.

En este caso, como en todo proceso de resumen de información al seleccionar menos de nueve factores se pierde información. La decisión debe equilibrar la pérdida de ésta con la ganancia en simplicidad. Los datos del cuadro 6 justifican seleccionar los primeros tres factores que dan cuenta de 80% de la varianza de "las variables explicativas".

Para interpretar el sentido de los factores seleccionados se puede realizar el examen de los pesos que afectan a las variables en cada combinación lineal:

CUADRO 7

Pesos de cada variable en cada uno de los factores seleccionados

<i>Variables</i>	P_1	P_2	P_3
X_1	0.028	0.600	0.110
X_2	0.128	0.482	0.323
X_3	0.488	-0.175	0.043
X_4	0.496	-0.167	0.092
X_5	0.495	-0.027	0.062
X_6	0.215	0.143	-0.548
X_7	-0.270	0.179	0.582
X_8	0.123	0.543	-0.178
X_9	0.347	0.002	0.466

El primer factor presenta los pesos mayores en las variables X_3 , X_4 , X_5 , que representan a la estabilidad laboral, a la participación en las decisiones productivas y a la participación gremial respectivamente, que son los indicadores de la estructura productiva agraria. En consecuencia, podemos interpretar a P_1 , como un índice de estructura.

En la combinación lineal P_2 , son las variables X_1 , X_2 y X_8 , las que tienen las ponderaciones mayores. Este factor recoge el efecto de la juventud, la escolaridad y la alta exposición a los medios de

comunicación de masas, por lo que podríamos considerarlo como un índice de la dimensión sicosocial de la "modernización".

El tercero da mayor importancia a las variables de conexión del campo con el medio urbano y podría considerarse como un índice de la dimensión "penetración urbana".

Ahora bien, si se estandarizan las variables y se ponderan por sus correspondientes pesos factoriales, entonces se generan los puntajes de cada unidad de producción en cada uno de los factores.

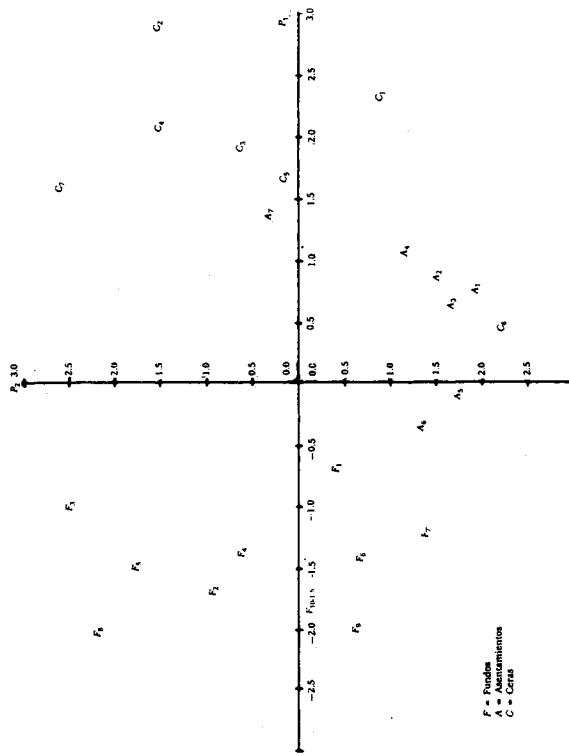
CUADRO 8

Distribución de cada factor por tipo de organizaciones productivas

<i>Unidades</i>	P_1	P_2	P_3
F ₁	-0.803	-0.382	0.138
F ₂	-1.830	0.976	-0.901
F ₃	-1.081	2.515	0.997
F ₄	-1.459	0.667	1.318
F ₅	-1.603	1.858	1.495
F ₆	-1.522	-0.642	0.455
F ₇	-1.377	-1.352	-1.495
F ₈	-2.085	2.222	-0.417
F ₉	-2.115	-0.608	-0.042
F ₁₀	-1.539	-0.029	0.384
A ₁	0.640	-1.856	0.356
A ₂	0.726	-1.446	0.492
A ₃	0.511	-1.640	1.206
A ₄	0.931	-1.099	0.108
A ₅	-0.245	-1.639	-0.460
A ₆	-0.495	-1.311	0.266
A ₇	1.193	0.320	-0.022
C ₁	2.216	-0.825	-0.729
C ₂	2.786	1.521	-2.142
C ₃	1.819	0.620	0.126
C ₄	2.012	1.573	-1.640
C ₅	1.567	0.144	-0.956
C ₆	0.317	-2.149	-0.811
C ₇	1.485	2.565	2.282

Otro camino para interpretar las combinaciones lineales consiste en la representación gráfica de los datos del cuadro 8, utilizando los factores como ejes coordenados. La gráfica que se incluye a continuación tiene como eje de abscisas P_1 y como eje de ordenadas P_2 .

