

264

**EL TRATAMIENTO DE LAS
VARIABLES EN ECONOMÍA:
Una revisión y un marco
conceptual**

**Oscar Millones y Jorge Bazán
Abril, 2008**

EL TRATAMIENTO DE LAS VARIABLES EN ECONOMIA: Una revisión y un marco conceptual

Oscar Millones*

Jorge Bazán**

Resumen

En este artículo se revisa el manejo del concepto de las variables en la investigación empírica en Economía. Se incorporan ideas desde otras áreas cuantitativas, y se ordenan en un marco conceptual. Se presenta una taxonomía de las variables y se discuten aspectos metodológicos que se derivan de ella. Consideramos que el reflexionar sobre la naturaleza de las variables desde el planteamiento de los objetivos de la investigación, y antes de vincularlas con el trabajo empírico, es una tarea que incluso debe ser parte de la formulación misma de teorías, hipótesis o ideas involucradas en el trabajo empírico económico.

Abstract

This article reviews the treatment of the concept of variables in empirical research in economics. It incorporates ideas from other quantitative areas, and are arranged in a conceptual framework. We present a taxonomy of the variables and discusses methodological issues that arise from it. We believe that the reflection on the nature of the variables since the definition of the study objectives, and ahead of linking them with the empirical work, is a task that even should be part of the formulation of theories, hypothesis or ideas involved in the empirical economic research.

* Profesor del Departamento de Economía, PUCP.

** Profesor del Departamento de Ciencias, PUCP.

EL TRATAMIENTO DE LAS VARIABLES EN ECONOMÍA: Una revisión y un marco conceptual¹

Oscar Millones
Jorge Bazán

1. INTRODUCCIÓN

La disponibilidad de más enfoques en la investigación (metodología cualitativa, metodología mixta), así como en métodos, técnicas, modelos y aplicaciones econométricas (por ejemplo. microeconometría, macroeconometría, econometría bayesiana), revelan un nuevo problema relacionado con la eficiencia en la utilización de recursos para el aprendizaje y la investigación en economía. Es decir, contamos cada vez con más literatura de gran implicancia metodológica empírica, pero disponemos de pocas guías orientadoras para optimizar su uso.

Se hace necesario desarrollar estrategias que ordenen “el panorama” generados por los temas mencionados, sobretudo para el estudiante del pregrado de economía que enfrenta la tarea de culminar sus estudios con una primera investigación. Este no es solo un problema local. Por ejemplo, Gagliardini (2006), menciona una creciente brecha entre los estudios de bachillerato y la investigación en general en la Economía y, Sevestre (2006), refiriéndose a la econometría en una perspectiva más amplia, señala la necesidad de una metodología para el uso de las técnicas econométricas así como comprender la lógica de los modelos econométricos y su lugar en la metodología de las ciencias.

Esta “metodología para uso de las técnicas” es un tema clave para la investigación económica y, sin embargo, la literatura econométrica, en

¹ Este documento refleja la interacción con una variedad de colegas, investigadores y estudiantes del Área de Economía, así como de los miembros que participan en la línea de investigación “Aspectos Metodológicos en la Investigación Económica y Social” (AMIES) de la Universidad. Agradecemos al Dr. Máximo Vega-Centeno, Profesor Emérito del Departamento de Economía por la revisión de la versión previa del documento y por las sugerencias hechas para la mejora del mismo. Tanto las deficiencias y errores remanentes en el documento son de responsabilidad de los autores.

general, la toma por dada. En los pocos casos en que el tema es abordado, a nuestro criterio, éstos son desarrollos parciales, incompletos o no están suficientemente difundidos. Aportes valiosos como las contribuciones de Zellner (1971), Leamer (1978), Malinvaud (1983), Hendry y Morgan (1995), artículos de la revista *Journal of Economic Methodology* y otros, aparentemente están aislados, referidos a líneas de aspectos específicos y necesitan ser integrados y complementados considerando los más recientes desarrollos en la econometría y de la metodología de la investigación.

Podemos definir esta estrategia o "metodología para uso de las técnicas" como un conjunto de conocimientos y habilidades que nos permiten pasar en forma lógica del planteamiento objeto de la investigación económica, al uso adecuado de métodos econométricos e información disponible, de tal manera que brinden aprendizajes significativos sobre los objetivos de la investigación. Esta tarea es amplia y es necesario abordarla sistemáticamente en foros y discusiones. En este trabajo tratamos un aspecto específico, que es la naturaleza de las variables y su uso en economía.

El objetivo de esta revisión es presentar una síntesis de cómo se ha manejado el concepto de las variables en la investigación en Economía, incorporando también algunas ideas desde otras áreas cuantitativas, y ordenando estos manejos en forma de un marco de análisis que llamamos marco conceptual. Consideramos que el reflexionar sobre la naturaleza de las variables desde el planteamiento de los objetivos de la investigación, y antes de vincularlas con el trabajo empírico, es una tarea que incluso debe ser parte de la formulación misma de teorías, hipótesis o ideas involucradas en el proceso de la investigación económica.

En la sección 2 consideramos la variedad de información y la necesidad de un marco conceptual para las variables. Luego, en la sección 3, se definen dimensiones básicas, especie de taxonomía de las variables. En la sección 4 a manera de discusión presentamos algunos comentarios implicados por la taxonomía de variables. Finalmente, a manera de

conclusión, en la sección 5, resumimos los aspectos más importantes del trabajo.

2. NECESIDAD DE UN MARCO CONCEPTUAL

2.1 Variedad de información y métodos

Con el progreso informático y estadístico, nos encontramos ante una variedad de información, sean estos microdatos o estadísticas agregadas, que se traducen en disponibilidad de bases de datos y de instrumentos o técnicas para su producción. Ejemplos son, las estadísticas publicadas por el Instituto Nacional de Estadística e Informática, INEI (Censos Nacionales, Encuesta Nacional a Hogares -ENAHO, encuestas continuas, cuentas nacionales y otros), las series publicadas por el Banco Central de Reserva (boletín semanal, reportes de inflación, de operaciones monetarias, proyecciones macroeconómicas, series mensuales), la información generada por los organismos públicos de los sectores productivos (estadísticas de producción mensual o anual), y de servicios (atención ambulatoria de salud o información sobre rendimiento educativo) y organismos reguladores (casos y litigios en imperfecciones de mercado -INDECOPI), sin mencionar fuentes privadas generadas de las investigaciones e instituciones sin fines de lucro y otras fuentes internacionales.

La investigación económica hace uso de esta información y produce estudios usando una variedad de técnicas econométricas que van en paralelo con el desarrollo de nuevos métodos y procedimientos que se ilustran en revistas especializadas sobre teoría, métodos y aplicaciones.

Por un lado, el desarrollo de técnicas econométricas, incluyendo el trabajo formativo, tiene ya casi un siglo de existencia (ver antecedentes en Hendry y Morgan, 1995, Christ 1994, Duo, 1997). Por otro lado, los textos formales econométricos, con esquemas ya definidos, solo aparecieron desde los inicios de los 60's (después de las contribuciones probabilísticas en la econometría de los 50's), con las primeras ediciones de textos como los de

de Lawrence Klein, Oskar Lange, Jack Johnston, entre otros, marcando un primer dinamismo en la proliferación de estos métodos.

En las últimas dos décadas, han aparecido en forma creciente nuevos temas en la econometría, tanto propios como incorporaciones y actualizaciones metodológicas provenientes de otras áreas cuantitativas. Por ejemplo, en el prefacio de la cuarta edición del texto de Johnston y DiNardo, 1997, se resumen seis importantes innovaciones aparecidas solo en esa década. Otros ejemplos de temas emergentes, puede encontrarse en el índice de textos como el de Greene (2008), y, sobretodo, también en las ediciones de las revistas de econometría².

Este desarrollo vertiginoso demanda urgentemente el delineamiento de una estrategia para su efectiva incorporación en la investigación empírica y en especial para la enseñanza econométrica. Con este rápido progreso se está haciendo ya inapropiado el mantenimiento de esquemas tradicionales tanto de los textos econométricos actuales como los esquemas para hacer frente a las desafiantes variedades de enfoques metodológicos que emergen.

Un tema importante en este desarrollo gira alrededor de la definición y de la naturaleza de las variables. El resaltar esta importancia es uno de los objetivos de esta revisión. Si además, un propósito en la investigación es el contraste empírico de planteamientos propuestos (la inferencia estadística en el sentido de Jeffreys 1957, 1966), el estudio de las variables es también indispensable para elaborar una estrategia en la elección del camino de la contrastación.

2.2 Variables y Medición

¿Qué es una variable? Daremos una definición heurística. Variable es una característica, factor o atributo de una unidad de análisis (individuos, hogares, firmas, objetos o sistemas) susceptible de ser medido o conceptualizado. También su mismo nombre lo define: es algo que puede esperarse varíe con el tiempo o entre unidades de análisis.

² Ver <http://www.feweb.vu.nl/econometriclinks/journals/#journals> por ejemplo

En este contexto, ¿qué es medir? Es asignar números o valores a las unidades de análisis para representar ciertas propiedades³. A manera de ejemplo, considérese un tema de investigación donde aparecen diferentes variables.

Ejemplo 1. Se plantea un estudio para entender por qué los estudiantes del bachillerato están demorando más de lo esperado en terminar su tesis. Algunos datos recogidos son el número de años dedicados a la especialidad, el promedio de rendimiento obtenido en la carrera, la pertenencia o no al tercio superior, la escala de pensiones de los estudiantes, el hecho de tener empleo o no, percepción sobre la calidad de la enseñanza medida con un test, el tipo de seminario de tesis llevado en el pregrado, etc.

