



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN ECONOMÍA
INSTITUTO DE INVESTIGACIONES ECONