

# ***Poder de mercado y estrategias: El método de máxima entropía generalizada***

Henry A. Lubel<sup>1</sup>  
Abril de 2016

---

## **Resumen**

*El presente trabajo hace una revisión de los principales conceptos vinculados a la estimación por máxima entropía y su aplicación al ámbito de la organización industrial. Se exponen las particularidades del método, y cómo puede utilizarse para la estimación directa de las estrategias en un juego a partir de un modelo multinomial de elección discreta. A tal fin, se propone una estructura de mercado donde la variable de decisión de las firmas es el precio bajo un esquema no cooperativo. Finalmente, se hace una aplicación del método al caso del mercado argentino de Kerosene.*

**Palabras clave:** Poder de mercado, Estrategias, Juegos no cooperativos, Máxima Entropía Generalizada

---

**JEL:** C13, C35, C72, D43, L13, L71

---

<sup>1</sup> Universidad Torcuato Di Tella – Departamento de economía. Tutor: Leandro Arozamena

# ***Poder de mercado y estrategias: El método de máxima entropía generalizada***

## *Tabla de contenidos*

### **I. INTRODUCCIÓN**

### **II. ANTECEDENTES EN LA ESTIMACIÓN DEL PODER DE MERCADO**

### **III. ¿CÓMO MEDIR EL PODER DE MERCADO?**

### **IV. EL PRINCIPIO DE MÁXIMA ENTROPÍA**

#### **CONCEPTO DE ENTROPÍA**

#### **ESTIMACIÓN POR MÁXIMA ENTROPÍA**

#### **MÁXIMA ENTROPÍA GENERALIZADA**

### **V. ESTIMACIÓN DEL PODER DE MERCADO**

#### **DEFINICIÓN DEL JUEGO**

#### **ADECUACIÓN ECONOMÉTRICA**

#### **PROBLEMA DE ESTIMACIÓN**

#### **FUNCIÓN DE BENEFICIO DESCONOCIDA**

#### **PODER DE MERCADO**

### **VI. APLICACIÓN: MERCADO ARGENTINO DE KEROSENE**

#### **CARACTERÍSTICAS GENERALES**

#### **ESTIMACIONES**

### **VII. CONCLUSIONES**

### **REFERENCIAS**

### **ANEXO 1: CONSISTENCIA DE LOS ESTIMADORES GME-NASH**

### **ANEXO 2: ESTIMACIONES DE ESTUDIOS PREVIOS DEL PODER DE MERCADO**

## I. INTRODUCCIÓN

La mayor parte de los trabajos empíricos que buscan determinar el poder de mercado ejercido por las firmas en una industria en particular se concentran en buscar el *Mark-up*<sup>2</sup> que establecen los *players* y las causas que los originan.

Los estudios bajo el enfoque *Estructura-Conducta-Desempeño*<sup>3</sup> intentan a partir de datos de corte transversal de industrias múltiples, establecer relaciones entre ciertas variables de estructura de mercado con variables proxy del poder ejercido, mientras que los estudios de la llamada nueva organización industrial, a partir de modelos micro fundados, estiman parámetros de conducta de un mercado particular. Luego, computan alguna medida de poder de mercado (e.g. índice de Lerner) construido con los precios y costos marginales que surgen del equilibrio del modelo planteado.

Sin embargo, los estudios tanto en uno como en otro enfoque, rara vez han intentado estimar las estrategias de las firmas de forma directa en los mercados analizados y cómo éstas se vinculan con el poder de mercado ejercido.

La razón de la ausencia de este tipo de estudios no es un resquicio intelectual, sino más bien se trata de una dificultad de carácter econométrico: La estimación de las estrategias de forma directa normalmente acarrea problemas donde la cantidad de parámetros a estimar es muy superior a la cantidad de observaciones disponibles (*Ill-Posed Problems*). Incluso aunque fuera posible lidiar con este problema, está la dificultad adicional al utilizar los métodos de estimación tradicionales, como el de máxima verosimilitud, que tienen poca flexibilidad (o directamente imposibilidad) para incorporar restricciones teóricas, como las que surgen por ejemplo, de las nociones de equilibrio de Nash presentes en la teoría de juegos, y que normalmente deben ser tenidas en cuenta al momento de trabajar en mercados de carácter oligopólico.

El método de estimación de máxima entropía generalizada (GME) basado en el principio de máxima entropía, que proviene de la teoría de la información (IT), permite lidiar con estos inconvenientes, además de otros como la no necesidad de asunción de una función de densidad conjunta particular, como lo es en el caso de la estimación por máxima verosimilitud (MLE).

La utilización del método para la estimación de problemas económicos (y más aún dentro de la organización industrial) ha sido poco explotada y explorada.

El método fue introducido en este ámbito por Golan, Judge, y Miller (1996), basándose en literatura disponible de la teoría de la información que tiene su origen en Shannon (1948) y su medida de entropía, y luego extendida por Jaynes (1957).

Ambos métodos están basados en la fórmula de Hartely, que se propone medir la cantidad de información necesaria para caracterizar un conjunto dado de elementos  $A = \{a_1, a_2, a_3, \dots, a_m\}$ . Tanto Shannon como Jaynes se proponen mensurar el grado de incertidumbre que refleja una distribución de probabilidad cualquiera sobre un espacio muestral, y cómo seleccionar la más desinformada de todas aquellas distribuciones que sean apropiadas en función de las observaciones y restricciones que el problema de interés imponga.

De este modo, en este trabajo se expondrá un procedimiento para estimar las estrategias (puras o mixtas) de un juego discreto (con dos enfoques distintos) que refleja un mercado oligopólico, y computar a partir de ellas, el poder de mercado ejercido por las firmas participantes sin tener que lidiar con los problemas econométricos mencionados.

---

<sup>2</sup> Más adelante se brindaran precisiones sobre este concepto.

<sup>3</sup> *Structure-Conduct-Performance*

En la sección II, se hará una breve reseña de los trabajos que han indagado en la estimación del poder de mercado en una o múltiples industrias, buscando de esta forma reflejar el estado del arte en la materia.

Luego, en la sección III, se establecerán postulados básicos acerca de qué se entenderá por poder de mercado.

A continuación, en la sección IV, se expondrá el principio de máxima entropía y el método de estimación de máxima entropía generalizada, las propiedades de los estimadores, algunos teoremas útiles, y algunas simulaciones de casos prácticos para mostrar su potencial y su relación con otros estimadores.

En la sección V, se propone un modelo que replica aquel propuesto por Golan, Karp y Perloff (1998) de un juego repetido para el caso de un duopolio donde las firmas participantes eligen su precio en cada etapa de un espacio de acciones, y cómo adaptarlo para transformarlo en un problema de elección discreta multinomial para ser estimado mediante GME. Se buscará formalizar con mayor precisión al modelo original, y se mostrará cómo calcular el poder de mercado a partir de las estrategias estimadas.

Finalmente, en la sección VI, se hace una aplicación del método en el mercado argentino del Kerosene, donde se computan las estrategias de los participantes de este sector, y el poder de mercado ejercido por cada una de las firmas.

Finalmente, en la sección VII, se esgrimirán las conclusiones.

## II. ANTECEDENTES EN LA ESTIMACIÓN DEL PODER DE MERCADO

El análisis sistemático de la *performance* de los mercados (la habilidad de éstos en producir beneficios para firmas o consumidores) y su relación con la conducta de los agentes, sitúa su inicio para la mayor parte de las reseñas históricas en Mason (1939, 1949), en donde se hace el primer intento serio por buscar de forma empírica una relación explícita entre aquellas a partir de un análisis pormenorizado de la industria en cuestión.

Más tarde, Bain (1951, 1956) generalizó y formalizó este enfoque conocido como estructura-conducta-desempeño (SCP en inglés) utilizando datos de corte transversal para industrias múltiples<sup>4</sup>.

Típicamente, estos estudios obtienen una medida de la *performance* de múltiples industrias y, otra sobre su estructura. Finalmente, se regresa la una con la otra, para luego determinar por qué hay industrias que tienen mejor o peor desempeño que otras.

Una de las medidas de *performance* más utilizadas (como proxy del poder de mercado ejercido por las firmas) suelen ser, en estos estudios, la tasa de retorno (sobre activos o patrimonio neto)<sup>5</sup>.

Los problemas de la utilización de esta medida son numerosos, y algunos autores se han ocupado de consignarlos<sup>6</sup>, aunque todos pueden resumirse en que se hace uso de hojas de balance para su cálculo, las cuales siguen principios contables basados en la legislación a la cual está sujeta la firma y prácticas generalmente aceptadas de la profesión contable, y no están construidas como la teoría microeconómica indica (por ejemplo, los costos no están valuados a su costo de oportunidad, o el capital y los activos no corrientes pueden estar valuados a valores históricos).

Otra de las medidas de *performance* utilizadas son el índice de Lerner, definido como

---

<sup>4</sup> Una reseña de trabajos posteriores puede encontrarse en Schmalensee (1989).

<sup>5</sup> Tal como aclara Schmalensee (1989) algunos autores trabajan con el retorno antes de impuestos, para evitar contaminar el estudio las particularidades y distorsiones del sistema impositivo, y otros con el retorno luego de impuestos para que estas sean consistentes con la hipótesis de beneficio cero bajo competencia perfecta.

<sup>6</sup> Ver, por ejemplo, la reseña de Perloff, Golan y Karp (2006).

$$L \equiv \frac{p-MC}{p},$$

donde  $p$ , y  $MC$  son el precio, y los costos marginales de mercado respectivamente.

Dado que el costo marginal es una variable generalmente no observada, suele usarse como proxy el costo medio variable, lo cual puede generar sesgos relevantes en la medición.

También, se ha utilizado como medida de performance la  $q$  de Tobin, definida como el valor bursátil de una compañía sobre su valor contable. Naturalmente, esta medida es atractiva si las firmas que intervienen en los mercados bajo análisis cotizan en los mercados de valores y, además, si lo hacen en una porción significativa de su capital integrante.

Por el lado de las medidas de estructura, las más utilizadas son medidas de concentración<sup>7</sup> como el índice HHI, definido como  $HHI = \sum_i^N s_i^2$  para las  $i=1\dots N$  firmas de un mercado siendo  $s_i$  la cuota de mercado de la firma  $i$ ; pero también, medidas de concentración como C4, C8, C20, etc.<sup>8</sup>

Suele criticarse de estas medidas el hecho de no suscribir al criterio de mercado relevante, es decir que no incluyen firmas productoras de sustitutos cercanos, y dan la apariencia de ser mercados muy concentrados, cuando en realidad se trata del mismo mercado pero con productos diferenciados. O bien, el caso contrario, donde se incluyen más productos de los que debieran en la definición de la industria, y los indicadores terminan reflejando que el mercado está más atomizado de lo que en realidad es.

Para definir el mercado relevante la mayor parte de los organismos y entes de defensa de la competencia, tienen en cuenta la sustituibilidad del bien, tanto desde el punto de vista de la demanda como de la oferta, así como el alcance geográfico donde dicho bien es ofrecido y/ demandado.

Más allá de esta crítica general que puede hacerse en cualquier estudio de organización industrial, el problema central de las medidas de concentración de la oferta, y de los estudios SCP en general, es que estas medidas son endógenas al fenómeno que se intenta explicar, por ejemplo: la tasa de ganancia en una industria puede afectar el grado de concentración haciendo más atractivo el ingreso de otras firmas. De esta forma, la causalidad es bidireccional y las estimaciones no resultan consistentes.

Incluso, Schmalensee (1989) va más allá, diciendo que la mayor parte de los autores están casi exclusivamente preocupados por los estadísticos  $t$  y  $f$  de la estimación, más que por encontrar variables instrumentales válidas que resuelvan el problema de la endogeneidad. En realidad, la posibilidad de encontrarlas es, desde el punto de vista teórico, muy escasa.

La base de este problema estriba en que la mayor parte de estos estudios se caracterizan por ser de carácter descriptivo y no siguen un enfoque estructural (*i.e.* no están micro fundados). Esto es, no buscan la conexión teórica que puede existir entre estructura, conducta y performance. Por lo tanto, cualquier interpretación de estos trabajos que escape a la simple observación de ciertos hechos estilizados o regularidades empíricas sería un error.