Si se revisa las variables del ejemplo se puede establecer que no todas ellas son del mismo tipo ni tienen igual jerarquía. Para caracterizar estas peculiaridades es necesario contar con un marco para entender mejor la naturaleza de estas variables, lo que se presenta a continuación.

Un Marco para el análisis de las variables

En Economía, los 4 criterios más aceptados para comprender la naturaleza de las variables, son:

- Criterio A: El nivel o Escala de Medición de las variables
- Criterio B: La Naturaleza de la Construcción de las variables
- Criterio C: El Rol de las variables en el tema en estudio
- Criterio D: El Diseño Muestral para las variables

El primer criterio, el nivel o escala de medición, se refiere a un conjunto de propiedades que distinguen las variables medidas de acuerdo a las operaciones matemáticas que pueden ser permitidas con ellas. La naturaleza de la construcción de las variables se refiere al hecho de que las variables consideradas son o no directamente observables y por tanto susceptibles de ser medidas con los datos directamente recolectados o

³ Por los objetivos de este trabajo, aquí no será necesario hacer referencia a la definición axiomática de números en la medición (ver J. Pfanzagl, 1971). *Theory of Measurement*. Physica Verlag, Würzburg, 2nd revised edition).

derivados o inferidos a partir de estos. El rol de las variables, lo distinguimos porque nos permite apreciar las variables de acuerdo a su rol funcional en el estudio. Finalmente, el efecto del diseño muestral sobre la medición de las variables se refiere a si las variables planificadas para ser usadas finalmente son obtenidas en la forma muestral esperada por el investigador en cuanto a su dimensiones en el tiempo, espacio y unidades de análisis.

Aquí es preciso establecer (ver sección 4) que estos criterios no son excluyentes o autónomos, por el contrario, es posible caracterizar una variable, combinando simultáneamente algunos o todos estos criterios con lo cual emergen muchas posibilidades a partir de ellos⁴. En las siguientes dos secciones discutiremos, con mayor detalle, los cuatro aspectos mencionados⁵.

3. CRITERIOS PARA CLASIFICAR LAS VARIABLES

3.1 Criterio A: El nivel o Escala de Medición de las variables

El tema de las escalas está tomando importancia en la literatura econométrica (ver por ejemplo el orden con que se presentan los temas en Scott y Freese 2001 o las escalas definidas en Joreskog 2001). Siguiendo el ejemplo 1, podemos notar que los valores de las variables consideradas no son de la misma naturaleza, como es el caso del promedio de rendimiento y el tipo de seminario de tesis. Según Stevens (1946), existen cuatro tipos de escalas o tipos de medición en las variables (o maneras de interpretar estos números), que son la escala Nominal, Ordinal, de Intervalo y de Ratio. La columna (3) de la Tabla 1 ilustra estos tipos de medición⁶.

⁴ Por otro lado, existen otras caracterizaciones que no se incluyen en este trabajo y son solo referidas mas abajo.

⁵ Aunque el orden de discusión de estos criterios no es importante se presentan en este orden por la aparente complejidad de los mismos.

⁶ Los lectores familiarizados con esta clasificación pueden continuar con la lectura del criterio (B).

Tabla 1: Propiedades y Ejemplos de Escalas de Medición de las variables

| Escala (1) | Propiedad (2) | | | Ejemplo 1 (3) |
|------------------|------------------|-----------|--------|---|
| | Orden | Distancia | Origen | |
| NOMINAL | NO | NO | NO | Tipo de seminario llevado en el pre-grado |
| ORDINAL | SI | NO | NO | Escala de pensiones del alumno |
| INTERVALO | SI | SI | NO | Calidad de la enseñanza |
| RATIO | SI | SI | SI | Años en la especialidad |

Propiedades operacionales en la medición

El tipo de escala de medición (columna 1) se determina por las propiedades operacionales (columna 2) que tienen los números asignados en la medición. Estas propiedades son: orden, distancia y origen.

- Orden: al medir una característica en un sujeto, asignados dos números, uno de ellos es igual, mayor o menor que el otro.
- Distancia: dado dos pares de números asignados en la medición, la diferencia entre ellos puede ser igual, mayor o menor que otras diferencias.
- Origen: en la medición de una variable, el cero tiene el significado cuantitativo de "cantidad nula" de la propiedad medida.

Escalas resultantes

Cuando aplicamos números a objetos (medimos) es importante saber si el número asignado cumple o no con estas tres propiedades (ver columna 2). Como se observa en cada fila, las escalas de medición se definen de acuerdo a si la variable, cumple o no con estas propiedades. Estas escalas resultantes entonces son Nominal, Ordinal, Intervalo y Ratio o razón⁷.

⁷ Para las propiedades formales puede revisarse Valdivieso (1991) y para las implicaciones prácticas en Economía el Anexo 1.

3.2 Criterio B: La Naturaleza de la Construcción de las variables

Desde el punto de vista de cómo se obtienen (construyen) los valores para las variables, éstas pueden ser observables (medibles a través de los datos) o constructos (derivados o inferidos de los datos).

Variables Observables

Las variables son observables si el valor obtenido o dato es una buena aproximación (o coincide exactamente) al concepto que se quiere medir. En palabras de Torgerson, hay una relación cercana, si no directa, entre el concepto teórico abstracto y su operacionalización en el mundo real (“the nature”, ver Torgerson 1958). Son ejemplos de observables, muchas de las que encontramos en la Encuesta Nacional de Hogares, allí tenemos Ingresos por trabajo, Pagos mensuales por crédito, Tipo de Ocupación, Tenencia de celular, Estado Civil, etc. En otras estadísticas oficiales podemos obtener variables como Compra de dólares del Tesoro Público, Exportaciones (en valores o cantidades), Tipo de cambio intercambiario promedio, etc.

Los valores de estas variables (los datos) se obtienen directamente en el sentido de que su elaboración o construcción no crea mayor problema y es relativamente fácil de obtener (aún cuando su determinación sea algo compleja como el ingreso). La facilidad de su obtención se debe a que no hay diferencia entre lo que se quiere medir (el concepto a medir, edad, por ejemplo) y lo que se obtiene como medición (información en el DNI, edad en una encuesta, por ejemplo).

Es cierto que a menudo hay problemas de medición de variables observables, y la exactitud de su medición depende del contexto de su uso. Por ejemplo, en algunos trabajos, si hablamos del ingreso personal o familiar, es probable que nos refiramos al ingreso monetario (no en especie) de la ocupación principal, obtenido en un tipo de ocupación (actividad principal) y en un periodo dado (mes). Pero cualquiera que sea la definición que finalmente escojamos, es posible obtener una medición satisfactoria de lo que se quiere obtener a través de los datos recogidos. El problema de los errores de medición se discuten en el Anexo 2.

Las variables observables, sin embargo, no constituyen el único tipo de variable para el cual prestamos atención o para el que desarrollamos una hipótesis o teoría. También estamos interesados en otro tipo de variables (constructos) que por su naturaleza no podemos (aunque quisiéramos) obtenerla directamente con medición inmediata, con los datos recolectados.

Variables Constructos⁸

Las variables constructos o latentes son las que se miden no directamente de los datos recolectados. Estas variables latentes se pueden medir a través de sus indicadores. Algunos ejemplos son:

Pobreza, variable que, dada su definición por el investigador, puede ser aproximada por algunos indicadores -observables como el Ingreso Familiar, Acceso a servicios básicos, Grado de Malnutrición, etc., pero como constructo, es diferente de estos indicadores.

Desempeño del estudiante en econometría, Capacidad de análisis económico (a los que podríamos aproximarnos por las notas en sus exámenes, en prácticas calificadas, etc.). Otros ejemplos son Tecnología, Productividad, Expectativas de inflación, Utilidad, Ingreso Permanente, Competitividad, Barreras arancelarias, Intangibles, etc.

Los constructos y el término "no observabilidad"

Se evitará aquí el uso del concepto de variables no-observables. Un problema con la etiqueta "no observables" es que sobre éste concepto se tienen diversas interpretaciones en su uso corriente en la literatura econométrica (ver Anexo 2). Ninguna de éstas otras interpretaciones sobre no observabilidad, del anexo 2, concuerda con la definición que se da en este documento sobre los constructos. La definición de constructo dada aquí es simple y tal como se ha definido arriba. Esta definición es clásica y se estableció originalmente en otras áreas fuera de la Economía. Concuerda con el concepto de "constructo" en Sociología. Concuerda también con su uso

⁸ Se evitará usar el nombre de "No-observable" aquí por las razones explicadas en el Anexo 2.

en Psicología, donde se le ha denominado variable latente⁹. Es decir, en un constructo o en una variable latente, el concepto se conoce pero es imposible medirla directamente de los datos porque carece (tal vez, por ahora) de contrapartida empírica, u operacionalidad empírica. Aún cuando existan los recursos (y las intenciones) para medirla, los constructos solo pueden ser derivados o inferidos a través de indicadores, o de las propiedades medibles que envuelve el concepto.

3.3 Criterio C: El Rol de las variables en el tema de estudio

Para el estudiante de bachillerato de Economía el concepto de rol debe ser familiar, por lo que pasaremos rápidamente esta sección. Estaremos obviando aquí otros temas relacionados al rol, por ejemplo, significado del término "variable independiente" en el marco de regresión (ver Koopmas 1937) o en el marco de ecuaciones simultáneas, desarrollados por la Comisión Cowles¹⁰ así como los debates sobre el tema de exogeneidad (causalidad a la Granger) de enfoques econométricos posteriores (ver Enders 2004, sección 5.8). Para una revisión reciente, aunque breve, del concepto relacionado al rol -"causalidad", aplicado en varias áreas que incluyen biometría, demografía, sociología, ciencias naturales y otros, ver Cox (2007).