Esta gráfica ordena, en relación al eje de abscisas, a los tres tipos de organizaciones productivas: los fundos se concentran hacia el extremo negativo del eje, los asentamientos alrededor del origen y los Ceras se concentran en los valores positivos.



GRAFICA 1

Distribución de las unidades productivas en función de los factores estructural (P_1) y sicosocial (P_2)

Los fundos, asentamientos y Ceras se mezclan en relación al eje de las ordenadas. Por ejemplo, en los valores altos de P_2 se encuentran fundos, asentamientos y Ceras que se caracterizan por presentar altas proporciones de agricultores modernos en la dimensión sicosocial (las tres formas de organizar la producción contienen un alto porcentaje de jóvenes, de educados y de miembros altamente expuestos a los medios de comunicación de masas).

Los resultados a que se ha llegado a través del análisis de componentes principales permite sostener que los datos reflejan la operación de los fenómenos estructural y modernización (en sus dos dimensiones: sicosocial y penetración urbana) que generarían la matriz de intercorrelaciones del cuadro 5.

Los índices construidos serían indicadores válidos de las determinantes de las migraciones. A través del análisis factorial se ha logrado establecer una correspondencia apropiada entre los dominios teórico y empírico.

7. La construcción de un modelo para las migraciones rurales en el origen del flujo: análisis de regresión.

Una vez construidos los índices hay que pasar a analizar el poder explicativo de las teorías que se manejan. Habría que jerarquizar el poder explicativo de los tres factores identificados, la naturaleza de la relación y cuál es el tipo de impacto que se esperaría sobre el volumen de migraciones por la variación de algunos de ellos.

La técnica apropiada para encarar estas inquietudes es el análisis de regresión, pero a diferencia de la aplicación que se hizo en la sección 5, esta vez las variables explicativas miden apropiadamente los conceptos teóricos y, por lo tanto, se espera que los cambios en las migraciones covaríen con ellas.

El paso de las variables a factores, además de resolver el problema de la concordancia entre los dominios conceptual y empírico, ha deshecho la fuente de la multicolinealidad; el método empleado usa entre sus criterios el que los factores sean independientes entre sí.

El análisis factorial ha resuelto los dos problemas básicos que habían llevado a dudar de los resultados de los análisis de regresión y de correlación: la relación lineal estrecha entre las variables independientes y la identificación de índices que corresponden a los conceptos teóricos. Despejados estos obstáculos queda abierto el camino para aplicar el análisis de regresión.

El ajuste del modelo de regresión lineal múltiple en que la variable dependiente es la migración potencial y en que P_1 , P_2 y P_3

simbolizan, respectivamente, los factores estructural, sicosocial y atracción urbana entregó los siguientes resultados:

$$Y_i = 10.58 - 5.45 P_1 + 4.48 P_2 - 0.35 P_3$$

$R^2 = 0.62$, $F = 11.01$ y los F asociados a cada factor fueron $F_1 = 20.55$, $F_2 = 12.48$ y $F_3 = 0.04$

Ya que las variables explicativas son todas independientes entre sí, el coeficiente de determinación indica que este modelo explica 62% de la varianza de las migraciones potenciales de las unidades de producción. El valor de F (general) permite concluir que no todos los coeficientes de regresión son estadísticamente iguales a cero.

El signo de los coeficientes de regresión muestra el efecto depresivo del cambio estructural sobre las migraciones y el alentador de los factores sicosociales. Los valores F asociados a los coeficientes indican que los efectos de P_1 y P_2 son estadísticamente distintos de cero a 5% y que un cambio unitario en P_1 disminuirá la migración potencial en 5.4% y uno de la misma magnitud en P_2 la aumentará en 4.5%. En cambio, el impacto del factor P_3 (atracción urbana) es estadísticamente igual a cero.

La carencia de poder explicativo de la penetración urbana se hace patente al ajustar el modelo con las variables explicativas P_1 y P_2 :

$$Y_i = -10.58 - 5.39 P_1 + 4.44 P_2$$

$R^2 = 0.62$, $F = 17.31$, $F_1 = 21.56$ y $F_2 = 13.07$.