Estos estudios, además, se concentran en describir las causas del poder de mercado y no el grado del poder ejercido, ni tampoco en como las estrategias elegidas por los jugadores lo afectan.

<sup>7</sup> También se han utilizado como medidas de estructura indicadores sobre las barreras de entrada como la intensidad del capital, tamaño mínimo eficiente, etc., pero no serán abordadas aquí.

<sup>8</sup> Market Share de las primeras  $n$  firmas

Es por esta razón que, a finales de los '80, el paradigma SCP ha caído en desuso y, en su lugar surgió la Nueva Organización Industrial empírica (NEIO). El procedimiento habitual de este enfoque, ya sea en versiones estáticas o dinámicas, es la de estimar ecuaciones o derivaciones de condiciones de primer orden resultantes de una problema de optimización apoyada en el principio de maximización del beneficio de las firmas (o de la función objetivo respectiva de los agentes económicos involucrados).

Dado que estos trabajos modelan el comportamiento estratégico de los agentes en un escenario oligopólico, se apela al herramental provisto por la teoría de juegos.

Toda modelización, por tanto, involucra necesariamente la especificación de un juego. Los parámetros conductuales del mismo serán estimados, por lo que todo el espectro competitivo (desde el monopolio puro hasta la competencia perfecta) son posibles resultados de estos ejercicios.

El objetivo final de estos estudios es el de estimar el poder de mercado y el costo marginal (asumidos como desconocidos) de forma simultánea y consistente.

Un típico set de ecuaciones es el siguiente:

Dado un mercado de un bien homogéneo con firmas idénticas (i.e. funciones de costos idénticas), la función de demanda inversa del mercado es:

$$p = p(Q, Z),$$

donde  $p$  es el precio,  $Q$  son las cantidades demandadas, y  $Z$  es un vector de variables exógenas como el ingreso o precios de bienes sustitutos<sup>9</sup>.

Definiendo al ingreso como  $I = p(Q, Z) Q$ , el ingreso marginal será

$$Img = p + \lambda P_Q(Q, Z)Q$$

donde  $\lambda$  es un parámetro de conducta introducido de forma ad-hoc, para lograr capturar el grado de competencia en el mercado<sup>10</sup>.

Si el costo marginal es  $Cmg(Q, w)$ , y la condición necesaria de optimalidad es  $Cmg = Img$ , entonces, con  $\lambda = 0$  se estará frente a un escenario de competencia perfecta ( $Cmg = p$ ), mientras que si  $\lambda = 1$  el ingreso marginal será el del monopolio. Con  $\lambda$  entre 0 y 1, se trata de un oligopolio<sup>11</sup>. Por ejemplo,  $\lambda = 1/n$  es el valor que adquiere el parámetro en una solución *a-lá* Cournot con  $n$  representando a la cantidad de firmas (idénticas) participantes.

La estimación (consistente) de estos parámetros, permitirá obtener una medida del costo marginal y consecuentemente el poder de mercado ejercido en la industria.

Lógicamente, como afirma Bresnahan, la mayor parte de estos estudios usan como método de estimación el de ecuaciones simultáneas.

En los modelos dinámicos, altos niveles de  $\lambda_t$  están asociados a períodos de cooperación, mientras que  $\lambda_t$  bajos están vinculados con períodos de guerras de precios (ver por ejemplo Porter, 1984) o de no cooperación.

Existe una variante a esta clase de modelos que también forma parte de la corriente NEIO que son los llamados modelos de variación conjetural: Los defensores de estos modelos argumentan que los modelos estructurales hacen asunciones teóricas y distribucionales demasiado fuertes (en el ejemplo anterior, la especificación de una forma funcional para la

<sup>9</sup> Naturalmente que en caso de que el problema involucre también un enfoque estructural para obtener la demanda del mercado, el ingreso y otras variables pueden volverse endógenas.

<sup>10</sup> Una interpretación más completa de  $\lambda$  puede encontrarse en Bresnahan (1989)

<sup>11</sup> Existe un debate acerca de si  $\lambda$  es una variable continua o no, y de cuál es su correcta interpretación, la cual queda fuera del ámbito de este trabajo.

demanda y la función de costos), y por tanto prefieren utilizar modelos de formas reducidas o no paramétricos que hagan menos asunciones o supuestos. Como contrapartida, estos modelos no podrán estimar el grado de competencia en un mercado determinado, aunque si testear hipótesis acerca de si el mercado es o no competitivo.

La mayor parte de estos modelos se basan en pruebas empíricas basadas en ejercicios de estática comparada: Por ejemplo, si el costo marginal de un producto se eleva como consecuencia de un incremento en la tasa de un impuesto, entonces, en el caso de un mercado competitivo, es de esperar que dicho incremento se de en la misma proporción en el precio.

La hipótesis de que los precios se muevan en la misma proporción que el costo marginal en un ejercicio de estática comparada está en la base de este tipo de modelos.

Un muy citado trabajo dentro de este enfoque es el de Hall (1988), que asumiendo una tecnología con retornos constantes a escala (que es lo esperable en un mercado competitivo), estudia si el precio y el costo marginal se igualan en el tiempo. A tal fin, se utilizan las variaciones en el producto y en las cantidades y precios de los insumos. De esta forma, Hall puede confirmar o refutar la existencia de mercados competitivos (y retornos constantes a escala de forma conjunta) sin hacer asunciones paramétricas sobre las funciones de costos, aunque también sin poder encontrar una medida concreta del poder de mercado ejercido.

Tal como menciona Bresnahan, tanto los modelos estructurales como los de forma reducida, estiman el poder de mercado o el grado de competencia para un *setting* determinado. Por lo que rara vez, las especificaciones antes mencionadas permiten la pregunta sobre cómo los cambios de estructura afectan la conducta y la performance. El mayor riesgo de los modelos estructurales, es justamente, especificar mal la estructura.

Ha habido algunos trabajos que han invertido la relación de los estudios SCP para estudiar este vínculo, y han establecido al precio como variable dependiente, y a medidas de concentración como independientes en una ecuación de forma reducida del precio. Estos estudios asumen a la concentración como exógena y buscan mostrar empíricamente lo que básicamente todas las teorías sobre el oligopolio dicen: que un mayor grado de concentración causa un *ratio* precio-costo mayor.

Sin embargo, como se mencionó previamente, dentro de la familia de modelos estructurales, sólo unos pocos estudios se ocuparon de estimar las estrategias del juego subyacente de forma directa (Bjorn y Vuong 1985, Bresnahan y Reiss 1991, Kooreman 1994, Tamer 2003).

Kooreman (1994) propone distintos juegos (Nash, Stackelberg, Optimalidad de Pareto, y un caso mixto Nash-Pareto) para juegos estáticos de dos jugadores y con dos acciones que luego estima con probits bivariados mediante máxima verosimilitud.

Bresnahan y Reiss (1990) extienden a múltiples jugadores la estimación de un juego estático a través de un sistema lineal de variables dummies endógenas utilizando el procedimiento de Heckman (1978).

Algunos intentos recientes, como Aguirregabiria y Mira (2007), intentan, mediante un procedimiento secuencial, estimar por un procedimiento de máxima verosimilitud modificado (pseudo máxima verosimilitud) un juego dinámico con un espacio de acciones discreto e información incompleta.

El método que se expondrá en las próximas secciones está dentro de este último grupo de trabajos, aunque a diferencia de ellos, no se hará ninguna asunción distribucional.

Un último comentario debe hacerse sobre la producción de estudios dentro de la temática a nivel local: hay pocos estudios empíricos previos (al menos publicados) para la República Argentina sobre poder de mercado, y prácticamente ninguno que modele un oligopolio a partir de un juego entre firmas. Coloma (1998) analiza el mercado de expendio de gasolina a partir

de estimaciones de un modelo estructural básico y Coloma (2002) estudia si hubo cambios en la estructura y poder de mercado ejercido antes y después de la fusión entre Repsol e YPF con el mismo marco teórico. Burdisso, Cattena y D'Amatto (2001) estiman el poder de mercado en el sector bancario argentino, a partir de un modelo de variación conjetural para el caso de firmas multiproducto para un panel de bancos seleccionados, y comparan los resultados obtenidos con los que arrojaría un análisis tradicional SCP. Lema y Barron (2003) realizan un análisis similar a Coloma (1998) para la industria láctea.

### III. ¿CÓMO MEDIR EL PODER DE MERCADO?<sup>12</sup>

Dadas  $n$  firmas productoras de un único bien homogéneo, con tecnologías idénticas, y una estructura de mercado oligopólica donde las firmas eligen las cantidades que producirán en un único período con información completa, las firmas maximizan su beneficio:

$$\max_{q_i} p(Q)q_i - cq_i$$

siendo  $Q = q_1 + q_2 + \dots + q_n$ ,  $p(Q)$  la función de demanda inversa del mercado, y  $c$  el costo medio de cada firma.

La condición de primer orden de este problema<sup>13</sup> es aquella que iguala el costo y el ingreso marginal, es decir:

$$Img = p + q_i \frac{dp}{dQ} = c = Cmg$$

Como todas las firmas son idénticas, en equilibrio, todas producen las mismas cantidades (i.e.  $q_1 = q_2 = \dots = q_n \equiv q$ ), por lo que en equilibrio,  $Q = nq$ . Con esto es posible reescribir la expresión anterior como:

$$L \equiv \frac{p - c}{p} = -\frac{1}{n\varepsilon} = -\frac{s}{\varepsilon}$$

donde  $\varepsilon = \frac{dQ}{dp} \frac{p}{q}$  es la elasticidad precio de la demanda, y  $s = \frac{q}{Q} = \frac{1}{n}$ .

Puede notarse que cuando  $n = 1$ ,  $L = -\frac{1}{\varepsilon}$  que es el caso usual del monopolio en donde la firma establece un ratio precio-costo marginal igual al inverso de la elasticidad-precio de la demanda. Por otro lado, debe notarse que,

$$\lim_{n \rightarrow \infty} L = \lim_{n \rightarrow \infty} -\frac{1}{n\varepsilon} = 0$$

Lo que significa que bajo competencia perfecta, el índice de Lerner es cero, o lo que es lo mismo, que el precio es igual al costo marginal.

De esta forma, queda demostrado, que el índice de Lerner es una medida apropiada del poder de mercado que se ejerce en una industria.

Levantando el supuesto de firmas idénticas, puede calcularse el índice de Lerner del mercado como un índice ponderado a lo largo de las  $i$  firmas por su respectiva cuota de mercado, y puede demostrarse que:

$$L = \sum_{i=1}^n s_i \frac{p - c_i}{p} = -\sum_{i=1}^n \frac{s_i^2}{\varepsilon} = -\frac{HHI}{\varepsilon}$$

<sup>12</sup> Esta sección está basada parcialmente en Golan, Karp y Perloff (2006)

<sup>13</sup>  $\frac{\partial(p(Q)q_i - cq_i)}{\partial q_i} = 0$



donde HHI es el índice de Herfindahl-Hirschman mencionado en la sección anterior. De esta manera, se observa de forma clara la relación directa que existe entre el poder de mercado, su grado de concentración (siempre partiendo de un esquema *a-lá* Cournot), y la demanda de los consumidores.

#### IV. EL PRINCIPIO DE MÁXIMA ENTROPÍA

##### Concepto de entropía

Como se mencionó en la introducción, el concepto de entropía que aquí se utilizará<sup>14</sup> tiene su origen en la teoría de la información (IT) desarrollada por Shannon (1948). La IT es un campo de conocimiento que se nutre de la matemática, la estadística, y la informática para brindar soluciones a problemas que se presentan en diversa áreas de conocimiento, entre ellas: teoría de comunicaciones, física estadística, teoría de redes, compresión de datos, reconstrucción de imágenes, ciencias de la computación, inferencia estadística y probabilidad e, incluso, en teoría de portafolios.

Shannon inauguró el concepto de entropía en el seno de la teoría de las comunicaciones argumentando que es posible reducir a cero la probabilidad de error en una transmisión de datos en donde la tasa de comunicación es inferior a la capacidad del canal de transmisión.

Parte de estos hallazgos son, por ejemplo, utilizados hoy en los algoritmos de corrección de errores para la lecto-escritura de discos ópticos (CD's, DVD's, etc.).