Desde nuestras primeras clases de Macroeconomía o Microeconomía, nos hemos encontrado con especificaciones tales como la función consumo, inversión, o ecuaciones de demanda. El rol de las variables es el papel que juega cada variable en cada una de estas especificaciones, de allí que se prefiere hablar de rol funcional, identificando variables dependientes e independientes.

Variables dependientes e independientes

Usemos por ahora las definiciones de la Comisión Cowles. Una variable es endógena (dependiente) si es explicada por otra(s). Y es exógena

⁹ Para mayor discusión revisar Borsboom, Mellenbergh y van Heerden (2003).

¹⁰ Los integrantes de la Cowles Commission creada por la década de los 30's, hicieron grandes contribuciones al desarrollo de la Econometría ("probability approach"). Los revisores de historia econométrica los identifican como los forjadores de los métodos para los Sistemas de Ecuaciones Estructurales ver por ejemplo Malinvaud (1983).

(independiente) si se incluye en la especificación del modelo pero no es explicada por otras¹¹. El siguiente es un ejemplo en el contexto uniecuacional.

Ejemplo 2. Se trata de indagar los determinantes del rendimiento educativo en Matemáticas para el 6to grado de Primaria en una escuela urbana. Se puede decir que asociados a este rendimiento hay factores o variables del estudiante (horas por semana dedicadas al estudio, ingreso familiar), del profesor (calificación, experiencia) o de la escuela (material educativo, recursos informáticos). Para este ejemplo el rendimiento en matemáticas es la variable endógena y las otras son exógenas.

Las variables dependientes en un estudio pueden ser independientes en otros estudios, y viceversa. Igualmente entre las variables independientes hay grados de jerarquías que han sido dadas en la literatura. Por ejemplo en las variables del estudiante, horas de estudio cumplen un rol directo en el tema, mientras que otras variables (tratamientos o factores) sociodemográficas como género o estatus económico pueden cumplir un rol de control en el análisis.

Una pregunta importante sobre el rol de las variables es si siempre hay rol para las variables o no. Esto se retomará en la discusión de la Sección 4.

3.4 Criterio D: El Diseño Muestral para las variables

Este criterio, a diferencia de los anteriores, se refiere más a los sujetos sobre quienes se miden las variables, que a las variables mismas. Por ahora solo mencionaremos los aspectos más conocidos en la literatura econométrica.

Frecuentemente en Economía la información ya está recogida y es relevante ver si el diseño empleado en la selección de las unidades de estudio sugiere algún tipo de análisis especial. En otros casos el investigador puede producir la información y dependiendo del objetivo de la investigación

¹¹ En la literatura se encuentra denominaciones como variables explicadas y explicativas; regresando y regresores, y otras.

y los recursos disponibles, se decide una manera de levantar la información y el tipo de análisis a efectuar. En ambos casos, conviene caracterizar a los datos que generan las variables (cualquiera sea la naturaleza de su construcción, escala de medición o su rol) respecto a la unidad de análisis (i), a la posible jerarquía identificada sobre las unidades de análisis (j) y a la ocasión utilizada en la producción de sus valores (t). En cada caso se han sugerido en la literatura una familia de modelos econométricos. Esto se puede sintetizar en la notación

$$Y_{i,j}(t)$$

Por ejemplo en un estudio sobre rendimiento escolar para el año 2004, en el que se recoge información sobre n estudiantes, y se incluye aspectos de m escuelas y también de las direcciones regionales educativas, si el estudio enfoca dos niveles de análisis, las unidades de análisis de primer nivel (alumnos) van de $i=1,2, 3, \dots n$; y las unidades de análisis de segundo nivel (escuelas) van de $j=1,2,3,\dots m$. Como éste es un estudio en 2004, $t = 1$, entonces $Y_{i,j}(t)$ es simplemente $Y_{i,j}$. Es decir, se trata que las variables se definen en corte transversal, cuya información se analizará a dos niveles jerárquicos (alumnos, y escuelas). Si se incluye otro nivel jerárquico, por ejemplo, las direcciones regionales, entonces tendríamos $Y_{i,j,k}(t)$

Las series temporales, a su vez, según la ocasión en que se las obtienen, pueden ser series de tiempo, es decir, medidas con intervalos comunes de tiempo (ver Enders 2004) o puede ser medidas en diversas ocasiones con longitud de intervalos de tiempo que son discrecionales al investigador (ver por ejemplo Bock 1975, Cap 7). Los estudios en el cual el intervalo de tiempo de recolección es variable, incluyen por ejemplo, estados de seguimiento de intervenciones sociales o en tratamientos específicos en salud (chequeos, dosis de tratamientos, vacunaciones) o educación (efectos de nuevos métodos de enseñanza, seguimiento de evaluación de estudiantes, etc), y desde el punto de vista más macroeconómico el estudio de ciclos económicos. En estos casos se definen las variables como medidas repetidas.

Un caso mixto lo constituyen los modelos de panel $Y_i(t)$ (ver por ejemplo Goldstein 2003) y otros modelos dinámicos y no paramétricos (ver Arellano 2003; Tong 1990; Anderson y Hsiao 1983).

Respecto a las variables en niveles (índices i, j en $Y_{ij}(t)$), recientemente se están incorporando familias de modelos a cada nivel de análisis pero simultáneamente, (ver West 2007). Una aplicación en los estudios de rendimiento educativo son los elaborados por la Unidad de Medición de la Calidad educativa del Ministerio de Educación, ver por ejemplo UMC (2001).

Adicionalmente para cada nivel de análisis jerárquico es posible definir especificaciones en términos de los niveles de variables (valores esperados) o en términos de otros momentos de la distribución de las variables.(por ejemplo, varianzas). Así se pueden generar especificaciones que incluyen, modelos de heteroscedasticidad condicional (GARCH, EGARCH, etc. ver Engel 1982; Alexander 2001), estado-espacio y otros (ver Fan y Yao 2005; Fan, Yao. y Cai 2003).

Igualmente si se quiere modelar distribuciones (por ejemplo distribuciones conjuntas), en un solo nivel jerárquico, existen modelos conocidos como Count Models (ver Cameron y Trivedi 1998), modelos de duración o hazards models (Goldstein 2003) o en el caso de variables independientes nominales y/o ordinales se sugieren una familia de modelos de asociación y de correlación (ver Fienberg y Meyer 1983; Goodman 1978, 1991).

3.5 Otras Taxonomías para Variables

Otros criterios de clasificación de variables que pueden ser mencionados son:

- Por la forma de distribución de las variables, podemos considerar variables con distribución paramétrica y variables con distribución no-paramétrica, o variables semi-paramétricas.

- Por la forma cómo se concibe un número de variables en un estudio - por ejemplo univariada, multivariadas, conjuntas, condicionales.
- Por el estado de su manipulación, variables primitivas y derivadas.
- De acuerdo al dominio en el que son definidas sus valores, por ejemplo, variables positivas, variables acotadas en un intervalo. Sobre este punto ver el Anexo 3.

4. DISCUSIÓN

4.1 Escalas de medición

Siempre ha habido controversia en definir una taxonomía apropiada para las variables en cuanto a su escala de medición. La escala de Stevens de la sección 3, fue una de las primeras propuestas, y ha sido criticada por algunos estadísticos, ver por ejemplo Velleman and Wilkinson (1993)¹². Sin embargo, para efectos de diseñar una estrategia de investigación en Economía, consideramos que la escala de Stevens es apropiada por ser simple (parsimónica), fácil de reconocer y suficiente para levantar temas metodológicos que igualmente se pueden levantar con otros esquemas.

¿Es mejor la variable de ratio?

Se ha comentado en la literatura cuantitativa que las variables medidas a nivel de ratio son "mejores" que las otras. Esta afirmación solo responde al aspecto del uso operativo de la información, cual es, lo que podemos hacer (en el sentido de operaciones matemáticas y estadísticas) con los números obtenidos en las diferentes escalas.

Las escalas de medición de Stevens sirvieron en un primer momento para llamar la atención de que en el trabajo estadístico, que implica operaciones con la información, no se incurriera en inconsistencias en el manejo de las variables. Así, las escalas de medición, al describir la naturaleza de la información en cuanto a las propiedades de los números que han sido

¹² En psicología, se sigue trabajando sobre la consistencia empírica de la escala de ratio y también su formalización desde un punto de vista axiomático, por ejemplo, ver el tema Psychophysical Scaling, en Gescheider (1997), Koornstra (2006), Thomas (2006), entre otros.

asignados a objetos (ver anexo 1), nos dicen las operaciones matemáticas posibles o permitidas con estas variables.

Atributos de los datos y revestimiento

El beneficio del conocimiento de las escalas de medición, en cuanto a su operacionalidad es, sin embargo, solo secundario. Para los objetivos de este documento las escalas de medición (y los otros criterios de la Sección 3), tienen el papel de guiar al investigador en la estrategia cuantitativa a embarcarse. La utilidad operacional matemática, sin embargo, nos ayuda solo al aspecto operativo del trabajo empírico, utilidad operativa que se relaciona más con el análisis unidimensional o cuando se trabaja con dos o pocas variables casi homogéneas.

El trabajo empírico en Economía, usualmente es complejo, multivariado o multidimensional o la combinación de todo esto. Por lo que el criterio operacional de la escala de medición es manejado también de igual forma en el marco de los modelos econométricos. Esto es abordado en la literatura con modelos de cierta complejidad que están siendo popularizados en textos como el de Greene (2008) o Scott y Freese (2001).

