



Universidad de Buenos Aires  
Facultad de Ciencias Económicas



**Instituto de Investigaciones en Administración,  
Contabilidad y Métodos Cuantitativos para la Gestión  
(IADCOM)**

*Directora: Profesora Emérita Dra. María Teresa Casparri*

**REVISTA DE INVESTIGACIÓN  
INTERDISCIPLINARIA EN MÉTODOS  
EXPERIMENTALES**

Año 1. Volúmen 1

**Revista de Investigación Interdisciplinaria en Métodos  
Experimentales. Vol. 1**

*Directora*

María Teresa Casparri

*Comité académico*

María Teresa Casparri

Pablo Fajfar

Adriana Fassio

Javier García Fronti

---

**Editor responsable:**

*Instituto de Investigaciones en Administración, Contabilidad y Métodos  
Cuantitativos para la Gestión*

Facultad de Ciencias Económicas de la Universidad de Buenos Aires

Av. Córdoba 2122

Ciudad Autónoma de Buenos Aires, Argentina

Contacto: iadcom@econ.uba.ar

## **Índice**

- 4      **Prefacio**
- 5      **Acerca de los autores**
- 8      **Una nota sobre el rol de los experimentos económicos en el diseño eficiente de nuevos mercados**  
*Javier García Fronti*
- 15     **Metodología de análisis de las desigualdades sociales**  
*Nora Mabel Lac Prugent*
- 31     **Las escalas como instrumento de medición de actitudes. Articulación con la investigación de mercados**  
*Ana Marsanasco*
- 42     **El diseño experimental en el contexto de la gestión: notas sobre su implementación en la evaluación de políticas públicas. Experiencias, alcances y limitaciones**  
*Adriana Fassio*  
*María Gabriela Rutty*
- 53     **La construcción de un tren trasnacional de alta velocidad y sus efectos sobre el territorio – un caso de economía espacial**  
*Verónica Caride*  
*Mercedes Ramos*
- 78     **Valor esperado concentrado (vec) Y simulación monte carlo en la evaluación de proyectos de inversión**  
*José Luis Pungitore*
- 91     **Una propuesta de índice de Actividad Científica y su validación en el campo de ciencias sociales y humanidades**  
*Julieta Barrenechea*  
*Javier García Fronti*  
*Javier Castro Spila*  
*Andoni Ibarra*

## **PREFACIO**

Esta publicación da cuenta de las temáticas abordadas en la I Jornada Interdisciplinaria de Métodos Experimentales aplicados a la Gestión y Economía, realizada el 3 de mayo de 2012, en el marco del Programa Interdisciplinario de Métodos Experimentales Aplicados a la Gestión y a la Economía (PIMEAGE), que se desarrolla en el Instituto de Investigaciones en Administración, Contabilidad y Métodos Cuantitativos para la Gestión (IADCOM) de la Facultad de Ciencias Económicas de la Universidad de Buenos Aires.

Cabe destacar que los métodos experimentales aplicados a la economía y a la gestión constituyen una línea metodológica poco explorada en la comunidad latinoamericana, por lo que esta propuesta de trabajo promueve el reconocimiento de nuestra Facultad como un espacio líder en la innovación de construcción del saber.

La publicación se desarrolla en tres ejes claramente delineados: a) la construcción de mercados eficientes a partir de metodologías experimentales; b) la reflexión sobre los alcances de los modelos experimentales aplicados a la evaluación de proyectos de inversión, transporte y sociales y c) la aplicación de métodos experimentales para la medición de problemáticas tales como la desigualdad social y las actitudes en la investigación de mercado.

Deseo agradecer a los autores que con su aporte han contribuido a ampliar este espacio en construcción e invito a colegas, investigadores y alumnos a sumarse a esta desafiante propuesta.

*Prof. Emérita Dra. María Teresa Casparri*

*Directora del IADCOM*

## **ACERCA DE LOS AUTORES**

### **Javier García Fronti**

Doctor de la Universidad de Buenos Aires. Profesor titular de Matemática para Economistas, Facultad de Ciencias Económicas (UBA). Subdirector del Centro de Investigación en Métodos Cuantitativos Aplicados a la Economía y la Gestión (CMA), Facultad de Ciencias Económicas (UBA).

### **Nora Mabel Lac Prugent**

Doctora. Licenciada en Estadística. Docente-investigador del Instituto de Investigaciones Económicas de la Escuela de Economía, Facultad de Ciencias Económicas y Estadística, Universidad Nacional de Rosario.

### **Ana Marsanasco**

Licenciada en Administración y Especialista en la Gestión de Pequeñas y Medianas Empresas por la Facultad de Ciencias Económicas de la Universidad de Buenos Aires. Doctoranda de la Universidad de Buenos Aires en Administración y docente de esta Casa de Estudios en las áreas de Matemática y Administración. Directora del Proyecto UBACyT "Aglomerados Productivos en la Argentina: un enfoque desde la gestión del conocimiento y la innovación". Profesional Analista en la Unidad de Planificación Estratégica y Gestión Interna de la Superintendencia de Riesgos del Trabajo. Autora de varios capítulos de libros, artículos en revistas nacionales e internacionales. Ha dictado cursos de postgrado y participado como expositora en congresos y jornadas de la especialidad y afines.

### **Adriana Fassio**

Doctora en la Universidad de Buenos Aires, Facultad de Filosofía y Letras orientación Antropología Social, Magister en Gerontología Social V Edición Iberoamericana, Universidad Autónoma de Madrid, y Licenciada en Sociología, Profesora de Enseñanza Secundaria Normal y Especial en Sociología, Facultad de Ciencias Sociales, Universidad de Buenos Aires. Docente de grado y posgrado. Directora del Programa Interdisciplinario de Métodos Experimentales Aplicados a la Gestión y a la Economía (PIMEAGE)

e investigadora del IADCOM en la Facultad de Ciencias Económicas de la UBA.

### **María Gabriela Ruty**

Doctora en Administración, con especialidad en evaluación. Docente de la cátedra de Sociología de las organizaciones. Investigadora del IADCOM en la Facultad de Ciencias Económicas de la UBA.

### **Verónica Caride**

Licenciada en Economía con Honores, Magister en Economía, Doctoranda y Docente de Grado y Posgrado de la Facultad de Ciencias Económicas, Universidad de Buenos Aires (UBA). Con estudios de posgrados en Buenos Aires y en el exterior (Italia y Holanda) en Desarrollo Local, Finanzas Aplicadas y Economía Espacial. Investigadora del Centro de Investigación en Métodos Cuantitativos Aplicados a la Economía y la Gestión, Facultad de Ciencias Económicas, (UBA).

### **Mercedes Ramos**

Licenciada en Economía de la Facultad de Ciencias Económicas, Universidad de Buenos Aires (UBA). Con estudios de posgrado en la Facultad de Ciencias Económicas, Universidad de Buenos Aires (UBA). Docente de Grado de Geografía Económica y Social Mundial en la Facultad de Ciencias Económicas, Universidad del Salvador (USAL).

### **José Luis Pungitore**

Doctor en Ciencias Económicas (orientación Administración, UBA). Profesor titular e investigador en la UNLZ y profesor titular en UADE y en maestrías y doctorados. Amplia experiencia en cargos gerenciales y directivos, públicos y privados. Autor de varios libros y artículos de su especialidad.

### **Julieta Barrenechea**

DEA en Gestión de la Ciencia y la Tecnología, Departamento de Lógica y Filosofía de la Ciencia, Universidad del País Vasco (UPV/EHU). Licenciada en Sociología de la UBA. Investigadora del Área de Investigación y Gestión de Redes de Ciencia, Tecnología e Innovación, Cátedra Sánchez-Mazas, UPV/EHU. Investigadora Asociada del Programa PIRNA, Instituto de Geografía, FFyL - UBA.

### **Javier Castro Spila**

Doctor Europeo en Filosofía de la Ciencia, Tecnología y Sociedad, Universidad del País Vasco (UPV/EHU). Ms. en Educación Superior Universidad de Palermo/UNESCO. Licenciado en Sociología de la UBA. Ha sido Investigador del Area de Investigación y Gestión de Redes de Ciencia, Tecnología e Innovación, Cátedra Sánchez-Mazas, UPV/EHU. Actualmente es Responsable del Area de Investigación y Redes de SINERGIAK Social Innovation Center, País Vasco e investigador colaborador en Ecole Supérieure des Technologies Industrielles Avancées ESTIA - Francia.

### **Andoni Ibarra**

Profesor titular de Filosofía de la Ciencia en la Universidad del País Vasco (UPV/EHU). Ha sido co-director de la Unidad Asociada de Estudios de Ciencia y Tecnología del CSIC-UPV/EHU y actualmente es coordinador de la Cátedra Sánchez-Mazas de la UPV/EHU, entidad dedicada a la investigación en temas de ciencia, tecnología e innovación. Es editor de la revista Theoria. Investigador Principal del grupo de investigación consolidado PRAXIS, dedicado al estudio de las teorías y prácticas representacionales de la ciencia, y a las transformaciones científico-tecnológicas de la sociedad y la cultura.

# UNA NOTA SOBRE EL ROL DE LOS EXPERIMENTOS ECONÓMICOS EN EL DISEÑO EFICIENTE DE NUEVOS MERCADOS

*Javier García Fronti*

## INTRODUCCIÓN

Jean-Baptiste Say (1841) fue uno de los primeros autores en afirmar que la economía es una ciencia experimental y debería dedicarse a la observación y a la recolección de los hechos a fin de señalar las regularidades causales. Sin embargo, dejó claramente a la economía fuera del laboratorio, reservado para las ciencias naturales. La economía recién entra en los laboratorios a principios de la década de los sesenta del pasado siglo (Smith, 1962). Desde su creación, la economía experimental ha tenido que justificar su validez fuera del laboratorio (Guala, 2002). Es importante notar que al definir las condiciones para la validez externa de un experimento económico, Vernon Smith (1989) establece que el laboratorio experimental debe ser real, contar con personas reales, bienes reales y dinero real. Una segunda característica de la economía experimental es su capacidad manipuladora, se interviene sobre la materia de diversas formas. Por ejemplo si el objeto es un coche a testear, la manipulación puede ser desde hacer un *focus group* con usuarios que abran y cierren puertas, hasta implicar la destrucción del objeto.

Esto lleva a un tercer aspecto importante de la economía experimental: su objeto. Según Vernon Smith, los experimentos son un medio para contrastar teorías económicas; su objeto son las teorías, no el comportamiento humano (Muniesa y Callon, 2007). El entorno experimental pretende imitar las teorías económicas, no la actividad económica concreta (Smith, 1994). El contexto es, por consiguiente, un "sistema micro-económico" controlado y estable donde los agentes económicos (caracterizados con funciones de utilidad) interactúan. El conocimiento así producido es legítimo en la medida en que permite contrastar hipótesis derivadas de las teorías económicas en cuestión. La generalización de los resultados experimentales se complica debido al localismo de los experimentos en cuestión, pero no es un impedimento (Burlando y Guala, 2005). Muchas veces el objetivo de un experimento es



convertir una teoría compleja en un conjunto explícito de reglas y comportamientos. Este vuelco hacia lo explícito ya estaba presente en los primeros experimentos de (Chamberlin, 1948). Los experimentos pueden llevarse a cabo en las aulas, dentro de sistemas informáticos o en un mercado real (Muniesa y Callon, 2007).

En resumen, los experimentos económicos tienen un impacto performativo sobre los mercados; los investigadores describen un objeto producido por ellos mismos. Si bien esto es claro cuando se trabaja dentro de un laboratorio, resulta más evidente en experimentos que utilizan un mercado real como su campo de testeo; al realizar el experimento en mundo real ya están construyendo el nuevo mercado. Por ejemplo, cuando se permite utilizar una medicina en un área controlada determinada del país se está testeando y creando a la vez. (Muniesa y Callon, 2007). En los mercados financieros muchas veces se experimenta con nuevos productos (por ejemplo, derivados exóticos) (MacKenzie y Millo, 2003). Es más, muchas veces las economías nacionales pueden convertirse en un experimento para probar una doctrina económica (Ghannadian y Goswami, 2004). Cada vez más, los experimentos económicos son un elemento fundamental en el proceso de construcción de mercados. Y, sobre todo, en el caso de los experimentos a escala real la controversia abre la posibilidad de ampliar el debate superando el enfoque meramente económico e incluir aspectos políticos, regulatorios y sociales.

Esta nota se propone introducir la problemática del uso de técnicas experimentales en el diseño de nuevos mercados. Para ello la primera sección analiza los elementos que interactúan en la constitución de un nuevo mercado, la segunda se focaliza en el diseño de mercados y su carácter performativo. Por último, la tercera sección plantea la problemática de la eficiencia de los mercados diseñados.

## **1. LA CONSTITUCIÓN DEL MERCADO**

El diseño de mercados es un proceso de mutua interacción; el proyecto necesita ser validado por la experimentación y ésta actúa sobre el primero (Roth, 2008). Ahora bien, estos experimentos pueden ser *in vitro* (desarrollados dentro de un laboratorio) *in vivo* (experimentos en el mercado real) coexistiendo en un proceso de permanente intercambio

(Muniesa y Callon, 2007). Para esto, deben existir redes que organicen y faciliten las relaciones entre ellos de manera de permitir el avance de los conocimientos teóricos sobre los mercados, por un lado, y los dispositivos materiales e institucionales, por el otro. Estos experimentos se han ido estableciendo progresivamente en los mercados y han permitido su mejoramiento permanente. En particular, en los nuevos mercados, donde todo tiene que ser inventado, ni los economistas, ni los agentes económicos habituales pueden diseñarlos sin ayuda. Tienen que cooperar y aceptar que otros actores están involucrados y que, en contextos de incertidumbre, el proceso de diseño debe consistir, necesariamente, en un largo proceso de ensayo y error (Callon, 2009).

La eficiencia de un experimento de mercado radica en organizar la discusión de los asuntos de interés que emergen de su propio funcionamiento, contemplando los desbordes (externalidades). Se deben establecer procedimientos para facilitar la evaluación de soluciones teóricas o prácticas a esos problemas. Este enfoque otorga centralidad a los experimentos y abre los debates; se presenta como un par problemático indivisible, lo económico y lo político (que muchas veces intenta ser excluido) (Callon, 2009).

En los mercados emergen cuestiones donde la incertidumbre es tal que no se sabe cuál es el enfoque adecuado para resolverlos. No es claro si el abordaje debe ser político, económico o tecno-científico. Callon sostiene que ni la economía ni la política ni la ciencia pueden ser consideradas como realidades acabadas y estancas. Callon centra su posición entre el constructivismo social (lo que se considera como el político, económico y científico es simplemente el resultado de un enfrentamiento entre grupos que luchan por imponer sus propios puntos de vista) y el esencialismo (hay una o más definiciones de la política, la economía y la ciencia, que proporcionan criterios objetivos que permiten decir, a priori, si un comportamiento, forma de pensar o dispositivo es político, económico y científico). Los mercados, en fase experimental, resaltan este proceso de reconfiguración conjunta. Este enfoque permite una problematización multidimensional, constituyendo una red de problemas. Es más, la configuración del mercado puede dar prioridad a un enfoque sobre otro y esto no es casual sino que es parte del diseño. Se puede encontrar, entonces, que la dimensión política quede reducida a su mínima expresión.

En el experimento aquéllos que diseñan e implementan los nuevos mercados deben contestar preguntas emergentes, tratando de no encerrarse en organizaciones existentes y permitiéndose innovar en busca de la solución. Incorporando esta problematización en el proceso de diseño se construye una economía política compleja. Asimismo, en este proceso, las ONG se convierten en socios legítimos e inevitables. Este mercado en proceso de diseño evoluciona transformando la economía, la política y la ciencia que lo rodea, por lo que experimentar nuevos mercados tiene un carácter performativo de la realidad y es una acción política. Los procedimientos que se diseñan en el mercado son dialógicos. Se debe permitir que todos los actores concernidos por el diseño y el funcionamiento de un mercado puedan expresarse y, luego, analizar las cuestiones y comparar las soluciones propuestas (Callon, 2009).

## **2. COMPLEJIDAD, EXPERIMENTOS Y DISEÑO**

Si la teoría de juegos va a ser un instrumento fundamental para el diseño, deberá contener no sólo el herramental formal para el desarrollo de conceptos a partir de modelos sencillos, sino también incluir las complicaciones de los mercados concretos. Alvin Roth (2002) señala dos tipos de complicaciones. Primeramente, las complicaciones en el entorno estratégico, en los resultados posibles y en las estrategias disponibles para los jugadores. En segundo lugar, las complicaciones en el comportamiento de los agentes económicos reales que no se comportan como meros maximizadores de beneficio.

El diseño debe anticiparse a cómo las personas se comportarán en el ambiente delineado. Asimismo, es importante asegurar la sustentabilidad del proyecto en el largo plazo, pero verificando su posibilidad de corto plazo. Tradicionalmente, el énfasis formal al analizar sistemas estuvo puesto en los equilibrios en infinito; es indispensable verificar la viabilidad del corto plazo para alcanzar ese equilibrio futuro. Los métodos computacionales cumplen un rol fundamental al permitir el análisis de juegos complejos, imposibles de resolver analíticamente. Los experimentos de laboratorio, por su parte, informan sobre cómo se comportará la gente cuando se enfrenta a estos ambientes que estamos diseñando y cómo adquieren experiencia.

### **3. MERCADOS EFICIENTES**

Roth sostiene que el funcionamiento eficiente de los mercados requiere (1) densidad (para atraer a un número suficiente de participantes), (2) superar los problemas de congestión que puede traer la cantidad de agentes operando (realizar transacciones con suficiente rapidez que permitan a los agentes tomar decisiones), y (3) que sea seguro y simple. Asimismo, debe contemplarse que algunas operaciones están moralmente excluidas, lo cual es una limitación a tener en cuenta al diseñar el mercado. Por último, el autor destaca el rol que tienen los experimentos en el diagnóstico y la comprensión de las deficiencias del mercado, en contrastar el éxito de un diseño y en la comunicación de resultados a los responsables políticos (Roth, 2008).

Tradicionalmente, se reconoce la importancia de la densidad de los mercados, pero no siempre se contemplaron los temas de la congestión, la seguridad y la sencillez. Una forma de evitar la congestión de los mercados es emplear una cámara de compensación centralizada para coordinar el mercado donde un algoritmo aprobado realice las asignaciones. Ahora bien, Roth (1984) demuestra que la evolución de los mercados puede hacer imposible que el algoritmo funcione eficientemente frente a transformaciones. La congestión es un problema, especialmente, en mercados en los que las transacciones son heterogéneas y las ofertas no se pueden hacer a todo el mercado. Aunque las operaciones se realicen rápidamente, al tener que dirigir la oferta hacia agentes particulares, se corre el riesgo de que otras oportunidades desaparezcan.

### **4. CONCLUSIÓN**

Los experimentos<sup>1</sup> son un ejemplo claro de la característica performativa de la economía. Un experimento es un crisol en el que las teorías, discursos, textos, prácticas, intereses y materiales interactúan.

---

<sup>1</sup> (Muniesa y Callon, 2007) presentan una distinción entre tres configuraciones esquemáticas de experimentación que se diferencian por su grado de apertura: el laboratorio, la plataforma, y el experimento *in vivo*.

Al diseñar un nuevo mercado el experimento juega un rol fundamental, tanto para testear ideas como para construir nuevas. Asimismo la eficiencia de este nuevo mercado estará fuertemente influenciada por el desarrollo de experimentos que permitan actuar sobre los problemas encontrados.

Por último, se postula la necesidad de una participación democrática en este proceso de diseño para contemplar los intereses de la sociedad en dicha construcción.

## REFERENCIAS BIBLIOGRÁFICAS

Burlando, RM y Guala, F (2005): "Heterogeneous agents in public goods experiments", *Experimental Economics*, vol. 8, no. 1, pp. 35-54.

Callon, M (2009): "Civilizing markets: Carbon trading between in vitro and in vivo experiments", *Accounting, Organizations and Society*, vol. 34, no. 3-4, pp. 535-48.

Chamberlin, EH (1948): "An experimental imperfect market", *The Journal of Political Economy*, vol. 56, no. 2, pp. 95-108.

Ghannadian, FF y Goswami, G (2004): "Developing economy banking: the case of Islamic banks", *International Journal of Social Economics*, vol. 31, no. 8, pp. 740-52.

Guala, F (2002): "Models, simulations, and experiments", *Model-based reasoning: Science, technology, values*, pp. 59-74.

MacKenzie, D y Millo, Y (2003): "Constructing a market, performing theory: the historical sociology of a financial derivatives exchange", *American journal of sociology*, pp. 107-45.

Muniesa, F y Callon, M (2007): "Economic experiments and the construction of markets", en D MacKenzie, F Muniesa y L Siu (eds), *Do economists make markets*, Princenton Univerisity Pr, Princenton, pp. 163-89.

Roth, AE (1984): "The evolution of the labor market for medical interns and residents: a case study in game theory", *The Journal of Political Economy*, vol. 92, no. 6, pp. 991-1016.

— (2002): "The economist as engineer: Game theory, experimentation, and computation as tools for design economics", *Econometrica*, pp. 1341-78.

— (2008): "What have we learned from market design?" *Innovations: Technology, Governance, Globalization*, vol. 3, no. 1, pp. 119-47.

Say, JB (1841): *Traité d'économie politique ou simple exposition de la manière dont se forment, se distribuent et se consomment les richesses*, vol. 9, Guillaumin.

Smith, VL (1962): "An experimental study of competitive market behavior", *The Journal of Political Economy*, vol. 70, no. 2, pp. 111-37.

— (1989): "Theory, experiment and economics", *The Journal of Economic Perspectives*, vol. 3, no. 1, pp. 151-69.

— (1994): "Economics in the Laboratory", *The Journal of Economic Perspectives*, vol. 8, no. 1, pp. 113-31.

# METODOLOGÍA DE ANÁLISIS DE LAS DESIGUALDADES SOCIALES

*Nora Mabel Lac Prugent*

## 1. INTRODUCCIÓN

Los datos provenientes de las encuestas a hogares son una base para la investigación, para testar teorías acerca del comportamiento de los hogares y para descubrir cómo el pueblo responde a los cambios en el medio económico en el cual viven. Desde fines del siglo pasado, en Argentina, estas encuestas constituyen la base para documentar, entre otros temas, la pobreza y la indigencia en el país.

En este trabajo se expone una revisión teórica de indicadores de pobreza y de distribución de ingreso frecuentemente utilizados señalando sus ventajas y limitaciones. En el intento de avanzar en la medición de la pobreza es necesario tener presente las propuestas de Foster, Greer y Thorbecke, quienes desarrollaron un conjunto de medidas que son descompensables. Para medir la desigualdad de una distribución dada de ingresos existen variados índices, entre ellos, el conocido coeficiente de Gini.

En la sección siguiente se presenta una breve descripción teórica de medidas de pobreza y desigualdad utilizadas. Posteriormente, se muestra su contenido empírico a partir de la Encuesta Permanente a Hogares (EPH), con aplicaciones en el Aglomerado Gran Rosario (AGR) desde el año 1999 hasta el año 2003, momentos en los cuales se dispone de información homogénea. Con el propósito de acercarse a la descripción de los habitantes del AGR en situación de pobreza e indigencia, se trabaja con el ingreso *per capita* familiar (IPCF) en términos de unidades de adulto equivalente.

## **2. MARCO TEÓRICO**

### **2.1. Definición y medición de la pobreza**

En esta sección se puntualizan indicadores que hablan de la extensión, la profundidad y la severidad de la pobreza. Al mismo tiempo, se hace mención a los problemas y a las decisiones metodológicas asociadas con su medición e identificación utilizadas en este trabajo. En la búsqueda de los factores determinantes de las variaciones ocurridas en los últimos años, se indagan características cuantitativas y cualitativas.

Según el Programa de Naciones Unidas para el Desarrollo (PNUD), la pobreza se refiere a la incapacidad de las personas de vivir una vida tolerable. Entre los aspectos más importantes pueden enunciarse: vida saludable, tener educación, vida decente, libertad política, respeto a los derechos humanos, seguridad personal, acceso a un trabajo remunerado, participación en la vida comunitaria, etc. Aspectos de la calidad de vida difíciles, si no imposibles de medir. Dado este impedimento el estudio de la pobreza se ha restringido a sus aspectos cuantificables.

Los métodos cuantitativos utilizados para su medición son dos: el método directo también denominado de las Necesidades Básicas Insatisfechas (NBI) y el método indirecto o enfoque del ingreso.

El método directo se refiere a la tenencia o no de determinados recursos. Las variables que considera son: hacinamiento, vivienda deficitaria, condiciones sanitarias, asistencia escolar y capacidad de subsistencia. Un hogar se considera con NBI cuando posee al menos una de estas características.

El método indirecto determina un monto mínimo con el cual se satisfacen un conjunto de necesidades básicas, Línea de la Pobreza (LP) e identifica a aquellos hogares o personas cuyos ingresos se ubican por debajo de esta línea. El enfoque del ingreso puede adoptar, a su vez, dos criterios: absoluto y relativo. El criterio absoluto fija un umbral normativo mínimo. Este método es el más difundido y utilizado en Argentina por el Instituto Nacional de Estadística y Censos (INDEC); la Comisión Económica para América Latina y el Caribe lo emplea para sus estimaciones regulares de la proporción de hogares pobres para los países de la región. El criterio



relativo compara con un estándar de la sociedad y postula que las necesidades humanas no son fijas y varían de acuerdo a los cambios sociales y a la oferta de productos en un contexto social determinado, dependiendo en última instancia del nivel de ingresos general. Se lo utiliza en países desarrollados.

### 2.1.1. *Indicadores de pobreza. Enfoque axiomático de Amartya Sen*

El problema de medición de la pobreza se asienta en primer término en la identificación de los pobres en el total de la población y en segundo término, en la construcción de índices a partir de la información disponible de los mismos. El primer problema involucra la selección de un criterio de pobreza; por ejemplo, el adoptado por el INDEC consistente en la definición de la *Línea de Pobreza*<sup>2</sup>, para delimitar aquellos que caen debajo de esta línea. El indicador más común para enfocar el segundo problema es simplemente contar  $q$ , el número de pobres y calcular la proporción respecto a la población total  $n$ . Esta proporción -*Head-Count Index*- llamada tasa de incidencia  $H = q/n$  es evidentemente un índice rudimentario y a pesar de sus limitaciones es ampliamente usado.

La tasa de incidencia es completamente insensible a la distribución del ingreso entre los pobres y viola los siguientes criterios necesarios para ser un buen indicador:

✓ Axioma de Monotonicidad: *ceteris paribus*, una reducción del ingreso de una persona debajo de la línea de pobreza debe incrementar la medida de pobreza.

✓ Axioma de Transferencia: *ceteris paribus*, una transferencia pura de ingresos de una persona debajo de la línea de pobreza hacia otra que es más rica, debe incrementar la medida de pobreza.