La eliminación del impacto no disminuyó el porcentaje de varianza explicada y aumentaron significativamente los valores de F_1 y F_2 como consecuencia de la ganancia en el número de grados de libertad.¹⁶

A pesar de que este modelo satisface los requerimientos de la estadística, no representa adecuadamente las hipótesis teóricas. Se ha planteado que los factores estructurales tienden a neutralizar el efecto positivo que tiene la modernización sobre las migraciones. El conjunto de relaciones en que participa el trabajador agrícola es un importante aliciente para permanecer dentro del área reforma-

¹⁶ Debido a que los componentes principales son independientes, los coeficientes de regresión de éstos dos últimos modelos deberían ser los mismos. Las pequeñas diferencias se deben a las aproximaciones en los cálculos.

da. Pero además, se ha señalado que hay una interacción entre ambas determinantes que no ha sido formalmente incorporada al modelo de regresión. Las ecuaciones lineales reflejan la actuación independiente de cada variable explicativa sobre la explicada, bajo el supuesto de que las restantes se mantienen constantes.

Con la intención de incorporar la interacción se propone un modelo que captura la interacción a través del producto aritmético entre las variables explicativas. El resultado del ajuste fue el siguiente:

$$Y_i = 10.49 - 4.30 P_1 + 4.49 P_2 - 1.73 P_1 P_2$$

$$R^2 = 0.68, F = 14.07, F_1 = 24.11, F_2 = 14.62 \text{ y } F_{12} = 3.49$$

La inclusión del término interactivo ha sido poco feliz; su efecto es estadísticamente igual a cero con 95% de confianza. La ganancia en explicación (aumentó el coeficiente de determinación) es ficticia porque se debe a que la correlación entre P_1 y el término interactivo ($P_1 P_2$) es de 0.463.¹⁷ El otro elemento que explica el aumento en R^2 es el número adicional de regresores.

Las medidas estadísticas analizadas llevan a concluir que se debería preferir el modelo que no incluye interacciones. Pero no suministra elementos suficientes para rechazar la idea de que la estructura y el factor sicosocial interactúan al incidir sobre las migraciones. En efecto, consideremos el cambio que provoca sobre la migración potencial una variación unitaria en el índice estructural, manteniendo constante el índice sicosocial:

$$\delta Y_i / \delta P_2 = 4.49 - 1.73 P_1$$

Esta ecuación muestra que el impacto de la estructura sobre las migraciones depende linealmente del grado de modernización de la unidad productiva: a niveles altos de modernización el efecto del cambio estructural es menor. Este resultado indica que si se expropia un fundo en que un alto porcentaje de sus trabajadores se pueden catalogar como sicosocialmente moderno, entonces el efecto de la reforma agraria tenderá a ser menor que si hubiese tenido una proporción menor de agricultores con ese atributo. Este resultado claramente no concuerda con las proposiciones teóricas: se ha propuesto una formalización que no refleja los desarrollos del dominio teórico.

¹⁷ Recuérdese que la existencia de correlación lineal entre las variables explicativas no autoriza a interpretar el coeficiente de determinación de la manera usual.

Interesa proponer una formalización que muestre el efecto diferencial de la dimensión sicosocial de la modernización en los fundos con respecto a las explotaciones del área reformada. Con este propósito se define una variable ficticia (también llamada muda o "dummy") que asume el valor 1 para los fundos y 0 para los asentamientos y Ceras y ha sido simbolizada como X . Como el argumento conceptual señala que la modernización sólo operaría en las explotaciones tradicionales (fundos) se construye la variable explicativa P_2X , que toma el valor P_2 para los fundos y 0 para los predios del área reformada. Con la variable estructural y esta variable explicativa, se ajustó un modelo de regresión lineal, obteniéndose la siguiente ecuación:

$$Y_i = 8.57 - 4.01 P_1 + 9.23 P_2X$$

$$R^2 = 0.72, F = 27.07, F_1 = 27.07 \text{ y, } F_{2X} = 25.02$$

Al comparar estos valores con los de los modelos anteriores se ve que todos ellos tienen mayor significación estadística. Además el R^2 se puede interpretar como el porcentaje de varianza explicada, ya que la correlación entre las variables explicativas es baja: 0.28.

Los resultados del ajuste permiten sostener que la interacción formalizada como el producto de P_2 por X , se adecúa mejor a los datos que el producto de P_1 por P_2 , por lo tanto, permiten depositar mayor confianza que en los modelos ensayados anteriormente. Por otra parte, esta ecuación representa fielmente las ideas teóricas.

Con el propósito de abundar en este último aspecto consideremos:

$$\delta Y_i / \delta P_2 = 9.23 X$$

Como X asume el valor 1 para los fundos y 0 para los asentamientos y Ceras:

$$\delta Y_i / \delta P_2 = 9.23, \text{ en el caso de los fundos y,}$$

$$\delta Y_i / \delta P_2 = 0, \text{ en el caso de los predios del área reformada.}$$

En el modelo que mejor ajusta a los datos se encuentra expresada la idea de que la dimensión sicosocial de la modernización es un factor explicativo de las migraciones sólo en las explotaciones agrarias tradicionales.