Por esta razón declarar un ranking de importancia entre las distintas escalas de medición (por ejemplo decir que es mejor tener variables de ratio que nominales) es en general arbitrario en el marco de los recientes esquemas cuantitativos. Tal vez esto era relevante en un tiempo de poca disponibilidad de modelos y técnicas cuantitativas, pero esto ya no viene siendo una restricción. Así pues ante la nueva variedad de modelos estadísticos disponibles podemos ahora ver las escalas de medición no como atributos fundamentales de los datos, sino solo como una especie de revestimiento que tiene la información (ver Velleman and Wilkinson, 1993). Este revestimiento, así como la naturaleza en cuanto a su rol, su construcción y el diseño muestral, son ahora igualmente importantes para ayudar a clarificar la estrategia del pase desde los objetivos de la investigación, hasta el manejo de los datos.

4.2 Los Constructos

Constructos en otras áreas

En otras áreas como en Sociología y Psicología, el tratamiento de las variables constructo ha sido metodológicamente más sistemático y explícito que en el caso de Economía.

En Sociología se ha experimentado una variedad de opciones metodológicas en las últimas décadas, especialmente en el área de investigación cualitativa (por ejemplo, solo para revisar una opción ver Strauss y Corbin 1990). Estas las contribuciones empíricas han estado muy atadas al desarrollo teórico. Por ejemplo, en un balance metodológico, Parsons señalaba como uno de los grandes retos para la Sociología de esos años, el poder desarrollar teorías que puedan ser sujetas a contrastación empírica tanto cuando se emplean variables "más observables" como cuando se trate de variables "fundamentales" donde la contrapartida empírica, es todavía desconocida (Parsons 1954, pag. 14-16).

En este contexto, y en una posición menos positivista de la que tuvo en 1937, en *The structure of Social Action*, Parson define tres tipos o niveles de aproximaciones metodológicas de acuerdo al nivel bajo, medio o alto en el que las categorías empíricas juegan o no un papel crucial en la teoría. Es interesante notar que en la alta pone como prototipo al trabajo en investigación histórica de Weber en Sinología e Indología, usando categorías teóricas como "Tradicionalismo" y en el que según su evaluación, se logra exitosamente aplicar el método comparativo en gran escala. Y señala en el nivel intermedio, al trabajo de Durkheim en el tema del suicidio (tasas) usando combinaciones satisfactorias de datos, con el uso de lo que entonces denominó modernos métodos estadísticos. Lazarsfeld ha sido también seguidor de esta última línea donde las variables fundamentales eran concebidas como clases (¿escalas nominales?), ver por ejemplo Lazarsfeld (1950). El nivel bajo ubica desarrollos teóricos que se basan en procedimientos operacionales meticulosos ("meticulous" ideal of operational procedure).

El desarrollo de la metodología empírica en Sociología ha sido variado, y que, poniéndolo en términos simples, va desde una posición interpretativista a una positivista sugiriendo estrategias diversas para el trabajo de campo. Para una revisión de esta gama, en la parte cualitativa, ver Travers 2001.

Por otro lado, desde una perspectiva cuantitativa, la Psicología es un área en que probablemente casi todas las variables de interés son latentes. Un ejemplo clásico ha sido el tema de la Inteligencia (su naturaleza, medición, desarrollo), que ha llenado debates y libros en la literatura en Psicología, ver por ejemplo los clásicos Spearman (1927), Thurstone (1938), Guilford (1977) por mencionar solo algunos. La variable inteligencia es, bajo cierta perspectiva, un conglomerado de otras variables observables. La discusión sobre su definición, desarrollo y posibilidad de medir es aún tema de interés, ver Sternberg (2003)¹³.

La Psicología ha brindado marcos empíricos que han sido muy útiles en otras disciplinas. Ejemplos son métodos de reducción de información, tales como Componentes Principales, Análisis Factorial en finanzas y recientemente en series de tiempo; la evaluación psicométrica de instrumentos de medición en propuestas como la Teoría de Respuesta al Item en educación; el tratamiento de variables categóricas con escalamiento multidimensional en áreas de la sociología, filosofía y política.

Constructos en Economía

El tratamiento en Economía de las ideas que solo se plasman en conceptos como constructos han jugado un papel clave en el desarrollo de la Teoría Económica desde sus comienzos. Por ejemplo en el tratamiento del concepto de utilidad o bienestar y en otros conceptos en la teoría de la demanda y oferta y en particular en el concepto de escasez. Las implicancias empíricas de distinguir entre constructos y observables ha sido un tema clásico en econometría desde los escritos de Wright (1929) y Robbins (1932), (en particular el problema de distinguir empíricamente entre la demanda y

¹³ Es interesante notar que hay otro grupo de variables latentes que han sido consideradas de mayor dificultad respecto a su medición como las Actitudes, Valores, Ansiedad (nivel de), Depresión, Autoritarismo, Modernismo, Liberalismo, etc.

cantidad tranzada). Sin embargo su reconocimiento formal como tal (variable latente), ha sido explícito solo desde los 70´s u 80´s¹⁴.

Más recientemente en la literatura económica, por ejemplo en temas de organización Industrial, se definen categorías latentes, como contexto político, controles verticales, ventajas y desventajas del costo de aprendizaje, cambios institucionales, flexibilidad del sistema regulatorio, barreras comerciales y otros, como conceptos de gran interés.

En un contexto más general y en el marco del desarrollo de teorías, Blalock ha mencionado que “ para aplicar (contrastar) dichas teorías con datos empíricos se debe, de alguna manera, enlazar los conceptos más abstractos con varios indicadores u operadores de investigación” (Blalock 1984, pag. 14). Evidentemente detrás de este deseo están los temas clásicos sobre identificación discutidos en la literatura econométrica desde la época de los forjadores como Phillips G. Wright y posteriormente por Jean Tinbergen entre otros, (ver excelente recopilación en Hendry y Morgan, 1995).

Podríamos concluir que, en cuanto a la construcción de las variables se refiere, en Economía es frecuente el uso de constructos toda vez que la postulación de teorías en su forma general, exige el uso de conceptos abstractos en la definición de las variables.

4.3 El rol de las variables

¿Siempre hay rol para las variables?

Si la respuesta fuese afirmativa, minimizaríamos la abundante y útil literatura empírica sobre análisis exploratorio. Dicho más generalmente, tal vez subestimaríamos la variedad de paradigmas metodológicos existentes para la investigación en Economía y en Ciencias Sociales. No abriremos este debate aquí pero solo resumiremos algunas ideas ampliamente aceptadas.

¹⁴ La literatura econométrica ya ha puesto explícito el tratamiento de variables latentes desde fines de los 70´s por Joreskog y Goldeberger —ver Aigner y Goldberger (eds., 1977)— en el marco de modelos estructurales lineales (LISREL). Esto ha sido complementado a nivel microeconométrico con los trabajos de Heckman y MacFadden en los modelos de elección discreta.

Las técnicas de investigación se pueden clasificar como técnicas exploratorias y confirmatorias. Las técnicas exploratorias se usan frecuentemente cuando poco se sabe sobre un tema de interés, especialmente en los inicios de un tema de investigación. Se puede identificar a las técnicas exploratorias como técnicas descriptivas que pueden aplicarse a poblaciones o muestras. Por lo que el tema de la inferencia no necesariamente es importante. Por el contrario las técnicas confirmatorias implican dominio del tema a tal punto que el primer paso del análisis confirmatorio es la presentación formal de hipótesis o teorías, las que van a ser sometidas a pruebas o contrastes en el trabajo empírico. En este sentido el concepto de muestra y población es relevante.

Hay una variedad de técnicas exploratorias que incluyen por ejemplo desde el análisis exploratorio mismo (Tukey, 1977) y (Benzécri et al., 1973; Nishisato, 1994), pasando por las técnicas de reducción de información (Hotelling, 1933, Anderson, 1958), funcional data análisis (Ramsay y Silverman 1997) hasta los llamados ejercicios de Data Mining (Hastie et al. 2001). Esta última técnica se caracteriza por asociarse con recolección masiva de datos, el uso intensivo de procesadores (multiprocesadores) e incluir algoritmos en sus procedimientos.

La búsqueda de especificación de modelos puede hacerse incluso con la misma información, algunas veces con técnicas de remuestreo –data que luego se usará para futuras contrastaciones empíricas más confirmatorias, ver por ejemplo Leamer (1978). El trabajo exploratorio y confirmatorio también pueden combinarse en un mismo estudio en forma secuencial como en Calvo et al. (1993).

En el análisis exploratorio no hay rol para las variables, al menos al inicio de la investigación. Esta característica de no rol para las variables es propicia para metodologías de investigación que son más cualitativas o menos positivistas en el lenguaje epistemológico (ver por ejemplo Greenacre y Blasius, 1983). Más recientemente se están considerando en la literatura metodologías mixtas de investigación (ver Tashakkori y Teddlie, 1998).

Una reconsideración para el Rol

Es interesante observar cómo el énfasis sobre el rol en el tratamiento de variables en Economía se ha movido como un péndulo. Esto se ve más claro en series de tiempo. Observamos el estudio de variables con rol -Real Business Cycle (ver los Barómetros de Harvard de los 20 y 30's en Lange 1964) que luego fueron calificados posteriormente como estudios de no rol (en la práctica los métodos del National Bureau of Economic Research fueron calificados por Koopmans a fines de los 40's como estudios de "medida sin teoría", ver Malinvaud, 1983). Luego hubo un tránsito al desarrollo con fuerte componente teórico (con rol) desde los 50's con el pensamiento - "probability approach"-de la Cowless Commission, pasando luego en los 70's por tratamientos más exploratorios, ver por ejemplo los trabajos iniciales de Box y Jenkins (1974) excluyendo las funciones de transferencias que solo fueron referidas entonces para aplicaciones en física (Cap 10), para volver últimamente a enfoques que son menos exploratorios. Dos ejemplos de esto último lo constituyen el concepto de cointegración (ver test en Engle & Granger, 1987) y la literatura reciente referida por Stock y Watson (2005, 2006) en Factor Dynamic Models.

Desde el llamado nacimiento de la econometría (que incluyen los desarrollos incipientes) hasta nuestros días se ha entablado una disputa, aparentemente innecesaria, entre estrategias exploratorias y confirmatorias. Por ejemplo, en el análisis de series de tiempo, se puede interpretar el "conflicto" de la alternativa de usar VAR no restringido versus Ecuaciones Estructurales como una disyuntiva de este tipo.

Sobre este aspecto, sin embargo, se está reconociendo (aunque implícitamente) en la literatura econométrica que las estrategias exploratorias vs confirmatorias son complementarias. En desarrollos recientes se está incorporando técnicas más exploratorias para combinarse en el trabajo econométrico. Ver por ejemplo Calvo et al. (1993), sección III, y la reciente estrategia para levantar indicadores para la economía americana en Stock (2006) que fue sugerida inicialmente por Geweke (1977) y después por Diebold y Rudebusch, (1996).

Este es un avance metodológico interesante en Economía, en parte debido tal vez a la influencia metodológica de otras áreas de las Ciencias Sociales (notablemente desde Sociología y Antropología) en donde no siempre el investigador plantea su trabajo con hipótesis previas, ver Travers (2001).

4.4. El diseño muestral

La Dualidad de las observaciones

En conexión con el diseño muestral de las variables hay algunos aspectos que no han sido incorporados sistemáticamente en la literatura econométrica. Se trata del carácter dual de los datos. Por dualidad de los datos se entiende aquí la peculiaridad de analizar los datos por el lado de las variables y también por el lado de los sujetos. Para visualizar mejor esta idea conviene considerar el aspecto físico de las variables, es decir, el tener los datos de las variables en un archivo de filas y columnas, como es usual, donde las columnas son las variables y las filas los sujetos.

Si miramos a los sujetos (filas), como especie de "variables transversales" podríamos interpretar los datos en filas como "variables especiales" generadas por el lado de los sujetos. Esta dualidad de los datos ha sido destacada en la literatura cuantitativa por Nishisato (1980, 1994) sintetizando, a su vez, 50 años de desarrollos en el análisis multivariado.¹⁵

La discusión de la naturaleza de los patrones que se pueden identificar desde las filas (sujetos) ha sido tratada en la literatura econométrica en forma parcial. Considerar la fila como un vector de observaciones implica tomar en cuenta parte o todas las variables (columnas) en la heterogeneidad presentada por los criterios A, B y C. Esto implica considerar métodos de manejo de variables que sean flexibles y que enfrenten a todo tipo de variables. Para algunos casos, por ejemplo si todas las variables son de ratio o de nivel intervalo, métodos como componentes principales han sido utilizados exitosamente en Economía (es el caso de Calvo et al. 1993). Sin

¹⁵ Para una breve introducción a estos métodos con ejemplos e ilustraciones ver Nishisato (1994a).

embargo cuando muchas de estas variables son nominales, los métodos tradicionales en econometría han sido limitados.

Por otro lado, trabajos como Weitzman (1998) o Nehring y Puppe (2002), que miran este lado de la información (sujetos) y que han sido discutidos en temas de conservación del medio ambiente, todavía están siendo planteados a nivel de proposición de teorías. En otros estudios más empíricos como por ejemplo, los temas de distribución de ingresos y convergencia espacial, se ha predominado un enfoque agregado y por lo tanto perdiendo información valiosa individual de los sujetos (ver críticas en Quah, 1996). Igualmente, los estudios de paneles por individuos enfatizan mucho la escala de medición de las variables del tipo ratio o intervalo perdiendo información rica en las otras escalas mencionadas arriba.

Sin embargo, por la propiedad de dualidad de los datos desarrollada en Nishisato (1980) es posible estudiar comportamientos en forma desagregada y agregada simultáneamente para una variedad amplia de tipos de variables, sean medidas en escalas del nivel nominal y/o otras escalas.

Además, con la disponibilidad reciente de software y hardware (de mayor velocidad y economía), y con la creciente disponibilidad de información al nivel nominal o categórico, la idea de descomponer la varianza de las filas y/o columnas, por algún método de estimación, por ejemplo, "singular value decomposition" (Stewart, 1973), promedios recíprocos (Nishisato, 1994a, apéndice 1), máximo-verosímil (Millones, 1991), bayesiano u otro, se podrá explorar con más flexibilidad el comportamiento de los sujetos o grupos de ellos.

La incorporación de estos métodos en temas de Economía es algo que puede sugerirse en la literatura econométrica. Por ahora estas aplicaciones pueden inspirarse en el análisis exploratorio hecho ya algún tiempo en otras áreas (ver por ejemplo, Lebart, Morineau y Warwick 1984, Bourdieu 1979; en sociología y filosofía, y más recientemente en Mockus y Corzo 2003 en política).

4.5 Implicancias en el trabajo empírico

El estudio de las variables es en sí un marco conceptual que está relacionado a la utilización adecuada de los métodos y modelos econométricos. Esta es la motivación para revisar este tema. Un área de interés es ver cómo la naturaleza de las variables sugiere el método de análisis empírico, entendiendo esto último como el uso de modelos econométricos u otras técnicas estadísticas disponibles. En este sentido la consideración de la naturaleza de las variables es útil para definir una estrategia de investigación, detectar usos inadecuados de métodos o especificaciones incorrectas econométricas y también clarificar los retos en el análisis de datos. Es decir, nos facilita organizar una agenda para pasar desde los objetivos de la investigación al recogimiento de la información, la organización de los datos y al análisis de las variables.

Algunos ejemplos de estrategia empírica, en técnicas usualmente usadas en la literatura econométrica (y sin considerar la parte muestral) son:

- La combinación rol, escala y construcción es útil para ordenar la especificación de los modelos lineales uniecuacionales. Por ejemplo, Variable dependiente nominal u ordinal observable con link de constructo en casos de Logit, Probit u otros.
- La combinación de rol con escala sin uso de constructos es útil para considerar especificaciones de ratio combinadas con nominales en casos de Análisis de Covarianza (ANCOVA) o Regresión con dummy.
- La combinación de rol con constructos nos inspira un marco para explorar una clase de métodos que son caracterizados por sistemas de ecuaciones que pueden incluir variables latentes: Análisis Factorial o Path-Análisis, Análisis Factorial Dinámico.

Finalmente el marco conceptual propuesto puede ser útil para evaluar procedimientos o desarrollos metodológicos adoptados en trabajos específicos de investigación o evaluar tendencias en el desarrollo de la

literatura cuantitativa. Por ejemplo se puede destacar que la econometría se ha ocupado mucho en técnicas confirmatorias (análisis de regresión, análisis de varianza, regresión logística, modelos probit, ecuaciones simultáneas, etc.) y relativamente poco de técnicas exploratorias. Cuando ha habido oportunidad para su desarrollo más sistemático, por ejemplo en el caso de los modelos VAR –vectores autoregresivos- sin restricciones, se ha minimizado su importancia como técnica exploratoria y se ha enfatizado su uso como técnica confirmatoria (ver SVAR – VAR estructural) en un ambiente de variables no estacionarias.

5. CONCLUSIONES: DELINEANDO ESTRATEGIAS

El propósito de este trabajo fue revisar cómo se ha manejado el concepto de variables en Economía proponiendo al mismo tiempo un marco conceptual para estas variables. Es también un intento de alentar a que este tema sea discutido en la literatura. En este sentido no es posible presentar conclusiones definidas. Sin embargo a manera de conclusión nosotros delineamos algunas implicancias que se desprenden de este marco y que son importantes de resaltar.

- La proliferación de métodos y técnicas econométricas y otras recientes hace necesaria delinear una estrategia que ordene el panorama para el uso óptimo de esta tecnología disponible.
- La consideración de la naturaleza de las variables es útil y podría decirse necesaria para formular un plan para el trabajo empírico. Debiera ser parte (si no el primer paso) en los trabajos de investigación. No se podría elegir una estrategia adecuada empírica si se desconoce la naturaleza de las variables.
- La visión tradicional que clasifica las variables en cuantitativas (continuas y discretas) y cualitativas es ya muy estrecha para enfrentar las alternativas de análisis de datos que existe en la literatura actual.

- Académicamente debe introducirse un tema formal (sea en forma de un curso o capítulo especial en libros de texto) de métodos o metodología que propicie el desarrollo de estos temas y así tender un puente entre teoría y verificación empírica.
- Existe un desbalance metodológico en el tratamiento de variables en Economía. Relativamente más uso de variables de ratio y en un esquema predominantemente confirmatorio. La Econometría, en parte ha contribuido a propiciar este desbalance en cuanto al conjunto de técnicas que ha puesto disponibles para la investigación.
- En el tratamiento de variables en áreas cuantitativas como Estadística, Econometría, se prioriza la atención en el aspecto de las distribuciones de las variables con relativa poca atención de explorar métodos que toman en cuenta la naturaleza de las variables mismas. Es decir, existe un desequilibrio entre los aspectos del rango y del dominio de las variables, con casi absoluta atención al rango.
- El desequilibrio mencionado favorece a un enfoque paramétrico con atención de solo algunas características de la variable, por ejemplo, las que se relacionan con su función de distribución. La consideración de la escala, el rol y la naturaleza de la construcción de las variables podrían minimizar estos potenciales sesgos.
- El proceder operativo y nemotécnico que induce el uso del software debe ser complementado, sino supeditado, con un uso más interpretativo de las variables. Esto puede lograrse con un marco más flexible donde la comprensión fundamental de la información o dato y del método de análisis sean tomados en cuenta. Obviamente estos aspectos se desprenden de la naturaleza de la problemática envuelta en la investigación.
- Hacer un uso responsable de software implica reconocer de antemano la naturaleza de las variables. El tener en cuenta este marco facilitará

un mejor uso del software y de los modelos econométricos disponibles en ellos.