Otra medida habitual es la llamada brecha de pobreza *PG -Poverty Gap-*. Es la media ponderada de la brecha de ingresos de todos los pobres con respecto a la línea de pobreza, referida a la población total. Por su

---

<sup>2</sup> Para mayor información, ver documentos en <http://www.indec.mecon.gov.ar> (ingresar por la **Búsqueda temática**, capítulo **Trabajo e ingresos**, submenú **Empleo y desempleo**, renglón **Metodología**)

definición suele considerarse como un indicador de la profundidad de la pobreza que satisface el axioma de monotonicidad pero viola el axioma de transferencia. También es insensible al número de personas –o porcentaje de personas- pobres que comparten una brecha de pobreza dada, desventaja fácilmente solucionable normalizándola como se verá más adelante.

La mayor contribución propuesta por Sen (1976) reside en su enfoque axiomático, en el que desarrolla una serie de axiomas, que debe ser cumplido por toda medida de pobreza. Dada una sociedad  $S$  con  $n$  personas, el conjunto de personas con ingresos no mayores que  $x$  es  $S(x)$ . Si  $z$  es la *línea de pobreza*, el nivel de ingresos en el que la pobreza comienza,  $S(z)$  es el conjunto de pobres y  $S(\infty)$  es la población total  $S$ .

La brecha de ingreso de una comunidad,  $g_i$ , para cualquier individuo  $i$  puede expresarse como  $g_i = z - y_i$ , la diferencia entre la línea de pobreza  $z$  y su ingreso  $y_i$ ; resultará no negativa para los pobres y negativa para los otros.

Otra medida, la brecha de ingresos *per cápita*, -llamada  $I$ - se obtiene normalizando la brecha de pobreza:

$$I = \sum_{i \in S(z)} g_i / qz$$
, donde  $g_i = z - y_i$  es la brecha de ingreso y  $q$  es la cantidad de pobres.

Mientras la tasa de incidencia ( $H$ ) calcula la proporción de personas bajo la línea de pobreza, la brecha de ingresos ( $I$ ) da la proporción promedio de carencia a partir del nivel de pobreza. Pero ni  $H$  ni  $I$  brindan información adecuada sobre la distribución exacta de los ingresos entre los pobres.

Sin embargo, en el caso especial en el cual todos los pobres tienen exactamente el mismo nivel de ingreso  $y^* < z$ , puede argumentarse que  $H$  e  $I$  juntas deberían dar adecuada información del nivel de pobreza. De este modo, suministrarán la proporción de personas que están por debajo de la línea de pobreza y la magnitud del déficit de ingresos. Para obtener

una simple normalización, en este contexto el índice de pobreza  $P$  será igual a  $H^*I$ .

✓ Axioma N, valor de pobreza normalizada: Si todos los pobres tienen el mismo ingreso, entonces  $P = H^*I$ .

### 2.1.2. El índice de pobreza derivado

El enfoque axiomático desarrollado determina un índice de pobreza unívoco. Ordenando las personas de acuerdo a un orden no decreciente de ingreso que satisfagan  $y_1 \leq y_2 \leq \dots \leq y_n$ , se puede enunciar la medida propuesta por Sen (1976) que incorpora la desigualdad entre los pobres y cuya definición es:

$$P_s = P_0\gamma^p + P_1(1 - \gamma^p),$$
 donde  $\gamma^p$  es el coeficiente de desigualdad Gini entre los pobres, calculado considerándolos como la población total. Hay que destacar que cuando no existe desigualdad entre los pobres,  $P_s$  se reduce a la brecha de pobreza  $P_I$ . Inversamente, cuando todos menos uno de los pobres no tiene nada,  $P_s = P_0$  y la medida de Sen coincide con la tasa de incidencia  $H$ . Generalizando,  $P_s$  es el promedio de  $P_0$  y  $P_1$  ponderadas por el coeficiente de Gini de los pobres.

A su vez, como  $P_s$  depende del coeficiente de Gini, comparte sus desventajas: no puede ser usada para descomponer la contribución de la pobreza en diferentes subgrupos, lo cual es muy informativo cuando se monitorean cambios en los niveles de pobreza.

La generalización de la medida de Sen propuesta por Kakwani (1980) incorpora el siguiente axioma que contempla la transferencia entre los pobres:

✓ Axioma de Sensibilidad a Transferencias: Si una transferencia de ingreso  $t > 0$  tiene lugar desde un hogar pobre con ingreso  $y_i$  hacia un hogar pobre con ingreso  $y_i + d$  ( $d > 0$ ), entonces la magnitud del incremento en la pobreza debe ser menor cuanto mayor sea  $y_i$ .

Retomando, sea  $\mathbf{y} = (y_1, y_2, \dots, y_n)$  el vector de ingreso familiar en orden creciente,  $z > 0$  la línea de pobreza predeterminada,  $g_i = z - y_i$  el

déficit del ingreso del  $i$ -ésimo hogar,  $g = g(y, z)$  el número de hogares pobres (no poseen un ingreso superior a  $z$ ) y  $n = n(y)$  el número total de familias, se define la medida de pobreza  $P(y, z)$ :

$$P(y, z) = \frac{1}{nz^2} \sum_{i=1}^q g_i^2 .$$

De acuerdo a la propuesta de Sen, la medida  $P(y, z)$  es la suma ponderada normalizada del déficit de ingreso de los pobres. A diferencia de la medida  $P$  de Sen, la cual adopta un esquema ponderado de rango ordenado;  $P(y, z)$  utiliza como ponderadores a los mismos niveles de déficit de ingreso. Así, la privación depende de la distancia entre el ingreso observado del hogar pobre y la línea de pobreza, no del número de hogares que se hallan entre un hogar dado y la mencionada línea.

A pesar de esta diferencia básica en los ponderadores, los axiomas que verifica la medida  $P$  de Sen pueden extenderse a  $P(y, z)$ . Por ejemplo, Sen propone que cuanto más pobre sea el hogar mayor debe ser el ponderador que satisface la medida  $P(y, z)$ . Además, Sen argumenta que los ponderadores deben basarse en la noción de privación relativa experimentada por los hogares pobres.

Volviendo a las propiedades básicas propuestas por Sen y la generalización de Foster, Greer y Thorbecke (1984) desarrollan una clase de medidas consistente ante las descomposiciones; es decir, si aumenta la pobreza para un subgrupo debe aumentar la pobreza total en alguna proporción respecto a la participación de dicho subgrupo en la población total. La nueva medida propuesta por estos autores es conocida en la literatura como medidas FGT que cumplen con el

✓ Axioma de Monotonicidad de Subgrupos: Sea  $\hat{y}$  el vector de ingresos obtenido de  $y$  al cambiar los ingresos de un subgrupo  $j$  proveniente de  $y^{(j)}$  a  $\hat{y}^{(j)}$ , donde  $n_j$  es invariante. Si  $\hat{y}^{(j)}$  es más pobre que  $y^{(j)}$ , luego  $\hat{y}$  debe tener también un mayor nivel de pobreza que  $y$ .

Mientras  $P(y, z)$  no satisface el axioma de Sensibilidad a Transferencias, el conjunto de medidas FGT puede generalizarse de la siguiente manera, para un parámetro  $\alpha$  no negativo:

$$P_{\alpha}(y, z) = \frac{1}{n} \sum_{i=1}^q \left[ \frac{z - y_i}{z} \right]^{\alpha}, \quad \alpha \geq 0 \quad \text{donde:}$$

$n$  = población total

$q$  = cantidad de pobres

$y_i$  = ingreso del  $i$ -ésimo individuo

$z$  = línea de pobreza

$\alpha$  = parámetro no negativo, medida de aversión a la pobreza

Si  $\alpha = 0$ ,  $P_0(y, z) = H$ , la tasa de incidencia o índice de extensión de la pobreza.

Si  $\alpha = 1$ ,  $P_1(y, z)$  es la brecha de pobreza, índice de profundidad de la pobreza.

Si  $\alpha = 2$ ,  $P_2(y, z)$  indica la severidad de la pobreza, un gran  $\alpha$  da un mayor énfasis a los más pobres.

La medida de pobreza  $P_{\alpha}$  satisface el Axioma de Monotonicidad para  $\alpha > 0$ , debido a que  $g_i$  aumenta cuando disminuye  $y_i$ . El Axioma de Transferencia se cumple para  $\alpha > 1$ , la convexidad estricta de  $P_{\alpha}$  en el vector ingreso de los pobres asegura que una transferencia de una familia pobre a otra familia pobre aumenta la medida de pobreza. Por último, el Axioma de Sensibilidad de Transferencia se cumple para  $\alpha > 2$ , ya que una transferencia de una familia pobre a otra familia que se encuentra sobre la línea de pobreza aumenta la medida de pobreza.

Así descriptas, las medidas de pobreza muestran sus diferentes dimensiones: la tasa de incidencia mide el *predominio* de la pobreza, la

brecha expresa la *profundidad* de la pobreza y las medidas sugeridas posteriormente -Sen y sus variantes- indican la *severidad* de la pobreza.

### *2.1.2.1. Una digresión acerca de la confiabilidad de las medidas de pobreza*

En muestras complejas resulta dificultoso evaluar la variabilidad de las estimaciones resultantes. El *bootstrap* es un método alternativo para calcular la variabilidad muestral, inventado por Efron (1979), muestrea repetidas veces a partir de la muestra, no a partir de la población. Cada remuestreo replica las estimaciones y realiza un seguimiento de los resultados a través de cada una de ellas. La variabilidad de estas estimaciones remuestreadas es luego usada para evaluar la variabilidad del estimador a través de las distintas muestras provenientes de la población.

El caso más sencillo es el muestreo simple aleatorio con un tamaño muestral de *n*- *hogares*. El *bootstrap* extrae repetidamente muestras de tamaño *n* a partir de la muestra *con reposición*. En cada replicación, el estadístico de interés –media, mediana, variancia- es calculado y almacenado. Después de *k* replicaciones, los *k* valores del estadístico son usados para calcular la medida de dispersión; por ejemplo: el desvío estándar como una medida del error estándar, o –necesariamente en los casos donde los momentos pueden no existir- los percentiles usados para estimar percentiles de la distribución muestral. El valor de *k* variará según la aplicación.

## **2.2. Los indicadores de la distribución de los ingresos**

Las características de la concentración de los ingresos se pueden estudiar desde la doble perspectiva de las curvas de concentración de la distribución -curva de Lorenz- y el índice de concentración, también llamado de Gini.

En realidad, la naturaleza del coeficiente de Gini revela que no se trata de un verdadero instrumento de medición de la concentración, más bien de una medida de desigualdad, quiere decir de la variación respecto a una distribución perfectamente igualitaria. No puede detectar movimientos de la concentración.

Un índice de desigualdad que se deriva de la curva de Lorenz es el *Coefficiente de Gini*. Esta medida no paramétrica por definición, toma valores cercanos a cero si la curva de concentración es cercana a la bisectriz del primer cuadrante. Mide el área existente por debajo de la recta de perfecta igualdad y por encima de la curva de Lorenz. Cuanto menor es esta área, mas igualitaria es la distribución y menor es el valor que toma. Entre las ventajas del coeficiente de Gini se puede mencionar que es invariante a la escala que se utiliza para medir el ingreso, satisface la condición de simetría -si dos individuos intercambian sus respectivos ingresos, manteniéndose iguales los demás, su valor no se altera- y cumple el principio de trasferencias de Dalton-Pigou. La esencia de este principio radica en que una transferencia de ingresos desde una unidad de mayores ingresos a otra de menores debe reducir la desigualdad, siendo este cambio capturado en una baja en el indicador que se esté utilizando. A pesar de sus desventajas, el coeficiente de Gini es ampliamente usado debido a que es fácil de interpretar en términos geométricos. En el intento de medir la desigualdad de una distribución dada de ingresos existen variados índices. Como a priori, no puede afirmarse cual es el mejor, se presenta un detalle de los índices utilizados en este trabajo.

El coeficiente de variación (CV) es el cociente entre el desvío y la media de ingresos de la población, asigna igual ponderación a las transferencias de ingresos realizadas en cualquier tramo de la distribución, generalmente expresado en porcentajes. El cociente entre la participación del último y el primer quintil (Q5/Q1) otorga igual ponderación a las transferencias realizadas en ambos extremos de la distribución. Su principal inconveniente es que puede no captar el efecto de transferencias realizadas entre los quintiles segundo a cuarto.

### **3. ANÁLISIS DE LOS RESULTADOS**

A pesar de las limitaciones ocasionadas por la fuente de información básica disponible se trata de ser lo más riguroso posible en los posteriores análisis, teniendo como aval la robustez de las mediciones y su relevancia estadística.

#### **3.1. Desigualdad y pobreza en el Aglomerado Gran Rosario**

Para determinar si el hogar es pobre o no, se calcula su composición en términos de unidades de adulto equivalente (UAE). Es decir, cuántas UAE representan los miembros del hogar y se calcula su IPCF en términos de UAE -consistente en dividir el ingreso total del hogar por el total de UAE que representa la estructura del mismo-. En adelante, para evitar confusiones, el ingreso *per capita* familiar en términos de unidades de adulto equivalente se denominará Ingreso Familiar Equivalente (IFE). Para ello, se aplica la escala de adulto equivalente oficial del INDEC.

Los cálculos del coeficiente de Gini sobre la distribución del ingreso familiar equivalente para todos los hogares y el coeficiente de variación indican un incremento de la desigualdad hasta 2002 para luego disminuir, revelando un comportamiento similar al del IPCF del AGR.

En el intento de avanzar en la medición de la pobreza es necesario tener presente las propuestas de FGT, quienes desarrollaron un conjunto de medidas que son descompensables y cuya expresión general se presentó anteriormente.

Recordando que la tasa de incidencia es la proporción de pobres quienes viven en hogares con ingresos menores que la línea de pobreza ( $y_i \leq z$ ), respecto del total de la población. Hay que destacar el aumento de la incidencia de la pobreza entre los años considerados con marcado pico en el período 2002- 2003, años de la última peor crisis de la historia argentina.

Una forma de continuar realizando ranking de situaciones de pobreza es mediante  $P_1$  y  $P_2$  que representan la evolución de la situación interna de los pobres calculados, de la forma habitual, como promedios ponderados de toda la población. La primera, puede ser también interpretada como una medida *per capita* del déficit total del nivel de bienestar individual bajo la LP; es la suma de los déficits dividido por la población.

La brecha de pobreza se incrementará por una transferencia desde un pobre a uno no pobre o desde un pobre a uno menos pobre quien luego se transformará en no pobre. Pero la transferencia entre los pobres no tiene efectos sobre la medida de pobreza, en esta situación se desearía considerar otras medidas. Es por ello, que en el cuadro N° 1 se incorpora la medida de Sen (1976) que remedia este inconveniente incorporando la



desigualdad entre los pobres y cuya definición es  $P_s = P_0\gamma^p + P_1(1-\gamma^p)$ , donde  $\gamma^p$  es el coeficiente de desigualdad Gini entre los pobres, calculado considerándolos como la población total.

Suponiendo los valores de las líneas de pobreza calculadas por el INDEC para cada período y año por mes por adulto equivalente, se estimaron la tasa de incidencia  $P_0$ , el índice de profundidad de la pobreza  $P_1$ , la medida FGT con  $\alpha= 2$  y la medida de pobreza de Sen. Se trabaja con la escala de adulto equivalente oficial, como en todos los cálculos anteriores. Posteriormente, para conocer qué tan confiables son los resultados obtenidos, se computaron los errores estándares para estas medidas que provienen de 100 replicaciones utilizando la técnica de *bootstrap*.

**Cuadro 1. Medidas de pobreza, Aglomerado Gran Rosario, ondas mayo. Entre paréntesis los errores estándares del *bootstrap***

| Año  | Tasa de incidencia, $P_0$ | Brecha de Pobreza, $P_1$ | Indice FGT, $P_2$ | Indice de pobreza de Sen, $P_s$ |
|------|---------------------------|--------------------------|-------------------|---------------------------------|
| 1999 | 0.1953 (.0153)            | 0.0686 (.0062)           | 0.0356<br>(.0042) | 0.0951 (.0083)                  |
| 2000 | 0.2431 (.0148)            | 0.1048 (.0085)           | 0.0666<br>(.0072) | 0.1455<br>(.0115)               |
| 2001 | 0.2743 (.0197)            | 0.1272 (.0106)           | 0.0865<br>(.0086) | 0.1769<br>(.0143)               |

|      |                |                |                   |                |
|------|----------------|----------------|-------------------|----------------|
| 2002 | 0.4559 (.0489) | 0.2377 (.0127) | 0.1696<br>(.0117) | 0.3207 (.016)  |
| 2003 | 0.4832 (.0259) | 0.2466 (.0164) | 0.1702<br>(.014)  | 0.3299 (.0201) |

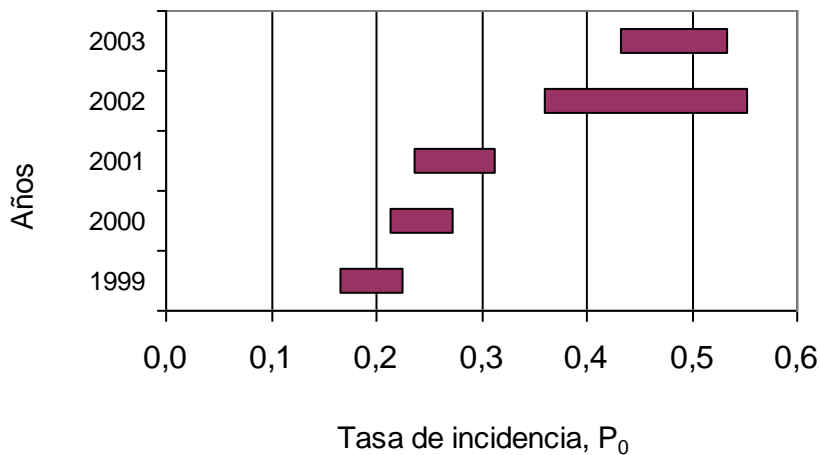
**Nota:** Se utilizó el Ingreso Familiar Equivalente, calculado sobre la base hogares

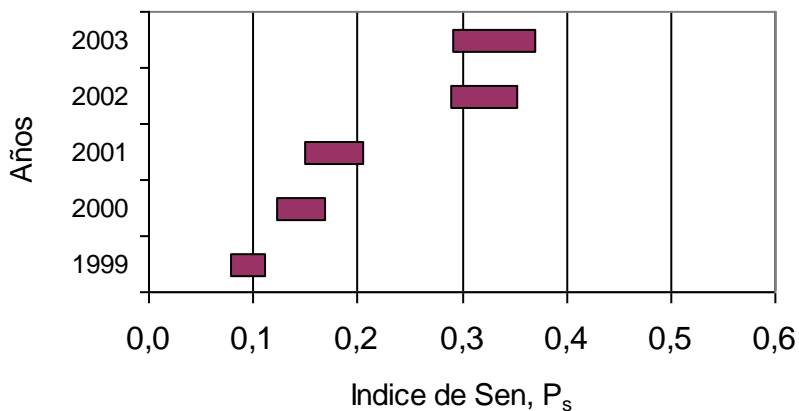
**Fuente:** Elaboración propia basada en EPH, INDEC-IPEC

Aunque el Promedio del Ingreso Total Familiar disminuye abruptamente desde mayo de 1999 hasta mayo de 2001, las medidas de pobreza se mantienen virtualmente inalteradas. Estas estimaciones proveen un buen ejemplo del hecho de que la pobreza no siempre se mueve en la misma dirección que los ingresos. Además, la tasa de crecimiento de  $P_0$  en el período es menor que la de  $P_1$  y a su vez la de  $P_1$  es menor que la de  $P_2$  reflejando la situación interna de los *nuevos pobres* que se fueron incorporando, en términos distributivos, son mucho peores que la que enfrentaban los *pobres anteriores*. Los valores del cuadro anterior se pueden visualizar en los siguientes gráficos, los cuales muestran claramente un cambio estructural a partir de la primera onda de mayo del año 2002. La construcción de los intervalos de confianza incluyen los errores estándares provenientes del *bootstrap*.

Esta propuesta, está desarrollada con la ayuda del análisis de algunos casos donde la pregunta de la medición puede relacionarse con dificultades tanto teóricas como empíricas. El vacío estadístico, que surge de la mala medición o la difícil medición de los fenómenos sociales, está estrechamente relacionado con el vacío de las ciencias sociales. Solamente la interacción entre los estadísticos y los científicos sociales puede aportar un progreso continuo de la ciencia social y un mejoramiento de la calidad de la estadística social. Es a las teorías sociales, a sus explicaciones y a sus interpretaciones, que los estadísticos deben dirigirse para afrontar los nuevos desafíos de la medición social.

**Gráfico 1. Intervalos de confianza de las medidas de pobreza utilizadas en el Aglomerado Gran Rosario. Ondas mayo 1999-2003**

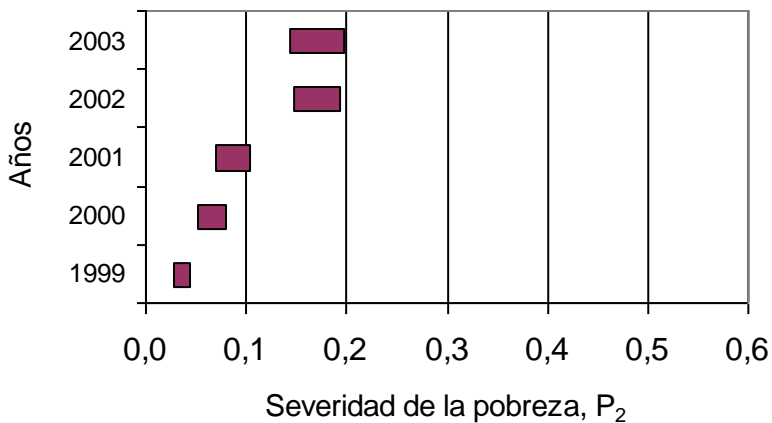
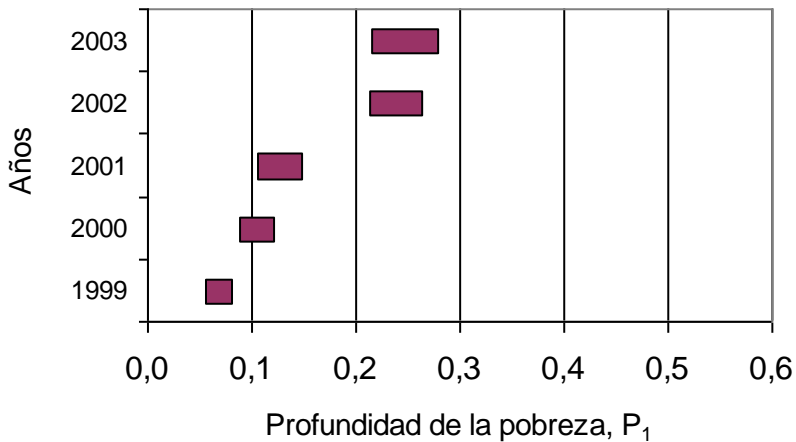




Fuente: Elaboración propia en base a EPH, INDEC-IPEC

#### 4. ALGUNAS REFLEXIONES

En esta exposición resulta evidente que las políticas puestas en práctica en el período considerado, no han logrado atenuar la falta de equidad en la distribución del ingreso y no han logrado frenar el aumento en los niveles de pobreza e indigencia de la población.



De acuerdo a la metodología aplicada y a la evidencia empírica se sugieren, para ulteriores trabajos, tener en cuenta las siguientes reflexiones:

- ✓ Tener siempre presente la población de referencia, la población de la cual fue extraída la muestra. Es frecuente encontrar estudios que hacen mención a la República Argentina y en verdad están reflejando la realidad solamente del Aglomerado Gran Buenos Aires.

- ✓ La necesidad de trabajar con valores expandidos cuando los datos provienen de encuestas muestrales.
- ✓ Analizar el fenómeno de las no respuestas especialmente en las variables sensibles como los ingresos percibidos.
- ✓ Debido a la poca precisión de los deciles de las variables relativas a los ingresos se aconseja trabajar a nivel de quintiles.
- ✓ Los valores de las Canastas Básicas Alimentaria y Total están calculados para el Gran Buenos Aires. El coeficiente de Engel se actualiza con la variación de los precios correspondientes del Índice de Precios al Consumidor, que no corresponden a cada región estadística.
- ✓ Recordar que el método de las NBI y hogares pobres puede ser interpretado como la medida de pobreza del jefe del hogar.

## REFERENCIAS BIBLIOGRÁFICAS

Altimir, O., Beccaría, L. y González Rozada M. (2002): "La distribución del ingreso en Argentina", 1974-2000 en *Revista de la Cepal*. 78: 55-85.

Deaton, A. (2000): *The Analysis of Household Surveys. A Microeconomic Approach to Development Policy*. Washington: The Johns Hopkins University Press (3ra. Edición).

Efron, B. (1979): "Bootstrap methods: another look at the jackknife" en *Annals of Statistics*. 7: 1-26.

Foster, J., Greer, J. y Thorbecke E. (1984): "A class of decomposable poverty measures" en *Econometrica*. Vol. 52, 3: 761-766.

Sen Amartya. (1995) *Nuevo examen de la desigualdad*, Madrid, Alianza Ed.

Sen, A. (1976), "Poverty: An Ordinal Approach to Measurement" en *Econometrica*, Vol. 44, 2: 219-231.

# **LAS ESCALAS COMO INSTRUMENTO DE MEDICIÓN DE ACTITUDES. ARTICULACIÓN CON LA INVESTIGACIÓN DE MERCADOS**

*Ana Marsanasco*

## **INTRODUCCIÓN**

Medir significa asignar numerales a objetos o acontecimientos de acuerdo con ciertas reglas (Stevens, 1951). La medición es posible porque existe un isomorfismo parcial entre las propiedades de la serie numeral y las operaciones empíricas que podemos realizar con los objetos. Esto es, algunas propiedades de la serie numeral pueden ponerse en correspondencia biunívoca con algunas propiedades de los objetos, mediante reglas semánticas.

En este sentido parece sencillo, por ejemplo, proponer llamar "X a la relación entre motivación en el trabajo y liderazgo del gerente", sin embargo, surge enseguida el interrogante sobre la connotación que poseen conceptos como "motivación en el trabajo" y "liderazgo".

Otros conceptos tales como "cantidad de empleados", "edad", "número de ausencias anuales", entre otros, representan características o atributos relativamente concretos que no presentan un problema para su definición y medición de forma directa. Con todo, no siempre es sencillo llegar a una definición satisfactoria de los conceptos cuando éstos son abstractos. La dificultad radica en idear definiciones que sorteen favorablemente "la prueba operacional del significado" (Steven, 1951). Es decir, otorgar significado al concepto en la realidad empírica.

Conceptos abstractos como los mencionados abundan en las investigaciones de las ciencias sociales y las técnicas científicas, tal como la Administración. Así, se puede tener la impresión de que la definición propuesta por Stevens (1951) no es tan adecuada en el marco de estas

disciplinas, ya que los fenómenos medidos son conceptos abstractos que no existen como entidades físicas.

Desde esta perspectiva, Hernández Sampieri y otros (1997) introducen en el concepto de medición la noción de indicadores: medir es un proceso que vincula conceptos abstractos con indicadores empíricos a través de la cuantificación y/o clasificación.

Aclarada la cuestión en torno a las nociones de medición, el problema se enfocaría en decidir cuáles son las reglas para asignar numerales a los conceptos abstractos.

Las reglas semánticas son básicamente arbitrarias, en el sentido que somos los investigadores quienes las creamos para nuestra conveniencia siempre y cuando sigan una regla lógica.

Una regla para asignar numerales a conceptos abstractos crea una escala o nivel de medición. La asignación se puede llevar a cabo de acuerdo con diferentes reglas que, a su vez, originan diferentes tipos de niveles: nominal, ordinal, de intervalo, o de razón. Las características de cada uno son ampliamente conocidas por lo que no profundizaremos sobre este tema. Aunque consideramos apreciable mencionar que la clasificación surge a partir de las comparaciones matemáticas que pueden aplicarse y, por consiguiente, las conclusiones (fruto del análisis de los datos) están limitadas por las propiedades matemáticas de cada nivel.

Las consideraciones hechas hasta aquí nos demuestran que medir un concepto implica elegir un nivel de medición, pero en la definición y medición no acaba la investigación. Es necesario, entre otras actividades adicionales, elegir y diseñar el o los instrumentos más adecuados para relevar los datos.

A los efectos de los objetivos de este trabajo, nos interesa considerar uno de los instrumentos de medición: las escalas, lo haremos a continuación.



## **1. ESCALAS COMO INSTRUMENTOS DE MEDICIÓN**

El término "escala" se emplea para denotar a instrumentos de medición, particularmente, los que miden actitudes, y será con esta connotación que recurriremos al término en esta exposición.

Las escalas de medición de actitudes son una de las herramientas más empleadas en las investigaciones de la conducta humana. Louis Leon Thurstone (1976) fue uno de los primeros en suscitar la medición de actitudes. En su trabajo "Attitudes can be measured" las define como el conjunto de inclinaciones y sentimientos, prejuicios o tendencias, nociones preconcebidas, ideas, temores y convicciones de un hombre respecto de un tema determinado.

Generalmente se acepta que las actitudes están conformadas por tres componentes relacionados: cognoscitivo, afectivo y de comportamiento.

El componente cognoscitivo representa la información que un individuo posee acerca de un concepto o fenómeno. Esta información no sólo incluye el conocimiento de la existencia del concepto sino también las creencias de sus características e importancia relativa.

El componente afectivo resume las emociones que un individuo expresa hacia un concepto o fenómeno. Cuando es posible escoger entre un número de alternativas, este componente se expresa en términos de preferencias por una alternativa sobre otra.

El componente de comportamiento se refiere a las expectativas de un individuo en relación con el comportamiento futuro respecto de un fenómeno. Refleja una predisposición a la acción.

De estos tres componentes, las escalas suelen medir el afectivo, principalmente, las preferencias. En estos casos, y a ello también se debe, los niveles de medición de sus ítems o reactivos son ordinales.

Ciertamente este tema abre el debate pues, si bien las mediciones son ordinales, es común trabajarlas como si fueran de intervalos. Pero tal supuesto no modifica el sentido estricto ordinal de la medición y, por lo

tanto, las interpretaciones de los resultados deben reconocer los límites fijados por propiedades matemáticas.

Es fácil encontrar un ejemplo de lo que estamos aquí aludiendo. La Escala de Likert es un instrumento de medición de actitudes de nivel ordinal y como tal no indica en cuánto es más favorable o desfavorable la actitud de una persona hacia el fenómeno o concepto medido. Así, si una persona obtiene 40 puntos esto nos informa que tiene una actitud más favorable que la persona que obtuvo 20 puntos, y no podremos concluir que la distancia es dos veces más grande.

Además de la Escala de Likert, existen otras escalas para medir actitudes. Entre ellas se destacan: la escala de intervalos de Thurstone; el Escalograma de Guttman; y el diferencial semántico de Osgood, Suci y Tannenbaum. El lector interesado en ampliar su información al respecto puede consultar la obra de Catalina Wainerman, C. *et alia* (1976) *Escalas de medición en ciencias sociales*.

De lo que hemos venido analizando, podemos deducir que las escalas son métodos indirectos para medir las actitudes a través de expresiones verbales (afirmaciones, proposiciones o juicios) sobre las cuales se infiere en qué grado un individuo posee los atributos que definen el fenómeno que mide la escala. Dicho en otras palabras, la relación entre los ítems o reactivos de la escala y la actitud que definen no es de certeza, sino de probabilidad. Situación que deriva en al menos dos cuestiones significativas: 1) la necesidad de enunciar un gran número de reactivos y, 2) entender las actitudes como señales y no como hechos de la conducta.

Estrechamente vinculado a lo anterior detectamos la limitación que también Thurstone (1976) observó: la medición de actitudes expresadas por opiniones no implica necesariamente una predicción de lo que hará una persona, ya que sólo reflejan una predisposición a la acción. Ante esta limitación, si bien no hay mediciones perfectas y es ciertamente imposible representar fielmente las actitudes humanas; pretendemos acercarnos lo más que podamos. Es así como el autor sugirió emplear las escalas de actitudes sólo en aquellas situaciones donde es razonable esperar que la gente diga la verdad acerca de sus convicciones u opiniones.

Indudablemente, no era interés de Thurstone (1976) determinar la coherencia entre opiniones y acciones porque su intención no fue predecir conductas manifiestas. La opinión verbal expresada por las personas sólo era útil como indicador de la actitud, independientemente de las creencias o conductas.

La coherencia entre opiniones y acciones, en cambio, sí fue estudiada por un grupo de científicos y filósofos de la Universidad de California (Berkeley) y la Universidad de Frankfurt: Theodor Adorno, Else Frenkel-Brunswick, Daniel Levinson y R Nevitt Sanford. El trabajo que realizaron fue impresionante y se plasmó en la obra *La personalidad autoritaria*. En ella detallan con gran rigurosidad y sencillez los lineamientos y estudios que derivaron en la construcción de la Escala F para medir las tendencias antidemocráticas. Seguidamente, pasamos a comentar los aspectos centrales de esta investigación.

## **2. EL TRASFONDO DE LA ESCALA F: CONSTRUCCIÓN DE ÍTEMS DELATORES**

En conformidad con los lineamientos teóricos de Thurstone (1976), Adorno y sus colaboradores (1965) sostenían que las opiniones, las actitudes y los valores se manifestaban a través de las expresiones verbales, pero, en cambio, creían que podían existir discrepancias entre lo que un sujeto dice y lo que realmente piensa: lo que se dice, y en menor grado lo que realmente se piensa, dependen en gran medida de la situación en que se encuentre el sujeto y del contexto en el que vive.

Conocer qué clase y con qué intensidad se presentaban las expresiones verbales y los prejuicios en un sujeto para llevarlo a la acción, revestía importancia práctica para la investigación que querían llevar a cabo.

La investigación comenzó estudiando la naturaleza del individuo potencialmente fascista. Los autores aclaran que el término "potencial" se introduce porque no han estudiado individuos declaradamente fascistas o que pertenecieran a organizaciones afines. El estudio se centró en este individuo porque en él subyace una estructura de personalidad tal que lo hace especialmente susceptible a la propaganda fascista.

Se respondieron los siguientes interrogantes: si existen individuos fascistas potenciales, ¿cómo son exactamente?, ¿cómo se forma su pensamiento antidemocrático?, ¿cuáles son sus fuerzas de organización interna? Si tales personas existen, ¿cuáles han sido los factores determinantes y el curso de su desarrollo?

Era fundamental, en este sentido, acceder a las tendencias profundas, subyacentes del individuo porque allí se encontraba el pensamiento y la acción potencialmente fascista ante situaciones críticas.

El marco teórico elaborado reunía, entre otros, lineamientos relacionados con la psicología y la estructura de la personalidad de Freud. Éstos llevaron a considerar que en la ideología intervenían tanto factores situacionales como de personalidad.

Los factores situacionales, en especial la condición económica y la pertenencia a los grupos sociales, se habían estudiado en recientes investigaciones. En tanto los factores internos referidos a la personalidad, no había recibido la suficiente atención.

A fin de medir las tendencias subyacentes de los individuos, diseñaron métodos tanto para describir y medir tendencias ideológicas como para revelar la personalidad, la situación y el medio social, que incluso abarcaron estudios clínicos profundos. Luego, elaboraron un cuestionario que incluía preguntas de hecho, escalas de opinión-actitud (que luego conformarían la Escala F) y preguntas proyectivas (de respuesta abierta).

Los datos se relevaron en Estados Unidos durante 1945. Los primeros grupos estudiados fueron los estudiantes universitarios. Sin embargo, las limitaciones de una muestra universitaria no sólo vería afectada la significatividad general de la investigación sino que, además, el nivel educativo relativamente alto reducía el número de individuos extremadamente prejuiciosos, y algunos de los factores que se suponían influían en el prejuicio (a partir de los estudios clínicos profundos realizados) estaban poco o nada presentes. Estas reflexiones llevaron a estudiar a otros grupos de sujetos: personas que pertenecían a grupos políticos, religiosos, laborales, de ingresos y sociales. Se estimó que estos

individuos reunían la mayor cantidad de variables sociológicas relevantes para el estudio.

Ahora bien, retengamos lo que a nosotros nos interesa. Desde el principio, para conseguir estimaciones cuantitativas de las tendencias ideológicas subyacentes, se diseñaron escalas de opinión-actitud. Los resultados de la prueba piloto y los análisis de correlación fijaron los ítems que finalmente integrarían la escala para medir las tendencias fascistas de la personalidad, denominada Escala F.

El desarrollo de esta escala, mencionan los autores, se realizó de dos maneras: primero formulando ítems que, aunque no tuvieran una conexión manifiesta con expresiones abiertamente antidemocráticas, estuvieran, no obstante, altamente correlacionadas con ellas. Y, en segundo lugar, demostrando que estos ítems "indirectos" eran realmente expresiones del potencial antidemocrático interno de la personalidad a partir de los estudios clínicos realizados. Luego, una persona se consideraría potencialmente antidemocrática en su estructura dinámica subyacente si hubiera demostrado acuerdo con la mayoría de los ítems de la Escala F.

De esta manera, entendemos que la creación y formulación de ítems "delatores" de tendencias subyacentes en la personalidad representa un aporte destacable que, por cierto, ha quedado desplazado por la importancia que originó la creación de una Escala para medir disposiciones a expresar ideas fascistas o sufrir su influencia.

Adorno y sus colaboradores se dieron cuenta de que a menudo los individuos no están dispuestos a aceptar hechos o situaciones que reflejen de modo negativo su autoimagen, o bien dar cuenta de los motivos para determinadas conductas o comportamientos. Y en estos casos suelen ofrecer una respuesta socialmente aceptada en el contexto que no revela las verdaderas actitudes. El fascismo es un tema sobre el que las personas no suelen estar dispuestas para hablar con absoluta franqueza.

Como forma conjugada de pensamiento y acción, en ningún momento se le comentó a los sujetos cuál era el propósito concreto de la

investigación, se les dijo, en cambio, que formaban parte de un "estudio de opinión acerca de diversas cuestiones de actualidad".

Ahora bien, dentro del amparo de la perspectiva teórica presentada, encuentran refugio muchos de los métodos actualmente empleados en la investigación de mercados. A esto nos referiremos en el siguiente apartado.

### **3. USOS POTENCIALES DE ÍTEMS INDIRECTOS EN LA INVESTIGACIÓN DE MERCADOS**

La investigación de mercados nos permite relevar datos que pueden ser de utilidad tanto para un emprendedor como para un empresario. En el caso del emprendedor, cuando se va a iniciar un negocio se necesita llevar a cabo una investigación de mercados porque se requerirá reunir datos, analizarlos y presentarlos en el plan de negocios a fin de demostrar su viabilidad.

El emprendedor podrá obtener algunos datos de fuentes secundarias, mientras que otros tendrán que ser recopilados de fuentes primarias utilizando algún instrumento para ello.

En el caso del empresario, la investigación de mercado suelen tener como objetivo recolectar datos para conocer a los clientes actuales o potenciales, analizar el lanzamiento de un nuevo producto o servicio, medir la satisfacción de los clientes, llevar a cabo pruebas de precios para conocer cuál es el mejor precio al que están dispuestos a comprar los clientes, testear la aceptación del nombre de un producto, conocer cuáles de las cualidades del producto o servicio a ofrecer son las más importantes para el cliente, detectar problemas actuales o potenciales del producto o servicio, entre otras.

En este sentido, la investigación de mercados se vale de un gran número de instrumentos para medir las actitudes de los clientes actuales o potenciales. Las técnicas proyectivas, las escalas y la observación son algunos de los métodos más empleados.

Las técnicas proyectivas, tales como la asociación de palabras, pruebas de frases incompletas, interpretación de dibujos, técnicas de la tercera persona, desempeño de roles, entre otras, se utilizan cuando el investigador sospecha que los entrevistados no responderán si se les formulan preguntas directas, o brindarán respuestas socialmente aceptadas.

Gran parte de las limitaciones de estas técnicas provienen de la ambigüedad en los resultados. De acuerdo con Aaker y Day (1990), el entrevistador cuenta con la flexibilidad para interpretar selectivamente los pensamientos y comentarios de los entrevistados y dar apoyo a una determinada perspectiva. En consecuencia, los resultados de las técnicas proyectivas proporcionan indicios e hipótesis sobre la realidad de la perspectiva del consumidor para una investigación mayor posterior.

A partir de la discusión sobre los ítems indirectos de la Escala F, podríamos hacer una suposición bastante acertada: la utilización de Escalas con ítems indirectos como instrumento complementario en el desarrollo de las técnicas proyectivas para medir actitudes, reduciría las limitaciones de esta técnica puesto que existe una metodología específica para su medición, o sea, se reduce la incidencia de la subjetividad del entrevistador.

Desde esta perspectiva, la aplicación de escalas haría posible representar de manera cuantitativa el grado que ocupa una marca, creencias o la imagen que ha construido un sujeto acerca de una organización, así como también analizar el comportamiento de los consumidores. En efecto, si bien la observación es la técnica principalmente aplicada en estos casos, la construcción de escalas con ítems indirectos se presenta como una alternativa plausible y, al mismo tiempo, complementaria en tanto se elaboren ítems delatores de las razones que conduzcan a comprar o no un determinado producto. Así, una vez más se puede preguntar aquí: ¿por qué utilizaríamos ítems indirectos? A causa de que es muy probable que la respuesta obtenida, si formulamos preguntas directas, no revelen las verdaderas actitudes de compra de los consumidores.

Consideremos ahora el siguiente ejemplo. Digamos que se quiere emprender un nuevo negocio y se desea transferir el conocimiento de las mejores prácticas de un competidor existente al negocio propio. En este contexto, la observación sería la técnica más adecuada para relevar datos del competidor en el local de ventas, ya que no sería viable encuestar a los clientes en el lugar.

Por ende, abrigamos la convicción de que lo expuesto no lleve a la conclusión de que hay métodos mejores que otros, sino que los hay adecuados e inadecuados de acuerdo a la investigación que se desea realizar.

#### **4. CONCLUSIONES**

Tal como adelantamos en la introducción, nuestro propósito en este trabajo se centró en analizar las escalas como instrumentos de medición de actitudes.

Con este objetivo, expusimos definiciones del concepto de medición así como sus limitaciones. Estas últimas nos condujeron a la presentación de los aspectos más relevantes de la investigación llevada a cabo por Adorno y sus colaboradores, no sólo como una propuesta superadora de tales restricciones sino, esencialmente, para revelar lo que, a nuestro entender, es uno de los principales aportes de la construcción de la Escala F: la elaboración rigurosa de ítems delatores de atributos representantes de actitudes subyacentes en la conducta del individuo.

Vinculado a lo anterior, para terminar, introdujimos contextos de investigación en los cuales la formulación de reactivos delatores o indirectos es una técnica adecuada para medir actitudes.

#### **REFERENCIAS BIBLIOGRÁFICAS**

Aaker, D. y Day, G. (1990): *Investigación de mercados*. Cali, Colombia, McGraw-Hill.



Adorno, T., Frenkel-Brunswik, E., Levinson, D. y Nevit Sanford, R. (1965): *La personalidad autoritaria*. Buenos Aires, Proyección.

Hernández Sampieri, R., Fernández Collado, C., Baptista Lucio, P. (1997): *Metodología de la investigación*. Buenos Aires, McGraw-Hill.

Likert, R. (1976): "Una técnica para la medición de actitudes", en Wainerman, C. (comp.), *Escalas de medición en ciencias sociales* (pp. 199-260). Buenos Aires, Nueva Visión.

Osgood, C., Suci, G. y Tannenbaum, P. (1976): "El diferencial semántico como instrumento de medida", en Wainerman, C. (comp.), *Escalas de medición en ciencias sociales* (pp. 331-370). Buenos Aires, Nueva Visión.

Stevens, S. (1976): "Matemáticas y medición", en Wainerman, C. (comp.), *Escalas de medición en ciencias sociales* (pp. 15-74). Buenos Aires, Nueva Visión.

Thurstone, L. (1976): "Las actitudes pueden medirse", en Wainerman, C. (comp.), *Escalas de medición en ciencias sociales* (pp. 15-74). Buenos Aires, Nueva Visión.

# **EL DISEÑO EXPERIMENTAL EN EL CONTEXTO DE LA GESTIÓN: NOTAS SOBRE SU IMPLEMENTACIÓN EN LA EVALUACIÓN DE POLÍTICAS PÚBLICAS. EXPERIENCIAS, ALCANCES Y LIMITACIONES**

*Adriana Fassio*

*María Gabriela Rutty*

## **1. LA EVALUACIÓN DE POLÍTICAS SOCIALES Y LA EVALUACIÓN COMO PARTE DE LA GESTIÓN**

Evaluación, desde nuestro punto de vista, no significa plantearse solamente los errores cometidos en procesos complejos como los de implementación de políticas sociales, sino por el contrario, nuestra visión es la de aprendizaje, es decir la de sistematizar las buenas prácticas y mejorarlas y de analizar lo que se considera que no resultó de acuerdo a lo esperado para superarlo. Es decir que la evaluación consistiría en un proceso a partir del cual se explora de manera sistemática y rigurosa el cumplimiento de las actividades y de la entrega de productos o servicios en función del logro de un cambio sostenible. La evaluación se constituye en una herramienta básica de la gerencia social adaptativa y orientada a resultados que intenta medir la generación de valor social, es decir, la mejora de la calidad de vida de la población sujeto de intervención. Es parte de un proceso integral de la gerencia de programas y organizaciones. Desde esta concepción la evaluación integral trascendería el paradigma meramente administrativo formal que solo indaga sobre el cumplimiento de las actividades y /o entrega de los productos y servicios (Mokate, 2003:2-9).

La información que se obtiene a través de la evaluación permite mejorar la efectividad cuantitativa de los programas, la calidad de los resultados, aumentar la eficiencia, adecuar los objetivos del servicio a las expectativas de la población objetivo, controlar la intervención de factores que dificultan la marcha o el logro y comprobar la relación entre el servicio

y los resultados buscados (Briones, 1999). Es decir, comprobar que los esfuerzos que se realizan están bien orientados y reajustar las acciones de ser necesario (función de retroalimentación de la evaluación).

Desde la década de los 90 en adelante el rol tradicional de la evaluación perseguía la indagación por los resultados, ocasionalmente los impactos, la rendición de cuentas (*accountability*) y las lecciones aprendidas.

Las lecciones aprendidas, se refiere al conocimiento ganado por medio de la reflexión sobre una experiencia o proceso a partir de la evaluación puesta en práctica. Permiten la identificación de factores de éxito, de deficiencias, y la resolución de problemas a través de nuevos cursos de acción. A su vez, mejorar la toma de decisiones futuras y de esta manera contribuir a la realización de Buenas Prácticas, entendidas como soluciones eficientes para resolver un problema.

Las tendencias actuales en evaluación a nivel mundial han planteado un cambio importante en las prioridades, los modelos de evaluación tradicionales de tipo experimental (*randomized scientific experimental trial*) ya no son vistos como los únicos apropiados. Hoy se enfatiza la necesidad de una mayor flexibilidad y métodos menos restrictivos para captar la complejidad de las iniciativas.

Se reconoce la existencia de un incremento en el conjunto de consumidores de la evaluación, lo que implica la atención de distintos intereses, no solo el de los financiadores de los proyectos. Entre los nuevos intereses se privilegia, además del logro de los resultados, la efectividad de las acciones y la evaluación como un proceso de *empowerment*. Desde esta mirada la evaluación es vista como un camino a través del cual grupos marginales y sin poder son capaces de adquirir habilidades e influencia a través de su compromiso en ella (Conlin y Stirrat, 2008).

Luego de muchos años, en los cuales los especialistas intentaron convencer a los distintos auditorios sobre la necesidad de la evaluación, hoy se observa el advenimiento de una corriente bastante crítica, donde algunos de los autores señalan como necesario, frente a una tarea de

evaluación, volver a emplear la estrategia de la estimación previa de su *evaluabilidad* (Dahler-Larsen, 2007). La principal cuestión en la estimación de la *evaluabilidad* no es si se puede hacer una evaluación, sino, si es *racional* hacerlo bajo las circunstancias del momento a la luz de las mejoras esperadas de ella.

La *evaluabilidad* hace referencia a si existe una descripción clara del programa (si no existe, es mejor usar los esfuerzos para clarificar el programa), si está bien implementado el programa (si ya se sabe que no es el caso, la evaluación será siempre negativa), si existe una buena teoría del programa (si no la hay, mejor clarificar el marco teórico subyacente), si existen metas plausibles y bien escritas (si no existen, el resultado de la evaluación es predecible), si hay datos relevantes accesibles (si no hay, los recursos de la evaluación podrían gastarse mejor en alternativas a la evaluación), si se han identificado oportunidades para mejorar el programa (si los usuarios no son capaces o no utilizan los resultados de la evaluación, no tiene sentido realizarla). La idea que alimenta la *evaluabilidad* es que la evaluación debe ser utilizada solamente cuando sea conveniente y cuando la información que produzca sea utilizada para "marcar una diferencia" en provecho del programa.

Uno de los aspectos importantes respecto de la estimación de la *evaluabilidad* es la comprobación de la existencia de disponibilidad de información para la realización de la evaluación. Este aspecto es vital a la hora de utilización de métodos experimentales en dicho proceso.

## **2. DISEÑOS EXPERIMENTALES**

Tanto la evaluación de procesos (monitoreo) como la evaluación de impacto (en el curso del proceso o al finalizar el mismo) de los proyectos/políticas sociales pueden desarrollarse a partir de la implementación de métodos experimentales (Cohen y Franco, 1997:120-127).

Se ha discutido bastante la viabilidad de su implementación, fundamentalmente a partir de la selección de dos poblaciones equivalentes, por lo menos respecto de los factores considerados

pertinentes para la medición de los efectos del proyecto. En esta dirección la aleatorización de las muestras distribuiría de modo uniforme los efectos no controlados sobre los resultados obtenidos (tanto en el grupo semejante experimental como en el de control).

En el marco de los métodos experimentales la evaluación de impacto requiere por lo menos de dos mediciones (antes y después) y controlar los efectos netos del proyecto realizado. Es decir que la función de la evaluación es la de establecer que los efectos producidos en la realidad a partir de la implementación del proyecto están invariablemente relacionados con sus acciones. Para ello es necesaria la lógica de la prueba que muestra como la teoría, a través de su puesta en juego a partir de un proyecto, modifica la realidad a partir del principio de variación concomitante. Desde esta perspectiva los autores afirman que se trabaja sobre la causa eficiente, es decir la condición necesaria y suficiente para la aparición de un evento. Los diseños experimentales, a partir de la prueba establecen la covariabilidad de variables, es decir de las causas que producen determinados efectos.

Hay validez en el experimento cuando se mide lo que se tenía planificado medir, es decir cuando las diferencias en la medición de las variables del objeto de estudio reflejan diferencias en el atributo medido. Encontramos dos requisitos: a) validez interna: medir suficientes casos de mi unidad de análisis y b) validez externa: la posibilidad de generalizar los resultados obtenidos, ambas relacionadas con el N muestral y la construcción de una muestra probabilística. Hay confiabilidad si repetidas observaciones del fenómeno arrojan datos constantes. Validez y confiabilidad se construyen a partir de principios lógicos diversos, aunque en la práctica la alta validez (repeticiones de las mediciones) podría constituirse en una aproximación a la confiabilidad.

### **3. SU APLICACIÓN EN LA EVALUACIÓN DE PROGRAMAS/ PROYECTOS SOCIALES**

Una de las razones fundamentales que dificulta la aplicación del modelo experimental es la aleatorización de los casos que conforman cada uno de los grupos, puesto que lograr esto las muestras deben ser mayores que

en otros diseños y por lo tanto más costosas. Por otro lado, en algunas oportunidades plantean dilemas éticos (quien dejar fuera y quien dentro; consideraciones respecto del respeto a los derechos humanos, el respeto al consentimiento informado, etc.), o no hay posibilidades de conformar ambos grupos.

En general lo más habitual es la puesta en juego de modelos cuasi experimentales de evaluación en los que los miembros de los grupos experimental y de control no son seleccionados aleatoriamente.

Si no existe una tendencia metodológica dominante, producto de la incertidumbre que caracteriza a las ciencias sociales y al campo de la acción social, a partir de los cuales conviven paradigmas diversos, en algunos casos hasta contradictorios, tampoco existe un único modelo de evaluación de los impactos o procesos de las acciones llevadas adelante. No se pueden trasladar modelos de evaluación provenientes de proyectos de inversión en obras de ingeniería (construcción o infraestructura) muy poco adaptables a los escenarios en los que los proyectos sociales tienen lugar y que no solamente han generado resistencias para su implementación sino que no han servido para medir el impacto de dichos proyectos de modo integral. A pesar de ello los organismos de financiamiento, en particular las agencias internacionales, los exigen (Niremberg, Brawerman y Ruiz, 2003<sup>a</sup>:41).

Sin embargo, el experimento es viable y pertinente en un marco evaluativo integral en el que a partir de la triangulación se proponga a) la superación de la brecha entre lo cuali y lo cuantitativo; b) la identificación de indicadores pertinentes, en función del marco conceptual y la variedad de intereses y perspectivas asociadas a los distintos actores intervinientes; c) la definición y manejo de flujos de información originada en el proceso de evaluación y el incentivo para su uso por parte de los actores relevantes (Mokate, 2003:14).

#### 4. ANÁLISIS DE CASOS: EL CASO DE LA EVALUACIÓN DE PROGRAMAS SOCIOEDUCATIVOS<sup>3</sup>

Cuando se aborda el diseño de una evaluación de programas socioeducativos, salvo que se trabaje con datos de mucha agregación, es necesario comprobar la disponibilidad de la información que se requiere antes de pensar en la implementación de estudios experimentales. Cuando se prevé la aplicación de una evaluación con la utilización de métodos experimentales, la experiencia de trabajo en el campo señala el empleo de una serie de recaudos previos. Una de las primeras cuestiones que se debe abordar es asegurarse que se posee la información suficiente para llevar a cabo esa tarea. Esta afirmación puede parecer sobranante, y no pertinente para evaluadores formados, pero es muy frecuente haber previsto contar con información que luego, cuando el evaluador llega al campo, no es posible relevar.

Muchos de los programas y proyectos de intervención en escuelas persiguen como objetivo lograr una mejora en la cobertura, la calidad del servicio educativo, o en la eficacia de la escuela<sup>4</sup>. En general en evaluación de programas educativos, no solo se pretende indagar sobre el logro de los objetivos, sino que también se pretende indagar si las estrategias puestas en marcha han sido las adecuadas, sobre los obstáculos encontrados, y la identificación de evidencia razonable sobre los resultados logrados por el programa.

Es frecuente que muchos de los programas que se implementan en el sistema educativo nacional y sistemas educativos provinciales se realicen con financiamiento mixto, (de origen nacional y externo). Es así que los organismos financiadores muchas veces son proclives a la utilización de técnicas de tipo experimentales para poder comprobar la incidencia del programa en la escuela, más aún, en muchos casos su utilización forma parte de los requisitos que integran los marcos contractuales exigidos para la evaluación.

---

<sup>3</sup> Para la elaboración de este punto fue consultada la Lic. Lilia Toranzos, experta en evaluación.

<sup>4</sup> Esto supone una mejora en los valores llamados de **rendimiento educativo**: *promoción, repitencia, retención, sobre edad, etc.*

Desde el planteo inicial pareciera posible y relativamente simple la utilización de grupo control a partir de la selección de escuelas similares que no hubieran participado del programa (*recibido el estímulo*), para comparar en tiempos distintos (momento 0 y momento 1 en un diseño antes -después) la ganancia de uno y otro grupo de escuelas y de esta forma evitar los problemas de *atribución*. Es decir, poder corroborar que efectivamente los cambios en la población destinataria se deben a la incidencia del programa y no debido a otro factor.

Sin embargo, a la hora de poner en práctica el procedimiento, el evaluador se enfrenta con algunos de los problemas característicos del campo de la educación que le dificultan la conformación de los grupos experimental y control:

a.- Las escuelas no se encuentran distribuidas homogéneamente en el territorio por lo tanto si se busca una escuela de características similares a una que se elige como escuela experimental, ya nos encontramos con la primera dificultad, las escuelas no pueden ser seleccionadas al azar.

b.- Si se selecciona la escuela más cercana, no necesariamente está integrada con la misma población, ni tiene el mismo tamaño, con lo cual no es posible desde el inicio pensar en una escuela "idéntica" a la experimental, aunque se puede inferir que una escuela cercana tenga al menos una población con características similares.

c.- Si los sujetos a evaluar son escuelas rurales, por ejemplo, el problema se agrava, porque las escuelas rurales se encuentran a bastante distancia una de otra, con lo cual lo único que se mantiene similar es la condición de ruralidad, pero no las otras variables de contexto que caracterizan la población.

d.- Si salvamos el tema de las distancias, de contexto, o tamaño y composición de la escuela y todavía mantenemos la idea de trabajar con grupo control, se le suma una nueva dificultad. Las escuelas públicas están atendidas por múltiples programas generados desde el estado y es raro que reciban el apoyo de un solo programa. Con lo cual el esquema de que un grupo de escuelas reciben el estímulo de un programa y otras no,



deja de ser exacto, pues todas reciben algún tipo de estímulo y en muchos casos múltiples.

e.- Si desestimamos esta contingencia, porque en resumidas cuentas si todas las escuelas reciben el estímulo y eso de alguna manera las homogeniza, nos encontramos con otra dificultad que tiene que ver con la información que se intenta recabar. Las escuelas poseen muchos registros: datos de los padres de los alumnos, de los alumnos, de la asistencia del alumno, de su desempeño escolar, etc., pero muchos de esos registros quedan en la escuela y luego no son utilizados. La mayoría de ellos se encuentran en formato papel. Es muy posible que la información que se precisa no esté disponible, pues la escuela la tiene pero nunca llegan al sistema de información que se encarga de recolectar las estadísticas escolares.

f.- Si al evaluador le interesa analizar el impacto positivo del programa en una mejora en los aprendizajes de los alumnos, puede recurrir a la información provista por los resultados de las pruebas ONE<sup>5</sup>. El Ministerio de Educación implementa periódicamente relevamientos nacionales cuyo objetivo es la evaluación de los aprendizajes de los alumnos y de los factores asociados. Estos relevamientos se realizan con el objeto de contar con información con el objeto de mejorar el desarrollo de políticas educativas. Sin embargo, a diferencia de lo que ocurre con las estadísticas educativas que se aplican a toda la población, estos datos no están disponibles para todas las escuelas (al menos en forma histórica), lo cual dificulta aún más, la tarea de constitución de los grupos control y experimental.

## **5. LIMITACIONES Y POTENCIALIDADES**

A partir de la segunda mitad del siglo pasado se ha planteado una muy fuerte crítica al uso de modelos de evaluación a partir de la cual se

---

<sup>5</sup> Estos relevamientos son llevados a cabo bianualmente por la Dirección Nacional de Información y Evaluación de la Calidad Educativa (DINIECE) del Ministerio de Educación de la Nación. Son pruebas que se aplican para conocer el grado de dominio que el alumno tiene sobre un recorte específico de contenidos y de capacidades cognitivas.

incorporan visiones más relativistas e inclusivas de los factores contextuales desde metodologías intersubjetivas. Por un lado, se ha dejado la concepción de la evaluación como una etapa final de la implementación de un proyecto social que conllevó a la utilización de modelos experimentales o cuasi experimentales puros que permitirían aislar los factores determinantes que produjeron los cambios esperados en la situación que motivó la intervención, y si esos cambios se debieron efectivamente a las acciones desplegadas (Mokate, 2003:3; Niremberg, Brawerman y Ruiz, 2003<sup>a</sup>:40). Por el otro, el amplio desarrollo de la planificación estratégica y de las líneas participativas de investigación acción, dieron lugar a reconocer que las líneas de planificación no son lineales y que los actores involucrados en la planificación y la gestión reinterpretan sus acciones, por lo que se considera el amplio margen de incertidumbre de las intervenciones sociales.

A partir de ello, lo que proponemos es la incorporación de modelos experimentales y cuasi experimentales a partir de una lógica de triangulación que implique un abordaje multi-método, multi-actores y / o multi-técnicas que promueva el análisis de una línea de acción socioeducativa en la que se incorpore el experimento como uno de los ejes de la evaluación social, es decir ya no como "el diseño evaluativo " sino como una alternativa en función de la comprensión de los efectos de una acción dada, acompañada por la visión de los propios actores sobre el proceso en función del aprendizaje en el marco de la gerencia social.

La aplicación del método experimental resulta de esta manera como una instancia de enriquecimiento y profundización de los procesos de evaluación. La utilización de la técnica deberá observar una serie de recaudos para lograr una aplicación exitosa, entre ellos, un cuidadoso planeamiento de su implementación, la previsión de un tiempo adicional previo al proceso efectivo de la evaluación a los efectos de obtener un buen conocimiento del campo y la comprobación cierta de la información que se dispone y de la que no. Adicionalmente, si es necesario, prever alternativas de relevamiento de información ad hoc.

Como alternativas frente a situaciones donde se advierte la existencia de muchos obstáculos para la aleatorización, puede ser interesante, inclusive, la utilización de métodos experimentales con un N pequeño, con

muestras constituidas de manera intencional, pero más sólido en la constitución de las muestras en relación con las exigencias del experimento, que luego pueda combinarse con otros métodos descriptivos en más cantidad de sujetos.

En síntesis, frente a las severas críticas sobre la utilización de métodos experimentales, más que descartar su aplicación, resulta conveniente explorar las posibilidades valiosas que el método ofrece reivindicando su rol en la generación de conocimiento válido, confiable y oportuno para la gestión.

## **REFERENCIAS BIBLIOGRÁFICAS**

Briones, G.(1999): "Evaluación Educativa" Formación de docentes en investigación educativa. Colombia, Convenio Andrés Bello.

Cohen E. y Franco R. (1997): *Evaluación de proyectos sociales*. Buenos Aires, Siglo XXI.

Dahler-Larsen, P. (2007): "¿Debemos Evaluarlo Todo? O De La Estimación De La Evaluabilidad A La Cultura De La Evaluación". *Evaluación De Políticas Públicas*. España, Mayo-Junio.. N.º 836 Ice. 93-104.

Feinstein, O. (2007): "Evaluación pragmática de políticas públicas" en *Revista ICE*. España, Mayo –junio 2007, nº 836.19-31.

Gómez Serra, M.(2004): *Evaluación de los servicios sociales*. España, Ed. Gedisa,.

Ministerio de Educación (2011): "Operativo Nacional de Evaluación 2010" en *Censo De Finalización.de la Educación Secundaria. Informe de Resultados*. Argentina.

Mokate, K. (2003): "Convirtiendo al monstruo en aliado: la evaluación como herramienta de la gerencia social" en *Documentos de trabajo del INDES*. Washington D.C.

Niremberg, O., Brawerman, J. y Ruiz, V. (2003a): *Evaluar para la transformación*. Buenos Aires, Paidós.

----- (2003b): *Programación y evaluación de proyectos sociales*. Buenos Aires, Paidós.

# LA CONSTRUCCIÓN DE UN TREN TRASNACIONAL DE ALTA VELOCIDAD Y SUS EFECTOS SOBRE EL TERRITORIO – UN CASO DE ECONOMÍA ESPACIAL

*Verónica Caride*

*Mercedes Ramos*

## INTRODUCCIÓN

El análisis de Economía Espacial presenta un elevado potencial de desarrollo para la Argentina que aún no ha sido explorado. La amplitud de su territorio nos invita a despertar el interés por esta novedosa temática y fue precisamente esto lo que motivó el desarrollo de este trabajo.

El objetivo del presente estudio es analizar los efectos de la inversión en infraestructura de transporte sobre la actividad económica en el territorio. Para poder llevar a cabo dicha tarea, realizaremos un análisis de simulación del efecto económico – espacial de la mencionada inversión. En particular, se evalúa desde la perspectiva de la economía espacial los efectos de la construcción de un tren de alta velocidad transnacional en Europa que comunique las ciudades de Berlín y Budapest, en una primera instancia, para posteriormente analizar su extensión hasta Bucarest.

En la *Primer Sección* comenzamos exponiendo la evolución histórica de los costos de transporte dentro de la teoría económica, para posteriormente presentar su rol dentro del Core Model de Krugman en la *Sección II*. A continuación, en la *Sección III*, analizamos el efecto económico-espacial de la construcción de un tren de alta velocidad transnacional desde la ciudad de Berlín a Budapest y luego desde ésta a Bucarest. El modelo utilizado para la estimación de dichos efectos es una extensión del "Core Model" de Krugman para R regiones, al cual se le agregan los costos de congestión. En la *Sección IV*, exponemos las posibles implicancias de este tipo de análisis económico – espacial para la

Argentina, presentando las características particulares del sistema ferroviario en este país. Para finalizar, se exponen las conclusiones que se derivan de los resultados obtenidos de la simulación aplicada.

## 1. LOS COSTOS DE TRANSPORTE

Mucho tiempo atrás, Adam Smith en su libro *La Riqueza de las Naciones (1776)* ya destacaba la importancia de la locación - cercana a los puertos - para reducir los costos de transporte "*so, it is upon the sea coast, and along the banks of navigable rivers, that industry of every kind naturally begins to sub-divide and improve itself*"<sup>6</sup>. Sin embargo los mismos no fueron incluidos en los modelos económicos hasta mucho tiempo después.

Los primeros aportes a las teorías de localización fueron realizados a principios del siglo XIX por el teórico alemán Von Thünen, quien incorporó la dimensión espacial en la teoría económica.

El modelo de Von Thünen<sup>7</sup>, intenta explicar la localización de los distintos tipos de producción agrícola, incorporando como variable fundamental la distancia al centro de la ciudad. Esta distancia era medida en términos de tiempo y costos de transporte, bajo el supuesto de un territorio espacialmente uniforme y mercados competitivos.

El autor incorpora la idea de la renta decreciente en función de la distancia. La distancia considerada es aquella existente entre la localización de la producción y el mercado. En base a esta distancia y la renta que de ella se deriva para cada cultivo es que se determina el uso del suelo. Como resultado del modelo queda determinado un ordenamiento de zonas circulares alrededor del mercado central, donde a mayor distancia más almacenable (menos perecedera) resulta la producción agropecuaria.

---

<sup>6</sup> Citado por Steven Radelet y Jeffrey Sachs (1998)

<sup>7</sup> Von Thünen (1826) "The Isolated State". English translation. Oxford: Pergamon.

En las zonas más lejanas respecto al mercado, la producción más rentable es la de productos almacenables como el maíz o el trigo. Este resultado se deriva de unas menores rentas para la tierra dados sus mayores costos de transporte y escasa pérdida del producto dada su condición de almacenable versus perecedera. El caso de las flores es justamente el opuesto, este sería un producto altamente perecedero y consecuentemente no podría recorrer altas distancias sin sufrir altas pérdidas de producto. Consecuentemente los cultivos de flores tenderán a pagar una mayor renta de la tierra pero menores costos de transporte al localizarse cerca de la ciudad.

En un desarrollo teórico posterior, otro economista alemán, Alfred Weber<sup>8</sup> incorpora al análisis de localización los costos de transporte derivados de la distancia hacia los recursos requeridos para la producción. Esta teoría es más abarcativa, dado que sistematiza las distintas posibilidades de ubicación de otras actividades como la industria.

A diferencia del modelo de Von Thünen, esta teoría no supone un ambiente espacialmente uniforme, sino que incluye la existencia de materias primas que ocupan un sitio fijo de localización. Siendo así, la localización de la producción depende no sólo de la distancia al centro de consumo sino también de la distancia a los recursos espacialmente localizados. Este autor considera además que los costos de la mano de obra y las economías de escala de las empresas pueden influir en su localización.

Por su parte la Teoría de Central Place, cuyo principal referente es W. Christaller<sup>9</sup>, intenta explicar la distribución de los centros urbanos entendidos como núcleos de servicios que abastecen a una población circundante. En esta modelización, se supone un espacio uniforme con distribución homogénea de la población que posee igual poder de compra y posibilidades de desplazamiento en todo el territorio. Entonces, debido a los costos de transporte, el precio de un producto vendido en la ciudad aumenta a medida que nos alejamos de ella y en igual sentido disminuye

---

<sup>8</sup> Weber, A. (1929): "Theory of the location of the industries". University of Chicago Press.

<sup>9</sup> Christaller, W. (1933): "Die zentralen Orte in Süddeutschland". Traducción inglesa: The central places of southern Germany, Prentice Hall, Englewood Cliffs.

el poder de compra de la población. Debido a la caída en el poder de compra de los consumidores, a partir de cierto límite será beneficioso para ellos adquirir la mercancía en otros centros urbanos más cercanos, de manera tal que los costos de transporte sean menores.

Este modelo produce una estructura jerárquica de lugares centrales basada en una jerarquía de funciones. Por un lado, está el centro de orden superior con muchas funciones especializadas, las que son abastecidas por aquellos bienes y servicios que requieren de mayores demandas iniciales. Por otro lado, hay un gran número de lugares centrales más pequeños con funciones de orden inferior que se abastecen de bienes o servicios que requieren menores demandas iniciales.

En línea con Christaller, A. Lösch<sup>10</sup> sostiene la existencia de centros de distintos tamaños que ofrecen bienes y servicios a regiones comerciales de distinto orden, pero además incorpora la idea que dentro de cada área de mercado un productor adquiere una ventaja monopólica. Esta ventaja se encuentra restringida por la competencia de los productores vecinos y consecuentemente los productores deben procurar encontrar la localización óptima que les brinde el máximo beneficio, considerando los costos de producción en localizaciones alternativas y el área de mercado que puedan controlar.

Un método alternativo al análisis locacional es el análisis regional, el cual utiliza como base las conclusiones de la teoría de la localización expuestas hasta aquí. Este tipo de estudios se focalizan en las condiciones económicas diferentes entre las regiones, para luego identificar la localización de las actividades económicas, su volumen y los flujos de recursos existentes entre distintas regiones. Dentro de este grupo de teorías se pueden ubicar a los modelos liberales, modelos estructuralistas y modelos como el de centro – periferia de P. Krugman (1991).

Los modelos liberales tuvieron su origen a partir de 1950. Los mismos sostienen que las diferencias en el desarrollo económico se deben a desequilibrios de corto plazo entre oferta y demanda dentro de cada

---

<sup>10</sup> Lösch A. (1940): "The Economics of Location". English translation: New Haven, Conn.: Yale Univ. Press.



región. Sin embargo, en el largo plazo el resultado sería diferente, ya que el crecimiento se extendería también a las regiones menos desarrolladas, siempre que las fuerzas del mercado no fuesen obstaculizadas. Dentro de los modelos liberales se encuentran los modelos basados en la exportación, los modelos basados en la desigualdad de la renta regional y los modelos de polos de desarrollo.

Por su parte, los modelos basados en la exportación, ponen el acento en el rol que tienen las exportaciones en el desarrollo de una región. Uno de sus principales exponentes es D. North<sup>11</sup>, quien señala cinco etapas de crecimiento. Luego de un estado de subsistencia los países se desarrollan y comienzan a exportar productos básicos. Posteriormente, en una tercera etapa, sigue la expansión y diversificación de las exportaciones. En la cuarta etapa las industrias migran hacia el mercado local y en la quinta y última etapa finalmente las industrias locales diversifican sus exportaciones.

Dos años después, en 1957 se publica el trabajo de G. Myrdal<sup>12</sup>, quien analiza la desigualdad en la renta regional y utiliza la idea de "causalidad circular acumulativa". De acuerdo con el autor, hay regiones que crecen y regiones estancadas en función de ciertas ventajas iniciales. Los flujos de bienes y factores fluyen hacia las regiones más dinámicas y consecuentemente la desigualdad en el crecimiento se reproduciría. Sin embargo, si bien en el corto plazo el efecto acumulativo de las economías externas produciría mayores beneficios a través de la concentración geográfica, en el largo plazo estas desigualdades se disipan a través del "efecto propagación". A partir de ese momento las demandas locales serían abastecidas por la producción local.

En cuanto a los modelos de polo de desarrollo podemos ubicar a F. Perroux<sup>13</sup>, quien señala que el crecimiento económico se manifiesta de

---

<sup>11</sup> North, D. C. (1955): "Location theory and regional economic growth". Journal of Political Economy. 69, 319-40.

<sup>12</sup> Myrdal G. (1957): "Economic Theory and Under-developed Regions". London: Duckworth

<sup>13</sup> Perroux, F. (1955): "Note sur la notion de pole de croissance". En I. Livingstone (comp.) (1979), Development Economic and policy. Selected Readings. Londres: Allen & Unwin.

forma heterogénea, concentrándose en lo que llama "polos de crecimiento". De esta manera, se puede distinguir una polarización a nivel nacional, regional y local. En su modelo, cada industria posee un peso relativo específico dentro de la economía como impulsoras del crecimiento. Las industrias impulsoras son las que atraen la instalación de otras industrias en la región, creando así un "polo de crecimiento" o un "centro de crecimiento".

En cuanto a la corriente estructuralista, considera los desequilibrios regionales como una característica estructural de los mercados capitalistas. De esta manera, el desarrollo de una región conduce al subdesarrollo de otras. Uno de los principales exponentes dentro de esta corriente es el argentino Raúl Prebisch<sup>14</sup>, quien desarrolla la teoría sobre la existencia de dos tipos de países: los de "centro" y los de "periferia". Los primeros se especializarían en la producción industrial presentando mayores niveles de productividad e ingresos y se abastecerían de la producción primaria ofrecida por los países de la periferia. Esta división internacional del trabajo se reproduciría en el tiempo generando una estructura la cual puede ser sólo modificada a través de políticas activas de industrialización. Por este motivo, los teóricos estructuralistas remarcan la importancia de la necesidad de industrialización en los países periféricos, a fin de poder captar los beneficios del comercio internacional y elevar progresivamente el nivel de vida de la población.

La distribución desigual de los beneficios del comercio se debería a que los precios relativos de los productos exportables por la periferia (productos básicos) vs. aquellos exportables por el centro (manufacturas) manifestarían una tendencia decreciente. Consecuentemente, observamos un deterioro secular en los términos de intercambio de los países menos avanzados. De no cambiarse la estructura, estos modelos predicen un desequilibrio persistente y el desarrollo económico de una región implica

---

<sup>14</sup> Prebisch, R. (1949): "El desarrollo económico de América Latina y algunos de sus principales problemas". Reeditado en El trimestre económico, Nro 249. Vol. LXIII, México, Enero - Marzo de 1996, pp. 175 - 247.

una profundización del desequilibrio histórico entre regiones más que una convergencia.

Esta tendencia decreciente se debería a un suceso fundamental: Existe una menor elasticidad en la demanda internacional de bienes primarios - en comparación con aquella de los manufacturados - y consecuentemente los progresos tecnológicos en estos mercados se traducen automáticamente en caídas de su precio. En palabras de Prebisch:

“Cuanto mayor sea la inelasticidad en la demanda de los productos exportables por la periferia, mayor será la proporción de los frutos que son transferidos al centro”

Los incrementos en productividad de los productos básicos, por lo tanto, no serían una vía para el desarrollo de los países de la periferia sino que servirían a los fines de acentuar más las diferencias entre esta y el centro. Consecuentemente, un cambio de estructura y políticas activas para lograrlos sería el único camino posible para el desarrollo.

Para finalizar el presente recorrido, debemos mencionar el modelo utilizado en el presente trabajo para hacer las simulaciones: el modelo de centro – periferia de P. Krugman (1991)<sup>15</sup>. Si bien el mismo presenta un patrón de centro – periferia, a diferencia de las teorías anteriores la convergencia o divergencia en el desarrollo entre los países o regiones dependería de los valores iniciales de los que este partiese, de los parámetros del modelo y del nivel de las fuerzas de aglomeración y dispersión que se generasen. En este sentido, se puede afirmar que este modelo permite la existencia de *path dependence*, de equilibrios múltiples y que pueden llevar a convergencia o divergencia, dependiendo del caso.

## **2. EL ROL DE LOS COSTOS DE TRANSPORTE EN EL CORE MODEL DE KRUGMAN (1991) DE ECONOMÍA ESPACIAL**

Los modelos de Economía Espacial han cobrado especial relevancia en los últimos años, principalmente desde que Paul Krugman

---

<sup>15</sup> Krugman, P. (1991):“Increasing returns and Economic Geography” MIT Press

ganara el premio Nobel en Economía en el año 2008 por su trabajo sobre esta temática. Al observar que la actividad económica no se encuentra distribuida aleatoriamente en el territorio, la Economía Espacial se proponen explicar quién, por qué y dónde se localiza la actividad económica. Estas novedosas temáticas poseen una amplia utilidad para el análisis de políticas - públicas y privadas - y se encuentran en pleno desarrollo en un mundo cada vez más globalizado. La geografía, lejos de ser un factor exógeno, parecería estar íntimamente ligada con las dinámicas de la actividad económica. Diversas fuerzas de aglomeración y dispersión determinarían la economía geográfica en el territorio e impactarían directamente en el desarrollo de las diversas regiones. Por consiguiente, comprender la dinámica detrás de las decisiones espaciales de los agentes resulta crucial para la toma de decisiones con implicancias territoriales, tanto en el ámbito público como privado.

En este marco teórico, los costos de transporte juegan un rol crucial ya que son los que hacen que la geografía o el espacio importen para las decisiones económicas. Como bien señalan Brakman, Garretsen y van Marrewijk (2009)<sup>16</sup>, sin costos de transporte no existiría geografía y el ejercicio de transformar modelos económicos en modelos de economía espacial se tornaría meramente irrelevante. Es precisamente incorporar a la geografía de un modo no trivial lo que el Core Model de Economía Espacial pretende y a continuación explicaremos de qué modo.

## **2.1. Tipos de costo de transporte**

Los costos de transporte pueden ser introducidos como un sector aparte o bien del modo en que Samuelson (1952) lo realiza a través de un concepto que él denomina *Iceberg Transport Costs*.

---

<sup>16</sup> Steven Brakman, Harry Garretsen y Charles van Marrewijk (2009). *The New Introduction to Geographical Economics*. Cambridge University Press. pp. 107

En el primero de los casos cada costo para alguien es una ganancia para otro, ya que los costos de transporte son el ingreso del sector transporte. A su vez, la decisión de locación de la empresa de transporte no juega un papel irrelevante y por consiguiente el modelo se complejizaría. Para comprender la segunda opción debiéramos primeramente comprender qué son los costos de transporte iceberg, para lo cual se definen a continuación.

***Costos de Transporte Iceberg:*** Este tipo de costos de transporte implican que una fracción del producto no llega a destino cuando los productos son transportados entre regiones. Esta fracción representa a los costos de transporte. Si suponemos que  $T$  representa el número de bienes necesarios a ser transportados para que una unidad llegue a destino por unidad de distancia entonces si  $T=1,07$  el costo de transporte es del 7% por unidad de distancia.

La introducción de los costos de transporte a la Samuelson, por lo tanto, resulta muy atractiva ya que no requiere de las complicaciones mencionadas en el primer caso, lo que simplifica ampliamente la modelización. Este es precisamente el tipo de costos que asume el Core Model de Krugman para modelar los costos de transporte.

## **2.2. Las distancias**

Según el Core Model de Krugman, las variables económicas responden a procesos que no son constantes en el espacio debido a la existencia de fuerzas de aglomeración y dispersión: las economías internas a escala, economías de urbanización (externalidades tipo Jacobs) y los costos de transporte. Estos últimos se definen no sólo por  $T$  - que representa a los costos por unidad de distancia (costos tales como la gasolina) - sino que deben tener en cuenta la locación de los agentes económicos y consecuentemente las distancias relativas. La información por lo tanto debe estar georreferenciada, es decir que cada punto en el

espacio debe contar con un par de coordenadas (x;y). Existen dos tipos alternativos para definir la locación en el espacio que describimos a continuación:

- 1- Las distancias se estiman en forma de líneas rectas o como líneas sobre las carreteras entre puntos del plano. Claramente esta simplificación puede resultar en una buena aproximación para los casos de espacio uniforme con distribución homogénea pero no para otros que se alejen de este tipo de geografía. Nótese, que cuando el terreno presenta relieves importantes, como podría ser una cadena montañosa como los Andes, dicha simplificación podría llevarnos a resultados equivocados.
- 2- Otra alternativa que se utiliza comúnmente es la de medir la distancia en términos de tiempo, caso en el cual sí se estarían contemplando las diferencias en la geografía como ser el relieve. Sin embargo, en dicho caso también se estaría realizando una simplificación ya que el tiempo de transporte depende del tipo de vehículo que se esté empleando.

Ninguna de las dos alternativas descritas es correcta o incorrecta sino que debe ser evaluada al momento de incorporarla dentro del modelo en cada caso en particular.

En el Core Model de Economía Espacial, las distancias se presentan de forma exponencial en los costos de transporte por unidad de distancia de modo que los costos de transporte quedan definidos de la siguiente forma:

$$T_{rs} = T^{D_{rs}} \text{ para } r,s=1,2$$

$$\text{Note que } T_{rs} = T_{sr} \text{ y que } T_{rr} = T^0 = 1$$

### 2.3. Cambios en los Costos de Transporte

Los costos de transporte han cambiado a lo largo de la historia por diversos motivos, entre ellos la revolución de los medios de transporte que comenzó en el SXIX. A partir de ese momento se aceleró la integración económica mundial de manera exponencial, alejando a los países cada vez más de la autarquía y llevándolos a la globalización. En el Core Model de Krugman, los cambios en los costos de transporte pueden ser de dos tipos: i) Cambios tecnológicos que afecten a todas las regiones ii) Cambios de distancias relativas entre dos puntos.

En el primero de los casos lo que se modifica es T y puede resultar, por ejemplo, de variaciones en el precio del combustible o de una nueva tecnología que reduzcan su consumo. En el segundo caso, lo que se modifica son las distancias relativas debido a inversiones en infraestructura, como ser la construcción de un tren de alta velocidad o la de un túnel que atravesase una montaña de considerable altura. Nótese, que al cambiar una distancia relativa otras distancias relativas en el mapa también se modifican, veámoslo en un ejemplo concreto.

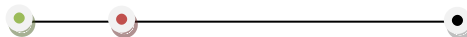
Supóngase que hay tres puntos alineados en el mapa, uno verde, otro rojo y otro negro.

**Gráfico 1**



Supóngase ahora que el punto rojo se desplaza hacia el punto verde, como vemos a continuación:

**Gráfico 2**



Como se puede observar en el Gráfico II, al cambiar la distancia relativa entre el punto rojo y el verde también se modificó la distancia relativa entre el punto rojo y el negro. Consecuentemente, al simular un cambio en una distancia relativa posteriormente se debe volver a reestimar toda la matriz de distancias relativas del mapa a fin de captar todos los posibles efectos.

## **2.4. El Equilibrio de Corto Plazo del Core Model de Krugman (1991)**

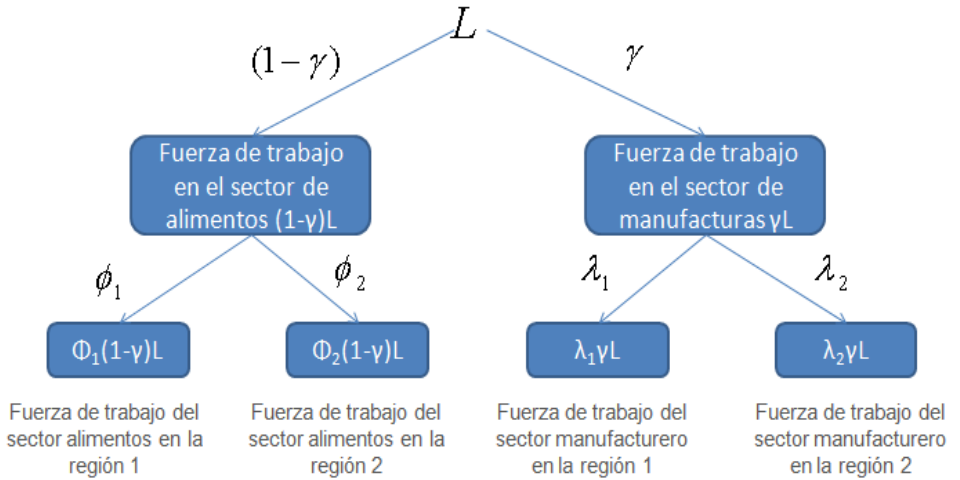
El modelo de Krugman supone dos bienes: manufacturas: M y alimentos: F, dos sectores: industria y agro y dos agentes que producen y consumen: granjeros y trabajadores del sector manufacturero. Los granjeros viven en zonas rurales y no pueden moverse porque su producción depende del factor tierra (inmovible). Los trabajadores del sector manufacturero, por su parte, se movilizan en busca de mejores salarios reales. Existen economías de escala en la industria (competencia imperfecta) y economías de escala constantes en el sector agropecuario (competencia perfecta).

En este modelo la demanda es endógena, ya que se localiza donde viven las personas, y los consumidores tienen lo que se denomina Love-for-variety (amor a las variedades) en el consumo de Manufacturas.

Suponiendo sólo dos regiones (Norte=1 y Sur=2), los parámetros del modelo serían los siguientes:

### **Gráfico 3**





En donde  $L$  es la cantidad de consumidores y trabajadores y los parámetros hacen referencia a proporciones.

El equilibrio de corto plazo se determina por el siguiente sistema de ecuaciones:

$$(i) \quad Y_r = \delta \lambda_r W_r + (1 - \delta) \phi_r$$

$$(ii) \quad I_r = \left[ \sum_{s=1}^R \lambda_s T_{rs}^{1-\varepsilon} W_s^{1-\varepsilon} \right]^{1/(1-\varepsilon)}$$

$$(iii) \quad W_r = \left[ \sum_{s=1}^R Y_s T_{sr}^{1-\varepsilon} I_s^{\varepsilon-1} \right]^{1/\varepsilon}$$


---

En donde  $Y$  es el Producto Interno Bruto (PIB),  $I$  es el índice de inflación,  $W$  son los salarios y  $\epsilon$  la elasticidad de demanda de bienes manufacturados.

Al observar las ecuaciones (i) a (iii) vemos que, dada la existencia de no linealidades, la relación entre las variables no es única y puede variar dependiendo del nivel inicial de las mismas. Menores costos de transporte podrían implicar un incremento de la competencia, lo que reduciría los precios de los bienes manufacturados incrementando el salario real y reduciendo la inflación local. Sin embargo, debido al efecto que cambios en el costo de transporte tienen en la distribución de los trabajadores manufactureros en el territorio y el consecuente efecto sobre las economías de aglomeración, esto podría no siempre ser así para todas las locaciones. Si existiesen sólo dos economías, una pequeña (en términos industriales) y otra grande y se viesen reducidos los costos de transporte, entonces en el largo plazo los trabajadores manufactureros podrían migrar hacia la economía de mayor tamaño con la consecuente pérdida de competencia local de los bienes manufacturados, mayores precios y caídas del salario real en la economía pequeña <sup>17</sup>. La no linealidad en las ecuaciones hace posible la existencia de múltiples equilibrios y consecuentemente la decisión óptima en términos de salario real e inflación dependerá del nivel del que partan las economías y del tamaño de cada una de ellas. A esta dependencia en la literatura se la denomina "Path Dependence".

Por otra parte, la existencia tanto de equilibrios estables e inestables también introduce en la discusión a cuestiones tales como el lock-in effect, en el cual la idea central es que políticas temporarias que afecten la movilidad de bienes y factores puedan tener efectos permanentes. En síntesis, un incremento o caída en los costos de transporte puede tener un efecto transitorio o permanente de incremento o decremento en el salario real local dependiendo del valor inicial del que partan las variables y

---

<sup>17</sup> Véase Steven Brakman, Harry Garretsen y Charles van Marrewijk (2009). *The New Introduction to Geographical Economics*. Cambridge University Press. pp. 221

consecuentemente no se pueden realizar recomendaciones de políticas de transporte atemporales.

### **3. UN CASO APLICADO – LA CONSTRUCCIÓN DE UN TREN DE ALTA VELOCIDAD BERLÍN – BUDAPEST - BUCAREST**

En esta sección se analiza el efecto económico-espacial de la construcción de un tren de alta velocidad transnacional en Europa: que comunicaría a las ciudades de Berlín y Budapest, en una primera instancia, para posteriormente analizar su tendido hasta Bucarest. El modelo a través del cual se estimarán dichos efectos es una extensión del Core Model de Krugman para R regiones al cual se le agregan los costos de congestión. Cabe destacar que, dado que la población rural dedicada al sector alimentos no puede movilizarse, los resultados se expresarán en términos de trabajadores manufactureros.

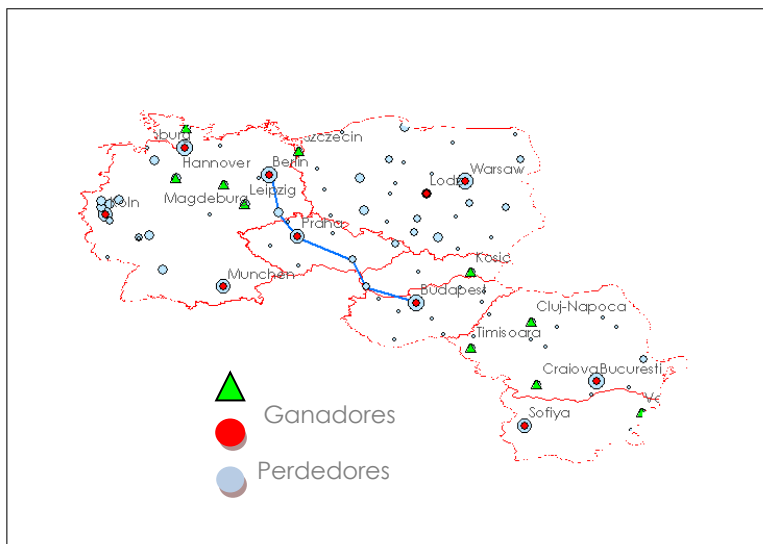
Para ello se redujeron los enlaces correspondientes de la matriz de distancia en 50%. Dado que la matriz de pesos espaciales es simétrica, cada modificación de distancias implicó un doble cambio en la matriz de pesos espaciales. Suponiendo que se redujeron las distancias entre los puntos  $r$  y  $s$ , en la Matriz deben modificarse tanto  $D_{rs}$  como  $D_{sr}$ . Posteriormente se volvió a ejecutar el modelo con esta nueva matriz de distancia, previa reestimación de las distancias relativas en todo el mapa. Finalmente se presentan los efectos de la política en las ciudades sobre el mapa a través de ArcGis, diferenciando entre ciudades ganadoras y perdedoras y se expresan los motivos de estos resultados.

#### **3.1. RESULTADOS**

Luego de hacer la simulación de la construcción de un tren de alta velocidad entre Berlín y Budapest, se volcaron en el *Mapa 1* los resultados obtenidos coloreándose con rojo aquellas ciudades perdedoras y con verde a las ganadoras. Como se puede apreciar en el mapa, las ciudades perdedoras son aquellas de mayor tamaño mientras que las ganadoras son ciudades pequeñas. Por lo tanto, se puede inferir razonablemente que las fuerzas de aglomeración son inferiores a las fuerzas de dispersión para las grandes ciudades en el nivel de aglomeración y costos de transporte

considerados. La creciente rentabilidad de la industria manufacturera es inferior a los mayores costos de congestión y por lo tanto, reducidas distancias hacen más conveniente para una industria trasladarse a una ciudad pequeña. Sin embargo, nótese que no cualquier ciudad pequeña crece sino principalmente aquellas ubicadas cerca de los puntos finales de la línea del tren, es decir Berlín y Budapest. Puesto que los costes de congestión no son tan altos en estas pequeñas ciudades y su locación es cercana a la demanda, la reducción de la distancia implica una reducción en los gastos de transporte y alienta a las empresas a establecerse allí. Estas ciudades están ahora cerca de la línea de tren y consecuentemente más cerca de otros lugares distantes, con la ventaja de tener bajos niveles de congestión en comparación con Berlín y Budapest. Como resultado, se vuelven más atractivas para la localización de las empresas y los trabajadores manufactureros. Por el contrario, Berlín y Budapest, así como ciudades situadas lejos de la línea del tren, evidenciarán pérdidas de población en el largo plazo.

**Mapa 1**



El efecto observado en el mapa es conocido como "Butterfly Effect" debido a la forma de mariposa en que se manifiesta la redistribución económica en el territorio. Ante una mejora en lo que se denomina *infraestructura en línea*, como puede ser la construcción de un tren de alta velocidad o de una carretera<sup>18</sup>, la relocalización económica se da principalmente alrededor de las ciudades de los extremos.

Una vez estimado el efecto de la construcción de un tren de alta velocidad entre Berlín y Budapest procedemos a estimar aquellos que se derivan de su extensión a Bucarest y compararemos los resultados entre las siguientes dos alternativas de política: 1) Construir un tren de alta velocidad que una Berlín y Budapest; 2) Construir un tren de alta velocidad que una Berlín y Bucarest, pasando por Budapest. Para ello, y para evitar errores de calibración, comparamos el resultado final de la política 1 con aquellos resultantes de realizar la extensión de la línea de tren hasta Bucarest. Por lo tanto, a continuación presentamos los resultados de la segunda política.

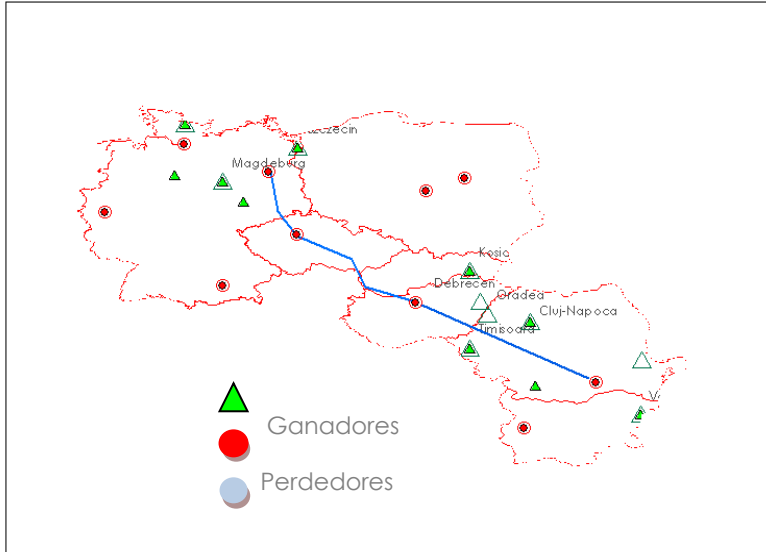
El tren de alta velocidad no sólo conectará Berlín y Budapest sino que ahora también a Budapest y Bucarest. En consecuencia, tenemos que reducir la distancia entre estos dos puntos en un 50% y, a continuación, volver a ejecutar el modelo. Así cambiamos las distancias desde Budapest a Bucarest y de Bucarest a Budapest y también corremos el algoritmo de la ruta más corta.

En el Mapa 2 se ilustran los resultados de inversión de la segunda alternativa de política en infraestructuras, a saber un tren que comunica a Berlín y Bucarest pasando por Budapest. Es evidente que la mayoría de las ciudades que ganaron desde el establecimiento de la línea de tren entre Berlín y Budapest también gana desde su extensión a Bucarest. La mayoría de las ciudades que perdieron desde la construcción de la línea de tren entre Berlín y Budapest también pierden a partir de la segunda política. Como resultado, podemos concluir que ambas alternativas de inversión llevarán a resultados similares.

---

<sup>18</sup> Véase Jan Oosterhaven y Thijs Knaap (2003). Spatial Economic Impacts of Transport Infrastructure Investment. University of Groningen.

## Mapa 2



Sin embargo, lo importante es no sólo identificar quien pierde y quien gana sino también en qué medida y en qué sentido lo hacen. Consecuentemente, debe analizarse la magnitud de cada pérdida y ganancia en cada caso en particular y comparar los resultados obtenidos. El cálculo de la diferencia absoluta producida por las dos políticas alternativas nos ayudará a determinar cuál de ellas resulta ser la más conveniente, dado ciertos objetivos de política buscados. En la Tabla I a continuación, se presentan dichos resultados.

Si el objetivo de la política fuese el de promover el desarrollo en el territorio, descomprimiendo la concentración en las ciudades grandes, entonces la segunda política bajo examen sería preferible. Esto se debe a que en este caso las ciudades más grandes perderían más y las más chicas ganarían más en conjunto en comparación con la primer alternativa de política. Por lo tanto, aplicando la opción larga de inversión en infraestructura la población y la actividad económica se distribuirían más homogéneamente en el territorio, promoviendo la cohesión social en el

mismo. Como resultado, las diferencias geográficas del bienestar tenderían a disminuir.

En conclusión, dados los objetivos de políticas propuestos, se recomendaría la construcción de un tren de alta velocidad que conecte Berlín y Bucarest, pasando por Budapest, en lugar de la opción más corta Berlín-Budapest.

**Tabla 1**

| Ciudad  | Cambio en % debido a la política 1 <b>(1)</b> | Cambio en % debido a la política 1 <b>(2)</b> | Diferencia de los valores absolutos de las columnas (2) – (1) <b>(3)</b> |
|---|---|---|--|
| Berlín  | -0,1026%                                      | -0,1033%                                      | <b>0,0007</b>  |
| Bucarest  | -0,0334%                                      | -0,1916%                                      | <b>0,1582</b>  |
| Budapest  | -0,0561%                                      | -0,0857%                                      | <b>0,0296</b>  |
| Hamburg   | -0,0672%                                      | -0,0814%                                      | <b>0,0143</b>  |
| Warsaw  | -0,0174%                                      | -0,0330%                                      | <b>0,0156</b>  |
| Munchen   | -0,0188%                                      | -0,0396%                                      | <b>0,0208</b>  |
| Praga   | -0,0217%                                      | -0,0465%                                      | <b>0,0248</b>  |
| Sofía   | -0,0278%                                      | -0,0487%                                      | <b>0,0209</b>  |
| Colonia   | -0,0263%                                      | -0,0516%                                      | <b>0,0254</b>  |
| Lodz  | -0,0233%                                      | -0,0480%                                      | <b>0,0247</b>  |
| <b>Total Perdedores (Sumatoria Columna 3)</b>             |   |   | <b>0,335</b>   |
| Szczecin  | 0,0419%                                       | 0,0398%                                       | <b>-0,0021</b>   |
| Varna   | 0,0155%                                       | 0,1257%                                       | <b>0,1102</b>  |
| Timisoara   | 0,1290%                                       | 0,0811%                                       | <b>-0,0479</b>   |
| Cluj-   | 0,0982%                                       | 0,0454%                                       | <b>-0,0528</b>   |
| Magdebur  | 0,1540%                                       | 0,2007%                                       | <b>0,0467</b>  |
| Kosice  | 0,0712%                                       | 0,0792%                                       | <b>0,0081</b>  |
| Kiel  | 0,1077%                                       | 0,1318%                                       | <b>0,0241</b>  |
| Craiova   | 0,0555%                                       | 0,2468%                                       | <b>NA</b>  |
| Hannover  | 0,0011%                                       | 0,2102%                                       | <b>NA</b>  |
| Leipzig   | 0,0047%                                       | 0,2010%                                       | <b>NA</b>  |
| <b>Total Ganadores (Sumatoria Columna 3)</b>              |   |   | <b>0,0863</b>  |
| <b>Total Ganadores y Perdedores (Sumatoria Columna 3)</b> |   |   | <b>0,4213</b>  |

## **4. IMPLICANCIAS PARA LA ARGENTINA**

Es interesante notar el potencial de desarrollo que en la Argentina tiene el análisis de la Economía Espacial, dada la amplitud de su territorio. La distancia que existe entre la Ciudad Autónoma de Buenos Aires y Río Gallegos es de poco más de 2.500 Kms al igual que la distancia que separa las ciudades de Amsterdam (Países Bajos) y Moscú (Rusia). En el último caso hay que atravesar tres países de considerable tamaño para llegar a destino: Alemania, Polonia y Bielorrusia.

La Argentina es el 8vo país más grande del mundo, detrás de India, con 2.780.400 Km<sup>2</sup>. Es 3 veces más grande que Bolivia, 16 veces que Uruguay y 100 veces que Haití. Pero su gran tamaño no sólo se destaca en Sudamérica sino a nivel mundial ya que Italia cabe 9 veces en su territorio, el Reino Unido 11, los Países Bajos 66 y Luxemburgo 1.075 veces.

En este contexto es que en la siguiente sección nos proponemos analizar la estructura del transporte ferroviario en Argentina. Dicho análisis implica el cuestionamiento sobre las causas del trazado, su composición y relevancia en la actualidad.

### **4.1. EVOLUCIÓN HISTÓRICA DEL TENDIDO FERROVIARIO ARGENTINO**

En 1857 se inauguró la primera línea de ferrocarril en Argentina, como iniciativa de empresas privadas y del estado. El origen de los capitales invertidos eran predominantemente ingleses y el avance del tendido vial siguió la lógica comercial del modelo agroexportador que se encontraba en desarrollo. Posteriormente, en 1948 durante el primer gobierno peronista, los ferrocarriles que estaban en manos británicas fueron estatizados.

La estructura urbana de Argentina es el resultado de un sistema económico centralizado del tipo centro periferia, conformando un "país abanico" en palabras de Alejandro Bunge. Por su parte, los centros regionales reproducen dentro de su área de influencia, el mismo sistema



de dominación por parte de la metrópoli regional concentrando el poder y centralizando las actividades económicas en Buenos Aires.

Esta estructura en forma de “abanico” fue la que dio origen al particular tendido de vías férreas en el territorio argentino. Cabe destacar que el sistema estaba diseñado para el transporte desde el interior del país desde y hacia la Ciudad de Buenos Aires únicamente, ya que las distintas líneas poseen diferentes trochas – ancha, angosta y media -, lo que impide la libre circulación de los vagones por toda la red.

Esta característica constituye un serio inconveniente para el aprovechamiento del material rodante, dado el perfil estacional de la mayor parte de las mercaderías transportadas (productos agropecuarios). Los vagones y locomotoras que no se utilizan en una región, donde por ejemplo la cosecha ha finalizado, podrían ser aprovechados en otras zonas, en período de cosecha, si los trenes fueran todos de la misma trocha.

En materia de inversión en infraestructura en la Argentina, cabe destacar que la falta de inversión en el sector ferroviario durante los últimos veinte años ha llevado a un deterioro tanto del material rodante como del tendido de vías. Ante este suceso, y dada la ausencia de un plan nacional integral y multimodal de transporte, el tendido de carreteras y rutas se realizó siguiendo un trazado paralelo a las vías férreas, constituyéndose una competencia entre el transporte automotor y el transporte ferroviario. De esta manera, el transporte ferroviario fue perdiendo importancia respecto del transporte mediante camiones y ómnibus. Entre los factores que causaron este resultado se encuentra la ausencia de políticas ferroviarias claras, y la falta de inversiones para el mantenimiento y modernización de la infraestructura y material rodante.

En la década de los noventa con la lógica de “reforma del estado” se realizó la división de los servicios en: urbanos de pasajeros, interurbanos de pasajeros y servicios de cargas. Los servicios urbanos de pasajeros y los de carga fueron concesionados, los servicios de pasajeros interurbanos fueron transferidos a las provincias. Este último hecho hizo que muchas provincias no pudieron sostener los costos y numerosos servicios fueron

cancelados definitivamente. En consecuencia, muchos trabajadores perdieron sus empleos y pequeñas localidades quedaron aisladas.

En la actualidad, el sistema ferroviario argentino metropolitano está compuesto por siete líneas con la mayor parte de su tendido en el Gran Buenos Aires. Actualmente tiene una extensión de 833 km de los cuales el 18% poseen electrificación. Según datos de la Secretaría de Transporte, la cantidad de viajes realizados en la Región Metropolitana durante el 2007 ascendió a 1.206.797. El 70,6% declara que el motivo del viaje es por trabajo y consecuentemente se destaca la importancia de este medio de transporte para la actividad económica de Buenos Aires y alrededores.

Ante este panorama del sistema ferroviario argentino, queda en evidencia la necesidad de realizar una planificación estratégica de infraestructura en transporte por parte del estado. La misma se impone como un objetivo prioritario de política de gobierno, con un alto impacto social y económico y un elevado potencial para incrementar la competitividad económica y la integración regional.

Resulta necesaria la adecuación del trazado y complementariedad de las redes así como de un plan de transporte nacional multimodal, que responda a las exigencias de la evolución de la economía, disminuyendo los recorridos y los costos de transporte. Es necesaria la incorporación de la dimensión territorial en la planificación económica y de localización en un país tan extenso como lo es la Argentina, que permita mejorar la accesibilidad a todas las áreas del territorio nacional así como a los países limítrofes.

## **5. CONCLUSIÓN**

En el presente trabajo hemos abordado la temática de localización desde una perspectiva de Economía Espacial, perspectiva novedosa y con un amplio potencial de desarrollo para la Argentina.

A través de un ejercicio de simulación se ha demostrado la relevancia de la Economía Espacial en la toma de decisiones de políticas en infraestructura. Específicamente, se ha analizado el efecto económico-

espacial de dos alternativas de política de inversión en la construcción de un tren de alta velocidad transnacional en Europa. La primera era la alternativa más corta, ya que se trataba de la construcción de dicho tren entre las ciudades de Berlín y Budapest. La segunda, por su parte, incluía no sólo el tendido de la red entre Berlín y Budapest sino que adicionalmente contaba con una extensión hasta Bucarest. El modelo a través del cual se estimaron dichos efectos es una extensión del Core Model de Krugman para R regiones, con costos de congestiónamiento.

La alternativa de política recomendada surge de los resultados de la simulación conjuntamente con el objetivo de política pública supuesto. Asumiendo que éste fuese el de promover el desarrollo en el territorio, se arribó a la conclusión de que la segunda política bajo examen sería preferible. Este resultado se deriva de que las ciudades más grandes perderían más y las más chicas ganarían más en la segunda alternativa de política en comparación con la primera. Por lo tanto, aplicando la segunda alternativa de política en infraestructura, la población y la actividad económica se distribuirían más homogéneamente en el territorio, promoviendo así la cohesión social en el mismo. Como resultado, las diferencias geográficas del bienestar tenderían a disminuir.

En conclusión, dados los objetivos de políticas supuestos, se recomendaría la construcción de un tren de alta velocidad que conecte Berlín y Bucarest - pasando por Budapest - en lugar de la opción más corta Berlín-Budapest.

Es importante destacar que queda pendiente para futuros trabajos la simulación de este tipo de políticas de inversión en infraestructura para la Argentina, dada la relevancia de la temática para el país así como del elevado potencial y la escasa difusión de este tipo de análisis en la toma de decisión política.

## **REFERENCIAS BIBLIOGRÁFICAS**

Brakman S., Garretsen H. y van Marrewijk C. (2009): *The New Introduction to Geographical Economics*. Cambridge University Press.

- Christaller, W. (1933): *Die zentralen Orte in Süddeutschland*. Prentice Hall, Englewood Cliffs, 1966.
- Jan Oosterhaven y Thijs Knaap (2003): *Spatial Economic Impacts of Transport Infrastructure Investment*. University of Groningen.
- Knox, P. y Agnew, J. (1994): *Geografía Económica Mundial*. Buenos Aires, Argentina. Ed. Docencia, pp. 109 – 169.
- Krugman, P. (1991): *Increasing returns and Economic Geography*. MIT Press.
- Losch, A. (1940): *The Economics of Location*. English translation: New Haven, Conn.: Yale Univ. Press, 1954.
- Myrdal, G. (1957): *Economic Theory and Under-developed Regions*. London: Duckworth, 1957.
- North, D. C. (1955): *Location theory and regional economic growth*. Journal of Political Economy. 69, 319-40.
- Perroux, F. (1955): "Note sur la notion de pole de croissance". En I. Livingstone (comp.) (1979): *Development Economic and policy*. Selected Readings. Londres: Allen & Unwin.
- Prebisch, R. (1949): "El desarrollo económico de América Latina y algunos de sus principales problemas". Reeditado en *El trimestre económico*, Nro 249. Vol. LXIII, México, Enero – Marzo de 1996, pp. 175 – 247.
- Radelet S y Sachs J. (1998): *Shipping Costs, Manufactured Exports and Economic Growth*. Columbia University.
- Rapoport, M. (2003): *Historia económica, política y social de la Argentina*. Buenos Aires, Argentina. Ed. Macchi, cap I y IV.
- Smith A. (1776): *An Inquiry into the Nature and Causes of the Wealth of Nations*. Edinburgh.
- Starret D. (1978): *Market Allocations Location Choice in a Model with Free Mobility*. Journal of Economic Theory, 17:21-37.

Von Thunen, J. K. (1826): *The Isolated State*. English translation. Oxford: Pergamon, 1966.

Weber, A. (1929): *Theory of the location of the industries*. University of Chicago Press.

# **VALOR ESPERADO CONCENTRADO (VEC) y SIMULACIÓN MONTE CARLO EN LA EVALUACIÓN DE PROYECTOS DE INVERSIÓN**

*José Luis Pungitore*

## **INTRODUCCIÓN**

Las decisiones entre diferentes alternativas de inversión en los proyectos públicos y privados generalmente se apoyan en herramientas tradicionales de proyección de flujos de fondos actualizados. En algunas de esas variantes, se compara el Valor Actual Neto (VAN) o la Tasa Interna de Retorno (TIR) de cada proyecto con sus oponentes, y la decisión se inclina hacia aquél con rendimiento más alto.

Bajo la concepción tradicional, el riesgo es incorporado a través de un adicional a la tasa de descuento utilizada para calcular el VAN del proyecto o de la exigencia de una TIR mínima más alta; tal decisión no es gratuita, pues por ese simple hecho se corre el riesgo de desechar proyectos interesantes y atractivos.

En proyectos que presentan estructura arbórea, con distintas alternativas o ramas con probabilidades asociadas, se utiliza ampliamente el criterio del Valor Esperado.

## **1. EL VALOR ESPERADO Y LOS SUPUESTOS SUBYACENTES**

La media o esperanza matemática "es el valor que se obtendría si repitiéramos muchas veces el experimento aleatorio y promediáramos todos los resultados obtenidos para la variable aleatoria" (García R. M., página 22). A los fines del presente trabajo, y siguiendo la línea de R. M. García, denominaremos Valor Esperado al valor que surge de la sumatoria

de la multiplicación de las probabilidades de cada alternativa integrante de un proyecto, por el importe de su respectiva recompensa (o pérdida).

Entonces, el Valor Esperado asocia los resultados con sus probabilidades de ocurrencia, introduciendo el concepto de riesgo en el análisis de proyectos de inversión; de su aplicación a varios proyectos que compiten por recursos escasos de inversión, resulta favorecido aquél con mayor valor.

El uso del criterio del Valor Esperado se encuentra ampliamente difundido, tanto en el campo académico como profesional, a pesar de haber sido criticado por diversos autores, entre los que podemos citar a J. P. Rheault, P. Pavesi y R. Pérez.

Sin embargo, las situaciones en que se lo utiliza están lejos de ser las ideales para las que fue pensado: en otras palabras, no es un criterio para ser utilizado en todo tiempo y lugar. Para que su aplicación en la evaluación de un proyecto determinado resulte de utilidad y pueda ser un indicador confrontable con similares de otros proyectos, deberían cumplirse las siguientes condiciones:

- El experimento (en nuestro caso, un proyecto) debería repetirse una gran cantidad de veces; esta condición (por ejemplo la instalación de una sucursal) no parece fácil de alcanzar para muchas empresas de nuestro medio: ya sean del país o del subcontinente. "En otras palabras, para decisiones idénticas y repetitivas el procedimiento de la esperanza matemática es formalmente válido. El problema es que en la vida real, las decisiones de este tipo son escasas o nulas". (Pavesi, página 228).
- "En la vida real, el decisor que ocupa una posición elevada en la empresa no acostumbra a usar la media (esperanza matemática), sino que tiende a asignar certeza (probabilidad = 1) cuando la probabilidad es elevada; o a utilizar el modo y tratarlo como si fuera 1". (Pavesi página 228)

Las desventajas que se observan en el Valor Esperado tienen que ver con que, si bien cada resultado se encuentra ponderado por su

probabilidad asociada, los efectos negativos y positivos se compensan, neutralizándose.

Pero el decisor podría no considerar (o, por su propia naturaleza, no estar en condiciones de considerar) "neutro" el hecho de tener probabilidades de ganar o de perder mucho dinero, tal como si en cada acción se jugara la vida de la empresa. En ese sentido, las empresas más pequeñas son más vulnerables a los fracasos, por lo que no es aventurado suponer que –en muchos casos- en caso de suceder significará su extinción lisa y llana.

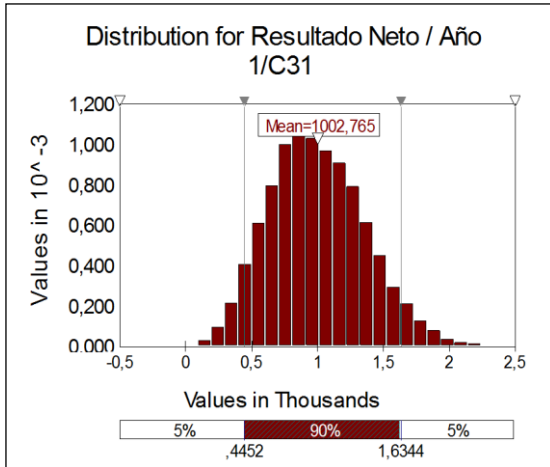
Por todo lo expresado anteriormente, existe espacio para proponer un indicador de proyecto que no contenga algunas de las características no deseadas del Valor Esperado. Y éste es el Valor Esperado Concentrado (VEC), que en este trabajo se presentará asociado con el análisis de riesgo por Simulación Monte Carlo.

## **2. UTILIDAD DEL ANÁLISIS DE RIESGO POR SIMULACIÓN MONTE CARLO**

En ambientes complejos y turbulentos resulta particularmente útil la utilización de la técnica de Simulación Monte Carlo. Bajo esta concepción, el resultado a obtener (salida) no será una cifra única, plurianual o no, según el caso, sino un histograma en el que se mapeen tanto el VAN, la TIR o las utilidades en un eje, con sus probabilidades de ocurrencia asociadas, en el otro eje.



**Figura 1. Histograma**



Sin embargo, para ser llevado eficazmente a la práctica, el método requiere que la organización cuente con datos de la gestión, adicionales a las que podría proporcionar el sistema contable, de forma tal de conocer el comportamiento de las variables de entrada.

Del análisis de la Figura 1 se puede advertir que la “síntesis” de cada histograma estará representada por el Valor Esperado de la variable de salida que se defina (media del histograma). Aunque justo es reconocer que una simple cifra, el Valor Esperado en este caso, no es capaz de contener la riqueza de información resultante del proceso de simulación.

### **3. UN EJEMPLO DE SIMULACIÓN EN DOS PROYECTOS**

Supongamos que existen dos proyectos, 1 y 2, muy similares, con los siguientes datos compartidos por ambos:

- Inversión Inicial: - \$ 100

- Flujo Positivo Proyectado del Período 1: \$90
- Flujo Positivo Proyectado del Período 2: \$ 110
- Flujo Positivo Proyectado del Período 3: \$ 135
- Flujo Negativo Proyectado del Período 1: - \$ 50
- Flujo Negativo Proyectado del Período 2: - \$ 55
- Flujo Negativo Proyectado del Período 3: - \$ 60

En la siguiente Figura 2 se muestran los flujos de fondos del proyecto 1, que a los fines de este trabajo resultan similares a los del proyecto 2, aunque difieran en el riesgo considerado, tal como se señalará seguidamente.

**Figura 2. Flujos del Proyecto**

|                        | Períodos |        |        |        |
|------------------------|----------|--------|--------|--------|
|                        | 0        | 1      | 2      | 3      |
| Inversión Inicial      | -100,00  |        |        |        |
| Flujos + Proyectados   |          | 90,00  | 110,00 | 135,00 |
| Flujos - Proyectados   |          | -50,00 | -55,00 | -60,00 |
| Flujo Neto del Período | -100,00  | 40,00  | 55,00  | 75,00  |
| Flujo Neto Acumulado   | -100,00  | -60,00 | -5,00  | 70,00  |
| Tasa de Descuento      | 10,00%   |        |        |        |
| VAN                    | \$ 38,17 |        |        |        |
| TIR                    | 28,36%   |        |        |        |

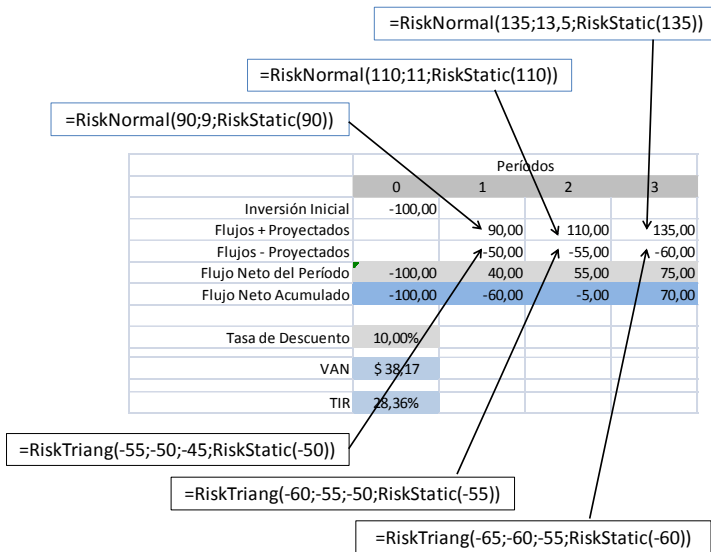
Por motivos de simplicidad, se ha omitido la consideración de factores adicionales de complejidad, como podrían ser los cambios en los precios relativos y la moneda homogénea.

Consideración del Riesgo: seguidamente incorporaremos el factor “riesgo” a los flujos de fondos de ambos proyectos; aunque luzcan aparentemente iguales, podremos apreciar que la variabilidad de sus flujos de fondos responde a distribuciones probabilísticas distintas, tal como se

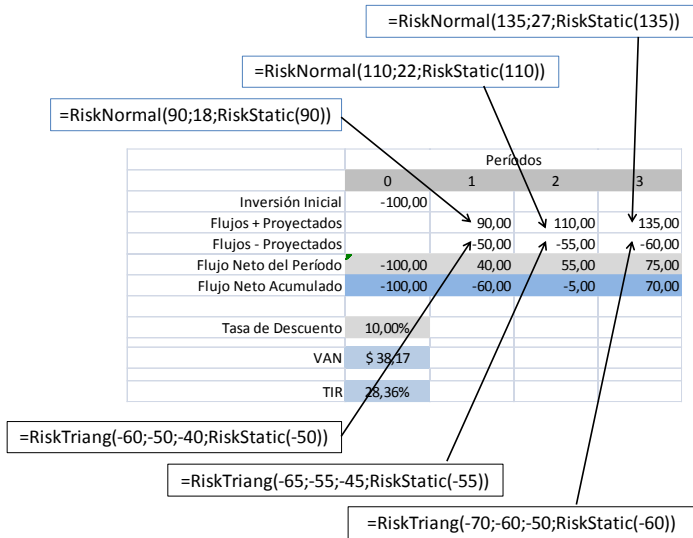
exponen en las Figuras 3 y 4 (y ello se debe a que el software utilizado en la ocasión, @Risk, ha sido configurado para devolver por pantalla el modo de la variable, en cada caso).

En las citadas Figuras 3 y 4 se observa que se han utilizado dos tipos de distribuciones: la Normal y la Triangular; aunque es de mencionar que de acuerdo con estudios que una organización podría realizar de sus datos, éstas bien podrían haber sido: Weibull, Log-normal, Exponencial, etc.

**Figura 3. Proyecto 1**



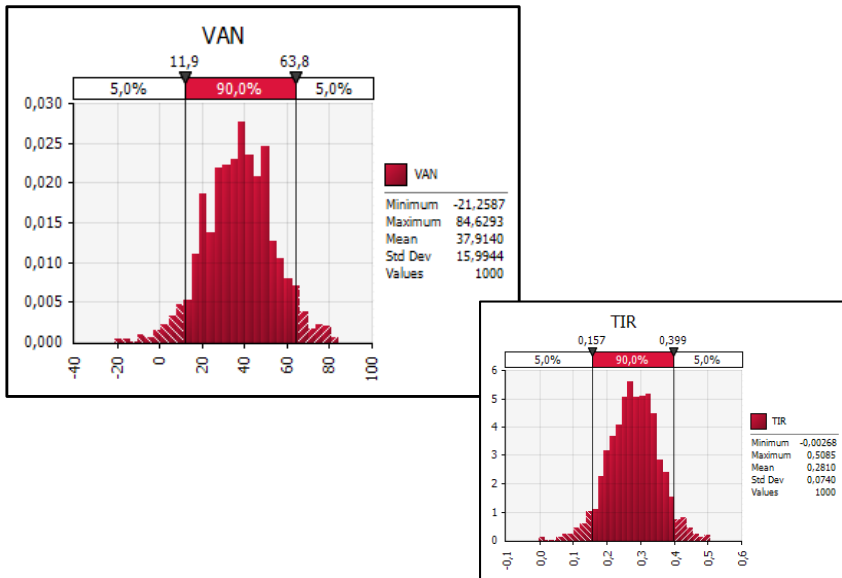
**Figura 4. Proyecto 2**



Al analizar los resultados de la simulación, realizada sobre la base de 1.000 iteraciones para cada proyecto (aunque muchas menos serían las efectivamente necesarias), se obtendrán los resultados expuestos en las Figuras 5 y 6:

**Figura 5. Histogramas (VAN y TIR) del Proyecto 1**

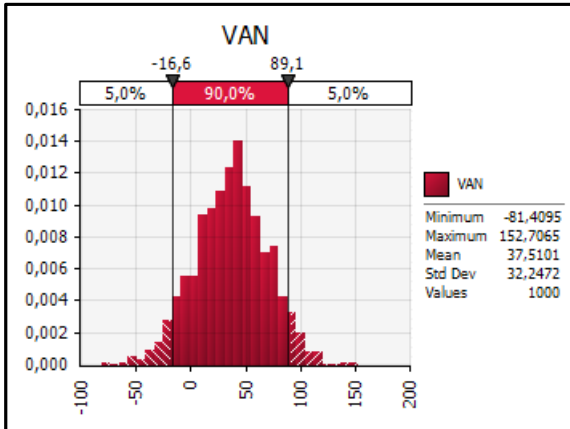
## Proyecto 1



A simple vista, nuevamente puede apreciarse que un histograma, comparado con un resultado único, despliega la dispersión posible de fondos (de acuerdo con las reglas definidas bajo cada distribución) que resulta en un conjunto de información mucho más rica que un simple número de VAN y TIR.

**Figura 6. Histograma (VAN) del Proyecto 2**

## Proyecto 2



#### 4. EL VALOR ESPERADO CONCENTRADO (VEC)

El concepto de VEC (Valor Esperado Concentrado) modifica al tradicional Valor Esperado dividiéndolo por el factor  $[1 + (\sigma / \mu)]$  con el objeto de darle menor peso específico a los flujos de fondos más distantes de la media.

$$VEC = \frac{VE}{1 + \frac{\sigma}{\mu}}$$

En su cálculo, intervienen:

- La Media o Esperanza Matemática  $\mu$ , que "es una medida de posición, porque únicamente nos indica el valor alrededor del cual

se agrupan los valores de la variable, pero no nos indica su densidad o concentración” (García, página 24).

- El Desvío Estándar  $\sigma$ , raíz cuadrada de la Varianza, que es una medida de dispersión.

El resultado es el Coeficiente de Variación  $\sigma / \mu$ , que le otorga al Desvío Estándar un sentido relativo, información que el dato puro y absoluto no está en condiciones de proporcionar.

Ambas herramientas, VEC y Simulación Monte Carlo, bien pueden conjugarse bajo un mismo enfoque que aporte un valor único de comparación entre alternativas (el VEC), que resultará un adicional y útil indicador sintético de los resultados de la simulación agrupados bajo un histograma, que le dará soporte.

**Figura 7. Valor Esperado Concentrado (VEC) del Proyecto 1**

| Proyecto 1                 |   |         |
|----------------------------|---|---------|
| Media --->                 |   | 37,9140 |
| Desvío Estándar --->       |   | 15,9944 |
| Coeficiente de Variación · | $\frac{15,9944}{37,9140}$                         | 0,4219  |
|                            | $\text{VEC} = \frac{\text{VE}}{(1 + \sigma/\mu)}$ |         |
|                            | $\frac{37,9140}{1,4219}$                          | 26,6651 |

**Figura 8. Valor Esperado Concentrado (VEC) del Proyecto 2**

| Proyecto 2                  |                                     |         |
|-----------------------------|-------------------------------------|---------|
| Media --->                  |                                     | 37,5101 |
| Desvío Estándar --->        |                                     | 32,2472 |
| Coefficiente de Variación · | $\frac{32,2472}{37,5101}$           | 0,8597  |
|                             | $VEC = \frac{VE}{(1 + \sigma/\mu)}$ |         |
|                             | $\frac{37,5101}{1,8597}$            | 20,1700 |

El VEC del proyecto 1 es mayor, y encuentra su fundamentación en el hecho de que los distintos valores que pueden asumir los proyectos se encuentran "corregidos" por un factor de actualización  $[1 + (\sigma / \mu)]$  que castiga la dispersión de los valores con relación a su media.

### 5. ¿PARA QUIENES PODRÍA RESULTAR DE UTILIDAD EL VEC?

En la Sección 2 quedó claro que el Valor Esperado, por sus limitaciones (útil sólo ante un gran número de experimentos, compensación de resultados positivos y negativos, etc.) no conjuga con la lógica de negocios u operativa de organizaciones de pequeño o mediano tamaño, ya que su adopción significaría asumir riesgos que podrían llevarlas a sufrir



pérdidas importantes, con consecuencias que derivarían en su extinción. Sin embargo el VEC, al "ajustar" los desvíos en relación a su distancia de la media de la distribución, atempera notablemente la compensación de resultados positivos y negativos extremos.

## **6. CONCLUSIONES**

Ante el uso indiscriminado del Valor Esperado, el que generalmente no es acompañado con aclaraciones sobre sus particulares debilidades como indicador válido de competencia entre proyectos por la asignación de recursos para su financiación, se presenta el Valor Esperado Concentrado (VEC) como nuevo indicador; en ese sentido, su uso, tímidamente, se está poniendo a prueba en toda ocasión en que se tiene oportunidad.

En nuestro país no son frecuentes los estudios específicos que describan cómo deciden los empresarios o los dirigentes de organizaciones en general, en situaciones como las que planteamos en este trabajo. Dentro de estos estudios, se debería presentar y poner en consideración la utilización del indicador VEC, de forma tal de comenzar a obtener opiniones propias de aquéllos que efectivamente toman decisiones.

Por otro lado, y dado que la utilización de Simulación Monte Carlo profesionaliza la forma de evaluar el riesgo propio de cada proyecto, de agregársele el VEC a su tratamiento para aprovechar así el aporte de sus resultados (salidas), trazaríamos un camino eficaz para asignar probabilidades a las distintas alternativas de un proyecto, o a proyectos distintos.

## **REFERENCIAS BIBLIOGRÁFICAS**

García, Roberto M. (2004): *Inferencia Estadística y Diseño de Experimentos*. Buenos Aires, EUDEBA.

Hertz, D. B. (1979): *Risk analysis in capital investment – HBR Classic*. Boston, Harvard Business Review.

Pavesi, P. F. J., Bonatti, P. y Avenburg, D. (2004): *La Decisión*. Buenos Aires, Editorial Norma.

Pungitore, J. L. (2003): "Planeamiento Económico y Financiero" en *Contextos Complejos y Turbulentos*. Buenos Aires, Editorial Osmar Buyatti.

Pungitore, J. L. (2009): "Simulación Monte Carlo en la Proyección de Estados Contables y Flujos de Fondos" en *Revista Desarrollo y Gestión Nro. 113*. Buenos Aires, Editorial Errepar .

Pungitore, J. L. (2010): "Evaluación de Proyectos en Condiciones de Riesgo: una Alternativa al Criterio de Valor Esperado" en *Revista Desarrollo y Gestión Nro. 134* - Buenos Aires, Editorial Errepar.

Rheault, J. P. (1980): *Introducción a la Teoría de las Decisiones*. México, Editorial Limusa.

# UNA PROPUESTA DE ÍNDICE DE ACTIVIDAD CIENTÍFICA Y SU VALIDACIÓN EN EL CAMPO DE CIENCIAS SOCIALES Y HUMANIDADES

*Julieta Barrenechea*

*Javier García Fronti*

*Javier Castro Spila*

*Andoni Ibarra*

## INTRODUCCIÓN

Los Ranking académicos internacionales que permiten comparar universidades a nivel global se han generalizado en los últimos años (Buena-Casal et al., 2007). Uno de los más destacados es el publicado por The Institute of Higher Education, Shanghai Jiao Tong University (Liu and Cheng, 2005) siendo su criterio metodológico predominante la publicación en revistas científicas.

Asimismo se han realizado clasificaciones a nivel regional europeo (Merisotis, 2002) y también a nivel de país como por ejemplo el CHE en Alemania (Federkeil, 2002) o el RAE en Reino Unido (Eccles, 2002). En las mencionadas clasificaciones, las publicaciones académicas tienen un rol cada vez más importante a la hora de evaluar la calidad de investigación en diferentes contextos académicos. Por ejemplo, (Bence and Oppenheim (2004) describen la evolución de su importancia en la evaluación gubernamental de la investigación en universidades en Reino Unido (RAE). Si bien formalmente se trata de un proceso de evaluación de pares, los autores remarcan el rol hegemónico de las publicaciones en el proceso de evaluación.

La empresa *Thomson Reuters*<sup>19</sup>, por medio de su base de datos "Science", se ha transformado en el árbitro de las revistas científicas a nivel mundial. Su sitio Web describe el Informe de citas de revistas que ellos

---

<sup>19</sup> <http://thomsonreuters.com/>

producen como "un medio sistemático y objetivo para evaluar de manera crítica las revistas más importantes del mundo"<sup>20</sup>. Su uso extendido otorga a la empresa un increíble poder sobre la comunidad académica y sobre los organismos de financiación. Isagani Cruz (2008) es uno de los exponentes de quienes proponen deconstruir esta situación hegemónica y critica tres conceptos fundacionales de su lógica argumental: a) sostiene que la metodología no es objetiva, b) que su evaluación de impacto no es correcta y c) que no es global pues tiene en cuenta principalmente las publicaciones en idioma inglés.

En el caso de España por ejemplo, Moya-Anegón *et al* (2004) realizaron un informe en el cuál sistematizan información sobre la producción científica del país con el fin de ofrecer un instrumento de política pública. Sin embargo su metodología utiliza como información base solamente las publicaciones internacionales indexadas por la empresa *Thomson Reuters* y como consecuencia gran parte de la producción científica ha quedado fuera de su consideración.

En el caso específico de las ciencias sociales y humanidades (CSH) la evaluación con foco en las publicaciones indexadas en Thomson Reuters es mucho más problemática. Estudios como los de (Becher, 1993), (Kivik, 2003); (Leeuwen, 2006); (Archambault, et al. 2006), ponen de relevancia la heterogeneidad que existe entre las disciplinas de CSH en sus modos de organizar la investigación y en los patrones de producción y distribución de resultados.

A partir del VII Programa Marco Europeo se asigna una mayor importancia relativa a las CSH ya que se incluyen líneas de investigación específicas para estas disciplinas que dejan así de ser subsidiarias de las Ciencias Básicas y Tecnológicas, como ha venido sucediendo en los anteriores Programas Marco.

En este nuevo escenario europeo, es un tema de agenda el problema de contar con indicadores para evaluar la actividad de investigación en CSH, puesto que no existen claros consensos sobre cuáles son los parámetros

---

<sup>20</sup> <http://ip-science.thomsonreuters.com/es/productos/jcr/>

más indicados para captar y valorar las actividades de investigación en estos campos disciplinarios. A pesar de este déficit, existen amplios acuerdos en relación con lo insuficiente y a veces inadecuado, que resulta aplicar al caso de las CSH criterios de evaluación elaborados para las Ciencias Básicas. Las principales dificultades conceptuales a la hora de evaluar las ciencias sociales han sido recogidas en (Ibarra et al: 2006) y se refieren a tres temas: a) las CSH tienen pautas de funcionamiento distintas en su actividad científica respecto de las ciencias experimentales (trabajo individual, mayor orientación hacia ámbitos locales, prácticas de colaboración menos frecuentes), b) la interpretación de sus resultados es dificultosa porque no se cuenta con criterios consolidados para valorar sus publicaciones, visibilidad, modalidades de colaboración y relevancia según sus ámbitos de contribución (local); c) los patrones o estándares de resultados y difusión son poco reconocido dentro de la lógica del status científico (exposiciones, artículos de revista locales, recensiones, exposiciones, etc.)

En función de lo anterior, si se reduce la complejidad de estos campos disciplinarios a unos pocos criterios de evaluación, se corre el riesgo de no evaluar lo que se debe evaluar o, simplemente, evaluar transfiriendo criterios desarrollados en otros campos científicos y aplicarlos al *modo de hacer* de las disciplinas de las CSH. Asumir esta heterogeneidad disciplinaria y su consiguiente complejidad evaluadora implica asumir el desafío de diseñar un modelo de evaluación que sea *integral*.

Con *integral*, en este trabajo nos referimos a dos planos: el primer plano se vincula con la idea de que a partir de una definición más inclusiva de *actividad científica*, será posible captar una mayor variedad en las prácticas y tipo de resultados que permitan procesar de manera analítica el funcionamiento de nuevos *patrones disciplinarios de investigación* (enfoque descriptivo), y al mismo tiempo, establecer de manera sintética niveles o estándares de productividad y calidad (enfoque evaluativo). El segundo plano está relacionado con el hecho de que la calidad de la actividad científica depende cada vez más de la capacidad de interactuar y mantener vinculaciones efectivas con una gama más amplia y compleja de esferas y agentes sociales. Una evaluación *integral* supone entonces que a la hora de evaluar un centro de conocimiento, se necesita tener en cuenta un escenario más diverso que el estrictamente científico-académico. Es

necesario desarrollar e implantar estrategias metodológicas que integren de forma complementaria las valoraciones que sobre la actividad científica tiene un conjunto más amplio de agentes sociales. El concepto de *calidad relacional* que articula esta propuesta integral de evaluación, está basado en la idea de que, en las condiciones actuales y frente a las exigencias que se plantean al sector científico-académico, la producción, distribución y uso de conocimiento científico alcanzará mejores estándares de calidad en la medida en que se garanticen buenas condiciones relacionales para estos procesos.

