

## **APROXIMACIONES ECONÓMICAS Y ECONOMÉTRICAS PARA LA PROBLEMÁTICA EDUCATIVA <sup>1</sup>**

**Walter SOSA ESCUDERO**

PhD en Economía - Universidad de Illinois  
Profesor Asociado y Director de la Maestría en Economía - Universidad de San Andrés  
Profesor Titular de Econometría - Universidad Nacional de La Plata

**DOCUMENTO DE TRABAJO N° 17**

---

<sup>1</sup> Transcripción revisada de mi presentación en el Seminario Permanente de Investigación de la Escuela de Educación de la Universidad de San Andrés. Resalto el carácter divulgatorio de la misma, con el objetivo de acercar a los alumnos de posgrado en Educación algunas experiencias mayoritariamente metodológicas del proceso de producción científica, en este caso, desde la econometría. A fines de mantener el carácter informal de la presentación decidí relegar las citas y referencias a una sección final de sugerencias bibliográficas. Agradezco a la Dra. Catalina Wainerman la invitación a participar en el mismo así como sus comentarios. Yanina Azzolina contribuyó considerablemente a mejorar la transcripción textual de la cinta original. Los errores son de mi exclusiva responsabilidad.

**APROXIMACIONES ECONÓMICAS Y ECONOMETRICAS  
PARA LA PROBLEMÁTICA EDUCATIVA**

**Walter SOSA ESCUDERO**

PhD en Economía - Universidad de Illinois  
Profesor Asociado y Director de la Maestría en Economía - Universidad de San Andrés  
Profesor Titular de Econometría - Universidad Nacional de La Plata



Universidad de  
**San Andrés**

Conferencia pronunciada el 20 de Septiembre de 2003 en el ámbito del SEMINARIO PERMANENTE DE INVESTIGACIÓN de la Maestría en Educación de la UdeSA.

Coordinadora: Dra. Catalina Wainerman

**DOCUMENTO DE TRABAJO N° 17**

Julio de 2005

SERIE "DOCUMENTOS DE TRABAJO"  
ESCUELA DE EDUCACIÓN  
Universidad de San Andrés.

Directora de la Serie: Dra. Catalina Wainerman  
Responsables de edición: Lic. Annie Mulcahy  
Lic. Marina Larrondo

Para obtener ejemplares de la Serie dirigirse a:

Universidad de San Andrés  
Vito Dumas 284  
(B 1644BID), Victoria, Pcia. De Buenos Aires

ISBN 987-98824-0-7

Hecho el depósito que dispone la Ley 11.723  
Impreso en Argentina - *Printed in Argentina*  
Primera edición: Julio de 2005



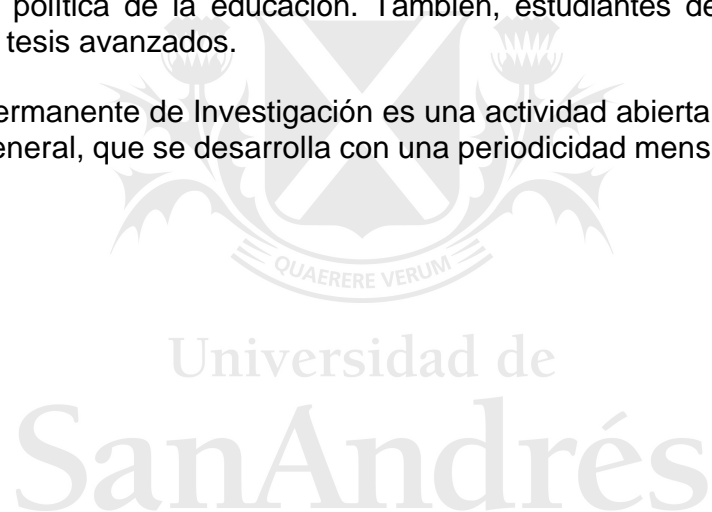
Universidad de  
**San Andrés**

## SERIE “DOCUMENTO DE TRABAJO” DE LA ESCUELA DE EDUCACIÓN

La serie de Documentos de Trabajo refleja parte de las actividades de la Escuela de Educación de la Universidad de San Andrés. Los documentos difunden conferencias dictadas en el ámbito del Seminario Permanente de Investigación de la Especialización y la Maestría en Educación con Orientación en Gestión Educativa, y del Doctorado en Educación. También difunden exposiciones de invitados especiales.

El Seminario Permanente de Investigación tiene el propósito de estimular la producción y difusión de la investigación en el campo de la Educación. Tiene, además, el objetivo de familiarizar a los estudiantes de posgrado con la producción de conocimiento riguroso en el área. Entre los expositores se cuentan, fundamentalmente, investigadores formados y, eventualmente, expertos en planificación y política de la educación. También, estudiantes de posgrado con sus trabajos de tesis avanzados.

El Seminario Permanente de Investigación es una actividad abierta a la comunidad educativa en general, que se desarrolla con una periodicidad mensual.



## INTRODUCCIÓN

Quería empezar agradeciendo a los organizadores de esta actividad, en particular a la Dra. Wainerman, la invitación a compartir algunas experiencias de mis investigaciones en economía.

Empiezo contando una parábola que me gusta porque ilustra qué siente un investigador en su tarea, y además nos va a servir para brindar un marco a algunas discusiones que vamos a desarrollar en esta charla. Andrew Wiles es uno de los matemáticos contemporáneos más conocidos, quizás no el más importante, pero es quien demostró el Teorema de Fermat, un aparentemente inocente problema matemático que tardó 360 años en ser resuelto. En una charla que dio hace un tiempo intentaba describir el proceso de investigación científica y contó la siguiente parábola, que me parece muy apropiada. La actividad de investigación se asemeja a lo siguiente: uno se mete en una especie de casa abandonada que está completamente en penumbras y empieza a tantear, empieza a usar sus sentidos, todos sus sentidos, toca, huele, va, viene, y en el proceso va armando un mapa de esta casa. Sabe donde están los muebles, las ventanas y conjetura, sospecha, se fija si las conjeturas que tiene son ciertas o falsas, tiene dudas, avanza, y a medida que pasa el tiempo uno se va armando una idea bastante interesante de cómo funcionan las cosas en esta casa en penumbras. Hasta que

en un momento encuentra la llave de luz y la levanta, y entonces en ese momento ve un montón de cosas, y, fundamentalmente aprende mucho acerca de la relación entre lo que uno pensaba acerca de lo que eran las cosas y lo que aparentemente son ... hasta que descubre otra puerta, abre y está de vuelta en penumbras.