La entropía es una medida de la incertidumbre media de una distribución de probabilidad sobre el espacio muestral que define su variable aleatoria asociada.

Formalmente, dada una variable aleatoria  $X$  que por definición es una función  $X: A \rightarrow R$ , donde  $A = \{a_1, a_2, \dots, a_n\}$  es un espacio muestral discreto, y dada una función de probabilidad puntual  $p(X = x)$ , la entropía es una medida que indica cuan informativa es la distribución de probabilidad  $p$  sobre el conjunto  $A$  (o lo que es lo mismo sobre la variable aleatoria  $X$ ).

**Definición** La entropía  $H(X)$  de una variable aleatoria discreta  $X$  está definida por

$$H(X) = - \sum_{x \in X} p(x) \log p(x)$$

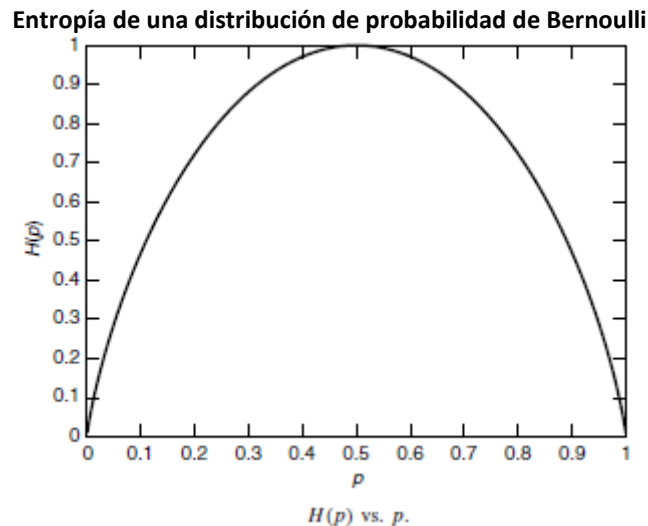
Por convención  $0 \log 0 = 0$  justificado en el hecho de que  $\lim_{p \rightarrow 0} p \log p = 0$ .

La entropía está medida en *nats*, que es el nombre que se le da a la unidad informacional cuando la entropía está expresada en logaritmo natural. Si el logaritmo estuviera expresado en base 2, entonces la unidad de medida serían *bits*.

Se trata de una función cóncava en  $p$ , que alcanza un mínimo cuando la distribución de probabilidad es una distribución degenerada ( $p(a) = 1$  con  $a \in X$ , y  $p(x) = 0 \forall x \in X \forall x \neq a$ ), es decir, que la incertidumbre es mínima cuando la probabilidad de ocurrencia de un estado particular es 1, mientras que alcanza un máximo cuando la distribución de probabilidad es uniforme, es decir cuando  $p(x) = \frac{1}{n} \forall x \in X$ . Lo que significa que la incertidumbre es máxima cuando todos los estados son igualmente probables.

<sup>14</sup> Existe una conexión entre el concepto de entropía que aquí se expone, y el esgrimido por la segunda ley de la termodinámica, pero no será tratado aquí.

Si se toma por ejemplo (tomado de Cover y Thomas 1991) una variable aleatoria binaria  $G$  que sigue una distribución de probabilidad de Bernoulli con parámetro  $p$ , la entropía tomará los valores que se observan en el gráfico, y que reflejan lo anteriormente mencionado:



Se observa entonces, que la entropía alcanza su máximo cuando  $p = 0.5$ , es decir cuando los 2 estados de  $G$  son igualmente probables. Cuanto más uniforme sea la distribución, más se incrementa la incertidumbre y por ende la entropía. Notar también que  $H(P) \geq 0$ .

Es importante mencionar, que la entropía se nutre únicamente de las probabilidades de una función de distribución, y no es necesario para su cálculo contar con los valores que toma la variable aleatoria.

Kullback y Leibler (1951) proponen una medida de entropía relativa que se define a continuación.

**Definición** La entropía relativa entre dos distribuciones de probabilidad  $p(x)$  y  $q(x)$  está definida como

$$D(p \parallel q) = \sum_{x \in X} p(x) \log \frac{p(x)}{q(x)}$$

Esta medida indica la ganancia de información sobre  $A$ , que refleja el conocimiento adicional que provee  $p$  en relación a  $q$ . De esta forma, si  $q$  fuera una distribución uniforme, entonces  $D(p \parallel q) = H(P)$ . Es si se quiere, una comparación de uniformidad de dos distribuciones de probabilidad (Golan, 2008).

La medida de entropía normalizada puede resultar útil para realizar comparaciones y diagnósticos. La misma se define como:

$$S(p) = \frac{[- \sum_{x \in X} p(x) \log p(x)]}{\log n}$$

Esta medida es una simple transformación de la entropía, en donde  $S(p) = 0$  refleja completa certidumbre, mientras que  $S(p) = 1$  refleja completa ignorancia o uniformidad.

Hay una gran cantidad de medidas que se derivan de las recién expuestas (ver, por ejemplo, la medida de entropía generalizada en Rényi 1970), y teoremas y propiedades que son de aplicación en distintas áreas. Se verá aquí sólo las que resultan de interés para el problema<sup>15</sup> a analizar.

### Estimación por máxima entropía

Una vez definida la forma de medir la incertidumbre de una distribución de probabilidad sobre una variable aleatoria, cabe preguntarse si es posible estimar una distribución de probabilidad discreta desconocida a partir de un número finito de momentos muestrales observados.

La respuesta a esta pregunta la da el método de máxima entropía (ME) propuesto por Jaynes (1957), que maximiza la entropía sujeta a los datos disponibles del problema, y selecciona de todas las distribuciones de probabilidad que cumplen con las restricciones impuestas, aquella que es más uniforme. Es decir, selecciona la distribución de probabilidad más desinformada posible.

En otras palabras, el método extrae la información disponible en los datos, pero con la cantidad mínima de asunciones a priori.

Formalmente, dados  $m$  momentos observados el problema ME puede ser escrito como:

$$\hat{p} = \operatorname{argmax} \left\{ H(p) \equiv - \sum_i p_i \log p_i \right\} \quad s. a.,$$

$$\sum_i p_i = 1$$

$$\sum_i h_j(X_i) p_i = b_j \quad j = 1 \dots m$$

Cuando hay información disponible *a priori*, por ejemplo una distribución de probabilidad  $q$ , se puede usar la medida de entropía relativa:

$$\tilde{p} = \operatorname{argmin} \left\{ D(p \parallel q) \equiv \sum_i p_i \log \frac{p_i}{q_i} \right\} \quad s. a.,$$

$$\sum_i p_i = 1$$

$$\sum_i h_j(X_i) p_i = b_j \quad j = 1 \dots m$$

Una aplicación sencilla de este método puede hacerse con el siguiente muy citado ejemplo del dado de 6 caras:

<sup>15</sup> Para más información ver Cover y Thomas (1991)

Un investigador conoce la media empírica (primer momento) luego del lanzamiento de un dado ordinario de 6 caras  $n$  veces ( $n$  grande). El investigador quiere saber, cuál es la probabilidad de que en el próximo lanzamiento salgan los números 1, 2, 3, 4, 5 ó 6.

Naturalmente, que el investigador sabe que la suma de dichas probabilidades es 1.

Consecuentemente, hay 6 valores desconocidos a predecir y solo 2 observados: La media empírica y la ley de cierre. Consecuentemente, hay infinitas distribuciones de probabilidad que están bien definidas y que satisfacen la media observada. Se trata, por tanto, de un problema indeterminado (*ill-posed problem*). El método de estimación por ME seleccionará aquella distribución que sea la menos informada posible:

$$\hat{p} = \operatorname{argmax} \left\{ H(p) \equiv - \sum_i^6 p_i \log p_i \right\} \text{ s. a.,}$$

$$\sum_i^6 p_i = 1, \sum_i^6 x_i p_i = y$$

donde  $y$  es la media empírica observada por el investigador.

Se han realizado simulaciones propias de este problema<sup>16</sup> para distintos valores de  $y$ . A continuación se muestra una tabla resumiendo los resultados<sup>17</sup>:

**Distribuciones de probabilidad estimadas para distintos valores de  $y$**

$\downarrow$ $y$	$x$ $\rightarrow$	1	2	3	4	5	6
1		1,000	0,000	0,000	0,000	0,000	0,000
1.5		0,664	0,224	0,076	0,026	0,009	0,003
2		0,478	0,255	0,136	0,072	0,039	0,021
2.5		0,348	0,240	0,165	0,114	0,079	0,054
3		0,247	0,207	0,174	0,146	0,123	0,103
3.5		0,167	0,167	0,167	0,167	0,167	0,167
4		0,103	0,123	0,146	0,174	0,207	0,247
4.5		0,054	0,079	0,114	0,165	0,240	0,348
5		0,021	0,039	0,072	0,136	0,255	0,478
5.5		0,003	0,009	0,026	0,076	0,224	0,664
6		0,000	0,000	0,000	0,000	0,000	1,000

Puede observarse en la tabla cómo los valores de las medias más extremas arrojan como estimación distribuciones de probabilidad cada vez más degeneradas, mientras que valores de medias más centradas muestran lógicamente distribuciones más uniformes.

En línea con este experimento, el teorema de concentración de la entropía (solo se enunciará aquí de modo informal), afirma que de todas las distribuciones que satisfacen las restricciones que los datos imponen (los momentos observados) una parte significativa está concentrada lo suficientemente cerca de la distribución de máxima entropía. O bien, que el subconjunto de las distribuciones que cumplen las restricciones que los datos imponen pero que tienen la más baja entropía son aquellas más atípicas.

<sup>16</sup> El código fue escrito en MATLAB. Existen algunos paquetes estadísticos que cuentan con la función de máxima entropía incorporada para ciertas aplicaciones específicas.

<sup>17</sup> Si bien no es necesario, imponer que las probabilidades deben ser mayores o iguales que cero ayudan a los algoritmos de optimización a encontrar más rápido y eficientemente la solución.

## Máxima entropía generalizada

En el método ME, los momentos que entran como restricción al problema deben respetarse con exactitud. Esto es deseable cuando la muestra es muy grande o está bien comportada. Sin embargo, en el ámbito de las ciencias sociales, las muestras son por lo general pequeñas o no se comportan de forma satisfactoria.

Por este motivo, el método de máxima entropía generalizada (GME) propuesto por Golan y Judge (1996) extiende el método de ME permitiendo la incorporación de restricciones estocásticas al problema. Consecuentemente se provee un marco en el cual las condiciones de momentos pueden cumplirse de modo aproximado (restricciones de momentos estocásticos). Además, este método permite la incorporación con facilidad de restricciones que provengan de la teoría económica adicionales a los momentos observados.

Los momentos estocásticos pueden introducirse de dos formas. Una posibilidad es incorporar a los errores (o ruido) con media cero en cada una de las ecuaciones de momentos. La otra, es tratar a cada una de las observaciones como un *composite* de señal y ruido.

A continuación se expondrá la segunda alternativa, mostrando como el método GME puede utilizarse para estimar un modelo de regresión lineal<sup>18</sup>.

Expresando el modelo de regresión lineal en forma matricial se tiene:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

Si se supone que hay  $T$  observaciones y  $K$  regresores entonces  $\mathbf{y}$  y  $\boldsymbol{\varepsilon}$  son vectores  $T$ -dimensionales, y  $\mathbf{X}$  es una matriz de dimensiones  $T \times K$ , mientras que  $\boldsymbol{\beta}$  es un vector  $K$ -dimensional. Se asume que el conjunto de parámetros  $\boldsymbol{\beta} \in B$  con  $B$  convexo.

A diferencia de mínimos cuadrados clásicos, en lugar de buscar estimaciones puntuales para  $\boldsymbol{\beta}$  cada  $\beta_k$  será visto como el valor medio de una variable aleatoria  $z$ . Asimismo, los errores no observados  $\varepsilon_t$  serán construidos como el valor medio de una variable aleatoria  $v$ .

Sin pérdida de generalidad, cada  $\beta_k$  estará acotado inferior y superiormente por  $\underline{z}_k$  y  $\overline{z}_k$  respectivamente. De esta forma, el conjunto  $B$  puede definirse como:

$$B = \left\{ \boldsymbol{\beta} \in R^k : \beta_k \in (\underline{z}_k, \overline{z}_k), k = 1, 2 \dots K \right\}$$

Dado  $z_k$  un vector  $M$ -dimensional  $z_k = (\underline{z}_k, \dots, \overline{z}_k)' = (z_{k1}, \dots, z_{km})' \quad \forall k = 1, 2 \dots K$ , y dada una distribución de probabilidad  $p_k$  definida sobre  $z_k$ , entonces, por construcción, se tiene que:

$$\beta_k = \sum_m p_{km} z_{km} \equiv E_{p_k}(z_k) \quad \forall k = 1, 2 \dots K$$

Cabe destacar que  $\boldsymbol{\beta}$  puede ser construido a partir de la media o de cualquier otra medida resumen de interés.

Análogamente, se asume que  $\boldsymbol{\varepsilon} \in V$  con  $V$  convexo y simétrico alrededor de cero. Por lo tanto, cada término de error puede ser redefinido como:

<sup>18</sup> El siguiente ejemplo sigue principalmente a Golan (2008)

$$\varepsilon_t = \sum_j w_{tj} v_j \equiv E_{w_t}(v_j) \quad \forall t = 1, 2, \dots, T$$

Es decir, los errores serán vistos como medias de extracciones aleatorias de una distribución de probabilidad con ponderaciones  $\{w_{tj}\}$ .

Tanto los espacios de soporte  $B$  como  $V$  son conjuntos discretos y acotados (aunque con algunas modificaciones puede trabajarse con espacios de soporte continuos).

Habiendo redefinido  $\beta$  y  $\varepsilon$ , el modelo lineal queda de la siguiente forma:

$$y_t = \sum_{k=1}^K \sum_{m=1}^M z_{km} p_{km} x_{tk} + \sum_j v_j w_{tj}$$

El objetivo es entonces estimar  $p_{km}$  y  $w_{tj}$  para luego poder construir los valores de  $\beta$  y  $\varepsilon$ . El método de estimación por GME entonces será:

$$\{\hat{p}, \hat{w}\} = \underset{p, w}{\operatorname{argmax}} \left\{ \begin{array}{l} \{H(p) + H(w)\} \\ - \sum_k \sum_m p_{km} \log p_{km} - \sum_t \sum_j w_{tj} \log w_{tj} \end{array} \right\} \quad \text{s. a.,}$$

$$y_t = \sum_{k=1}^K \sum_{m=1}^M z_{km} p_{km} x_{tk} + \sum_j v_j w_{tj}$$

$$\sum_m p_{km} = 1, \quad \sum_j w_{tj} = 1$$

Se trata de una función de pérdida que otorga mismos pesos a la predicción y a la precisión.<sup>19</sup>

Una vez estimado el vector  $\{\hat{p}, \hat{w}\}$ , se puede construir la estimación de los parámetros y el error:

$$\hat{\beta}_k = \sum_m \hat{p}_{km} z_{km}, \quad \hat{\varepsilon}_t = \sum_j \hat{w}_{tj} v_j$$

Lo interesante de este planteo, es que puede incorporarse fácilmente cualquier tipo de restricción de igualdad, desigualdad e incluso restricciones no lineales, como asunciones de distribuciones *a priori*, restricciones impuestas por la teoría económica (se hará esto más adelante), o algún tipo de presunción acerca de la matriz de varianzas y covarianzas de los estimadores.

Como ejemplo, se muestra a continuación el problema reformulado suponiendo que los errores tienen autocorrelación de primer orden:

<sup>19</sup> La función objetivo otorga el mismo peso a las entropías de  $p$  y  $w$ . Esto puede ser modificado si se dispone de información a priori que lo sugiera. Ver Golan, Judge, y Perloff (1996) para una discusión al respecto.

$$\{\hat{p}, \hat{w}\} = \underset{p, w}{\operatorname{argmax}} \left\{ - \sum_k \sum_m p_{km} \log p_{km} - \sum_t \sum_j w_{tj} \log w_{tj} \right\} \quad \text{s. a.,}$$

$$y_t = \sum_{k=1}^K \sum_{m=1}^M z_{km} p_{km} x_{tk} + \varepsilon_t$$

$$\varepsilon_t = \sum_j w_{tj} v_j \quad \text{para } t = 1$$

$$\varepsilon_t = \rho \varepsilon_{t-1} + \sum_j w_{tj} v_j \quad \text{para } t = 2, 3, \dots, T$$

$$\sum_m p_{km} = 1, \quad \sum_j w_{tj} = 1$$

siendo  $\rho$  el coeficiente de correlación de primer orden.

Lo interesante de este caso, es que no se fuerza al modelo a tener correlación. El método de estimación no generará un sesgo por asumir correlación *a priori*. Simplemente, seleccionará aquella estimación que se adapte a ello y que sea la menos informada posible.

En la siguiente sección, se mostrará como este método puede ser aplicado para estimar las estrategias de un juego y poder computar de forma directa el poder de mercado.

## V. ESTIMACIÓN DEL PODER DE MERCADO

### Definición del juego

Se describirá un juego (con algunas variantes) como aquel propuesto por Golan, Karp y Perloff (1998), en el cual dos firmas están estratégicamente interrelacionadas en un juego no cooperativo y estático de jugada simultánea que se repite en cada período, con un espacio de acciones discreto en donde la variable de decisión es el precio y las funciones de pago son el beneficio de cada una de las firmas.

Asimismo, se asume, sin pérdida de generalidad, que ambas firmas comparten el mismo espacio de acciones.

Más formalmente, un conjunto de jugadores  $I = \{i, j\}$  comparten un espacio de  $n$  acciones  $x = \{x_1, x_2, \dots, x_n\}$ , en donde la acción elegida por la firma  $i$  se denotará  $x^i_s$  (análogamente para la firma  $j$ ).

De igual modo, existe información privada en forma de una variable aleatoria  $\varepsilon_k^i$  que es observada por cada firma (y no por el economista).

Los posibles  $K$  valores que puede tomar esta variable aleatoria son iguales para ambas firmas (nuevamente sin pérdida de generalidad dado que las distribuciones de probabilidad si pueden ser diferentes para cada firma) y está dada por  $\varepsilon = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_K\}$ . Ambas firmas conocen las distribuciones de probabilidad de dicha variable, aunque no el economista.

En el estado  $k$ , la estrategia mixta (una distribución de probabilidad sobre el espacio de acciones, o un elemento del simplex  $\Delta(x)$ ) de la firma  $i$  es<sup>20</sup>  $\alpha_k^i = (\alpha_{k1}^i, \alpha_{k2}^i, \dots, \alpha_{kn}^i)$ , donde  $\alpha_{ks}^i$  es la probabilidad de que la firma  $i$  elija la acción  $x^i_s$  en el estado  $k$ . En caso que la firma  $i$

<sup>20</sup> En realidad, se trata de una estrategia condicional en  $k$ . La definición de estrategia implica un plan contingente completo, por eso el hecho de definir bajo el término "estrategia" un plan de acción frente a un particular estado de la naturaleza es un abuso consciente de la terminología.

elija una estrategia pura, entonces deberá cumplirse que para algún elemento  $\alpha_{kg}^i = 1$  con  $\alpha_{kg}^i \in \alpha_k^i$ , y que  $\alpha_{kf}^i = 0 \forall \alpha_{kf}^i \in \alpha_k^i \neq \alpha_{kg}^i$ .

El beneficio o función de pago de la firma  $i$  en el estado  $k$  será  $\pi_{rsk}^i = \pi^i(x_s^i, x_r^j, \varepsilon_k^i)$  donde  $r$  indexa la acción elegida por la firma  $j$ , y  $s$  indexa la acción elegida por la firma  $i$ .

Cada jugador elige  $\alpha_k^i$  para maximizar  $\sum_r \alpha_r^{-i} \pi_{rsk}^i(x_s^i, x_r^{-i}, \varepsilon_k^i)$  donde las expectativas están tomadas en base a las acciones rivales.

Se asume también, que las creencias de cada firma en relación a las probabilidades no condicionales  $\alpha^i = E_k \alpha_{ks}^i$  son correctas en el equilibrio, donde  $E_k$  es la esperanza matemática sobre la información privada  $\varepsilon$ .

Por lo tanto, el juego puede describirse con la cuaterna  $\Gamma_N = \{I, \Delta(x), \{\pi^i\}_{i \in I}, \varepsilon\}$

Una vez definidas las acciones en cada etapa del juego, los movimientos de la naturaleza (i.e.  $\varepsilon_k$ ), y las funciones de pago, es posible explicitar el equilibrio de Nash del juego.

**Definición** Un perfil de estrategias  $\alpha_k = (\alpha_k^i, \alpha_k^{-i})$  constituye un equilibrio de Nash del juego

$\Gamma_{Nk} = \{I, \Delta(x), \{\pi^i\}_{i \in I}, \varepsilon_k\}$  en estrategias mixtas, si  $\forall i \in I$  se cumple que:

$$\sum_r \alpha_r^{-i} \pi_{rsk}^i(x_s^i, x_r^{-i}, \varepsilon_k^i) \geq \sum_r \alpha_r^{-i} \pi_{rsk}^i(x_{s'}^i, x_r^{-i}, \varepsilon_k^i)$$

para todo  $x_{s'}^i \in x$ .

La condición de optimalidad recientemente descrita puede ser reescrita como sigue:

Si se define a  $Y_k^i$  como el beneficio esperado máximo para la firma  $i$  dado  $\varepsilon_k^i$ , entonces, la pérdida esperada de la firma  $i$  dado  $\varepsilon_k^i$  al elegir la acción  $x_s^i$  es:

$$L_{sk}^i \equiv \sum_r \alpha_r^{-i} \pi_{rsk}^i(x_s^i, x_r^{-i}, \varepsilon_k^i) - Y_k^i \leq 0 \quad (1)$$

Por definición de  $Y_k^i$  se desprende naturalmente que  $L_{sk}^i \leq 0$ .

Si además, es óptimo para la firma  $i$  elegir la estrategia  $\alpha_k^i$ , entonces debe cumplirse que:

$$L_{sk}^i \alpha_{sk}^i = 0 \quad (2)$$

Esta última ecuación sólo refleja que hay una probabilidad positiva de que la firma  $i$  elija la acción  $x_s^i$  solamente si los beneficios esperados de usar dicha acción igualan al beneficio esperado máximo  $Y_k^i$ .

### Adecuación económica

El objetivo es poder estimar con series de tiempo de precios, cantidades, y otras variables de la demanda y el costo, las estrategias de las firmas en el juego recién especificado.

Para ello se debe notar, en primer lugar, que las ecuaciones (1) y (2) involucran a la variable no observada  $\varepsilon_k$ , por lo que las ecuaciones no pueden ser usadas como restricciones de forma directa y deben ser redefinidas. Si se define a  $Y^i \equiv E_k Y_k^i$  y  $\pi_{rs}^i \equiv E_k \pi_{rsk}^i$  y se toma la esperanza matemática sobre las ecuaciones (1) y (2), se tiene:



$$\sum_r \alpha_r^{-i} \pi_{rs}^i(x_s^i, x_r^{-i}, \varepsilon_k^i) - Y^i \leq 0 \quad (3)$$

$$\left( \sum_r \alpha_r^{-i} \pi_{rs}^i(x_s^i, x_r^{-i}, \varepsilon_k^i) - Y^i \right) \alpha_s^i + \delta_s^i = 0 \quad (4)$$

donde<sup>21</sup>  $\delta_s^i = cov(L_{sk}^i, \alpha_{sk}^i) \geq 0$ .

Si los errores  $\varepsilon_k$  estuvieran correlacionados para algunas de las firmas, o bien la información no fuera privada, las ecuaciones (3) y (4) deberían ser modificadas.

En un contexto de racionalidad acotada, y la existencia de potenciales errores de medición en las variables utilizadas como insumo, la ecuación (4) debe ser tratada como estocástica incorporando a la misma un error aditivo de media cero. Sin embargo, dicha ecuación ya dispone de un término aditivo  $\delta$  por lo que no sería posible distinguirlo del error sugerido.

Por lo tanto,  $\delta$  será el único término de error que incluirá la ecuación, siendo en todo caso un *composite* de ambos términos.

Una vez definidas las restricciones teóricas, lo que resta para poder realizar la estimación GME es establecer las restricciones que surgen de los momentos observados.