- Reconsiderar el “papel” del rol de las variables nos invita a reflexionar en opciones más exploratorias en el trabajo de investigación y considerar a éste como un proceso. En las etapas iniciales de la investigación es de esperar que no hay rol para las variables. Esto es un factor que debe estar ponderando mucho en los estudiantes del nivel de bachillerato.
- La discusión del rol per se, nos lleva a reconsiderar metodologías más flexibles en la investigación, a descubrir similitudes, semejanzas o diferencias en los procesos históricos metodológicos. Es necesario definir un balance en el estudio de las variables. La problemática en Economía es heterogénea y múltiple, y debe basarse en un concepto amplio de la investigación.
- El desaliento por parte del investigador en el desarrollo de hipótesis o teorías que produce el no contar con información, puede ser explicado por la desconexión entre constructos y observables. En su extremo, esto lleva algunas veces al abandono de una estrategia empírica en la investigación.
- Existe literatura cuantitativa y cualitativa para muchos de estos casos que puede resultar útil al investigador antes que forzar el uso de metodologías inapropiadas para los objetivos de la investigación propuestos.
- Sobre este último punto, el análisis exploratorio de datos y la aproximación bayesiana, entre otras estrategias de investigación, juegan un papel fundamental en llenar estos vacíos.

Finalmente, en cuanto al trabajo empírico en Economía, sin estrategias que aclaren los criterios básicos para seleccionar y hacer uso de la literatura econométrica actual, el trabajo econométrico solo quedará reducido, en el

mejor de los casos, a la aplicación de un conjunto de "recetas cuantitativas", caracterizado por solo demandar habilidades o logros educacionales que pertenecen (según algunas taxonomías - Guilford, 1977 o Anderson y Krathwohl, 2001) a la obtención de objetivos de bajo nivel, no correspondiendo con los objetivos educacionales que se desean para la educación superior.

Referencias Bibliográficas

- Aigner D.J. y A.S. Goldberger (eds.)
(1977) *Latent Variables in Socioeconomic Models*. Amsterdam: North-Holland.
- Alexander C.
(2001) *Market Models. A guide to Financial Analysis*. New York: John Wiley & Sons, LTD.
- Anderson, L. y Krathwohl, D. A
(2001) *Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman, 2001.
- Anderson T. W.
(1958) *An Introduction to Multivariate Statistical Analysis*. A Wiley Publication in Mathematical Statistics. New York: John Wiley.
- Anderson T.W. y Hsiao C.
(1982) Formulation and estimation of dynamic models using panel data. *Journal of Econometrics*, 18: 47-82.
- Arellano, M.
(2003) *Panel Data Econometrics*. New York: Oxford University Press.
- Benzecri J.P. et al.
(1973) *L'Analyse Des Données: II. L'Analyse des Correspondences*. Paris Dunod.
- Berry S., Levenson J. y Pakes A.
(1995) Automovile prices in market equilibrium. *Econometrica*. Vol. 66. No. 4, July. pp: 841-890.
- Birnbaum M. y Sutton S.
(1992) *Scale Convergence and Utility Measurement*. California State University, Fullerton
- Blalock H. M. Jr.
(1984) *Construcción de teorías en ciencias sociales. De las formulaciones verbales a las matemáticas*. México: Trillas.
- Bock R.D.
(1975) *Multivariate Statistical Methods in Behavioral Research*. New York: McGraw Hill.
- Borsboom, D., Mellenbergh, G.J. y van Heerden. J.
(2003) The theoretical status of latent variables. *Psychological Review*, 110, 203-219.

- Bourdieu P.
(1979) *La distinction: critique sociale du jugement*. Paris:Minuit.
- Box y Jenkins
(1974) *Time Series Análisis*. Forecasting and Control. San Francisco: Holden Day.
- Calvo G., Leiderman L. y Reinhart C.
(1993) Capital inflows and real exchange rate appreciation in Latin America. -Section III: Econometric Analysis. *International Monetary Fund. Staff papers*, Vol. 40 No. 1, March 1993.
- Cameron A. C. y Trivedi P.K.
(1998) *Regression Analysis of Count Data*. Cambridge: Cambridge University Press.
- Christ, C.
(1994) The Cowles Commission's Contributions to Econometrics at Chicago, 1939-1955," *Journal of Economic Literature* 32, 1, 30-59.
- Cox D.R.
(2007) Applied Statistics: A review. *The Annals of Applied Statistics*, 1 pp: 1-16.
- Diebold, F. y G. D. Rudebusch
(1996) Measuring Business Cycles: A Modern Perspective, *Review of Economics and Statistics* 78, February, 67-7.
- Duo, Qin
(1997) *Formation of Econometrics*. Oxford University Press.
- Enders W.
(2004) *Applied Econometric Time Series*. (2nd ed.). Danvers MA: John Wiley & Sons, Inc.
- Engle R. F.
(1982) Autoregressive conditional heteroscedasticity with estimates of the variance of UK inflation, *Econometrica*, Vol. 50, pp. 987-1007.
- Engle R. F. y Granger C. W. J.
(1987) Co-integration and error correction representation, estimation, and testing. *Econometrica* 55: pp:251-276.
- Fan J. y Yao Q.
(2005) *Nonlinear Time Series*. Springer.
- Fan J., Yao Q. Cai Z.
(2003) Adaptive varying-coefficient linear models, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 65 (1), 57-80.

- Fiemberg S. y Meyer M.
 (1983) Loglinear models and categorial data analysis with psychometric and econometric applications. *Journal of Econometrics*, 22, pp: 191-214.
- Gagliardini P.
 (2006) Challenges in the Teaching of Econometrics: The Lesson of Pietro Palestra. Journee Pietro Balestra, Geneve, 30th November 2006. University of Lugano and Swiss Finance Institute.
- Gescheider G.
 (1997) "Psychophysics: The Fundamentals" Lawrence Erlbaum Assoc., ISBN: 0-8058-2281-X.
- Geweke J.
 (1977) "The Dynamic Factor Analysis of Economic Time-Series Models" En D.J. Aigner and A.S. Goldberger (eds.), *Latent Variables in Socioeconomic Models*. pp: 26-383. Amsterdam: North-Holland.
- Goldstein, H.
 (2003) *Multilevel Statistical Models*. London: Edward Arnold.
- Goodman, L. A.
 (1978) *Analyzing Qualitative / Categorical Data: Log-linear Models and Latent Structure Analysis*. Cambridge, MA: Abt Books
- (1991) Measures, models, and graphical displays in the analysis of cross-classified data. *Journal of American Statistical Association*, 86, 1085-1111.
- Granger C.W.J. y Mizon G.P. (eds.)
 (2005) *Advanced Texts in Econometrics*. Oxford: Oxford University Press.
- Greenacre M. y Blasius J. (eds.)
 (1983) *Correspondence analysis in the Social Sciences. Recent developments and applications* . London: Academic Press.
- Greene W.
 (2008) *Econometric Analysis*, 6th Edition. New Jersey: Prentice Hall.
- Guilford J.P.
 (1977) *La naturaleza de la inteligencia humana*. Buenos Aires: Paidos.
- Hastie T., Tibshirani R., Friedman J.
 (2001) *The elements of statistical Learning*. New York: Springer-Verlag.
- Hausman J. y Wise D.
 (1977) Social Experimentation, truncated distributions and efficient estimation. *Econometrica*, 45, pp. 919-938.

- Heckman J.
(1979) Sample selection bias as specification error. *Econometrica*, 47, pp. 153-161.
- Hendry D. y Morgan M.
(1995) *The foundations of Econometric Analysis*. Cambridge: Cambridge University Press.
- Hotelling H.
(1933) Analysis of a complex of statistical variables into Principal Components. *Journal of Educational Psychology*. 24, 417-41, 498-520.
- Jeffreys H.
(1957) *Scientific Inference* (2nd ed.) Cambridge: Cambridge University Press.
- (1966) *Theory of Probability* (3rd ed.). Oxford: Clarendon
- Johnston J. y DiNardo J.
(1997) *Econometric methods*. (4th ed.).. New York: McGraw-Hill.
- Joreskog K.
(2001) Structural Equation modeling with LISREL. University of Upsala, Department of Statistics, Sweden. Cursillo Universidad de Chile, Dpto. De Matemáticas.
- Lange O.
(1964) Introducción a la Econometría. Cap. 1. Mexico D.F.: Fondo de Cultura Económica.
- Lazarsfeld, Paul F.
(1950) "The Logical and Mathematical Foundations of Latent Structure Analysis", Chapter 10 en Stouffer et al. (1950). Ver también Chapter 11, "Some Latent Structures".
- Leamer, Edward E.
(1978) *Specification searches: ad hoc inference with nonexperimental data*. New York: J. Wiley.
- Lebart L., Morineau A. y Warwick K.
(1984) *Multivariate descriptive statistical analysis*. New York: John Wiley and Sons.
- Lohr S.
(2000) *Muestreo: Diseño y Análisis*. México DF: Internacional Thompson Editores.