Lo que me gusta mucho de esta parábola es que, desde un punto de vista científico, este andar a tientas es una cuestión metodológica, no es andar a tientas tocando y haciendo cualquier cosa sino que existe un proceso de reelaboración científica. Esta forma de producir conocimiento, de relacionar lo que parece que son datos con lo que uno cree que está detrás de los mismos, es un proceso *metodológico*, y posiblemente sea uno de los aspectos fundamentales que distingue al conocimiento científico de cualquier otro.

De lo que voy a hablar en esta charla es de la producción de conocimiento científico en educación y economía, voy a contar un poquito qué significa esto de andar a tientas en la teoría y como esto, eventualmente, se puede aprovechar positivamente. Cuando digo que voy a hablar de las relaciones que puede haber entre la economía y la educación no voy a hablar de las relaciones entre el Ministro de Educación y el de Economía, sino de cómo ciertos problemas y metodologías interactúan entre estas dos disciplinas, ya sea porque tienen problemáticas similares o porque tienen formas de razonar parecidas. Remarco una cosa que decía la Dra. Wainerman en la presentación, yo no soy un economista de la educación, mi especialización es en econometría y parte de la charla consiste en contarles qué es un econometrista y cómo esta disciplina puede contribuir al análisis de la problemática educativa.

Comencemos por ponerle título a las cosas. A falta de mejor título esta charla se llama "aproximaciones económicas y econométricas para la problemática educativa". Tengo un colega que dice que el tamaño del título de un trabajo es inversamente proporcional a su calidad. Cuando uno empieza una investigación no quiere dejar nada afuera pero a medida que se aprenden los verdaderos alcances del trabajo, el título se acorta. No sé si esto se aplica a

trabajos terminados, pero si esto se aplicase a esta charla, no lo veo como un buen comienzo.

Los objetivos de esta charla son esencialmente metodológicos, no voy a dar una clase en donde voy a explicar cosas de economía o de economía y educación. El objetivo general consiste en presentar algunas contribuciones temáticas y metodológicas de la economía y de la econometría a la problemática educativa. Como medio y como fin vamos a discutir algunos problemas empíricos y metodológicos que enfrentan este tipo de investigaciones. También me gustaría, como objetivo tangencial, motivar el uso de herramientas cuantitativas y difundir algunas fuentes de datos disponibles, más allá de la discusión compleja de qué significa “datos” y qué significa “disponibles”. Estos son, más o menos, los objetivos de esta charla.

#### UN EJEMPLO: LA EDUCACIÓN COMO INVERSIÓN

Quizás uno de los problemas educativos más estudiados desde la economía se refiere al rol de la educación como mecanismo para mejorar el bienestar de la gente, definido de una manera todavía muy amplia. Esto no implica que este sea el único rol de la educación, ni siquiera el más importante, hay otros roles. No es que la economía hace un planteo de la educación como mecanismo para generar bienes y bienestar a la gente, no hay un intento reduccionista, no estoy diciendo qué es lo importante y qué no, sino que estamos focalizando en uno de los roles de la educación al solo efecto de aislarlo como fenómeno científico. Esta cuestión de aislar fenómenos es crucial para el desarrollo de una investigación. Lo que voy a intentar argumentar es que posiblemente la contribución más importante de la economía al área educativa sea fundamentalmente metodológica.

La perspectiva de la educación vista como inversión que eventualmente reditúa en una mejora en el bienestar –insisto, como fenómeno científico- tiene que ver con que todo aquel que pensó que era importante aprender algo para mejorarse la vida estaba viendo a la educación, o por lo menos a una parte del problema educativo, como un problema de inversión. Como sucede con casi todas

las discusiones económicas, Adam Smith se ocupó quizás un tanto tangencialmente de este tema en su Teoría de los Sentimientos Morales, en donde plantea el problema del análisis costo-beneficio de la educación. Esta literatura permanece en un estado latente hasta la década del '60 del siglo pasado, de la mano de economistas como Schultz, Becker o Mincer, quienes ven al proceso educativo como una inversión en capital educativo, la cual surge de sopesar beneficios y costos de estas actividades.

La base empírica de esta relación es sugerente. Los datos indican una fuerte relación positiva entre niveles educativos alcanzados e ingresos laborales. ¿Qué quiere decir esto? Nada más que a la gente que tiene más educación "le va mejor". Lo cual es compatible con que si uno estudia va a ganar más plata, pero también con que si uno tiene más plata va a poder estudiar más. Entonces, a menos que hagamos un esfuerzo metodológico serio, estamos muy cerca de caer víctimas de la falacia de la causalidad, léase, de tomar esta correlación entre ingresos y educación como mecanismo para validar el hecho de que la educación mejora los ingresos de la gente. Conviengamos que esta empiria, en donde el que tiene más educación tiene ingresos más altos, es *sugerente*, y es lo que posiblemente motive este tipo de investigaciones en donde se busca estimar el efecto causal de la educación sobre los ingresos. Si fuese cierto que la gente que más se educa gana más dinero, entonces convengamos que si el bienestar pasa por tener más dinero, una forma de promover el bienestar de la gente es promover la educación. Pero estamos muy lejos de eso, todavía nos falta un trecho muy largo para ir de esta correlación de ingresos-educación a establecer, verificar o cuantificar una relación de causalidad.

Déjenme hacer un pequeño comentario sobre esto, que también tiene que ver con un punto metodológico que va a aparecer más adelante. Quizás la relación entre educación e ingresos se refiera a la relación subyacente entre educación y las "cosas que están relacionadas con los ingresos". ¿Qué quiere decir esto? Que la gente más rica tiene más posibilidades de estudiar, que la gente que más estudia se hace más rica, que la gente que tiene menos restricciones financieras es la que más se educa y varias cosas más. En definitiva, la relación educación-



ingresos puede ser un resumen un tanto burdo de una realidad posiblemente más compleja.

Muchas de las problemáticas empíricas y metodológicas que enfrentan este tipo de investigaciones tienen que ver con que si la educación es una inversión en capital humano, ésta difiere de cualquier otra inversión, dado que una persona no puede utilizar la inversión en capital humano como colateral. Yo podría convencer a un banco de que voy a hacer una inversión edilicia y pedirle un préstamo para comprar un edificio, y si tengo problemas en devolver el préstamo, permitir que el banco se quede con el edificio. Lamentablemente esto no sucede con las inversiones en educación individual, no me puedo ofrecer a mí mismo como colateral, no puedo decirle a un banco que se quede con mi Maestría en Educación si no pudiese devolver el préstamo. Esto introduce una complicación adicional que dificulta notoriamente el estudio entre la educación y los ingresos.

Una visión alternativa, de la década del setenta, enfatiza el rol de la educación no para incrementar la productividad de las personas, sino como mecanismo para adquirir una señal de mercado. ¿Qué significa adquirir una señal? Resulta que hay abundante literatura que indica que mas allá de la educación formal, los individuos difieren notoriamente en sus productividades en base a características que pueden agruparse bajo el nombre de “habilidades” - quizás genéticas o innatas, pero no necesariamente-. Es cierto que la educación mejora la productividad de la personas, es cierto que uno aprende cosas y que el mercado posiblemente las valore, pero convengamos que lo que eventualmente pueden observar las empresas (los empleadores) es una mezcla de estas habilidades específicas con las adquiridas en el proceso de educación formal. Resulta que también hay alguna evidencia, un tanto más complicado de comentar, acerca de que los individuos más hábiles (definido en un sentido muy general) son más “fáciles de educar”, les es un poquito menos costoso avanzar en el sistema educativo, tiene que ver con la (sana) envidia que uno siente cuando el compañero listo de la clase lee en una hora lo que al resto nos toma cuatro, es en ese sentido que estoy diciendo que a la gente más hábil le resulta “más barato” educarse. El problema es que las firmas no observan la habilidad específica de los

individuos sino directamente la educación (nivel alcanzado, título, etc.). Ciertamente hay un proceso de aprendizaje, la empresa necesita conocernos, pero hasta que no nos conoce lo único que uno puede poner arriba de la mesa son “señales”, “tengo esta educación” o “no la tengo”. Una de las cosas que dice esta literatura es que uno de los roles que cumple la educación formal es vender la señal. ¿Qué significa eso? A igualdad de restricciones financieras, si las cosas funcionan bien, la gente más hábil se educa más porque les es menos costoso (en los términos discutidos anteriormente), y la gente menos hábil se educa menos porque es más caro. Entonces una forma que tiene la empresa de distinguir a los más hábiles de los menos hábiles es mirando su título.

## CARACTERÍSTICAS DE LA INVESTIGACIÓN EMPÍRICA EN ECONOMÍA Y CIENCIAS SOCIALES

### Formalismos e inexactitud

En general, cuando hablamos de fenómenos sociales no podemos darnos el lujo de hacer planteos exactos, dado que estamos trabajando con fenómenos muy complejos. Lamentablemente, muchas veces la falta de formalidad en una disciplina es percibida como una limitación del sujeto que la investiga, pero, ciertamente, el precio de la formalidad, muchas veces, es dejar muchas cosas afuera que son relevantes y/o difíciles de cuantificar o formalizar. Los que trabajamos con disciplinas sociales convivimos con relaciones que no son exactas, y es crucial, a los efectos de avanzar en una investigación empírica, tener una percepción clara de cuáles son las ventajas y limitaciones de operar en este tipo de contextos. Pongamos como ejemplo una regularidad que los economistas llamamos “ley de demanda”, que postula que si el precio de un bien baja y cualquier otro determinante de la demanda permanece inalterado, una persona debería comprar, en general, más de ese artículo. No tiene sentido práctico refutar la ley de la demanda argumentando que una vez un amigo compró más de un bien cuyo precio aumentó. Las leyes sociales no pretenden este grado de especificidad

sino que se refieren a nociones más generales. Consecuentemente, la verificación o refutación empírica de las teorías sociales tiene que darse en un contexto empírico en donde estas imprecisiones “tolerables” sean tenidas en cuenta explícitamente.

Inobservables, incuantificables y lo aleatorio como estrategia de modelización

Otro tema con el que tenemos que lidiar en una disciplina social cuantificable tiene que ver con la existencia de una variedad de factores que por variadas razones no son cuantificables u observables. A modo de ejemplo pensemos en el rol que la inteligencia puede tener sobre la generación de ingresos. Cualquier intento de enfrentar empíricamente esta relación pasa por cuantificar esta noción de inteligencia y, previamente, de definirla. Las interminables discusiones metodológicas, tanto en la psicología como en disciplinas afines, pueden rápidamente llevar a un investigador a omitir esta variable de la lista de potenciales factores explicativos de los ingresos de las personas. En un estudio que relaciona los ingresos de los individuos con sus características observables (como su educación, la de sus padres, su edad, etc.), si los individuos difieren en sus inteligencias y las mismas no son observadas, esta variable pasa a engrosar la lista de factores omitidos. La econometría y la economía aplicada le ponen un nombre bien sofisticado a esta colección de factores omitidos: hablamos de *heterogeneidades no observables*. La estrategia empírica usual es tratar a los factores no observables como si fuesen factores aleatorios que están presentes en la relación de interés, y uno de los principales puntos de cualquier investigación causal consiste en garantizar que estos han sido tratados correctamente desde un punto de vista metodológico.

Podríamos discutir durante horas si es que realmente existen factores aleatorios que afectan al ingreso –más allá de que existen debates filosóficos que cuestionan la naturaleza misma de lo aleatorio- pero desde un punto de vista metodológico, cuando decimos que un factor es aleatorio, muchas veces nos referimos a factores que eventualmente son aleatorios o que se comportan como

si lo fueran. Filosóficamente, la diferencia entre cosas realmente aleatorias y omisiones deliberadas es abismal, pero metodológicamente la diferencia es muy menor. Entonces, a menudo, en este tipo de análisis hay cosas que se consideran “aleatorias” pero esto ocurre a falta de mejor nombre. En general estas reciben el nombre de heterogeneidad no observable, que hace referencia a cosas que pueden o no ser puramente aleatorias pero que están afectando el ingreso. El carácter aleatorio de ciertas magnitudes, entonces, obedece a una estrategia de modelización, en donde lo omitido es tratado como aleatorio.

### La inhabilidad de experimentar

Cuando se tiene la posibilidad de hacer un experimento uno puede aislar, a través del llamado diseño experimental, qué cosas se mueven y cuáles no. Piensen en el tipo de experimento que uno diseñaría para cuantificar el efecto que posee una droga sobre un resultado como la temperatura corporal. Podríamos pensar en asignar a distintas ratas de laboratorio distintas dosis de esta droga y luego medir la temperatura. A los efectos de que los datos (las mediciones dosis-temperatura) sean informativos acerca del efecto causal de la dosis sobre la temperatura, deberíamos garantizarnos, entre varias cuestiones, que las dosis fueron asignadas en forma independiente de las características de las ratas y que ningún otro factor relevante actuó sobre la temperatura, de modo que las discrepancias en las temperaturas se deban a diferencias en las dosis. En una visión muy burda de este ejemplo, podría suceder que el instrumento de medición de la temperatura esté sujeto a un error de medición (un termómetro poco preciso, redondeos arbitrarios, etc.), de modo que estas “aleatoriedades” deben ser explícitamente reconocidas en el análisis. En este contexto, lo que hemos hecho es *aislar* el efecto de la dosis de cualquier otro, a través de un diseño experimental correcto.

Hagamos el ejercicio intelectual de pensar en un experimento acorde para nuestro problema de la influencia de la educación sobre los ingresos. Podríamos pensar en ir a un hospital y decir que queremos aislar el efecto de la educación

sobre los ingresos, de cualquier otro factor relevante. Encerrar a varios recién nacidos -con el espíritu de programas de TV como Gran Hermano- asignar aleatoriamente niveles educativos para cada niño (uno va a hacer seis años del colegio primario, otro va a hacer siete, otro va a cursar el secundario hasta cuarto año, otro terminará el secundario, otro irá a la universidad, etc.). Luego los enviamos a un mercado laboral, persiguiéndolos de modo que ningún otro factor extraño interfiera en el mecanismo de generar ingresos (podríamos citar el espíritu del film *The Truman Show*), y si todo se realiza correctamente, la discrepancia entre los ingresos debería ser informativa acerca de la discrepancia en la educación. Lo relevante de este razonamiento -que en el mejor de los sentidos es ridículo y en el peor, éticamente aberrante- es que pone en evidencia las enormes dificultades que enfrenta una disciplina que ve limitada la posibilidad de experimentar.

#### Datos observacionales

Alguien podría conjeturar sobre la posibilidad de hacer una encuesta en la que se le pregunte a la gente sobre su ingreso y su educación, entre otras cosas, con el fin de estudiar esta relación. A los fines de pensar la respuesta, volvamos a nuestras pobres ratas y pensemos sobre la posibilidad de reemplazar las ratas de laboratorio con datos provenientes de un hospital (veterinario) en donde históricamente existen datos de personas que llevaron a sus ratas-mascotas para que les suministren una droga para curarlas de sus males (de alta temperatura). El problema con estos datos es que las dosis no fueron asignadas en base a ningún diseño experimental sino, justamente, en función del efecto que uno quiere medir (quizás se les dio una dosis más alta a ratas con mas temperatura!).

A este tipo de datos se les llama *datos observacionales* y la pregunta metodológica crucial -que subyace a la mayoría de las disciplinas sociales, en particular a la economía- es cómo proceder para el análisis de datos que no son generados por un experimento sino por otro tipo de mecanismos que no

contemplan las necesidades metodológicas de un investigador que quiere estudiar relaciones de causa y efecto.

Afortunadamente existen varias situaciones simples en donde los datos observacionales se comportan de una manera muy similar a los datos experimentales. Supongamos que estamos interesados en estudiar el efecto del género sobre la educación. Podríamos observar pares sexo-educación con la confianza de que la asignación de géneros, si bien no ocurre en un laboratorio, ocurre en un contexto que produce géneros que se parece sospechosamente a un método experimental: podemos tomar el sexo de una persona como si fuese asignado experimentalmente. A fines de considerar un ejemplo tonto, piensen que el razonamiento se cae si una persona se da cuenta de que es más fácil ganar plata siendo mujer y por eso se cambió de sexo (podríamos citar el film *Tootsie*, para continuar con nuestras analogías cinematográficas).

La teoría y la estadística al rescate

Descartada la ruta experimental -no por lógicamente incoherente sino por ser éticamente aberrante o de imposible implementación práctica- no nos queda otra que buscar caminos alternativos. Resulta que la econometría, a lo que yo me dedico, es una rama de la economía que enfatiza el uso de métodos estadísticos para explorar la naturaleza empírica de las relaciones económicas, y que internaliza explícitamente las características metodológicas específicas de la economía, tales como el carácter no experimental de los datos o la inherente inexactitud de las "leyes" económicas, como discutiésemos anteriormente. Creo que es importante remarcar que la imposibilidad de contar con métodos experimentales no es necesariamente una limitación personal de los investigadores en economía, educación o sociología. Van a escuchar por todas partes que una fuerte crítica que se les hace a los científicos sociales, en especial a los economistas, tiene que ver con la gran variedad de supuestos ridículos que hacemos. Pero no es que hacemos supuestos por una deficiencia metodológica de nuestra disciplina, sino porque "no nos queda otra"; la única forma que tenemos

de saltar las dificultades provenientes de la inhabilidad de implementar experimentos consiste en aislar fenómenos a través de una *percepción* acerca de cómo funcionan las cosas, la cual distingue entre factores relevantes, poco relevantes y nada relevantes, y que establece alguna noción de causalidad entre los mismos, o sea, el rol de una teoría. En definitiva, el rol que en una disciplina experimental le cabe a un diseño correcto del experimento, en una disciplina no experimental le cabe a una buena teoría y a un buen análisis de la información de base, lo que nos lleva a un importante tópico.

### Datos e información

En un diseño experimental correcto -casi por construcción- los datos son informativos para la pregunta en cuestión; piensen, por el contrario, qué podríamos aprender del efecto de una droga sobre la temperatura si un asistente descuidado le asigna a todas las ratas la misma dosis! Lamentablemente, en un contexto no experimental no tenemos ninguna garantía de que los datos sean realmente informativos acerca de la pregunta de interés. Peor aún, ciertos datos pueden ser informativos para algún problema pero no para otro.

Déjenme darles un ejemplo relacionado con el Censo Universitario de 1994. En ese año se hizo una encuesta en donde se preguntaron varias cosas a todos los estudiantes universitarios de todas las universidades públicas, tales como cuál era su status en la carrera, cuántas materias había aprobado, etc. También se hacían preguntas demográficas, por ejemplo, cuál era la educación del padre, dónde vivía, etc. La pregunta es si estos datos son informativos para estudiar el rol de ciertos factores observables de la persona, como la educación de su padre o madre, en la "performance" universitaria, ¿cuál es el problema con esta base de datos?, datos no faltan. Estamos hablando de todos los estudiantes universitarios de todas las carreras de todas las universidades públicas del país. Más aún, la Universidad de Buenos Aires es la quinta universidad más grande del mundo. Resulta que tenemos una gran cantidad de información, no sólo con respecto al caso argentino sino que tenemos mucha información en general. Pero vamos a



pensar más seriamente, supongamos, muy groseramente, que “performance” significa cantidad de materias aprobadas por año, esto es, si un alumno hizo solo un año de carrera y aprobó seis materias, su “performance” es seis materias por año. La misma que una persona que estuvo tres años y aprobó dieciocho materias, su “performance” es seis materias por año. Ahora bien, ¿quiénes son las personas que están incluidas en este censo? Si el censo fue realizado en 1994, están los ingresantes de 1994 y los de 1993 -o sea, son alumnos que están en primer año o terminando primero y empezando segundo, donde su “performance” es simplemente el número de materias que aprobaron ese año. Eso es un dato de “performance”, pero no muy preciso, por ejemplo, una persona ese año aprobó seis materias porque primer año era “fácil” pero esto no necesariamente es un buen predictor sobre lo que va a ser su “performance” durante el resto de la carrera.

O sea que, en definitiva, dentro de esta muestra tenemos una gran cantidad de gente cuya “performance” podría llegar a estar bien medida pero, en general, se trata de una medida muy imprecisa. En relación a los que ingresaron en el '93, la información de los que ingresaron en el '92 puede estar relativamente mejor medida dado que ya han estado bastante tiempo en la universidad. Pero, ¿qué pasa con las personas que ingresaron en el año 1988?, ¿cuál es el problema que tienen estas personas?: no están correctamente representados en este censo, porque los mejores alumnos de la clase, a la fecha del censo ya no están más en la universidad, no fueron encuestados. Pensemos en una persona que empezó a estudiar la carrera de contador público en el año 1988, que proviene de una familia con mucho dinero y que pudo dedicarle mucho tiempo a su carrera, o que es muy inteligente o una combinación de ambas. Este alumno hizo la carrera en cinco años, y cuando se realizó el censo ya no estaba en la Universidad. El censo se encontró con sus compañeros que tuvieron una “performance” menor, además de todos aquellos que ingresaron por ejemplo en 1980, y hace catorce años que están en la universidad, estas personas no son representativas de su compañeros. Esta muestra, entonces, que tiene muchísimos datos, al momento de querer aplicarla para estudiar el problema en cuestión contiene relativamente poca



información. De hecho el estudio que nosotros llevamos adelante estaba basado estrictamente en gente que ingresó en 1990, o sea que en definitiva tiramos un montón de datos para preservar información, para que no se nos mezclen cosas que son informativas acerca del problema con cosas que no lo son. El objetivo del estudio, en este caso, era cómo ciertos factores socio-demográficos afectan a la “performance” de los estudiantes.

La razón por la cual me gusta dar este ejemplo es, por un lado, para mostrar que no por el hecho de tener datos uno tiene información. Por otro, para mostrar qué significa tener información, con respecto a qué significa tener datos. Dato es algo que tiene que ver esencialmente con el problema que uno está estudiando. Si cambia el problema, los datos posiblemente no cambian, pero la información cambia por completo. Debemos tener cuidado con la relación que hay entre datos e información. En general, lo que determina qué tipo de información necesito es el problema que estoy estudiando, para ciertos problemas esto puede ser muy informativo y para otros no tanto. Eso les va a pasar frecuentemente cuando hagan investigación empírica: piensan que tienen muchos datos pero son poco informativos. Uno tiene que mirar los datos desde la perspectiva del problema que uno tiene, no desde un punto de vista meramente administrativo.

Otro comentario que quiero hacer con respecto a la relación entre datos e información tiene que ver con la sistematización de los datos: que los datos existan y que no estén sistematizados en la forma en la cual uno los necesita, es prácticamente lo mismo que que no existan. Yo sé que hay una base de datos en la Universidad de San Andrés donde hay una gran cantidad de información para que yo aprenda, por ejemplo, historias educativas de los alumnos. Podríamos ver cómo es la evolución de un estudiante que empieza con un proceso de admisión y termina como licenciado, lo que yo quiero son historias educativas. Lamentablemente los datos no están sistematizados para armar este tipo de historias, tomaría mucho trabajo al sistema construir esta nueva base de datos desde esta perspectiva, si bien es cierto que los mismos “existen”. A mis alumnos les cuento el siguiente ejemplo. Supongamos que ustedes tienen una biblioteca con cien libros de cien páginas cada uno, y yo les pido que me presten un libro. La

cantidad de material que estoy pidiendo es cien páginas. Ahora supongamos que yo les pido la página 14 de todos los libros que tienen. Fíjense que la cantidad de información que yo estoy solicitando es la misma que en el caso anterior, pero requiere un esfuerzo de sistematización completamente distinto, quizás tan elevado que a los fines prácticos sea equivalente a una situación en donde los datos ni siquiera existen. En síntesis, el dato y la sistematización del dato son dos cosas completamente distintas. Muchas veces ustedes dicen: “Yo sé que en el Ministerio de Educación tiene que estar la edad de los maestros”, pero si esta no está como ustedes necesitan que esté, es como si no estuviese, y no es falta de voluntad de la gente que organiza los datos, los objetivos que tienen ustedes como investigadores no tienen por qué coincidir con los objetivos de la gente que hace tareas administrativas.

#### RETORNOS A LA EDUCACIÓN, REVISITADOS.

Permítanme contarles algunos resultados para que puedan ver cómo se utilizaron los datos en un ejemplo concreto y algunas de las dificultades metodológicas discutidas anteriormente. Una pregunta que mucha gente hace es la siguiente: si la educación es una inversión rentable, ¿cuál es su rentabilidad? Lo que uno quiere saber cuando compra un activo cualquiera, desde un auto hasta una acción, es cuál es su rentabilidad. En términos de la educación, en principio, porque es relevante para las elecciones individuales (cuánto estudiar, qué estudiar y dónde hacerlo), pero también resulta relevante para decisiones agregadas o para la implementación de políticas públicas. Existe la percepción, a veces fundada, otras veces no, de que si el “retorno a la educación” es muy alto es porque no se ha invertido lo suficiente en este sector. Entonces, resulta que si la educación rinde mucho es porque está subexplotada. En un país con un bajo nivel de educación, la gente que tiene niveles muy altos de educación tiene retornos exageradamente altos. La preocupación por este tema no es meramente una cuestión individual de saber cuánto rinde la educación, sino una problemática agregada de gobiernos que intentan promover la educación para, a través de ella, mejorar el bienestar de la gente.

El problema empírico que enfrentamos a los efectos de medir el rendimiento de la educación es, como era de esperar, que no tenemos experimentos ni datos diseñados específicamente para este propósito. Elementalmente nos gustaría conocer *historias* de ingreso y educación, o sea, perseguir gente en el tiempo y que nos cuente su historia educativa: cómo fue el proceso a través del cual se determinó la educación, si tuvo que ver con decisiones personales, si tuvo que ver con limitaciones financieras de la familia, y la historia de sus ingresos. Técnicamente, nos gustaría contar con datos *longitudinales* (o de *panel*) en donde distintos individuos son observados en varios períodos. Este tipo de encuestas son de muy difícil implementación práctica, en particular porque resulta muy costoso re-entrevistar a una misma persona a lo largo de un período tan extenso como el relevante para medir la pregunta que nos interesa. Ciertamente, las encuestas longitudinales de largo plazo son todo un desafío para el análisis empírico en nuestro país.

Lo que tenemos, en definitiva, son datos de *corte transversal*, es decir, encuestas para un período en particular en donde se entrevistan varios individuos. La Encuesta Permanente de Hogares (EPH), que lleva a cabo el INDEC, es una suerte de compromiso entre una encuesta propiamente de panel (en donde los mismos individuos son re-entrevistados en distintos períodos) y una de corte transversal, en donde la encuesta es replicada en el tiempo para individuos esencialmente distintos. La misma tiene una estructura de *panel rotativo* en donde en cada período, un subgrupo de la muestra es reemplazado, quedando los restantes por un período relativamente corto.

Un punto metodológico importante es que la EPH no es una fuente de información diseñada para los fines de nuestro estudio, si bien contiene información que es relevante para los mismos. La EPH es una encuesta social, que realiza el INDEC en forma permanente, con el objetivo principal de producir estadísticas de empleo/desempleo, pero que pregunta varias cuestiones relacionadas con estas variables tales como los niveles de educación alcanzados, los ingresos, etc.. Insisto en que esta encuesta no está diseñada para ilustrar sobre el retorno a la educación, ni siquiera está diseñada para resolver problemas

científicos, tiene otros propósitos. Un vicio metodológico frecuente consiste en criticar una fuente de información porque no se ajusta a las restricciones y requerimientos de un plan de investigación particular. Muchas veces uno quiere utilizar una fuente de información para un fin para el cual no fue diseñada, y ese es uno de los peores problemas que uno encuentra permanentemente en la práctica. Resulta que la fuente de información fue diseñada para una cosa y uno la quiere usar para otra. En síntesis, cuando uno se pone a trabajar con datos que no fueron generados para el propósito que uno tiene, es esencial formular y responder una enorme variedad de preguntas metodológicas acerca de cómo se encuestó, qué se preguntó, por qué se preguntó, etc. , de modo de garantizar que los datos sean realmente informativos para nuestro problema.

Por ejemplo, en este problema es fundamental entender qué significa ingreso, para esto es muy importante saber cómo se formularon las preguntas relacionadas con esta variable, leer la pregunta que hicieron para la investigación. Entonces cuando uno trabaja con fuentes de información que no son relevadas por el investigador tiene que hacer varias preguntas metodológicas a la persona que las releva, hay que tener muchísimo cuidado con esto.

Ahora quisiera contarles algunos resultados. Habiendo salteado las numerosas dificultades empíricas discutidas anteriormente, la información de base está conformada por pares de ingreso y educación (además de otra información relevante), para un número muy grande de individuos. Técnicamente, el análisis procede a través de un modelo estadístico que permite relacionar ambos conceptos, usualmente algún modelo de regresión, muy posiblemente no lineal. Este es quizás el problema empírico más estudiado por la econometría aplicada, cualquier intento de ser exhaustivos en esta dirección nos agotaría inútilmente el tiempo disponible en esta charla; en esta dirección, no puedo más que invitarlos a leer la muy abundante literatura disponible.

Un resultado que se encuentra a menudo para nuestro país es que el retorno promedio de educación para los hombres es cercano a 8.7 por año y para mujeres es 9.3, es decir, un poquito más alto para estas últimas. Esto significa que un año de educación tiende a elevar los salarios promedio más o menos en nueve

por ciento para las mujeres y ocho por ciento para los hombres. Este es un resultado bastante coherente con respecto a las investigaciones disponibles. Es mucho menos rentable que en otros países de Latinoamérica, Argentina tiene un retorno cercano a diez por ciento anual, que es una cifra creíble, pero por ejemplo el caso de países mas pobres como Jamaica o Costa de Marfil, este número se eleva rápidamente a veinte por ciento. En general, en países que tienen escolaridades muy altas, como el caso de Suecia, los retornos a la educación son muy bajos.

Quisiera contarles otro problema metodológico relacionado con esta cuestión y con algunas de mis investigaciones. En algún momento de esta charla, hubo una discusión en relación a cómo interactúan los sistemas formales de educación con otros sistemas no tan formales, el tema es cómo interactúan algunas habilidades que -a falta de mejor nombre- llamaremos “psicológicas”, con procesos educativos. Este es, ciertamente, un problema muy viejo. Básicamente la pregunta tiene que ver con cómo algunas habilidades específicas y no observables del individuo interactúan con la educación formal, en la determinación de la productividad de una persona y, eventualmente, en sus ingresos. Existe alguna evidencia de que los individuos más “hábilés” tienden a tener retornos a la educación un poco mayores. Hay una evidencia, que nosotros no refutamos, que dice que la gente más hábil aprovecha el sistema educativo mucho más que la gente menos hábil. El problema que tiene esto, como dificultad metodológica, es que resulta que un aspecto no deseable de la educación -no deseable para algunos e importante para otros- es que puede aumentar las discrepancias entre la gente. Es cierto que le mejora la vida a todo el mundo pero también es cierto que a cierta gente se la mejora más que a otra. Lo que hicimos nosotros fue plantearnos cómo hacer para controlar diferencias biológicas no observables de cuestiones formales observables; e hicimos lo que muchísima gente hace en este tipo de estudio que es trabajar con una muestra de gemelos. Trabajamos con una encuesta de 723 pares de gemelos. Si lo hacemos con hermanos que no son gemelos vamos a terminar comparando ingresos de personas que tienen características observables diferentes y características genéticas diferentes, van a

estar interfiriendo otras variables. La “única forma” de que esto no suceda es si comparamos ingresos de personas que tienen información genética muy similar.

Utilizamos datos de una encuesta muy grande que hizo la universidad de Princeton. Me hubiera gustado mostrarles cómo se hace esta encuesta, fue muy gracioso, esto puede ocurrir en Estados Unidos nada más. Nosotros lo que hicimos fue comparar las discrepancias en los ingresos enfatizando el hecho de que para esta muestra las personas provienen de pares con información genética idéntica. Básicamente lo que queríamos hacer era aislar la información genética y, después de hacer un montón de cosas, con una tecnología un tanto sofisticada y esotérica, evaluar el impacto de la educación y las habilidades individuales en los ingresos. El estudio sugiere que, efectivamente, la educación le rinde relativamente más, en términos de ingresos, a la gente más hábil. Insisto, el resultado no es importante, lo que es importante es el proceso metodológico tortuoso que nos llevó al mismo. Tuvimos que utilizar datos muy específicos, la naturaleza del problema genera cosas muy concretas y la metodología estadística es marcadamente no estándar. Tiene muchas críticas, pero a mí me preocupan las decisiones metodológicas que llevan al resultado, garantizando que estas sean perfectamente reproducibles, y bien descritas. Así estamos garantizando que los resultados que obtenemos son una consecuencia de esa metodología y no un capricho de los investigadores.

## CONCLUSIONES

A modo de conclusiones, la contribución de la economía en el área de la educación tiene que ver con establecer ciertas pautas metodológicas que son relativamente estándar en la economía y que se pueden trasladar, no sin conflictos, a la problemática educativa. Enfatizo el hecho de que el análisis econométrico-estadístico muchas veces intenta contrarrestar las enormes limitaciones derivadas de la inhabilidad de experimentar. Hay una rama de la economía muy reciente, que se llama economía experimental y que es realmente promisoría, pero convengamos que es un camino muy largo para recorrer. En general los principales desafíos tienen que ver con cómo adecuar los datos que

están disponibles a los efectos de que los mismos sean informativos y relevantes para los problemas que nos hemos planteado. Ojalá la investigación empírica fuese un mundo en donde los datos están aquí, las teorías allí y los métodos acá, y existiese un manual que dijese “para estos datos haga tal o cual cosa”, “para esta pregunta me faltan estos datos”, o “esta técnica se puede usar en tal o cual contexto”. Lamentablemente este triángulo en donde se desenvuelven las técnicas, los datos y las teorías, implica que hay que interactuar mucho. Pensar que tenemos una teoría sólida y que los datos los vamos a encontrar a la vuelta de nuestra casa, o pensar que porque me han regalado muchos datos los voy a poder usar sin mayor análisis, es de una irresponsabilidad o inocencia muy grande. Lo que hace un investigador que trabaja responsablemente es moverse inteligente y metódicamente en este espacio en donde la teoría, la técnica y la información interactúan.

Bueno eso es todo lo que yo quería contar, les agradezco la atención dispensada.

## SUGERENCIAS BIBLIOGRÁFICAS

Debido al carácter informal de la reunión, decidí no mezclar referencias en el texto, y agrupar algunas recomendaciones en esta sección para aquellos interesados en ahondar en los temas tratados en esta charla. Muchas cuestiones metodológicas de la economía están tratadas en detalle en Blaug (1980). Klimovsky (1994) es una muy buena referencia general para cuestiones metodológicas en disciplinas sociales, y el trabajo reciente de Navarro (2005) se ocupa de varias cuestiones metodológicas específicas de la econometría. En relación al potencial de la economía experimental, ver Kagel y Roth (1997). Sobre el problema de retornos a la educación, la literatura es abismalmente extensa, una revisión reciente es Card (2001) y una más general, aunque no tan actualizada es Willis (1986). La referencia a Adam Smith proviene de Belfield (2001). Sobre las cuestiones de inteligencia y educación ver Devlin et al. (1997). Un buen texto introductorio de econometría básica, que combina hábilmente cuestiones técnicas

con metodológicas e interpretativas es Wooldridge (2002). Una colección de artículos que versan sobre cómo los economistas han lidiado con varias problemáticas de la educación es Becker y Baumol (1995). Varios ejemplos utilizados se derivan de proyectos de investigación en los que estuve directa o indirectamente involucrado. Arias, Hallock y Sosa Escudero (2001) es un estudio sobre heterogeneidades en los retornos a la educación usando técnicas de regresión por cuantiles en una muestra de hermanos gemelos. Sosa Escudero (2004) y Di Gresia, Fazio, Porto, Ripani y Sosa Escudero (2004) se refieren a los determinantes de la “performance” de los estudiantes de universidades públicas. La anécdota de Andrew Wiles esta tomada del apasionante libro de Singh (1997, pag. 350).



Universidad de  
**San Andrés**



## REFERENCIAS BIBLIOGRÁFICAS

Arias, O., Hallock, K. y Sosa Escudero, W. (2001) "Individual heterogeneity in the returns to schooling: instrumental variables quantile regression using twins data", *Empirical Economics*, 26:1, 7-40.

Becker, W. y Baumol, W. (eds) (1995) *Assessing Educational Practices. The Contribution of Economics*, MIT Press, Cambridge.

Belfield, C. (2001) *Did Adam Smith know everything about the economics of education? Or nothing?*, mimeo, Teachers College, Columbia University.

Blaug, M. (1980) *The Methodology of Economics, or How Economists Explain*, Cambridge University Press, Cambridge.

Card, D. (2001) "Estimating the Returns to Schooling: Progress on Some Persistent Econometric Problems", *Econometrica*, 69:5, 1127-60.

Devlin, B., Fienberg, S., Resnick, P. y Roeder, K. (eds.) (1997) *Intelligence, Genes, and Success: Scientists Respond to The Bell Curve*, Wiley, New York.

Di Gresia, L, Fazio, M., Porto, A., Ripani, L. y Sosa Escudero, W. (2005) "Rendimiento y Productividad de los Estudiantes. El Caso de las Universidades Públicas Argentinas", en Porto, A. (editor), *Economía de la Educación Universitaria: Argentina-Brasil-Peru*, Editorial de la Universidad Nacional de La Plata, La Plata.

Kagel, J. y Roth, A. (1997) *The Handbook of Experimental Economics*, Princeton University Press, Princeton.

Klimovsky, G. (1994), *Las Desventuras del Conocimiento Científico*, AZ editora, Buenos Aires.

Navarro, A. (2005) "Reflexiones sobre el estado actual de la metodología de la econometría", en Asociación Argentina de Economía Política (editores), *Avances Recientes en Econometría*, Buenos Aires, en prensa.

Singh, S. (1999) *El Último Teorema de Fermat*, Editorial Norma, Bogotá.

Sosa Escudero, W. (2003) *Heterogeneous effects of the observed determinants of college "performance"*, mimeo, Universidad Nacional de La Plata, La Plata.

Willis, R. (1986) "Wage determinants: a survey and reinterpretation of human capital earnings functions", en Ashenfelter, O. y Layard, R. (eds.), *Handbook of Labor Economics*, Vol. 1, Cap. 10, Elsevier Science, Amsterdam.

Wooldridge, J. (2002), *Introductory Econometrics: A Modern Approach*, 2ª edición, South-Western College Publishers, New York.



Universidad de  
**San Andrés**

OTROS DOCUMENTOS DE LA SERIE  
"DOCUMENTO DE TRABAJO" DE LA ESCUELA DE EDUCACIÓN

- Nº 1 OIBERMAN, Irene Beatriz. **La creación de un sistema de información educativa para la gestión.** Mayo de 2001; 20 páginas.
- Nº 2 ABDALA, Félix. **Los excluidos del sistema educativo argentino: resultados preliminares de una investigación.** Septiembre de 2001; 14 páginas.
- Nº 3 GALLART, María Antonia. **La escuela como objeto de investigación.** Septiembre de 2001; 8 páginas.
- Nº 4 FISZBEIN, Ariel. **Análisis institucional de la descentralización educativa: el caso de los países de Europa Central.** Agosto de 2002; 14 páginas.
- Nº 5 CARUSO, Marcelo. **La relación pedagógica moderna: cultura y política de la didáctica.** Octubre de 2002; 21 páginas.
- Nº 6 TEDESCO, Juan Carlos. **Una política integral para el sector docente.** Mayo de 2003; 11 páginas.
- Nº 7 WAINERMAN, Catalina. **Un siglo en los libros de lectura en la escuela primaria.** Junio de 2003; 14 páginas.
- Nº 8 BEECH, Jason. **El discurso de la era de la información y la educación en Argentina y Brasil.** Noviembre de 2003; 16 páginas.
- Nº 9 BAQUERO, Ricardo. **La educabilidad como problema político. Una mirada desde la psicología educacional.** Noviembre de 2003; 20 páginas.
- Nº 10 GAINZA, Violeta. **La educación musical entre dos siglos: del modelo metodológico a los nuevos paradigmas.** Noviembre de 2003; 15 páginas.
- Nº 11 AGUERRONDO, Inés. **Ministerios de Educación: de la estructura jerárquica a la organización sistémica en red.** Junio de 2004; 13 páginas.
- Nº 12 FINOCHIARO, Alejandro. **UBA con Estado Nacional: un estudio de caso sobre la autonomía universitaria.** Junio de 2004; 17 páginas.
- Nº 13 KESSLER, Gabriel. **Trayectorias escolares de jóvenes que cometieron delitos contra la propiedad con uso de violencia.** Noviembre de 2004; 26 páginas.

N° 14 PALAMIDESSI, Mariano. **El cambio del currículum para la escuela primaria a lo largo de un siglo (1880-1980)**. Noviembre de 2004; 31 páginas.

N° 15 CARLI, Sandra. **Infancia, Cultura y Educación en las décadas del 80 y 90 en Argentina**. Marzo de 2005; 17 páginas.

N° 16 RAVELA, Pedro. **Estudio de los factores institucionales y pedagógicos que inciden en los aprendizajes en escuelas primarias de contextos desfavorecidos en Uruguay**; Junio de 2005; 39 páginas.



Universidad de  
**San Andrés**