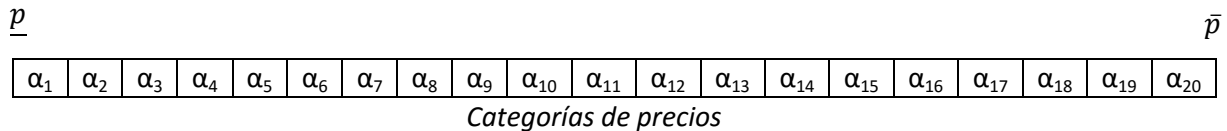
Para ello, queda pendiente determinar cuál será el espacio de acciones relevantes.

Siendo que como se estableció previamente ambos jugadores tendrán el mismo espacio de acciones, el criterio será el siguiente:

De las series de precios observadas para ambas firmas se buscará el valor mínimo, y el valor máximo y se los reducirá y expandirá respectivamente en un 10%.

Luego, con estos valores como cota inferior y superior ( $\underline{p}$  y  $\bar{p}$ ) se dividirá el espacio de acciones en veinte intervalos iguales. El objetivo es entonces computar para cada jugador el vector de probabilidades  $\alpha$  que cada jugador asigne al espacio de acciones.

Gráficamente<sup>22</sup>:



El número de intervalos elegido es lo suficientemente grande como para evitar cualquier potencial pérdida de información en la estimación. En cambio, los resultados de la estimación son poco sensibles al incorporar una mayor cantidad de intervalos de lo necesario, por lo que no hay peligro allí.

De este modo, queda definido el espacio de acciones (discreto) para cada jugador.

Luego, los precios observados en cada período para cada firma pueden ser posicionados en alguna de las categorías establecidas.

Sea  $n_s^i$  la cantidad de veces que la acción  $x_s^i$  es observada en las T observaciones. De esta forma, la frecuencia relativa observada será  $n_s^* \equiv \frac{n_s}{T}$  (por simplicidad notacional se suprime el superíndice de la firma). Luego, para cada firma, se iguala la frecuencia observada a la probabilidad correspondiente del vector de estrategia  $\alpha$  más un término de error:

<sup>21</sup>  $E(XY) = E(X)E(Y) + COV(X, Y)$

<sup>22</sup> En Golan, Karp y Perloff (1998) se estima un juego en donde las variables de decisión son 2 ( precio y publicidad), por lo que el diagrama puede ser expuesto a modo de grilla bidimensional

$$n_s^* \equiv \frac{n_s}{T} = \alpha_s + e_s \quad (5)$$

donde el término de error  $e_s \in [-1, 1]$ .

Se ha obtenido así, las restricciones estocásticas de momentos provenientes de la muestra necesarias para incluir la información muestral en la estimación por GME.

En efecto, en caso de no incluir el término de error  $e_s$  la restricción se cumpliría con exactitud y se estaría aplicando el método de ME. De todos modos, con  $e_s = 0$  la solución al problema sería trivial.

Debe remarcar también, que estimar  $\alpha_s$  es estimar la probabilidad de que una acción sea (o no) observada. Es decir, se trata de un modelo de variable dependiente binaria. Realizar esta estimación por métodos tradicionales equivaldría a hacer una estimación de un probit o logit multinomial por máxima verosimilitud. Como se mencionó previamente, GME no requiere hacer asunción distribucional alguna, y permite la incorporación de restricciones con facilidad.

### Problema de estimación

Habiendo obtenido las restricciones, es posible maximizar la entropía sujeta a las mismas.

Es factible realizar dos ejercicios de estimación: uno, en el cual se incorpore únicamente la información de la muestra (*Soft Data*) en donde solamente se incorpora como restricción a la ecuación (5); o bien, un segundo ejercicio en el cual se incluye también, las restricciones teóricas (3) y (4) basadas en la noción de equilibrio de Nash (*Hard Data*).

En ambos casos, la función objetivo para hacer la estimación por GME, será la suma ponderada de las entropías de  $\alpha$  y  $e$ .

Al incrementar el peso de  $e$  se aumenta la precisión de la estimación (se reduce el error cuadrático medio de las estimaciones de  $\alpha$ ), mientras que incrementar el peso de la entropía de  $\alpha$ , permite mejorar la predicción, dado que se hace una mejor asignación de las observaciones en su respectiva categoría de precio.

Por simplicidad, y al carecer de una teoría que indique cuáles deben ser los pesos a asignar, se establecerán los mismos pesos para ambas entropías.

Debe recordarse, que los argumentos de las medidas de entropía deben ser probabilidades.

Por definición, las  $\alpha$  son probabilidades, sin embargo,  $e \in [-1, 1]$ . Por lo tanto, para poder construir una entropía para los errores, el vector  $e$  debe ser reparametrizado.

Al igual que como se hizo con los  $\beta$  en la sección IV, se define un espacio de soporte para  $e$ .

En este caso, las cotas del espacio de soporte vienen dados por los límites naturales de  $e$ .

El espacio de soporte será entonces un conjunto  $v = [v_1, v_2, \dots, v_M]$  de dimensión  $M \geq 2$ , y simétrico alrededor de 0 dentro del intervalo  $[-1, 1]$ .

Cada  $e_s^i \forall i \in I$  tiene su vector de probabilidades asociado  $w_s^i = [w_{s1}^i, w_{s2}^i, \dots, w_{sM}^i]$  con  $0 \leq w_{sm}^i \leq 1$  y  $\sum_m w_{sm}^i = 1$ . De esta forma, cada elemento de error es reparametrizado como:

$$e_s = \sum_m v_m w_{sm}$$

De este modo, el espacio de errores se reconvirtió de  $[-1, 1]$  a  $[0, 1]$ . La ecuación (5), puede ser reescrita como:

$$n_s^* \equiv \frac{n_s}{T} = \alpha_s + \sum_m v_m w_{sm} \quad (6)$$

La estimación así planteada permite recuperar M-1 momentos de la distribución desconocida de los errores. Cuanto mayor sea M, más momentos de dicha distribución pueden

ser estimados. Golan, Judge y Miller (1996) realizan experimentos Monte-Carlo en donde verifican que espacios de  $M > 3$  reducen de forma muy escasa el error cuadrático medio de la estimación, mientras que las estimaciones difícilmente cambian para  $M > 7$ .

### Estimación con datos de la muestra (Enfoque Soft Data)

El problema GME para cada firma será:

$$\{\hat{\alpha}, \hat{w}\} = \underset{\alpha, w}{\operatorname{argmax}} \left\{ \begin{aligned} & \{H(\alpha) + H(w)\} \\ & \equiv - \sum_s \alpha_s \log \alpha_s - \sum_s \sum_m w_{sm} \log w_{sm} \end{aligned} \right\} \quad \text{s. a.,}$$

$$n_s^* = \alpha_s + \sum_m v_m w_{sm}$$

$$\sum_s \alpha_s = 1$$

$$\sum_m w_{sm} = 1 \quad s = 1, \dots, n$$

El Lagrangiano del problema es:

$$\begin{aligned} \mathcal{L}(\lambda, \rho, \eta) = & - \sum_s \alpha_s \log \alpha_s - \sum_s \sum_m w_{sm} \log w_{sm} + \sum_s \left[ \lambda_s \left( n_s^* - \alpha_s - \sum_m v_m w_{sm} \right) \right] \\ & + \rho \left( 1 - \sum_s \alpha_s \right) + \sum_s \left[ \eta_s \left( 1 - \sum_m w_{sm} \right) \right] \end{aligned}$$

Una vez obtenidas las estimaciones, puede recuperarse el término de error:

$$\hat{e}_s = \sum_m v_m \hat{w}_{sm}$$

El Hessiano de este problema es definido negativo, por lo que la solución es globalmente única.

### Estimación con restricciones teóricas (Enfoque Hard Data)

Si a la optimización anterior se incorporan las restricciones (3) y (4) se pueden estimar las estrategias del juego incluyendo además de la información de la muestra la noción de equilibrio de Nash propuesta.

Debe tenerse en cuenta 2 cuestiones:

En primer lugar, las ecuaciones (3) y (4) involucran parámetros a estimar de las diferentes firmas involucradas<sup>23</sup>. Consecuentemente, la optimización debe realizarse de forma conjunta involucrando las entropías y restricciones de ambas firmas.

<sup>23</sup> El hecho de que esto sea así es el correlato matemático lógico del comportamiento estratégico interdependiente que sugiere la teoría de juegos, y en particular la noción de equilibrio propuesta.

En segundo lugar, es necesario reparametrizar  $\delta_s$ , dado que al igual que en el caso de  $e_s$ , debe ser estimado, pero su rango natural no está en el intervalo  $[0,1]$ , condición necesaria para ser incluida en la función de entropía.

El espacio de soporte será, entonces, un conjunto  $v^\delta = [v_1^\delta, v_2^\delta, \dots, v_H^\delta]$  de dimensión  $H \geq 2$ , y simétrico alrededor de 0 con su correspondiente distribución de probabilidad desconocida  $\omega^\delta$  de tal modo que:

$$\sum_h \omega_{sh}^\delta = 1$$

$$\delta_s = \sum_h v_h^\delta \omega_{sh}^\delta$$

A diferencia de  $e_s$  en este caso, no hay una cota inferior y superior. Para establecer estos límites se sigue "la regla de los tres sigma" o *Three-sigma Rule* (Pukelsheim, 1994), es decir, se toma como cota inferior y superior el desvío estándar empírico de los precios observados para cada una de las firmas multiplicado por tres. Por ejemplo, para  $H=3$ , si el desvío de los precios observado para la firma  $i$  fuera 26.7, entonces  $v_i^\delta = [-80.1, 0, 80.1]$ . El desvío estándar empírico se calcula computando para cada observación el valor medio del intervalo que le corresponde al precio observado dentro del espacio de acciones definido previamente.

Ahora la ecuación (4) puede ser reescrita como:

$$\left( \sum_r \alpha_r^{-i} \pi_{rs}^i(x_s^i, x_r^{-i}, \varepsilon_k^i) - Y^i \right) \alpha_s^i + \sum_h v_h^{i\delta} \omega_{sh}^{i\delta} = 0 \quad (7)$$

Para simplificar la notación se hace un *stack* de los parámetros a estimar de cada firma:

$$\alpha = \begin{pmatrix} \alpha^i \\ \alpha^j \end{pmatrix}, \quad w = \begin{pmatrix} w^i \\ w^j \end{pmatrix}, \quad \omega = \begin{pmatrix} \omega^{i\delta} \\ \omega^{j\delta} \end{pmatrix}$$

Luego, el estimador GME-NASH puede obtenerse resolviendo el siguiente problema:

$$\{\hat{\alpha}, \hat{w}\} = \underset{\alpha, w}{\operatorname{argmax}} \left\{ \begin{array}{l} \{H(\alpha) + H(w) + H(\omega)\} \\ \equiv -\alpha' \log \alpha - w' \log w - \omega' \log \omega \end{array} \right\}$$

sujeto, para  $i=1,2$  a la restricción de momentos (6), y a las restricciones teóricas (3) y (7) :

$$n_s^{i*} = \alpha_s^i + \sum_m v_m^i w_{sm}^i$$

$$\sum_r \alpha_r^{-i} \pi_{rs}^i(x_s^i, x_r^{-i}, \varepsilon_k^i) - Y^i \leq 0$$

$$\left( \sum_r \alpha_r^{-i} \pi_{rs}^i(x_s^i, x_r^{-i}, \varepsilon_k^i) - Y^i \right) \alpha_s^i + \sum_h v_h^{i\delta} \omega_{sh}^{i\delta} = 0$$

y sujeto a las leyes de cierre de cada distribución:

$$1' \alpha = 1, \quad 1' w = 1, \quad 1' \omega = 1$$

En el Anexo de este trabajo se demuestra de forma sencilla cómo todos estos grupos de estimadores son consistentes.

### Función de beneficio desconocida

Si no se conocen las funciones de beneficio de las firmas, las mismas deben ser estimadas.