El Índice Directo de Actividad Científica (IDAC) que se presenta en este artículo forma parte de los resultados de un estudio desarrollado durante el año 2005 referido a la evaluación y el papel de las Ciencias Sociales y Humanidades de la UPV/EHU y a su inserción en el contexto de innovación de la Comunidad Autónoma del País Vasco (CAPV) <sup>21</sup> y cuyos principales resultados se han sistematizado en Barrenechea et al (2008) .

En el marco de dichos estudios se ha desarrollado una estrategia evaluativa que plantea un modelo de *Programa Integral de Evaluación de la Actividad Científica en CSH* organizado en cuatro fases consecutivas y articuladas que suponen triangulación metodológica.

Green et al (2001: 30) señalan que la triangulación metodológica en evaluación permite reducir la incertidumbre y comprender mejor los fenómenos estudiados. Tomando como referencia los trabajos de Cook (1985) y Greene et al. (1989) señalan 4 dimensiones en las que la triangulación metodológica es un aporte importante para la evaluación científica: a) mejora la validez y la credibilidad de las inferencias, b) ofrece una visión más comprensiva y contextual de los resultados y sus impactos, c) permite un entendimiento más profundo sobre la actividad que se está evaluando, d) aumenta la conciencia respecto de la diversidad de valores que están en juego en la actividad científica y es más inclusivo.

---

<sup>21</sup> El estudio citado ha sido desarrollado por el equipo del Área de Investigación en Gestión de Redes de Ciencia, Tecnología e Innovación, de la Cátedra Sánchez-Mazas de la Universidad del País Vasco/Euskal Herriko Unibertsitatea.

La primera fase, incluye un proceso de *evaluación por indicadores* que tiene el valor de un autodiagnóstico sobre la base de un modelo de evaluación integral de la actividad científica en CSH. En la segunda fase, se propone un proceso de evaluación por panel de expertos (*peer review*), que es complementario de la evaluación por indicadores, y que tiene el valor de una evaluación externa. Lo que se plantea para esta segunda fase está a medio camino entre una evaluación institucional (que valora condiciones académicas e institucionales de la investigación), y una evaluación de la calidad de la investigación (que valora los resultados obtenidos por los procesos de investigación según parámetros disciplinarios). La tercera fase, incluye el diseño y aplicación de indicadores relacionales que den cuenta de la conectividad académica y social de los investigadores. Se desarrolla a partir de una estrategia de consulta a investigadores y agentes sociales, y apunta a identificar en términos descriptivos, evaluativos y prospectivos la oferta académica y la demanda social efectiva y potencial, con el fin de fortalecer dinámicas de conectividad y colaboración entre los centros de conocimiento y su entorno social relevante (mediato e inmediato). La cuarta fase está orientada a la implementación de mecanismos que favorezcan acuerdos institucionales al interior de la comunidad científica en relación con la mejora de condiciones para las actividades de investigación.

El calificativo de *integral* por el que se apuesta en este trabajo, expresa dos niveles de análisis que hacen a la producción, distribución y uso del conocimiento científico. Por una parte, se vincula con la definición de actividad científica y expresa la necesidad de considerar de manera sistémica un conjunto más amplio de dimensiones que forman parte de la actividad científica, y que complementan a aquéllas que se centran estrictamente en la obtención de productos y en los mecanismos y soportes que se ponen en juego para la difusión de los resultados de investigación. En cambio en la noción integral de actividad científica, se incorporan elementos como la formación de recursos humanos, la participación en redes académicas, la trayectoria del personal investigador, la consideración de productos menos ortodoxos, etc.

En el caso de las CSH asumir una perspectiva integral y aplicar como enfoque la triangulación metodológica se hace muy necesario en tanto

que como se ha mencionado, bajo la etiqueta de las CSH se aglutinan un conjunto heterogéneo de disciplinas, cuyos modos de producir y distribuir conocimiento son diferentes y por ello resulta difícil que se reduzcan a unos cuantos criterios estandarizados de evaluación. A esta situación de heterogeneidad disciplinaria se suma el hecho de que el debate y las técnicas que se aplican a la evaluación de la actividad científica de las CSH está lejos aún de lograr consensos y de contar con herramientas sistematizadas. En este contexto paradigmático y técnico, una aproximación integral para la evaluación de la actividad científica tiene también objetivos exploratorios: se propone captar el funcionamiento de patrones disciplinarios de investigación y, en ese marco, reunir elementos que permitan en el mediano plazo establecer niveles de calidad.

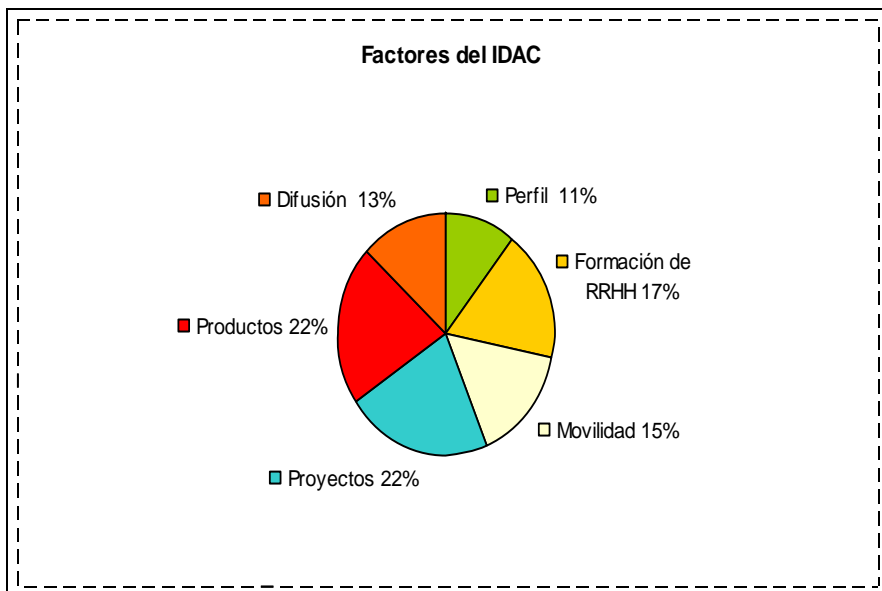
El Índice IDAC que se presenta a continuación, forma parte de las técnicas que se aplican en la fase II del Programa Integral de Evaluación desarrollado en el marco del estudio mencionado y para su cálculo se utiliza la base de datos que resulta de la aplicación de la encuesta de evaluación de la actividad científica elaborada también para este estudio. El índice ha sido diseñado para ser calculado en función de la ponderación de los factores e indicadores que componen el modelo de indicadores de evaluación. En este artículo se incluyen también los resultados de la consulta de validación con expertos que se ha aplicado en el marco del estudio mencionado y que recoge el análisis de las propuestas de ponderación para los indicadores del modelo

El IDAC es una medida síntesis que permite cuantificar en un solo valor la actividad científica de investigadores o grupos para un período determinado. Sin embargo, el IDAC se compone de 6 índices parciales que permiten medir la actividad científica en los factores que lo componen: Perfil, Formación de Recursos Humanos, Movilidad, Proyectos, Productos y Difusión de la Investigación. En este sentido, ofrece la posibilidad de analizar perfiles investigadores así como identificar patrones de actividad científica según la combinatoria que resulte de estos 6 factores. El índice puede también ser utilizado como herramienta de gestión por objetivos.



## 1 MODELO CONCEPTUAL DEL IDAC

En el marco de la Fase I del Programa Integral de Evaluación desarrollado en el estudio mencionado, se incluye la aplicación y cálculo de un índice de actividad científica que ha sido diseñado en función de la ponderación de los indicadores de evaluación que componen el modelo de evaluación por indicadores. El Índice Directo de Actividad Científica (IDAC) permite cuantificar en un solo valor la actividad científica de investigadores o grupos para un período determinado. El IDAC se compone a su vez de 6 índices parciales que permiten medir la actividad científica en los factores: Perfil, Formación de Recursos Humanos, Movilidad, Proyectos, Productos y Difusión de la Investigación. En el siguiente gráfico se presentan los seis factores y sus ponderaciones que expresan su contribución relativa a la actividad científica considerada globalmente:



Los seis factores están compuestos por un conjunto de indicadores que integran los distintos niveles ponderados en los que se desagrega el índice, hasta llegar al nivel más empírico compuesto por datos que se

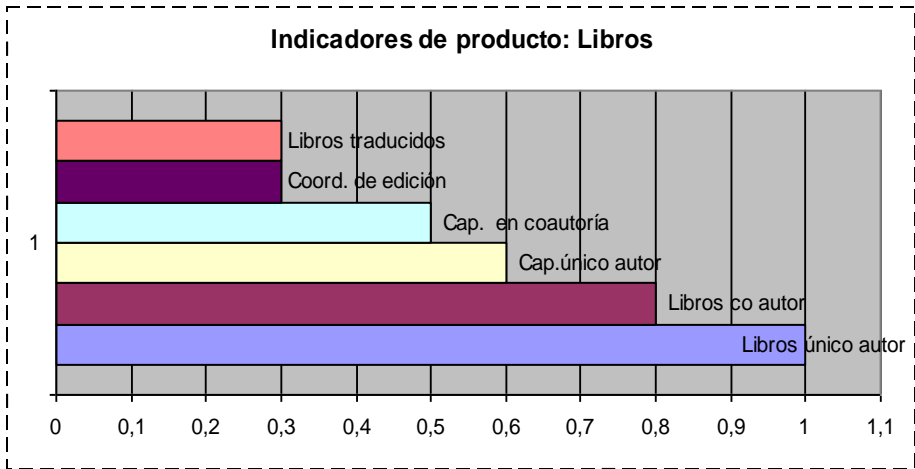
recogen mediante una medición basada en la aplicación de un cuestionario a los investigadores individuales. El número de niveles de desagregación varía según cada factor siendo el mínimo 2, como en el caso de *movilidad académica*, y 4 el máximo, como en el caso del factor *productos* cuyo indicador revistas se desagrega hasta un cuarto nivel al que corresponde una ponderación P4.

El sistema de ponderaciones del *modelo conceptual* del índice responde en todos sus niveles (factores, indicadores, ítem) a conocimientos y supuestos teóricos previos, y combina dos modalidades:

1. Ponderaciones porcentuales.
2. Valoración por equivalencias.

1 – En el caso de las ponderaciones porcentuales el valor atribuido representa la aportación de cada uno de los factores o indicadores al total. Por ejemplo, en el caso de los 6 factores, el porcentaje atribuido a cada uno de los factores equivale a la aportación de cada factor a la actividad científica considerada como un todo. Se representa en el Gráfico 1.

2 – En el caso de la valoración por equivalencias el indicador que es valorado con "1" es el referente. El resto de los indicadores del grupo son traducibles, es decir, guardan equivalencia con este indicador. Por ejemplo, en el caso de los productos *libros* el concepto *libros único autor* = 1 y libros coautor = 0.8, de modo que un libro coautor equivale a 0.8 libros único autor. Conceptualmente lo que se calcula en ese grupo de indicadores y que se obtendrá como resultado, es un valor que represente la *cantidad de equivalentes de libro único autor* que el investigador o grupo ha producido en el período estudiado. El gráfico de barras permite visualizar las equivalencias.



Cada una de las modalidades de ponderación supone una exigencia distinta respecto de la actividad investigadora que desarrollan los investigadores o grupos. En el caso de la ponderación porcentual, para poder alcanzar la máxima puntuación los investigadores deben cumplir con todos los factores o conceptos que conforman el grupo ponderado. En el caso de la valoración por equivalencia cumplir con todos los indicadores del grupo no es un requisito para obtener la máxima puntuación. Esta última modalidad de ponderación se aplica a indicadores de dos tipos: (a) indicadores acumulativos y/o excluyentes por definición, como en el caso del nivel de formación donde si se es doctor no se necesita ser master y se es indefectiblemente licenciado y (b) indicadores en los que no se considera una exigencia razonable que un investigador tenga que cumplir con todos los indicadores del grupo, por ejemplo, no se considera obligatorio que un investigador produzca traducciones o reseñas además de capítulos de libro como único autor. Aunque aquel investigador que tenga una buena producción y en conceptos diversos seguramente alcanzará una mejor puntuación en el grupo de indicadores.

## 2 METODOLOGÍA PARA EL CÁLCULO DEL IDAC

La información que se toma en consideración para el cálculo del IDAC proviene de la aplicación de un cuestionario a investigadores individuales. Los valores que se toman como datos se llamarán "v" y pueden ser cualidades (ej. ser o no doctor), cantidades (ej. número de artículos en coautoría) o porcentuales (ej. intensidad investigadora). En el caso de los indicadores con ponderación porcentual, estos valores deben estandarizarse en un número que varía entre 0 y 1 antes de ser ponderados, de modo que para estos casos al valor más empírico o que proviene de un nivel de cálculo anterior se lo llamará "q", y se llamará "v" al valor estandarizado. Se estandariza en función del valor máximo (M) obtenido en ese indicador por la población bajo estudio diferenciada según áreas disciplinarias. El algoritmo de cálculo del IDAC se basa entonces en la suma ponderada de valores estandarizados en función de máximos que resultan de la muestra o población estudiada.

El IDAC resulta de la suma ponderada de los subíndices de sus factores y oscila entre 0 y 100. Los subíndices de cada uno de los factores en que se descompone la actividad científica resultan de la suma ponderada de las distintas variables convenientemente estandarizadas que lo integran y también oscilan entre 0 y 100. La ponderación de los valores obtenidos en cada indicador constituye la forma canónica del término que integra la fórmula de cálculo del índice IDAC y de cada uno de los subíndices que lo componen, y es igual a  $(v^{ni} p^{ni})^{22}$ .

La estandarización de las variables se realiza en distintos momentos, según se trate de indicadores con ponderación porcentual o indicadores con valoración por equivalencia, y toma en cuenta el valor máximo (M) obtenido por la población bajo estudio diferenciada según áreas disciplinarias.

Dependiendo del número de niveles en que se desagrega cada uno de los factores de actividad científica, se necesitará realizar distinto número de pasos para obtener cada subíndice de factor. El cálculo se inicia multiplicando el "v" obtenido en los indicadores de mayor nivel de desagregación por su valor o ponderación, siendo 4 el máximo nivel y 2 el

---

<sup>22</sup> Más abajo puede verse la definición completa de la fórmula.

mínimo. Todos los pasos del cálculo del IDAC respetan el algoritmo del índice de modo que se obtienen subíndices para cada grupo de indicadores de distinto nivel de desagregación.

$$S^n = \begin{cases} \sum_{i=1}^{kn} v^{ni} P^{ni} \\ \sum_{i=1}^{kn} \frac{v^{ni} P^{ni}}{M^n} \end{cases}$$

1) El subíndice (S) del indicador n es igual a la sumatoria de los términos de la forma  $V P$ .

2) El subíndice (S) del indicador n es igual a la sumatoria de los términos de la forma  $v P$ , dividida por el máximo valor de sumatoria obtenido por la población bajo estudio y perteneciente a la misma área disciplinaria.

Donde:

$n$  es el número entero positivo que representa el nivel (Ej. Perfil  $n=1$ )

$S^n$  es la puntuación (expresada en porcentajes) obtenida en el nivel  $n$

$i$  es el número entero que varía entre 1 y  $k$  (Ej. para el indicador Grado académico: Doctor 1.1, Master 1.2, Licenciatura 1.3, Diplomatura 1.4)  $i= 1$  y  $k= 4$

$k$  define el número de términos que componen el indicador  $n$

$M$  es el máximo valor obtenido por los investigadores de la disciplina en el nivel  $n$ .

## 2.1 INTERPRETACIÓN DEL ÍNDICE

Por su construcción el índice IDAC mide la posición relativa de un investigador o un grupo (departamento, instituto, etc) con respecto al rango de la población estudiada y para el período medido, y adoptará un valor igual a 100 en el caso en que el investigador o grupo presente el valor máximo de la población y el período estudiado, y 0 si presenta el valor mínimo.

## 2.2 PROCEDIMIENTO DE CÁLCULO DEL IDAC

Tanto en el caso de indicadores con ponderación porcentual como en los que tienen valoraciones por equivalencia, el método de ponderación es directo. Sin embargo el cálculo de los subíndices para los grupos de indicadores de uno y otro tipo difiere en algunos pasos que se describirán a continuación.

## 2.3 PROCEDIMIENTO DE CÁLCULO PARA INDICADORES CON PONDERACIÓN PORCENTUAL

El subíndice (o puntuación) obtenido por un investigador para un grupo de indicadores con ponderación porcentual es el resultado de realizar los siguientes pasos:

- a) Obtener  $v =$  para cada indicador que compone el grupo, estandarizar los valores obtenidos en el nivel empírico o en el nivel de desagregación anterior en números entre 0 y 1.
- b) Ponderar  $v =$  multiplicar el  $v$  obtenido en cada indicador por su ponderación.
- c) Sumar los términos del indicador: sumar el total de términos ( $v^{ni}$   $P^{ni}$ ) que componen el indicador.

### a) Obtención de $v$ por estandarización

Para los valores que se han llamado  $q$  y que se expresan en cantidades o porcentajes obtenidos por un investigador en un indicador con ponderación porcentual, el procedimiento de estandarización consiste en tomar como referencia el máximo valor obtenido en ese indicador para el área disciplinaria que corresponda y en el caso de la aplicación/evaluación que se está analizando, y dividir el valor por el máximo obtenido en el área disciplinaria. Aun cuando las mediciones se realicen para períodos superiores a 1 año (ej 5 años), los máximos serán los promedios anuales.

$$v^{ni} = \frac{q^{ni}}{Mx^{ni}}$$

2.4  $Mx^{ni}$

#### b) Ponderación de $v$

La ponderación de los valores obtenidos en cada indicador constituye la forma canónica del término que integra la fórmula de cálculo del índice IDAC y de cada uno de los subíndices que lo componen. Deberá realizarse este cálculo para cada indicador  $n$  desde su máximo nivel de desagregación hasta el mínimo (el máximo nivel posible es 4 y el mínimo 1); y para todos los términos de cada indicador, es decir, para  $i$  desde 1 hasta  $k$ .

$$(v^{ni} \cdot p^{ni})$$

#### c) Sumatoria de los términos

Para cada grupo de indicadores con ponderación porcentual que integra el modelo conceptual del IDAC es posible calcular un subíndice realizando la sumatoria de todos los términos  $(v^{ni} \cdot p^{ni})$  que lo componen.

$k_n$

$$S^n = \sum_{i=1}^{k_n} v^{ni} \cdot p^{ni}$$

Cada subíndice calculado será llamado  $v$  cuando se sube un nivel de indicador, y se comporta como un valor a ser ponderado para el cálculo de un subíndice de menor nivel de desagregación de un indicador con valoración por equivalencia.

Cada subíndice calculado será llamado  $q$  cuando se sube un nivel de indicador y se comporta como un valor a ser estandarizado para el cálculo de un subíndice de menor nivel de desagregación de un indicador con ponderación porcentual.

## 2.5 PROCEDIMIENTO DE CÁLCULO PARA INDICADORES CON VALORACIÓN POR EQUIVALENCIA

El subíndice obtenido por un investigador para un grupo de indicadores con valoración por equivalencia es el resultado de realizar los siguientes pasos:

- a) Ponderar  $v$ : multiplicar el  $v$  obtenido en cada indicador por su ponderación.
- b) Calcular las unidades ( $U$ ): realizar la sumatoria del total de términos ( $v^{ni} P^{ni}$ ) que componen el indicador.
- c) Estandarizar las unidades: dividir el resultado de la sumatoria de los términos, sobre el valor del máximo valor en unidades obtenido en el mismo indicador por la población bajo estudio perteneciente a la misma área disciplinaria.

a) Ponderación de  $v$ .

La ponderación de los valores obtenidos en cada indicador constituye la forma canónica del término que integra la fórmula de cálculo del índice IDAC y de cada uno de los subíndices que lo componen.

$$(v^{ni} P^{ni})$$

b) Cálculo de unidades ( $U$ )

El resultado de la sumatoria de los términos que componen un grupo de indicadores con valoración por equivalencia representará el total de la producción científica realizada por el investigador o grupo en relación con ese ítem medido en la unidad de equivalencia correspondiente. Se denominará unidad  $U$  seguido por el número del indicador correspondiente.



kn

$$U^n = \sum_{i=1}^n v^{ni} p^{ni}$$

c) Estandarización de Unidades.

Para todos los grupos de indicadores con valoración por equivalencia, el subíndice se obtiene una vez que se han estandarizado los resultados de la sumatoria de términos o "unidades" respecto del máximo valor en unidades obtenido en ese mismo grupo de indicadores por la población bajo estudio que pertenece al mismo campo disciplinario.

$$S^n = \frac{U^n}{M^n}$$

Cada subíndice calculado será considerado  $v$  cuando se sube un nivel de indicador y se comporta como un valor a ser ponderado para el cálculo de un subíndice de menor nivel de desagregación.

Cada subíndice calculado será llamado  $q$  cuando sube un nivel de indicador y se comporta como un valor a ser estandarizado para el cálculo de un subíndice de menor nivel de desagregación de un indicador con ponderación porcentual.

### **3 RESULTADOS DE LA CONSULTA DE VALIDACIÓN**

En el marco del estudio desarrollado para el caso de la UPV/EHU, con el fin de validar el modelo de indicadores y las ponderaciones asignadas para el cálculo del IDAC, se ha realizado una consulta a expertos e investigadores de CSH tanto del ámbito de la Comunidad Autónoma del País Vasco (CAPV), como del ámbito estatal y europeo. El fin ha sido recoger sus opiniones sobre los indicadores diseñados y evaluar los valores de ponderación asignados a los distintos factores e indicadores que se incluyen en el índice de actividad científica.

La consulta ha consistido en el envío de un cuestionario electrónico a una muestra de expertos en evaluación y temas de educación superior y a un pequeño número de investigadores de CSH en su calidad de usuarios de procesos de evaluación.

Los consultados debían establecer un valor de ponderación en porcentaje para cada uno de los indicadores que componen el IDAC. La ponderación de cada indicador debía guardar relación con la ponderación del resto de los indicadores del mismo grupo, de modo que la suma de las ponderaciones de los indicadores del mismo grupo debía ser 100<sup>23</sup>. El cuestionario ha incluido la posibilidad de realizar comentarios abiertos por escrito y si fuera necesario señalar la *no aplicabilidad* o la *no pertinencia* de los indicadores propuestos. Los consultados no tuvieron acceso a los valores de ponderación asignados en el modelo que se estaba validando.

Se han obtenido 28 respuestas válidas que corresponden a una población conformada por: a) 11 investigadores de universidades españolas y del Reino Unido no expertos y vinculados a las áreas de educación, economía, ciencias políticas, historia, derecho, biblioteconomía, periodismo y psicología; b) 17 expertos de los cuáles: 9 forman paneles de evaluación para las áreas de CSH de la FECYT (Fundación Española para la Ciencia y la Tecnología), 4 participan de la RESUP (*Projet de Réseau d'Étude et de Recherche sur l'Enseignement Supérieur-Francia*); 1 pertenece al TASTI (*Research Group for Science, Technology and Innovation Studies – Finlandia*), 2 son especialistas en educación superior de la Universidad de Bath (Reino Unido) y, por último, un especialista en filosofía de la ciencia de la Universidad Autónoma de Barcelona (Cataluña).

La consulta ha tenido un carácter exploratorio. No se trata de una muestra representativa en relación con la pertenencia disciplinaria de los consultados, pero aún así el número de expertos que han respondido es importante (17) si se tiene en cuenta que han sido seleccionados en

---

<sup>23</sup> Un grupo de indicadores está conformado por todos los indicadores en los que se desagrega un indicador o factor del nivel de desagregación anterior.

función de su experiencia en evaluación<sup>24</sup>. De este modo, la técnica aplicada y el tamaño y la composición de la población que ha respondido al cuestionario permiten extraer conclusiones y realizar análisis cualitativamente significativos a partir de las respuestas obtenidas.

A continuación se analizan los resultados obtenidos para el nivel principal de indicadores que son los 6 factores que componen la actividad científica de acuerdo a la forma en que se la ha definido para el cálculo del IDAC. Se toma como referencia el valor de ponderación que se ha asignado a cada factor en el modelo del IDAC, y se presentan para cada caso: a) todos los valores de ponderación asignados por los consultados, b) la frecuencia de respuestas que se han obtenido para cada valor y c) el porcentaje que representan las frecuencias sobre el total de respuestas. Además se ha realizado el cálculo de la media, la mediana y el desvío estándar que se analizan junto con los resultados obtenidos.

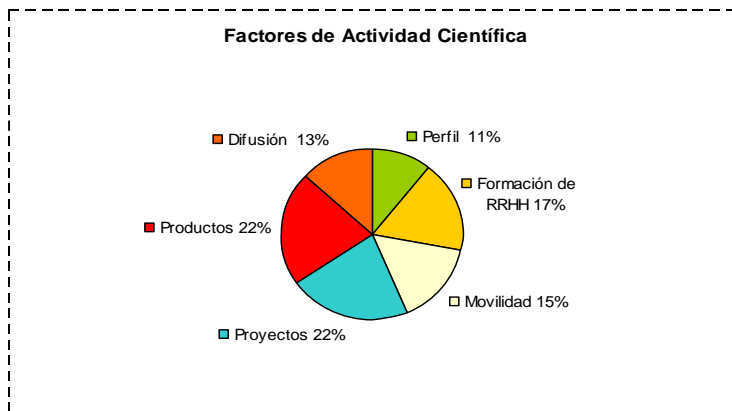
### *Factores de Actividad Científica*

En el Gráfico 14 se presentan los seis factores de actividad científica que componen el cuerpo del modelo de evaluación por indicadores y del modelo de cálculo del índice de actividad científica (IDAC). Los valores de ponderación que se les ha asignado expresan la contribución relativa de

---

<sup>24</sup> Cuando se utiliza la técnica de evaluación por pares académicos el criterio principal para la selección de los expertos es su perfil académico profesional y su experiencia en la temática a evaluar, esto significa que la representatividad de la muestra se basa en criterios cualitativos más que cuantitativos. La evaluación basada en el juicio de pares corresponde al modelo de investigación académica disciplinaria; la interpretación de los resultados de este ejercicio de validación debe encuadrarse en este contexto. Teniendo en cuenta que el consenso académico no es el único marco de relaciones sociales en que se constituye el conocimiento, y asumiendo una visión más interactiva de la ciencia, si se quisiera realmente ponderar incorporando la valoración y las prioridades sociales respecto de los factores de actividad científica, debería complementarse este estudio con otros procedimientos que sometan la validación del modelo al juicio de otros agentes sociales.

cada uno de los factores a la actividad científica de un investigador, o de una unidad de investigación considerada globalmente.