La diferencia básica entre las dos maneras de formalizar la interacción que se ha presentado radica en que la primera, a diferencia de la segunda, es simétrica. Representarla a través de un producto

aritmético implica que el efecto de lo estructural sobre las migraciones dependerá del "nivel" de modernidad de los agricultores y viceversa. Se introdujo la variable ficticia para romper la simetría.

Sólo en este último modelo hay concordancia entre la estructura lógica del planteamiento teórico y el modelo matemático. La hipótesis que se contrasta es asimétrica, malamente puede representarse por una expresión matemática simétrica.

8. Conclusiones

La íntima vinculación entre los diversos ámbitos de la actividad social dará origen, en la investigación social empírica, a matrices de datos con altas intercorrelaciones. En este trabajo hemos mostrado las consecuencias que tiene este hecho para el análisis estadístico de la información.

El análisis de correlación simple no sólo presenta serias limitaciones sino que normalmente induce a extraer consecuencias erróneas de los datos debido a la carencia de control. El control del efecto de las otras variables que oscurecen los efectos netos de una variable sobre otra, se realiza en el momento de generar los casos o bien a través del control estadístico para lo cual es necesaria la incorporación de aquéllas al análisis.¹⁸ En este escrito sólo nos hemos preocupado por la segunda de estas formas.

Debido a las dificultades que nacen de la aplicación del modelo de correlación simple, se decidió usar el modelo de regresión lineal múltiple, que permite tanto controlar el efecto de las otras variables, vía la incorporación de ellas al modelo, como pasar de los coeficientes de regresión a los coeficientes de correlación parcial. Las relaciones lineales estrechas entre las variables de la matriz de datos originó el problema de multicolinealidad que conduce a una estimación sesgada de los parámetros. Por otra parte, se tomó conciencia de que las variables del modelo de regresión planteado no reflejaban los conceptos teóricos.

Esta doble dificultad se resolvió aplicando análisis de componentes principales que garantiza la ortogonalidad de los factores, resolviéndose, de este modo, el problema técnico. Sin embargo, la superación del otro problema no es automática porque en su aplicación se presenta la dificultad de asignar contenido teórico a los factores matemáticamente determinados. Para esta tarea se analizaron los

¹⁸ Un tratamiento sistemático sobre el problema del control en ciencias sociales se encuentra en Arthur Stinchcombe, *La construcción de teorías sociales*, Ediciones Nueva Visión, Buenos Aires, 1970, pp. 47-49.

pesos factoriales y la representación gráfica de los puntajes de los factores en cada unidad, y se concluyó que el primero se puede entender como un índice de estructura, el segundo como uno que da cuenta de la dimensión sicosocial de la modernización y el tercero como un índice que representa la dimensión de la penetración de las normas y valores urbanos sobre el campo.

Una vez identificados los factores interesaba estudiar el papel que juegan como determinantes de las migraciones potenciales. Con este propósito se ajustó un modelo de regresión en que las variables independientes fueron los factores en lugar de las variables originales. Sin embargo, un modelo de regresión lineal múltiple de este tipo no da cuenta del efecto conjunto (interacción) de las determinantes de la modernización y del cambio estructural sobre las migraciones potenciales. La discusión acerca de la simetría de la interacción sólo tiene sentido al examinar la correspondencia entre el modelo y el enunciado de que la modernización opera como factor explicativo de las migraciones sólo en los contextos agrarios tradicionales.

En cuanto a la teoría de las migraciones, esta investigación señala que el volumen de migrantes potenciales es acicateado por el proceso de modernización de los agricultores y desalentado por cambios en la estructura agraria que implican su participación no sólo en el proceso de producción sino en el destino de la sociedad local y nacional. El cambio de estructura neutralizaría el efecto de la modernización.

El hecho de que la penetración urbana no haya jugado un papel significativo en la determinación de las migraciones potenciales se podría explicar a partir de los criterios utilizados para seleccionar las unidades de observación y de la definición de la variable dependiente: la selección del lugar en que se realizó la investigación controló el impacto de la penetración de la ciudad sobre el campo, y la variable dependiente se compone de corrientes migratorias potenciales rural-rural y rural-urbana.

Por último, hay que dejar en claro que estos resultados deben ser relativizados porque están afectados por la falacia temporal, tema que queda fuera de los límites de este trabajo.

*Cuanto más pesada sea la carga, más a ras de tierra
estará nuestra vida, más real y verdadera será.*

*Por el contrario, la ausencia absoluta de carga
hace que el hombre se vuelva más ligero que el aire,
vuele hacia lo alto, se distancie de la tierra, de su ser
terreno, que sea real sólo a medias y sus movimientos
sean tan ligeros como insignificantes.*

Milan Kundera: La insoportable levedad del ser.