Para ello, debe estimarse una función de demanda  $q^i = q^i(p^i, p^j, z)$ , y una función de costos  $C^i(q^i, z_c)$ . La especificación para estas funciones debe ser determinada previamente por el investigador<sup>24</sup>. La estimación de los parámetros de dichas funciones puede hacerse por cualquiera de los métodos tradicionales, o bien, reparametrizar todos los parámetros a estimar para que estén dentro del intervalo [0,1]. Luego se pueden incorporar funciones de entropía en la función objetivo para cada uno de ellos, y estimar todos los parámetros del problema de forma simultánea.

### Poder de mercado

Una vez obtenidas las estimaciones para las estrategias de las firmas y también sus costos, puede calcularse el índice esperado de Lerner definido como:

$$L^e = E \left[ \frac{p^i - c^i}{p^i} \right] = \sum_s^n \alpha_s^i \left[ \frac{p_s^i - c^i}{p_s^i} \right]$$

Es importante mencionar que en presencia de información privada, el índice de Lerner recién explicitado puede ser negativo, incluso con beneficios esperados positivos. Para demostrarlo, debe partirse del beneficio esperado de la firma  $i$ ,  $\sum_r \alpha_r^j \pi_{rs}^i = \sum_r \alpha_r^j (p_s^i - c^i) q^i(p^i, p^j)$  donde  $q^i(\cdot)$  es la función de demanda de la firma  $i$ , y por simplicidad se asume una función de costos constantes  $c^i$ .

Si se toma la esperanza matemática del beneficio esperado con respecto a las acciones  $r$  de la firma  $j$ , el beneficio esperado puede ser reescrito como  $(p_s^i - c^i) Q^i(p^i, p^j)$  donde  $Q^i(p^i, p^j) \equiv \sum_r \alpha_r^j q^i(p_s^i, p_r^j)$  es la cantidad demandada esperada por la firma  $i$  al tomar la acción  $s$  (i.e. eligiendo el precio  $p_s^i$ ). Reemplazando esta expresión en la ecuación (4) se tiene:

$$((p_s^i - c^i) Q^i(p^i, p^j) - Y^i) \alpha_s^i + \delta_s^i = 0$$

Dividiendo esta ecuación por  $\alpha_s^i$ , para un  $\alpha_s^i$  estrictamente positivo, se tiene que el índice esperado de Lerner de la firma  $i$ , condicional a haber elegido la acción  $s$  es:

$$L_s^e = \frac{p_s^i - c^i}{p_s^i} = \frac{Y^i - \delta_s^i}{p_s^i Q^i(p^i, p^j)}$$

Debe apreciarse, que, como se mencionó previamente, con información privada y no correlacionada  $\delta_s^i \geq 0$ .  $\delta_s^i \equiv cov(L_{sk}^i, \alpha_{sk}^i) = 0$  únicamente cuando es óptimo elegir la

<sup>24</sup> Notar, que este método es por tanto, vulnerable a las mismas críticas que hacen los defensores de los modelos no paramétricos de variaciones conjeturales a los modelos de tipo estructural.

acción  $s$  con probabilidad positiva para cualquier valor de la realización de  $\varepsilon_k$ . Sin embargo, si para algún valor de  $\varepsilon_k$ ,  $\alpha_{sk}^i = 0$ , entonces,  $\delta_s^i > 0$ .

Si este último fuera el caso, y los beneficios esperados  $Y^i$  fueran cercanos a cero por tratarse de un mercado muy competitivo, entonces el índice de Lerner esperado podría ser negativo.

Por tanto, esta medida puede ser mala para medir el poder de mercado. La razón de ello, es que está siendo alterada por el desconocimiento del econometrista acerca de la información privada  $\varepsilon_k$  que es solo observada por las firmas.

Debe notarse, que desde el punto de vista de la firma, al conocer  $\varepsilon_k$ ,  $\delta_s^i = 0$ , por lo que la firma, puede calcular su propio índice de Lerner.

Para salvar este problema desde el punto de vista del econometrista, puede removerse el término  $\delta_s^i$  de  $L_s^e$ . Tomando expectativas sobre las acciones  $s$  se tiene:

$$L^{*e} = \sum_s \left[ \alpha_s^i L_s^e + \frac{\delta_s^i}{p_s^i Q^i(p^i, p^j)} \right]$$

Luego, este indicador del poder de mercado carente de unidad de medida tendrá el mismo signo que el beneficio esperado de la firma  $i$ .

## VI. APLICACIÓN: MERCADO ARGENTINO DE KEROSENE

### Características generales

En esta sección se aplicará el modelo recién descrito al mercado argentino mayorista de Kerosene. Se estimará un modelo en donde 2 firmas, YPF y AXION, eligen el precio.

En el período a considerar ambas firmas participan en un 93% del total producido, teniendo YPF el 74% y AXION el 19%<sup>25</sup>. Dado que se trata de un mercado mayorista o B2B, los canales de comercialización son variados. Asimismo, ambas firmas operan en todo el país.

Es posible que las firmas no compitan de la misma forma a lo largo del territorio y canales, pero aquí se asumirá por obvia necesidad de simplicidad y parsimonia que si lo hacen.

El producto no tiene un grado de diversificación significativo por lo que la hipótesis de bien homogéneo es bastante realista.

El ingreso y salida de firmas se ha mantenido constante a lo largo del tiempo, y no afecta de forma significativa los shares de mercado de las compañías. Las importaciones podrían eventualmente jugar un rol relevante, aunque es una porción no tan grande del total de Kerosene comercializado.

Por otro lado, la hipótesis de elección simultánea podría también ponerse en tela de juicio, y proponer en cambio un esquema de tipo líder-seguidor. En próximas investigaciones puede también ponerse a prueba esta hipótesis.

Por último, el mercado hidrocarbúrico argentino no ha estado exento de regulaciones en los últimos años. En opinión del autor esto no anula el comportamiento estratégico de las firmas, sino que simplemente añade una restricción adicional sobre el espacio de acciones

<sup>25</sup> Oum, Zhang, y Zhang (1993) optan por remover de la muestra aquellos períodos temporales donde se verifica que una de las firmas tiene una cuota de mercado superior al 90%, por considerar que en dichos períodos el comportamiento es de carácter monopólico. Si bien en la presente muestra hay períodos en donde dicha situación se presenta, el autor no considera oportuno hacer dicha omisión de información en pos de otra hipótesis comportamental, además de que dicha omisión implica la asunción implícita de la necesidad de modelar un juego de tipo dinámico. Aun así, dicha remoción puede intentarse en futuros trabajos.

disponibles en cada período. La imposición de regulaciones ya sean *de jury* o *de facto* no son más que restricciones adicionales que pueden adicionarse a los problemas de optimización explicitados en secciones anteriores. Dado que como se desarrolló previamente, las restricciones son tratadas como estocásticas, estos “ruidos” sumado al hecho de que las firmas toman decisiones en un marco de racionalidad acotada, son parcialmente consideradas.

El set de datos utilizado se compone de variables de precios y cantidades y otros insumos para estimar la demanda y el costo para las firmas involucradas. Los datos fueron obtenidos de la Secretaría de Energía de la República Argentina. Son de carácter mensual totalizando 109 observaciones comprendidas en el período (2005:10; 2015:5). Para la estimación, se consideraron los precios con impuestos, y los mismos fueron deflactados por el índice general de precios al consumidor<sup>26</sup> publicado por el Instituto de Estadísticas de la Provincia de San Luis. Los precios fueron ponderados adecuadamente por su importancia de comercialización en las distintas regiones y en los distintos canales de venta.

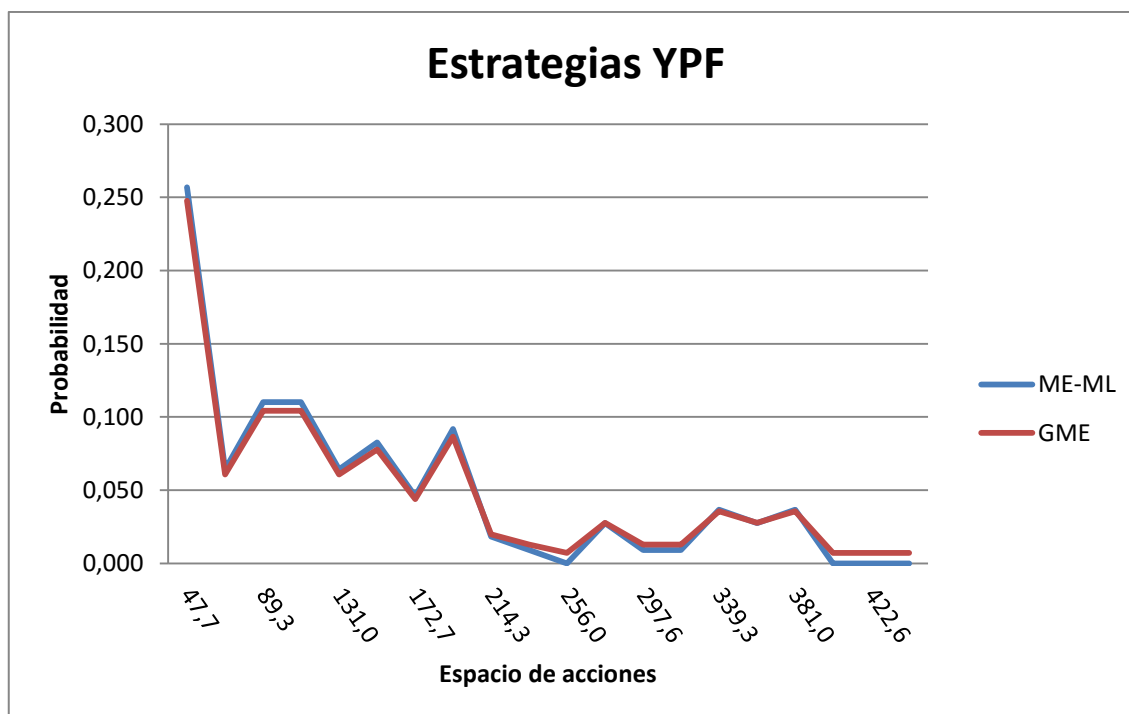
## Estimaciones

Las estimaciones se realizaron en MATLAB y se han establecido espacios de soporte para los errores de tamaño 7. Se ha comprobado en todas las estimaciones las condiciones de segundo orden encontrando Hessianos definidos negativos en los distintos ejercicios de maximización realizados, garantizando de este modo que se trata de máximos globales.

Debe recordarse además, que el espacio de acciones fue dividido en veinte intervalos.

### Enfoque Soft Data

Se comparan los resultados de las estimaciones para ambas firmas con restricciones estocásticas (GME), y determinísticas o frecuencias relativas observadas (ME-ML):



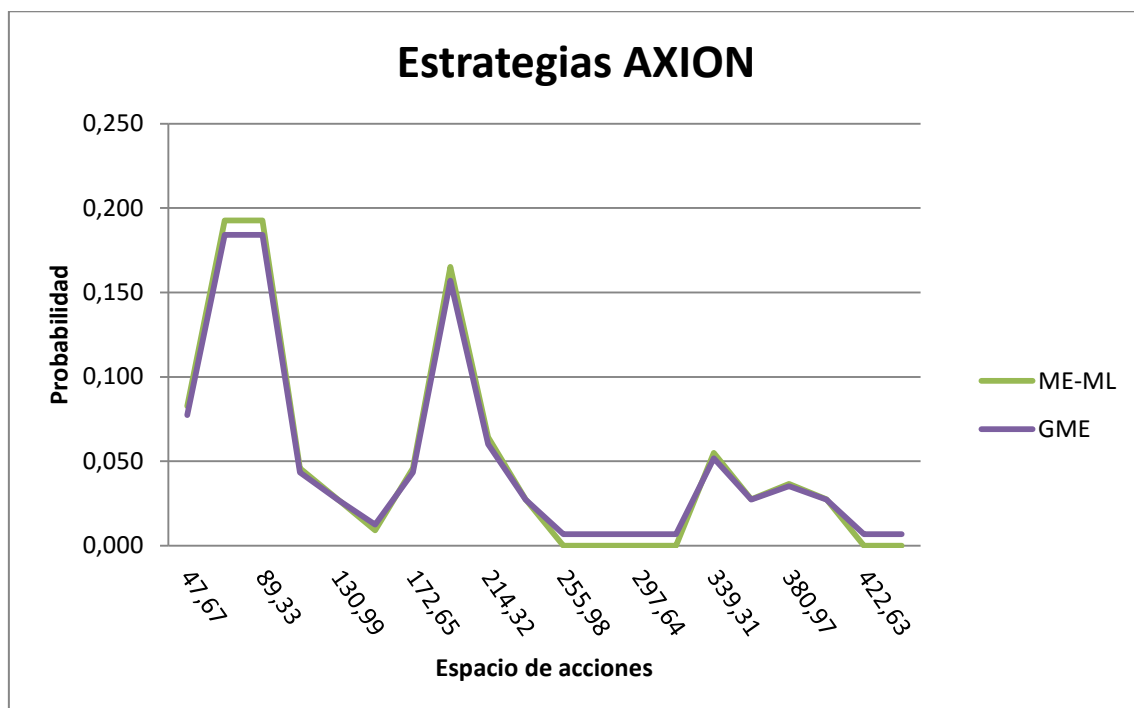
<sup>26</sup> Si bien se está analizando un mercado mayorista, se utiliza un índice de precios minorista por no contar con un índice de precios mayorista fiable. Más allá de eso, se verificó que en el mediano plazo los índices de precios minoristas y mayoristas convergen.

En el caso de YPF, ambas distribuciones resultan muy parecidas (La correlación supera el 99%), la estimación por GME muestra más uniformidad (incertidumbre). Prueba de ello es que la medida de entropía normalizada  $S(\alpha)$  es 0.8 para la distribución ME-ML, mientras que la estimación por GME alcanza 0.84. Este resultado es lógico dado que la restricción de momentos estocástica permite diferir con más flexibilidad de las frecuencias observadas. Al maximizar la entropía se empuja a la distribución a ser más uniforme.

Puede testearse si ambas distribuciones son o no idénticas, mediante un test de cociente de entropía (*ratio entropy test*)<sup>27</sup>. El mismo arroja un valor de  $87.52 > \chi_{41,0.05}^2$  lo que verifica que ambas distribuciones son estadísticamente similares.

Lo mismo ocurre con las estrategias de AXION: ambas distribuciones ME-ML y GME tienen una correlación superior al 99%. Nuevamente el test de cociente de entropía arroja que las distribuciones son estadísticamente similares.

La medida de entropía normalizada  $S(\alpha)$  es de 0.773 para la distribución ME-ML (frecuencias relativas), y de 0.827 para el caso GME. Se verifica mayor uniformidad en este último caso.



Si bien las distribuciones en el caso AXION tienen apariencia más “picuda”, las medidas de entropía normalizada indican que ambas firmas tienen una distribución similarmente informativa acerca del espacio de acciones sobre la que están construidas.

La interpretación de las distribuciones encontradas es la misma que la de cualquier estrategia mixta en un contexto de optimalidad de Nash: Es la elección randomizada de cada jugador sobre el espacio de acciones que le garantiza el mayor valor esperado bajo la hipótesis de que el resto de los jugadores tienen el mismo comportamiento.

Visto de otro modo, si el juego se repitiera infinitamente siempre bajo las mismas condiciones, los jugadores elegirían una acción específica en una proporción igual a la probabilidad encontrada para cada una de ellas.

<sup>27</sup> Se trata de un test asintótico donde se comparan las entropías en su óptimo. El mismo se calcula como  $2[H^* - H^{**}]$ , y tiene una distribución asintótica chi-cuadrado con la cantidad de restricciones del problema de optimización como grados de libertad. Para más información sobre la derivación de este test, ver Kullback (1957).



De todos modos, la interpretación conductual de porque los jugadores eligen estrategias mixtas (y el modo en que lo hacen), ha sido siempre controvertida y ha recibido distintas explicaciones a lo largo de los años.<sup>28</sup>

### *Enfoque Hard Data*

Para incorporar las restricciones teóricas (3) y (7) es necesario obtener estimaciones para los parámetros de las funciones de demanda y costo, una vez establecidas sus especificaciones respectivas.

En el caso de las funciones de demanda, se ha optado por simplicidad por estimar sus parámetros mediante regresiones lineales simples por mínimos cuadrados.

La especificación fue para cada firma del tipo:

$$q_t^i = a_i + b_i p_t^i + c_i p_t^j + \mu_t$$

Se han realizado transformaciones logarítmicas a las variables para lograr que los signos sean los que la teoría económica indica, significatividad individual y global, y una bondad de ajuste lo suficientemente buena<sup>29</sup>.

Las estimaciones para los coeficientes de demanda de YPF son  $a_1 = 17.92$ ,  $b_1 = -0.92$  y  $c_1 = 0.001$ , mientras que para AXION son de  $a_2 = 12.29$ ,  $b_2 = -0.007$  y  $c_2 = 0.008$ .

Dado que los costos son desconocidos, los mismos son estimados de forma simultánea dentro del marco del problema de optimización como restricciones adicionales. Para ello se asumió rendimientos constantes a escala, y dependiente de los precios del salario real medio<sup>30</sup> y del barril de crudo<sup>31</sup>, siendo estos productos los principales insumos para la producción del Kerosene, además por su puesto del capital, que aquí se considerará fijo. Como los parámetros de la función de costos no están necesariamente dentro del intervalo  $[0,1]$ , los mismos fueron reparametrizados mediante la *Three-Sigma Rule*, y sus distribuciones de probabilidad fueron incluidas en la entropía. Se han impuesto restricciones adicionales de no negatividad para estos parámetros. Los coeficientes que acompañan al salario real resultaron muy cercanos a cero sugiriendo que los salarios tienen poco impacto en el costo medio y marginal del

<sup>28</sup> Ver por ejemplo Aumann (1985)

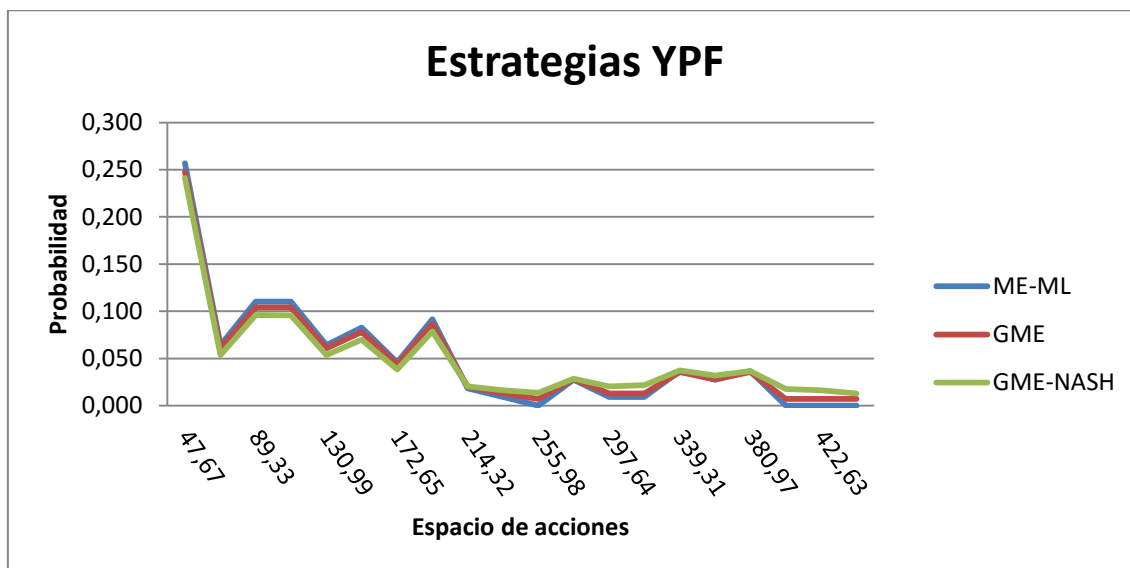
<sup>29</sup> La especificación y la adición de variables adicionales así como las transformaciones logarítmicas modifican los resultados, por lo que estudios en mayor profundidad son necesarios para determinar la mejor especificación para este mercado.

<sup>30</sup> El salario real medio se calculó mediante una serie de salario bruto en pesos argentinos obtenida del Ministerio de Trabajo para la actividad "Fabricación de productos de petróleo". La misma fue deflactada por el índice de precios de la provincia de San Luis que resulta un buen sustituto para los índices de precios publicados por el INDEC y que están sujetos a cuestionamientos. Si bien se trata de índices de carácter minorista, se pudo verificar que los mismos convergen con los mayoristas en un breve período de tiempo.

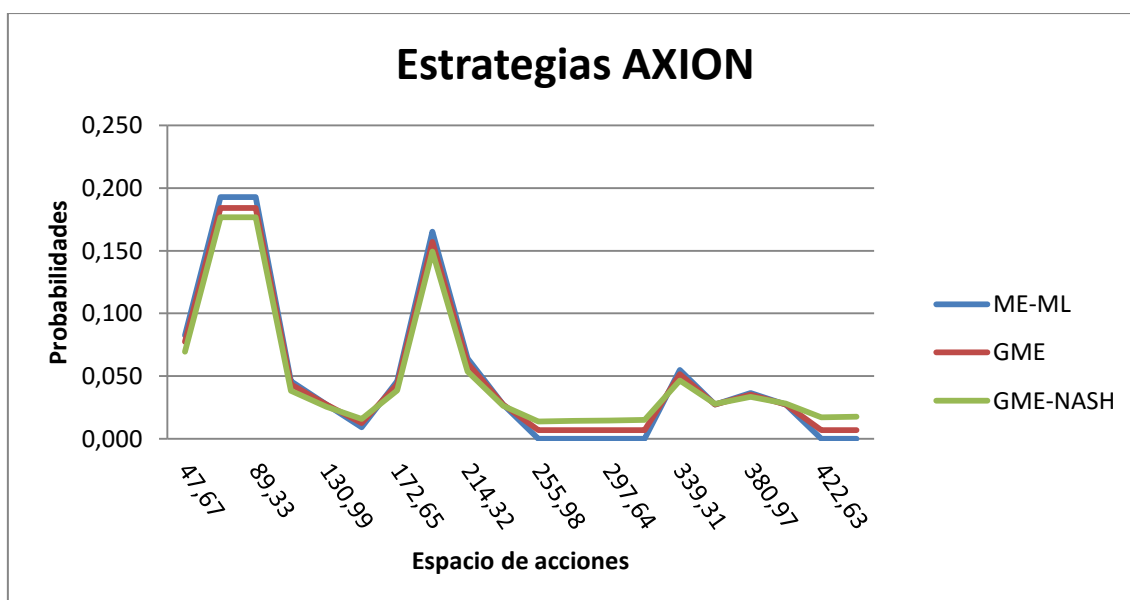
<sup>31</sup> Se han obtenido promedios mensuales de la cotización del barril WTI. Para homogeneizar las unidades de medida fue transformado a litros, y luego convertida a pesos argentinos utilizando una serie de tipo de cambio promedio publicada por el Banco Central de la República Argentina. Finalmente fue deflactada por el índice de precios de la Provincia de San Luis.

Si bien el tipo de cambio considerado no es necesariamente la cotización por la cual las firmas involucradas lo obtienen (El precio de la divisa que se obtiene para importación no es necesariamente la misma que la cotización oficial del Banco Central, aunque es una medida lo suficientemente cercana). Asimismo, el crudo producido localmente cuyo precio de referencia suele ser el de tipo Escalante producido en el Golfo de San Jorge en las Provincias de Chubut y Santa Cruz, suele diferir, del precio del crudo de importación, y además es muchas veces producido por las mismas firmas aquí consignadas, por lo que la integración vertical, hace pensar que pueda sea necesario endogeneizar en el modelo el mercado de *upstream*. Por simplicidad, estos aspectos son aquí dejados de lado.

Kerosene, mientras que los del petróleo crudo son 0.43 para ambas firmas<sup>32</sup>. A diferencia de lo encontrado por Golan, Karp y Perloff (1998, 2000) las restricciones teóricas no modifican de forma significativa las estrategias.



Esto podría sugerir, que asumiendo que las especificaciones de demanda y costos fueran correctas, y que el comportamiento es de elección simultánea, las decisiones no son de carácter estratégico, y las firmas toman sus decisiones sin observar las de sus rivales. La medida de entropía normalizada es para YPF de  $S(\alpha) = 0.878$ , mientras que para AXION es de  $S(\alpha) = 0.864$ , lo cual que refleja que en ambos casos las estimación GME-NASH arroja distribuciones aún más suaves o uniformes que las estimaciones ME-ML y GME.