### *Análisis de resultados*

#### **Factor 1: Perfil del PDI**

| <b>Tabla 1: Medidas estadísticas Factor 1</b> |         |
|---|---------|
| 1 - Ponderación del modelo                    | 11%     |
| 2 - Media                                     | 14,96 % |
| 3 - Mediana                                   | 10      |
| 4 - Desvío estándar                           | 9,92    |

**Tabla 2: Resultados obtenidos para el Factor Perfil**

| <b>Ponderaciones asignadas</b> | <b>Cantidad de respuestas</b> | <b>Porcentaje de respuestas</b> | <b>Porcentaje acumulado</b> |
|--------------------------------|-------------------------------|---------------------------------|-----------------------------|
| 5%                             | 4                             | 14%                             | 14%                         |
| 6%                             | 1                             | 4%                              | 18%                         |
| 8%                             | 1                             | 4%                              | 21%                         |
| 10%                            | 9                             | 32%                             | 54%                         |
| 15%                            | 4                             | 14%                             | 68%                         |
| 20%                            | 5                             | 18%                             | 86%                         |
| 25%                            | 1                             | 4%                              | 89%                         |
| 30%                            | 2                             | 7%                              | 96%                         |
| 50%                            | 1                             | 4%                              | 100%                        |
| <b>Total</b>                   | <b>28</b>                     | <b>100%</b>                     |                             |

Con respecto al factor de actividad científica "Perfil del PDI" la media de las ponderaciones asignadas (14,96) supera en casi un 5% a la ponderación propuesta en el modelo del IDAC (11%). Si bien el mayor porcentaje de respuestas 32,14% corresponde a la asignación de 10% como valor de ponderación (la más cercana a la asignada en el modelo del IDAC), se debe tener en cuenta que el desvío estándar es significativo 9.92 y que existe una diferencia de casi 5% entre los valores de la media y la mediana. El segundo lugar en el orden de las ponderaciones asignadas

corresponde al valor de ponderación 20% con 17,86 % de las respuestas y el tercer lugar a los valores de ponderación 5% y 15% con un 14% de respuestas cada uno. De acuerdo a los datos obtenidos y a la aplicación de medidas estadísticas, es posible concluir que existe un bajo nivel de consenso respecto de la ponderación que debe asignarse a este factor para la evaluación de la actividad científica.

## **Factor 2. Formación de recursos humanos.**

| <b>Tabla 3: Medidas estadísticas Factor 2</b> |        |
|---|--------|
| 1 - Ponderación del modelo                    | 17%    |
| 2 - Media                                     | 11,29% |
| 3 - Mediana                                   | 10     |
| 4 - Desvío estándar                           | 5,52   |

**Tabla 4: Resultados obtenidos para el Factor Recursos Humanos**

| <b>Ponderaciones asignadas</b> | <b>Cantidad de respuestas</b> | <b>Porcentaje de respuestas</b> | <b>Porcentaje acumulado</b> |
|--------------------------------|-------------------------------|---------------------------------|-----------------------------|
| 0%                             | 2                             | 7%                              | 7%                          |
| 5%                             | 4                             | 14%                             | 21%                         |
| 8%                             | 1                             | 4%                              | 25%                         |
| 10%                            | 9                             | 32%                             | 57%                         |
| 15%                            | 8                             | 29%                             | 86%                         |
| 18%                            | 1                             | 4%                              | 89%                         |
| 20%                            | 3                             | 11%                             | 100                         |
| <b>Total</b>                   | <b>28</b>                     | <b>100</b>                      |                             |

En este caso tanto la media 11,29% como la mediana 10% se distancian del valor de ponderación propuesto en el modelo por 6 y 7 puntos respectivamente. Aunque sólo representan un 7% de las respuestas, es notable la existencia de casos en que se ha asignado ponderación nula a este factor de actividad científica. Si se agregan con los que han asignado el menor valor 5%, se concluye que el 21% de los consultados otorgan un peso muy débil a la formación de recursos humanos como factor de evaluación de la actividad científica. Sin embargo, el desvío estándar no es alto y si se suman los porcentajes, puede observarse que el 65% de las respuestas se concentran en un rango de valores comprendido entre 10% y el 18%, además de que el máximo valor propuesto es 20%. Se podría concluir que, para evaluar la actividad científica de las CSH, es aceptable asignar al factor *Formación de Recursos Humanos* una ponderación intermedia entre 10% y 18% .

### Factor 3 Movilidad

| Tabla 5 : Medidas estadísticas Factor 3 |        |
|---|--------|
| 1 - Ponderación del modelo              | 15%    |
| 2 - Media                               | 12,79% |
| 3 - Mediana                             | 12,5%  |
| 4 - Desvío estándar                     | 6,05   |

| Tabla 6: Resultados obtenidos para el Factor Movilidad |                        |                          |                      |
|--|------------------------|--------------------------|----------------------|
| Ponderaciones asignadas                                | Cantidad de respuestas | Porcentaje de respuestas | Porcentaje acumulado |
| 0%   | 1                      | 4%                       | 4%                   |
| 5%   | 4                      | 14%                      | 18%                  |
| 8%   | 1                      | 4%                       | 21%                  |
| 10%  | 8                      | 29%                      | 50%                  |
| 15%  | 7                      | 25%                      | 75%                  |
| 20%  | 6                      | 21%                      | 96%                  |
| 25%  | 1                      | 4%                       | 100%                 |
| <b>Total</b>   | 28                     | 100%                     |                      |



Los datos indican que la valoración de los consultados sobre el peso del factor *Movilidad* se acerca bastante a la ponderación que se ha propuesto en el modelo del IDAC. La media y la mediana casi no difieren (12,79% y 12,50% respectivamente) y el desvío estándar no es muy alto, lo que indica que no hay una importante dispersión respecto del valor medio asignado por los consultados. Sumando los porcentajes puede verse que el 54% de las respuestas se acumulan entre los valores 10% y 15% (29% y 25% respectivamente de porcentaje de respuestas). Por otra parte, la amplitud del rango de las valoraciones propuestas es menos amplia ya que varía entre 0% y 25%. De modo que, de acuerdo con esta consulta, sería aceptable asignar a la movilidad un valor de ponderación intermedio entre 10% y 15% para evaluar la actividad científica de las CSH.

#### **Factor 4: Proyectos**

| <b>Tabla 7: Medidas estadísticas Factor 4</b> |        |
|---|--------|
| 1 - Ponderación del modelo                    | 22%    |
| 2 - Media                                     | 18,36% |
| 3 - Mediana                                   | 15%    |
| 4 - Desvío estándar                           | 10,96  |

En el caso del factor *Proyectos*, los resultados obtenidos difieren en forma significativa del valor de ponderación propuesto en el modelo del IDAC. La media se acerca un poco más 18%, pero la mediana indica que el 50% de los casos han asignado un valor de ponderación igual o menor al 15%. Por otra parte, el desvío estándar es alto, llegando casi al 11%. El 81% de los casos se acumulan en el rango que abarca el 10% y el 20%, pero se trata de una importante amplitud de rango. Al mismo tiempo, aunque son pocos

casos, considerados en forma agregada resulta que el 11% de los consultados han otorgado valores de ponderación muy altos al factor *Proyectos* 40% y 50% . Del mismo modo que en el caso del factor *Perfil* es posible afirmar que existe un muy bajo consenso con respecto al valor de ponderación que debe asignarse al factor *Proyectos* para evaluar la actividad científica de las CSH.

## Factor 5: Productos

|                            |        |
|----------------------------|--------|
| 1 - Ponderación del modelo | 22%    |
| 2 - Media                  | 29,54% |
| 3 - Mediana                | 25%    |
| 4 - Desvío estándar        | 14,05  |

| <b>Ponderaciones asignadas</b> | <b>Cantidad de respuestas</b> | <b>Porcentaje de respuestas</b> | <b>Porcentaje acumulado</b> |
|--------------------------------|-------------------------------|---------------------------------|-----------------------------|
| 10%                            | 3                             | 11%                             | 11%                         |
| 15%                            | 1                             | 4%                              | 14%                         |
| 20%                            | 7                             | 25%                             | 39%                         |

|              |    |      |      |
|--------------|----|------|------|
| 22%          | 1  | 4%   | 43%  |
| 25%          | 3  | 11%  | 54%  |
| 30%          | 2  | 7%   | 61%  |
| 35%          | 3  | 11%  | 71%  |
| 40%          | 4  | 14%  | 86%  |
| 50%          | 2  | 7%   | 93%  |
| 60%          | 2  | 7%   | 100% |
| <b>Total</b> | 28 | 100% |      |

El factor *Productos* muestra ser el más polémico a la hora de asignar un valor de ponderación para evaluar la actividad científica de las CSH. De hecho, el debate más importante en el campo de la evaluación científica se sitúa en este dominio. Se trata del factor con la mayor amplitud de rango de los valores propuestos ya que varían entre 10% y 60%. Por otra parte, el desvío estándar también es muy alto 14,05% por lo que, según esta consulta, queda expresado que no hay consenso en la valoración de este factor para evaluar la actividad científica de las CSH.

Si bien la media (29,54) reflejaría que existe una valoración alta respecto de este factor como medida de evaluación de la actividad científica (supera casi en un 8% al valor de ponderación propuesto en el modelo, 22%), el hecho de que la amplitud de rango y el desvío estándar sean tan importantes, desmerece el peso interpretativo de la media. De todos modos, el valor de ponderación "20%" acumula el mayor porcentaje de respuestas (25%), que resulta aún más significativo si se tiene en cuenta la dispersión de las respuestas obtenidas. Se puede destacar que este valor de ponderación se acerca bastante al propuesto en el modelo (22%),

aunque la principal conclusión con respecto a este factor es la falta de consenso para establecer un criterio de ponderación del factor *Productos*.

### **Factor 6: Difusión**

**Tabla 12: Resultados obtenidos para el Factor Difusión**

| <b>Ponderaciones asignadas</b> | <b>Cantidad de respuestas</b> | <b>Porcentaje de respuestas</b> | <b>Porcentaje acumulado</b> |
|--------------------------------|-------------------------------|---------------------------------|-----------------------------|
| 0%                             | 2                             | 7%                              | 7%                          |
| 5%                             | 2                             | 7%                              | 14%                         |
| 8%                             | 1                             | 4%                              | 18%                         |
| 10%                            | 9                             | 32%                             | 50%                         |
| 15%                            | 6                             | 21%                             | 71%                         |
| 18%                            | 1                             | 4%                              | 75%                         |
| 20%                            | 6                             | 21%                             | 96%                         |
| 30%                            | 1                             | 4%                              | 100%                        |
| <b>Total</b>                   | <b>28</b>                     | <b>100%</b>                     |                             |

| <b>Tabla 11: Medidas estadísticas Factor 6</b> |        |
|--|--------|
| 1 - Ponderación del modelo                     | 13%    |
| 2 - Media                                      | 13,07% |
| 3 - Mediana                                    | 12,5%  |
| 4 - Desvío estándar                            | 6,68   |

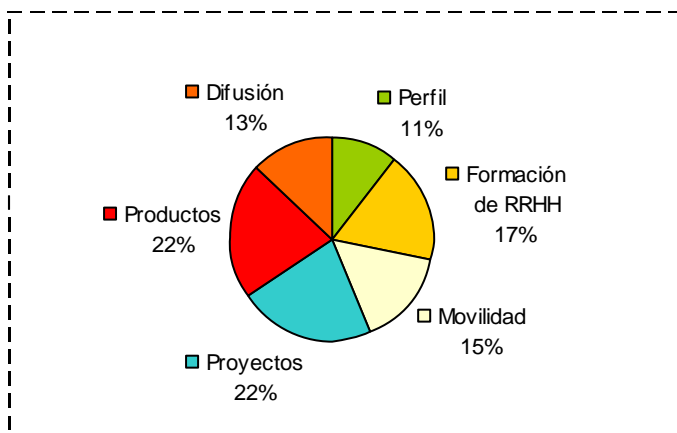
En el caso del factor *Difusión* la media y la mediana se acercan mucho al valor de ponderación asignado en el modelo del IDAC. El desvío estándar no es muy alto y la amplitud del rango de respuestas es moderada. Si se agregan, puede verse que el 53% de las respuestas se acumulan en los valores 10% y 15%, y que los porcentajes de respuestas en el resto de las categorías de ponderaciones asignadas no superan el 7%. El desequilibrio está dado por la acumulación del 21% de respuestas en el valor de ponderación 20%. A partir de estos resultados podría pensarse que el valor asignado en el modelo (13%) goza de un consenso significativo para ser aplicado a la evaluación de la actividad científica de las CSH, aunque podría ajustarse hacia un valor un poco mayor para dar respuesta a esta acumulación de frecuencias en el valor 20%.

### **Síntesis de resultados**

A continuación se presentan en forma comparada el Gráfico 15, que representa el modelo que se ha sometido a validación y muestra los valores de ponderación que se han asignado a los 6 factores de actividad científica para el cálculo del IDAC, y el Gráfico 16, que representa las medias de los valores de ponderación asignados por los expertos e investigadores consultados para cada uno de los 6 factores de actividad científica.

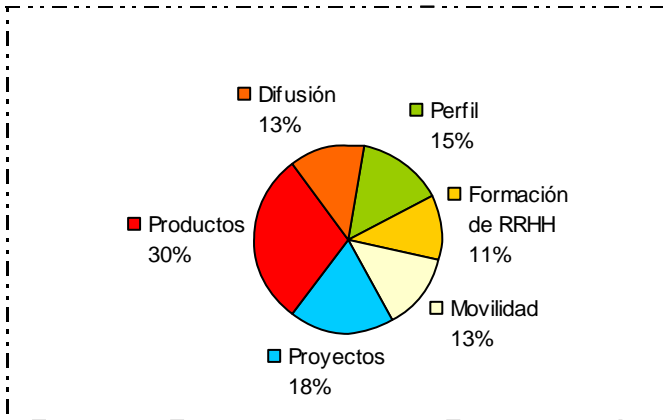
El análisis que se propone en función de las medias obtenidas en la consulta aporta nuevos elementos. Sin embargo, para mayor riqueza y precisión en las decisiones metodológicas que puedan tomarse, el mismo debe ser complementado con los análisis expuestos con anterioridad y que reflejan los resultados que se han obtenido para cada factor donde las medias son relativizadas en función de las frecuencias acumuladas y el desvío estándar.

### **Ponderación de factores del modelo de cálculo del IDAC**



## Ponderación de Factores según consulta de validación

| Orden de factores y diferencia porcentual interna en cada modelo |             |                       |                          |             |                       |
|--|-------------|-----------------------|--------------------------|-------------|-----------------------|
| Modelo IDAC  | Ponderación | Diferencia porcentual | Resultados de validación | Ponderación | Diferencia Porcentual |
| 1º<br>Proyectos /Productos                                       | 22%         | +5%                   | 1º Productos             | 30%         | +12%                  |
| 2º<br>Formación de RRHH  | 17%         | +2%                   | 2º Proyectos             | 18%         | +3%                   |
| 3º<br>Movilidad  | 15%         | +2%                   | 3º Perfil                | 15%         | +2%                   |
| 4º<br>Difusión   | 13%         | +2%                   | 4º<br>Movilidad/Difusión | 13%         | +2%                   |
| 5º Perfil  | 11%         |                       | 5º Formación de RRHH     | 11%         |                       |



En el caso de los factores *Perfil* y *Productos*, las medias de los valores de ponderación asignados por los consultados superan a los valores asignados en el modelo del IDAC siendo más significativa esta diferencia en el caso del factor *Productos*. La ponderación del factor *Difusión* coincide en ambos casos y para los factores *movilidad*, *formación de recursos humanos* y *proyectos* se les ha asignado valores de ponderación más bajos que los del modelo, donde la diferencia en el caso de *Formación de RRHH* es la más notoria.

Por otra parte, si se ordenan los factores según el orden de importancia asignada y se calcula la diferencia porcentual de los puntajes asignados con respecto al factor inmediatamente anterior, se obtiene que las principales diferencias que se han planteado entre el modelo del IDAC y el esquema que surge de la validación son: a) la ordenación de los factores especialmente en el caso de *Perfil* y *Formación de recursos humanos*, y b) la notoria diferencia en la valoración porcentual asignada al factor *Productos* (30%). La diferencia porcentual con respecto a *Proyectos* (12%) señala un orden de prioridad que contrasta con las diferencias porcentuales planteadas en el modelo. Para el resto de los factores aunque varíen los valores propuestos, las diferencias porcentuales se mantienen en el mismo orden, entre 2% y 3%.



### 3.1 CONCLUSIONES DE LA VALIDACIÓN

Si bien la muestra de expertos e investigadores consultados no es representativa en términos de disciplinas de CSH, es posible extraer conclusiones generales de este estudio que permitan revisar y analizar con mayor profundidad los valores de ponderación que se han propuesto para el cálculo del Índice Directo de Actividad Científica. Sería por tanto necesario realizar estudios específicos que permitan revisar los indicadores del modelo de evaluación y ajustar las ponderaciones del IDAC según las particularidades que presenta la actividad científica en las distintas disciplinas de CSH.

La consulta sobre el valor de ponderación de los indicadores ha aportado elementos interesantes en función de dos ejes principales: a) analizar y revisar el valor de ponderación que se ha asignado en el modelo a cada uno de los 6 factores que componen la actividad científica, y b) analizar y visualizar el grado de consenso que existe entre expertos e investigadores de CSH en relación con el peso que se debe asignar a cada uno de los factores para evaluar la actividad científica de las CSH.

De acuerdo a los resultados obtenidos, el factor *Productos* muestra ser el más polémico a la hora de asignar un valor de ponderación para evaluar la actividad científica. De hecho, el debate más importante en el campo de la evaluación científica se sitúa en este dominio. También en el caso del factor *Perfil* aunque la media se aproxima al valor propuesto, ha habido valoraciones altas por parte de algunos consultados. Ambos factores representan los aspectos en los que tradicionalmente se ha focalizado la evaluación científica. La propuesta metodológica que se ha diseñado en el marco del proyecto de investigación, tanto para la evaluación de la actividad científica como para la ponderación de los factores en el IDAC, asume la existencia de este debate y pretende contrarrestar la habitual focalización en los productos de la actividad científica a la hora de evaluarla. Para ello, tal como puede observarse en el Gráfico 3 (pp.,51), se propone un modelo que incorpora otros factores y que distribuye de una forma menos disonante los valores de ponderación de dichos factores. Con esta propuesta se pretende dotar de una mayor complejidad al concepto de actividad científica y dar valor y estatus a un conjunto más amplio de actividades y productos que los que tradicionalmente se han valorado. La sobrevaloración del *Perfil* conlleva una concepción un tanto más

individualista de la actividad científica. En el caso de los *Productos*, si bien se admite su importancia, se le asigna el mismo valor de ponderación que a *Proyectos* como una forma de valorizar los procesos y, al mismo tiempo, dar cuenta de que hay muchos aspectos de los procesos que se llevan a cabo en la práctica científica que no son plausibles de sistematización en publicaciones indizadas especialmente en el caso de las CSH.

Finalmente, este ejercicio de validación ha permitido reflexionar en torno a propuestas concretas de valorización de la actividad científica para el caso de las CSH. Se trata de una temática que requiere continuar con la reflexión, al mismo tiempo que con la acumulación de casuística y resultados que permitan estabilizar criterios y estándares de referencia para la medición y evaluación de las disciplinas específicas de CSH.

## **CONSIDERACIONES FINALES**

El punto de partida de este trabajo ha sido considerar que la evaluación de la actividad científica es una actividad que crecientemente compromete a las instituciones y a los grupos de investigación bajo un número más diverso de modalidades, instancias y entornos. La ampliación de las misiones y de los contextos de interacción de las instituciones científicas, ha ampliado al mismo tiempo el campo y los públicos frente a los cuales éstas deben rendir cuentas. Ello conlleva una mayor complejidad, debido a la diversidad de los criterios que participan en la valorización de la producción científica. Por otra parte, las exigencias actuales de los sistemas de Ciencia y Tecnología suponen alcanzar estándares de calidad en un contexto donde la competitividad necesita combinarse con la integración, la colaboración y la coordinación a fin de garantizar la circulación de conocimiento, investigadores y tecnologías.

La propuesta de evaluación que se ha desarrollado en el proyecto en el cual se inscribe el desarrollo de este índice asume que la triangulación de metodologías es un requisito ineludible, si se tiene en cuenta la diversidad de exigencias y problemáticas con las que se enfrenta la producción científica en la actualidad. Combinar metodologías, técnicas y modelos permite complementar enfoques y tipos de resultados como ocurre en el modelo aquí se propone y que incluye enfoques como evaluación por

indicadores, evaluación por pares, análisis de pertinencia y consenso institucional.

De este modo, el calificativo de *integral* utilizado en este trabajo se ha referido a varios planos: en primer lugar, al modo en que se ha definido y operacionalizado el concepto de actividad científica; en segundo lugar, a la propuesta de incorporar elementos relacionales en el tratamiento de la información y en el estudio de la calidad científica y, finalmente, se ha referido también al proceso evaluativo que se propone: un Programa de Evaluación diseñado con el múltiple objetivo de captar, evaluar y apoyar las dinámicas que organizan patrones de investigación.

La propuesta de evaluación integral ha sido desarrollada teniendo como objeto a las CSH y, en particular, teniendo presente la necesidad que existe en estos campos disciplinarios de identificar estándares de producción, patrones de funcionamiento y dinámicas de vinculación. En este sentido, la heterogeneidad y la escasa estandarización de criterios de evaluación que caracterizan la situación actual de las CSH, más que como un obstáculo, en este caso se han tomado como una ventana de oportunidad para reflexionar y elaborar propuestas que, por su carácter exploratorio y por el conjunto de aproximaciones que integra, no deja de tener interés para ser adaptada a la evaluación de otros campos disciplinarios. Hoy en día la importancia de la vinculación con el contexto y la conectividad social, no son una preocupación exclusiva de las CSH que han sido más clásicamente asociadas a la idea de *pertinencia*. La fórmula *calidad-productividad* resulta cada vez más insuficiente en términos evaluativos y por supuesto también, en términos explicativos.

El concepto de *calidad relacional* tal como hemos querido presentarlo aquí, incluye como dimensiones intrínsecas a la calidad científica un conjunto amplio de actividades pero principalmente, incluye en su definición a los modos de organización y entre ellos, los modos de establecer vinculaciones efectivas con una gama más amplia de esferas y agentes sociales en el marco de los procesos de producción científica. Con el fin de incorporar aspectos relacionales en el estudio de la calidad científica, el modelo de evaluación integral que aquí se ha presentado avanza en tres planos: a) en el plano *metodológico* introduce mecanismos de consulta y validación del propio modelo que incluyen criterios de valoración de la actividad científica de distintos agentes sociales (expertos y usuarios de los

sistemas de evaluación), b) en el plano *analítico*, incorpora un conjunto amplio de dimensiones e indicadores para dar cuenta de la integralidad de la actividad científica, y propone un tratamiento relacional de la información que resulta más exigente en términos explicativos que su tratamiento como indicadores aislados; y finalmente, c) en el plano de la *pertinencia*, donde a partir de la consulta a investigadores de CSH y a potenciales usuarios de la investigación, se busca testar su posible receptividad, expectativas y necesidades con el fin de orientar estrategias y ofrecer elementos que permitan estructurar la demanda y la oferta de conocimientos científicos.

En definitiva, la hipótesis que se mantiene detrás de esta propuesta metodológica integral, es que la producción, distribución y uso de conocimiento científico alcanzará buenos estándares de calidad, en la medida en que se garanticen buenas condiciones relacionales para estos procesos, especialmente en un contexto en que la búsqueda de la excelencia académica está asociada a lograr altos grados de integración y conectividad del sistema científico entre sí y con sus entornos. En este sentido, se espera que el esfuerzo realizado por diseñar una propuesta integradora de enfoques y aproximaciones, resulte una aportación que de lugar a intercambios productivos en el entorno diverso de agentes que se ven involucrados en procesos de valoración de la actividad científica.

## **REFERENCIAS BIBLIOGRÁFICAS**

ARCHAMBAULT, E. , VIGNOLA-GAGNE, E., COTE, G., LARIVIERE, V. , GINGRAS, Y. (2006). "Benchmarking scientific output in the social sciences and humanities : The limits of existing databases". *Scientometrics* 68 (3) 329-342.

BARRENECHEA, J., CASTRO SPILA, J. E IBARRA, A. (2008) Calidad relacional y evaluación integral de la actividad científica en ciencias sociales y humanidades: propuesta metodológica e indicadores. Estudios en Ciencia, Tecnología e Innovación nº 3. Cátedra Sánchez Mazas. Servicio Editorial de la UPV/EHU. País Vasco.

BECHER, T. (1993). "Las disciplinas y la identidad de los académicos". *Pensamiento Universitario*, 1, 56-77.

BENCE, V. & OPPENHEIM, C. (2004) The role of academic journal publications in the UK Research Assessment Exercise . Learned Publishing, 17, 53-68.

BUELA-CASAL, G., GUTIÉRREZ-MARTÍNEZ, O., BERMÚDEZ-SÁNCHEZ, M. P. & VADILLO-MUÑOZ, O. (2007) Comparative study of international academic rankings of universities. *Scientometrics*. 71, 349-365.

COOK, THOMAS D. (1985) 'Postpositivist Critical Multiplism', in R.L. Shotland and M. M. Mark (eds) *Social Science and Social Policy*, pp. 21–62. Thousand Oaks, CA: Sage.

CRUZ, I. R. (2008) Challenging ISI Thomson Scientific's Journal Citation Reports: Deconstructing "Objective," "Impact," and "Global". *portal: Libraries and the Academy*, 8, 7-13.

ECCLES, C. (2002) The Use of University Rankings in the United Kingdom. *Higher Education in Europe*, 27, 423-432.

FEDERKEIL, G. (2002) Some Aspects of Ranking Methodology—The CHE-Ranking of German Universities. *Higher Education in Europe*, 27, 389-397.

GREENE, J. BENJAMIN, L. Y GOODYEAR, L. (2001). The Merits of Mixing Methods in Evaluation. *Evaluation*, 7(1), pp.25-44.

GREENE, J., CARACELLI, V. Y GRAHAM, W. (1989) 'Toward a Conceptual Framework for Mixed-Method Evaluation Designs', *Educational Evaluation and Policy Analysis* 11: 255–74. Greenwood, Davydd and Morten Levin

IBARRA, A., CASTRO, J. Y BARRENECHEA, J. (2006) "Acta del taller: Indicadores para la evaluación de la actividad de investigación en Ciencias Sociales y Humanidades" en Ibarra, A., Castro, J. y Barrenechea, J. *La Evaluación de la Actividad Científica en Ciencias Sociales y Humanidades*.

Estudios en Ciencia, Tecnología e Innovación nº 1. Cátedra Sanchez Mazas. Servicio Editorial de la UPV/EHU. País Vasco

KYVIK, S. (2003). "Changing trends in publishing behaviour among university faculty, 1980-2000". *Scientometrics*, 58 (1) 35-48.

VAN LEEUWEN, T. (2006). The application of bibliometric analyses in the evaluation of social science research. Who benefits from it, and why it is still feasible. *Scientometrics*, 66 (1), pp.133-154.

LIU, N. C. & CHENG, Y. (2005) The Academic Ranking of World Universities. *Higher Education in Europe*, 30, 127-136.

MERISOTIS, J. P. (2002) On the Ranking of Higher Education Institutions. *Higher Education in Europe*, 27, 361-363.

MOYA-ANEGÓN, F., CHINCHILLA-RODRÍGUEZ, Z., CORERA-ÁLVAREZ, E., MUÑOZ-FERNÁNDEZ, F. J. & NAVARRETE-CORTÉS, J. (2004) Indicadores bibliométricos de la actividad científica española:(ISI, Web of science, 1998-2002)