- Malinvaud E.
 (1983) Econometric Methodology at the Cowles Commission: Rise and Maturity. Abstracted from the Cowles Fiftieth Anniversary Volumen. (conference held at Yale University on June 3-4, 1983 to commemorate the Cowles research group's fiftieth anniversary. Ver <http://cowles.econ.yale.edu/archive/reprints/50th-malinvaud.htm>)
- McFadden D.
 (1974) The measurement of urban travel demand. *Journal of Public Economics*, 3, 1974, pp. 303-328.
- Millones O.
 (1991) Dual Scaling in the framework of the association and correlation models under Maximun Likelihood. Disertación de tesis doctoral, Universidad de Toronto. Canada.
- Mockus A. y Corzo J.
 (2003) Dos caras de la convivencia. Cumplir acuerdos y normas y no usar ni sufrir violencia. *Análisis Político* (Bogotá) No. 48, Enero/abril 2003.
- Morgan, M.
 (1991) The History of Econometric Ideas. Cambridge University Press
- Mosteller, F. y Tukey, J.W.
 (1977) *Data Analysis and Regression*. Reading, Mass: Addison-Wesley Nie.
- Myoung-Jae Lee
 (2005) Micro-econometrics for policy, program, and treatment effects. In Granger C.W.J. and Mizon G.P. (eds.) *Advanced Texts in Econometrics*. Oxford: Oxford University Press.
- Nehring K. y Puppe C.
 (2002) A theory of diversity. *Econometrica*, 70, pp:1175-1178.
- Nishisato S.
 (1980) *Analysis of categorical data: Dual Scaling and its applications*. Toronto: University of Toronto Press.
- (1994) *Elements of Dual Scaling. An introduction to practical Data Analysis*. New Jersey: LEA Publishers.
- (1994a) *Dual Scaling in a Nutshell*. Toronto: MicroStats.
- Novales A.
 (1993) *Econometría*. 2da. Ed. Madrid: Mc Graw Hill.

- Parsons T.
(1954) *Essays in Sociological Theory*. Revised edition. Toronto: Collier-Macmillan Canada, Ltd.
- Quah D.
(1996) Regional Convergente Cluster across Europe, *European Economic Review*, 40, pp: 951-958.
- Ramsay, J.O. y Silverman B. W.
(1997) *Functional data analysis*. New York: Springer.
- Robbins L.
(1932) *An essay on the nature and significance of economic science*. London: Macmillan.
- Scott W. y Freese J.
(2001) *Regression Models for Categorical Dependent Variables using Stata*. Texas: StataCorp.
- Sevestre P.
(2006) L'enseignement de la micro-économétrie. Quels atouts? Quelles difficultés? Université de Paris 1-Panthéon Sorbonne. Communication à la Journée en Mémoire de P. Balestra. Département d'Econométrie. Université de Genève. 30 novembre 2006. (Travaux en cours)
- Shepard R. N.
(1962) The analysis of proximities: Multidimensional scaling with an unknown distance function. I. *Psychometrika*, 27, 125-140.
- Spearman C.
(1927) *The abilities of man. Their nature and measurement*. Toronto: The MacMillan Company of Canada, Limited.
- Sternberg R.J. Lautrey J. Lubart T.I.
(2003) Where are we in the field of intelligence, how did we get here and where are we going. En Sternberg et al. (Eds.) *Models of Intelligence. International perspectives*. Chap. 1. American Psychological Association, Washington DC.
- Stevens S. S.
(1946) On the theory of scales of measurement, *Science*, Vol. 10, 103, 677-680.
- Stewart F. M.
(1973) *Introduction to Linear Algebra*. Princeton: Van Nostrand.
- Stock J. H.
(2006) Forecasting and Now-Casting with Disparate Predictors: Dynamic Factor Models and Beyond FEMES 2006 Meetings Beijing

- Stock, J.H. y M.W. Watson
 (2005) An Empirical Comparison of Methods for Forecasting Using Many Predictors. Research funded by NSF grant SBR-0214131.
- (2006) Implications of Dynamic Factor Models for VAR Analysis, NBER Working Paper # 11467.
- Stouffer, Samuel A., et al.
 (1950) Measurement and Prediction, Volume IV de *The American Soldier: Studies in Social Psychology in World War II*. Princeton University Press. Reprinted 1973 by Peter Smith, Gloucester MA.
- Strauss A. y Corbin J.
 (1990) *Basics of Qualitative Research. Grounded Theory Procedures and Techniques*. Newbury Park: Sage Publications.
- Tashakkori A, Teddlie C.
 (1998) *Mixed Methodology: Combining Qualitative and Quantitative Approaches*. Thousand Oaks, Calif: Sage Publications.
- Thurstone L.
 (1938) *Primary Mental Abilities*. Chicago: University of Chicago Press.
- Tobin J.
 (1958) Estimation of Relationships for Limited Dependent Variables, *Econometrica*, 26, 24-36.
- Tong H.
 (1990) *Non-linear Time Series. Dynamic A dynamic approach*. Oxford University Press.
- Torgerson W.
 (1958) *Theory and Methods of Scaling*. New York: Wiley.
- Travers M.
 (2001) *Qualitative Research Through Case Studies*. London: Sage Publications.
- Tukey, J.W.
 (1977) *Exploratory Data Analysis*. Addison-Wesley
- UMC
 (2001) Efectos de la escuela en el rendimiento en Lógico-Matemática en cuarto grado de primaria. Boletín No.8 Unidad de Medición de la Calidad Educativa, UMC, Ministerio de Educación.
- Valdivieso S. L.
 (1991) Escalas de Medición. *Pro Matemática: Vol V, Nos. 9-10*, pp:51-65.

- Velleman P. y Wilkinson L.
(1993) Nominal, Ordinal, Interval, and Ratio Typologies are Misleading.
The American Statistician, 47, 1, pp: 65-72.
(ver http://en.wikipedia.org/wiki/Level_of_measurement)
- West B.T.
(2007) *Linear mixed models: a practical guide using statistical software*.
FL: Chapman & Hall
- Weitzman M.
(1992) On diversity. *Quarterly Journal of Economics* , 107, pp:363-405.
- Wright P.G.
(1929) Review of Schultz, Statistical Laws of Demand and Supply.
Journal of American Statistical Association, vol 24, pp: 207-215.
- Zellner A.
(1971) *An introduction to Bayesian Inference in Econometrics*. New
York: John Wiley & Sons, INC.

Anexo 1: Propiedades en Escala de Stevens.

Escala Nominal.- Esta escala no cumple con ninguna de las propiedades mencionadas (ver columna 2 de Tabla 1). Por ejemplo, la variable Nacionalidad. En un evento internacional en la que asisten representantes de varios países latinoamericanos, se toma información sobre los participantes. Es posible que a cada persona le corresponda el código de cada país y se le otorgue el código respectivo. Para un participante, este código (09, por ejemplo) es el valor de la variable Nacionalidad y esta variable es nominal pues los números (códigos) no cumplen con las tres propiedades mencionadas. El 09 sólo identifica como, por ejemplo, guatemalteco y este número no indica cantidad ni ordenación. Su diferencia con otros códigos tampoco indican distancia. Por otro lado el 00 no indicaría el origen (ausencia) del concepto de Nacionalidad.

Escala Ordinal.-Sea la Tabla 2, información de cinco comercializadores de una empresa textil X respecto al valor de ventas en una de sus líneas en el mes de Agosto 2007:

Tabla 2 Reporte de Comercialización de la empresa X, mes de Agosto

| Tienda (1) | Ranking (2) | Ventas (miles S/) (3) |
|-----------------------|------------------------|----------------------------------|
| B | 1 | 16 |
| A | 2 | 10 |
| D | 3 | 6 |
| C | 4 | 3 |
| E | 5 | 1 |

Observar que la columna (2) del Ranking es una variable ordinal, mientras que la última columna es una variable de ratio. Considerar la comparación de las tiendas C y A. En términos de ranking de ventas (columna 2) no se puede decir que la "Baja performance o grado de ineficiencia en ventas" de C es el doble que A. Esta idea de distancia solo se puede aplicarse a la última columna que indica las ventas mismas en miles de soles.

Escala de Intervalo.-En esta escala se asigna valores a objetos usando una escala cuyos intervalos se "desea" sean comunes o constantes. En esta escala el valor de cero es sólo referencial.

El ejemplo clásico de esta escala es temperatura. Por ejemplo, refiriéndonos al tiempo (temperatura) en centígrados, "cero grados centígrados" no significa ausencia de temperatura. Se dice que en esta escala el cero es arbitrario. (recordar que en otra escala tal temperatura es 32 grados).

Dado que el cero es arbitrario, todas las comparaciones como múltiplos o proporciones son por lo tanto arbitrarias. Así por ejemplo, comparando las temperaturas promedio de Julio y Diciembre, no podemos decir que "en Diciembre hizo el doble de calor que el Julio".

Las encuestas de opinión o de percepciones o basadas por la suma de preguntas del tipo "Diga Ud. en una escala de 1 al 5 el grado de acuerdo....etc.", se miden en esta escala porque podemos establecer que la distancia entre 25 y 20 es la mitad de la que hay entre obtener un puntaje entre 30 y 20. Este tipo de variables son típicas en Psicología y en los estudios de investigación de mercados.

Escala de ratio o de razón.-Son el tipo de variables que en Economía, comúnmente se conocen como cuantitativas, ejemplo, precios, ingresos, consumo, exportación, etc. Así si el precio del combustible subió de 8 a 16 soles el galón, podemos decir que el precio se ha doblado en el periodo considerado. Notar que en la escala de ratio el cero puede interpretarse como no precio (gasolina gratis).

Otro grupo de estas variables se obtienen por mediciones experimentales como conteos, o en Sociología como las frecuencias en tablas de contingencia.

Anexo 2 Diversas interpretaciones de “no observabilidad”

Un problema con la etiqueta “no observables” es que en la literatura econométrica hay diversas interpretaciones de lo que se entiende por “no observabilidad”, interpretaciones que siendo válidas no conviene usarlas aquí para los objetivos de este trabajo. Algunos ejemplos de estas interpretaciones son:

No observabilidad muestral. En ocasiones, la naturaleza de un proceso de recogida de datos hace que solo se “observen” las variables correspondientes a una unidad económica solo si el valor de la variable excede cierto umbral. En algunos casos este hecho es por decisión del investigador, por ejemplo, en un estudio sobre performance del sistema educativo, cuando sólo estudiamos a los que desaprobaron la prueba de rendimiento en matemáticas del 6to grado en una prueba nacional escolar. En otros casos para ciertos estudios el valor de la variable se trunca en un valor o umbral, por ejemplo, en un estudio del comportamiento del consumidor de bienes durables. Los que no compraron bienes durables su gasto en el bien es cero y solo se “observa” su gasto para los que entran al mercado de bienes durables. La no observabilidad en estos dos ejemplos se refiere a que para algunas unidades de estudio el valor de la variable no existe o está truncada. La propiedad de las variables de generar valores según entren o no en la muestra o según exista o no valores para rangos de valor de la variable, es un tema de escala mixta de variables que algún momento se ha mencionado en este trabajo.

Variables excluidas o no reveladas. La literatura econométrica también se refiere a variables “no observables” cuando en un estudio, el concepto es conocido y medible pero no es revelado por el sujeto agente económico, ya sea porque no se dispone de esa información en el momento de recoger la información (una encuesta, un catálogo con especificaciones técnicas) o por que la unidad analizada no quiere revelar voluntariamente su valor. Por ejemplo, en este último caso, hay características encubiertas de un producto o servicio ofrecido (ver Berry et al., 1995) que conviene no revelar, aún

cuando sean medibles, para lograr cierta ventaja en una transacción. Típico comportamiento de mercados imperfectos¹⁶.

El lector se dará cuenta que estas variables encubiertas pueden ser tanto observables o no, según nuestra definición. Otros ejemplos en esta interpretación son los ingresos altos en un estudio sobre evasión tributaria, potencia en caballos de fuerza en especificidades del motor de un auto. En estos casos el concepto es observable en el sentido de este trabajo, sin embargo formalmente incorporado en la literatura como "no observable", pues desde el punto de vista del econometrista el dato aunque importante, no está disponible.

Variables no controladas experimentalmente. Es el típico caso presentado en los modelos de panel (ver por ejemplo efectos latentes o no observados según la definición en Novales 1993, sección 15.1). En el caso de estudios de panel se menciona que hay variables que no son recogidas en el estudio de las unidades de análisis. Esto resulta perjudicial en un estudio de corte transversal por cuanto se está excluyendo información relevante en el análisis, sobre todo cuando se demuestra que las unidades de estudio son muy heterogéneas en estas variables no recogidas. Los modelos de panel son una manera de minimizar o anular este sesgo de exclusión de variables al tener información de las unidades de análisis en varias ocasiones. Luego la existencia de la heterogeneidad mencionada no resulta ser un problema pues está siendo "experimentalmente" controlada en el tiempo en el recojo de la información de cada unidad estudiada.

La literatura econométrica asociada a los estudios de Panel, ha llamado variables no observables a este tipo de variables cuyos efectos son usualmente excluidos en estudios de corte transversal, y que son causa de una heterogeneidad adicional no incluida o descuidada en el análisis. (Ver

¹⁶ En Berry et al. (1995) la realización como dato observado del valor de una variable, puede ser oculta técnicamente o intencionalmente, como es el caso de las variaciones en la calidad de un producto observadas solo por el productor pero no por el consumidor o por el econometrista. La incorporación de esta posibilidad en el análisis es importante por cuanto ayuda a establecer los procedimientos apropiados para eliminar los sesgos potenciales en las estimaciones de los parámetros.

aplicación de este concepto en marco del análisis de diseños experimentales en la discusión de efectividad de programas sociales en Myoung-Jae Lee, 2005). Sin embargo el lector podrá rápidamente identificar que estas definiciones son distintas a la dada en este trabajo.

Errores de medición, errores en las variables o regresores estocásticos.-

El error de medición se define como la diferencia entre el verdadero valor de la variable de interés y lo que actualmente se mide ("lo observado"). Este tratamiento es común en los textos econométricos actuales. Sin embargo la variable sujeta a error de medición usualmente se refiere a variables observables, como ingresos, importaciones, etc. Por lo que difiere en este sentido de la definición de este trabajo. Por supuesto que hay tratamientos de no observabilidad en este sentido usando el concepto de constructo, ver Aigner y Goldberger (eds.) (1977), sin embargo esta correcta interpretación no se ha mantenido en la literatura consistentemente por otros autores. El tratamiento del esquema de errores de medición en la literatura econométrica se ha concentrado en sus implicancias de sesgos en los trabajos de estimación de parámetros. El objetivo ha sido prevenir al investigador de las implicancias de los errores de medición, típicamente ligadas al modelo de regresión lineal múltiple. La razón es que someter variables con errores de medición a la estimación por Mínimos Cuadrados Ordinarios (MCO) da estimación de parámetros que son sesgados e inconsistentes y por ello requieren otros métodos de estimación.

Anexo 3 Escalas mixtas de medición. Otras escalas.

Maddala (1983) en su clásico estudio econométrico sobre variables dependientes limitadas, menciona desde la introducción los casos de variables truncadas, censuradas y dummies.

Modelos Truncados y Censurados

Cuando se considera el efecto de escala de la medición de las variables a menudo surge el problema de si los valores de las variables planificadas para ser usadas finalmente son obtenidas en la forma esperada por el investigador o de manera truncada o censurada.

Muchas de las variables de ratio en Economía tienen valores que son truncados o censurados, por ejemplo, generan solo valores no negativos o concentrados en un solo valor. Ejemplos clásicos del primer caso son precios e ingresos monetarios. Igualmente este caso se presenta cuando según el objeto de la investigación, se selecciona una sub-población de la población muestreada. Aquí es posible que las variables no son estudiadas en todo su rango, por ejemplo, en un estudio de estudiantes con bajo rendimiento (cuyo promedio es menor que 8).

James Tobin en 1958, trabajando el tema de gasto en bienes durables, sugirió inicialmente la necesidad de estudiar variables que representan casos mixtos de variables cualitativas y variables restringidas a ser no negativas. Pero esta idea solo fue madurada, en Economía, en los 70 y 80 (ver Malinvaud 1983) con los trabajos de McFadden, Heckman y otros (ver McFadden 1974; Hausman y Wise 1977; Heckman 1979).

Una de las consecuencias de este tipo de variables (mixto en escalas) es que afecta otra característica de las variables que es no muy mencionada en este trabajo, cual es su función de distribución. Como puede intuirse, la mixtura mencionada tiene un efecto sobre el criterio D, o efecto de diseño de muestreo. Esto se explica así:

En el caso de variables cuyos datos son censurados las discontinuidades de algunos valores para la variable de interés solo afectan la forma de su distribución. En el caso de variables cuyos datos son truncados la variable de interés pierde un segmento de sus valores. Esto no significa que necesariamente se pierde información (al no aparecer un rango de sus valores, cabe ver la posibilidad de ver si también se eliminan valores de la variable independiente dando la impresión que se pierde información). El problema es importante resolverlo pues induce a posibles sesgos en las estimaciones. La existencia de estos sesgos depende del control que se tenga sobre el diseño muestral de la investigación.

Si en el diseño del estudio esta exclusión se planea de antemano por los mismos objetivos de la investigación, entonces la exclusión es voluntaria, o en términos estadísticos, el universo concuerda con el marco muestral. Aquí la población objetivo está siendo adecuadamente representada por la población muestreada. En este caso, aunque la distribución de la variable está truncada, no hay sesgo de selección ni se pierde información. Pero si el diseño del estudio no planea de antemano esta exclusión, como es frecuente en el caso en el uso de fuentes secundarias de información, entonces existe la posibilidad de sesgo de selección, es decir, hay factores que se asocian a la variable estudiada que no son controlados por el investigador y es probable que las estimaciones hechas no sean las correctas¹⁷.

Otras escalas de medición

Otras clasificaciones para las escalas de las variables son las presentadas por Mosteller y Tukey (1977), con categorías como Nombres; Grados (niveles ordenados, medio, bajo, alto); Rangos (números de 1 a X); Fracciones (frecuencias relativas); Conteos, (frecuencias absolutas), Cifras (números reales no negativos), Balances (valores sin límites positivos o negativos). Shepard y otros (1962) definen escalas transformadas de por ejemplo ordinales a ratio, en una clasificación como Métricas y No métricas

¹⁷ Este comentario excluye otros sesgos denominados "ex ante" producido porque los sujetos investigados eligen voluntariamente participar en el estudio, pero esto está fuera del muestreo convencional, ver otros casos de sesgo muestral en Lohr (2000).

(Escalamiento multidimensional, ver también Nishisato 1980). Maddala (1983, p 14) por su parte define dos tipos de variables en el trabajo econométrico: a) categóricas y b) no categóricas. Las primeras se dividen a su vez en no ordenadas (ocupación), secuenciales (nivel educativo) y ordenadas (ingresos en clases, actitudes en preferencias).

Para modelos estructurales, y teniendo en cuenta el problema de datos censurados, se han propuesto escalas modificadas (ver por ejemplo Joreskog 2001, Scott, and Freese 2001) que se adaptan mejor a las hipótesis y proposiciones teóricas de la literatura económico –social y psicológica.

Finalmente, es interesante notar que en el marco de la teoría de elección, existe un nexo entre el tema de las escalas con la teoría de la utilidad, ver Birnbaum and Sutton (1992) para mayores detalles.