Finalmente, los índices de Lerner arrojan valores de 0.49 para YPF, y 0.36 para AXION en la estimación por GME-NASH, y con valores similares para ME-ML y GME) sugiriendo que el poder de mercado ejercido por YPF es mayor que el ejercido por AXION. El índice de Lerner de

<sup>32</sup> Para obtener estimaciones de los coeficientes consistentes con la teoría económica fue necesario adicionar restricciones de signo y rango adicionales.

la industria surge de ponderar los índices de cada firma por su cuota de mercado en el período considerado. De este modo, el poder de mercado ejercido en la industria es de 0.463.

Simplemente, a modo de comparación, en el Anexo 2 de este trabajo hay un sumario de estudios que reportan poder de mercado bajo diversas metodologías, compilado por Perloff, Karp y Golan (2007), donde se deduce que el poder de mercado aquí consignado está en valores cercanos a la media de dichos estudios.

## VIII. CONCLUSIONES

En este trabajo se ha mostrado una forma alternativa a las tradicionales para estimar de forma directa las estrategias de las firmas y computar a partir de ellas el poder de mercado ejercido.

El método de máxima entropía generalizada, permite de una forma sencilla incorporar restricciones teóricas y obtener estimaciones consistentes de problemas con mayor cantidad de parámetros a estimar que observaciones. Asimismo, el método evita hacer asunciones distribucionales o de otro tipo, permitiendo obtener de este modo estimaciones lo menos informadas posibles y utilizando únicamente las señales que los datos envían.

Pueden realizarse eventualmente experimentos de sensibilidad al tamaño de muestra, así como también diversos ejercicios de inferencia como *tests* con la hipótesis nula de competencia perfecta ( $P = Cmg$ ), o también la verificación de si las firmas utilizan la misma estrategia, y si las firmas maximizan sus beneficios esperados en cada período sujeto a las restricciones de Nash entre otras.

Golan, Karp y Perloff (2000) estiman las estrategias en el mercado de Colas con aproximadamente el mismo marco teórico aquí presentado, aunque permiten que las estrategias varíen en el tiempo con la demanda y el costo, y permiten que cada firma elija su estrategia de un espacio de acciones con dos variables de decisión: Precio e inversión en publicidad.

Más allá de esta pequeña variante, más estudios son requeridos. En rigor, poder adaptar el marco teórico recientemente desarrollado para estimar juegos dinámicos es un área apasionante y con grandes aplicaciones.

Debe tenerse en cuenta, que pese a las bondades del método, el mismo no deja de ser susceptible a las mismas críticas que se hacen desde los enfoques no paramétricos o de variación conjetural: El método requiere definir una especificación para las funciones de demanda y costos involucradas u otras hipótesis de comportamiento sobre los agentes involucrados.

Si bien, se ha utilizado el método de máxima entropía generalizada para la estimación de las estrategias de un juego, el mismo puede ser utilizado para cualquier tipo de estimación econométrica<sup>33</sup>.

---

<sup>33</sup> Existe incluso toda una literatura acerca de la entropía y su relación con los procesos de Markov, y otras aplicaciones en el análisis de series de tiempo. Ver por ejemplo Golan (2008).

## REFERENCIAS

- Aguirregabiria, V. y P. Mira (2007). Sequential simulation-based estimation of dynamic discrete games. *Econometrica* 75 (1), 1-53.
- Aumann, Robert (1985). What Is Game Theory Trying to Accomplish? *Frontiers of Economics*, editado por K. Arrow and S. Honkapohja, Basil Blackwell, Oxford, 1985, pp. 28-76.
- Bain, Joe S. (1951). Relation of Profit Rate to Industry Concentration: American Manufacturing, 1936–1940. *Quarterly Journal of Economics* 65:293–324.
- Bain, Joe S. (1956). Barriers to New Competition. Cambridge, Mass.: Harvard University Press.
- Bjorn, P. A. y Quang H. Vuong (1985): "Simultaneous Equations Models for Dummy Endogenous Variables: A Game Theoretic Formulation with an Application to Labor Force participation" *California Institute of Technology working paper* 537.
- Bresnahan, Timothy F. (1989): "Studies of Industries with Market Power," En Richard Schmalensee and Robert Willig, eds., *Handbook of Industrial Organization*, New York: North Holland, 1989.
- Bresnahan, Timothy F., y Peter C. Reiss. (1991). "Empirical Models of Discrete Games." *Journal of Econometrics* 48:57–81.
- Burdisso, T., Catena, M., y D'Amato, L. (2001). Bank competition in Argentina: 1997-1999. En *VI Jornadas de Economía Monetaria e Internacional (La Plata, 2001)*.
- Coloma, G. (1998). Análisis del comportamiento del mercado argentino de combustibles líquidos. En *Anales de la XXXIII Reunión Anual de la Asociación Argentina de Economía Política*.
- Coloma, G. (2002). The effect of the Repsol-YPF merger on the Argentine gasoline market. *Review of Industrial Organization*, 21(4), 399-418.
- Cover, Thomas M., y Joy A. Thomas. (1991). *Elements of Information Theory*. (New York: John Wiley & Sons).
- Golan, A. (2008), "Information and Entropy Econometrics — A Review and Synthesis". *Foundations and Trends in Econometrics*, 2, Nos. 1–2, 1–145.
- Golan, A., y Judge, G. (1996). A Maximum Entropy Approach to Empirical Likelihood: Estimation and Inference. University of California ARE Working Paper.
- Golan, Amos, George Judge, y Douglas J. Miller (1996): *Maximum Entropy Econometrics: Robust Estimation with Limited Data*, New York: John Wiley & Sons.
- Golan, Amos, George G. Judge, y Jeffrey M. Perloff. (1996). "Recovering Information from Multinomial Response Data." *Journal of the American Statistical Association* 91:841–53.
- Golan, Amos, Larry S. Karp, y Jeffrey M. Perloff (1998): "Estimating a Mixed Strategy: United and American Airlines". *University of California, Berkeley, Department of Agricultural and Resource Economics Working Paper*.
- Golan, Amos, Larry S. Karp, y Jeffrey M. Perloff (2000): "Estimating Coke and Pepsi's Price and Advertising Strategies." *Journal of Business & Economic Statistics* 18:398–409.
- Golan, Amos, y Jeffrey M. Perloff (2002): "Comparison of maximum entropy and higher-order entropy estimators." *Journal of Econometrics* 107.1 :195-211.
- Hall, Robert E (1988). "The Relationship Between Price and Marginal Cost in U.S. Industry." *Journal of Political Economy* 96:921–47.

- Heckman, James (1978). "Dummy Endogenous Variables in Simultaneous Equation System." *Econometrica* 46:931–60
- Jaynes, E. T. (1957): "Information Theory and Statistical Mechanics," *Physics Review*, 106, 620-630.
- Kooreman, Peter (1994): "Estimation of Econometric Models of Some Discrete Games," *Journal of Applied Econometrics*, 9, 255 - 68.
- Kullback, J. (1959): *Information Theory and Statistics*, New York: John Wiley & Sons.
- Kullback, S., y Leibler, R. A. (1951). On information and sufficiency. *The annals of mathematical statistics*, 22(1), 79-86.
- Lema, D., Barrón, E., Brescia, V., y Gallacher, M. (2003). Organización económica de la empresa agropecuaria: especialización, incentivos y escala en las explotaciones pampeanas. *Revista Argentina de Economía Agraria (Argentina)*.
- Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). "Microeconomic theory "(Vol. 1). New York: Oxford university press.
- Mason, Edward S. (1939). "Price and Production Policies of Large-Scale Enterprise." *American Economic Review* 29 suppl.:61–74.
- Mason, Edward S. (1949). "The Current State of the Monopoly Problem in the United States." *Harvard Law Review* 62:1265–85.
- Oum, Tae Hoon, Anming Zhang, y Yimin Zhang. (1993). "Inter-firm Rivalry and Firm Specific Price Elasticities in Deregulated Airline Markets." *Journal of Transport Economics and Policy* 27:171–92.
- Perloff, Jeffrey M., Larry S. Karp, y Amos Golan (2007). "Estimating market power and strategies." *Cambridge University Press*.
- Pukelsheim, F. (1994): "The Three Sigma Rule," *American Statistician*, 48, 88-91.
- Shannon, C. E. (1948): "A Mathematical Theory of Communication," *Bell System Technical Journal*, 27, 379-423.
- Schmalensee, Richard. (1989). "Inter-Industry Studies of Structure and Performance." In Richard Schmalensee and Robert Willig, eds., *Handbook of Industrial Organization*, pp. 951–1009 (New York: North Holland).
- Tamer, Elie. (2003). "Incomplete Simultaneous Discrete Response Model with Multiple Equilibria." *Review of Economic Studies* 70:147–67.
- Tirole, J. (1988). *The theory of industrial organization*. MIT press.

## ANEXO 1: CONSISTENCIA DE LOS ESTIMADORES GME-NASH

### Supuestos

- 1) Una solución GME-NASH  $\{(\alpha), (w), (\omega)\}$  existe para cualquier tamaño de muestra  $T$ .
- 2) El valor esperado de los términos de error es cero y su varianza finita.
- 3) El valor verdadero de cada parámetro desconocido está en el interior del espacio de soporte definido para el mismo.

### Demostración

- i) Los estimadores ME-ML de  $\alpha$  son las frecuencias observadas. Por lo tanto, por ley de los grandes números, debe ser cierto que con  $T \rightarrow \infty$   $plim(\hat{\alpha}_T) = \alpha$
- ii) El estimador GME se reduce al ME-ML cuando los espacios de soporte para los términos de error son  $-\frac{1}{\sqrt{T}}$  y  $\frac{1}{\sqrt{T}}$  y  $T \rightarrow \infty$ . Esto se puede demostrar a partir de la formulación del problema dual, y de cómo la restricción adicional del problema GME tiende a cero.
- iii) Al incorporar las restricciones (3) y (7), aún existe una solución por el supuesto 1). Luego, por ii), el estimador GME-NASH es consistente.

## ANEXO 2: ESTIMACIONES DE ESTUDIOS PREVIOS DEL PODER DE MERCADO

	Industry	$L = (p - MC)/p$	Market power
Gollop and Roberts (1979)	Coffee roasting (dominant firm)	0.06	
Just and Chern (1980)	Tomato harvesting		Yes
Sumner (1981)	Cigarette	0.50	
Appelbaum (1982)	Textile	0.07	
	Tobacco	0.65	
Lopez (1984)	Canadian food processing	0.50	
Roberts (1984)	Coffee roasting (largest firm)	0.06	
Sullivan (1985)	Cigarettes <sup>a</sup>		
Cottrell (1986)	Vermont food industry		Yes
Suslow (1986)	Aluminum	0.59	
Ashenfelter and Sullivan (1987)	Cigarettes <sup>b</sup>		
Slade (1987)	Retail gasoline	0.10	
Lopez and Dorsainvil (1990)	Haitian coffee		Yes
Baker and Bresnahan (1988)	Breweries		
	Coors 1962–1982	0.745	
	Pabst 1962–1982	0.058	
	Anheuser-Busch 1962–1975	0.312	
	Anheuser-Busch 1975–1982	0.110	
Roberts and Samuelson (1988)	Cigarettes		Yes
Karp and Perloff (1989b)	Rice Export (largest estimate)	0.11	
Azzam and Pagoulatos (1990)	Meat	0.46	
	Livestock	1.1 <sup>c</sup>	
	Composite meat processing	0.74 <sup>d</sup>	
Schroeter and Azzam (1990)	Beef	0.55	
	Pork	0.47	
Buschena and Perloff (1991)	Philippines coconut oil (post 1974)	0.89	
Wann and Sexton (1992)	Grade pack pears	0.15	
	Fruit cocktail	1.41	
Durham and Sexton (1992)	Tomatoes		No
Deodhar and Sheldon (1995)	German bananas	0.26	

<sup>a</sup> Not very competitive, but equivalent to at least 2.5 equal-size Cournot firms

<sup>b</sup> Not highly noncompetitive

<sup>c</sup> Monopsony power

<sup>d</sup> Combined market power in the output and factor markets

Sources: Bresnahan (1989), Carlton and Perloff (2005).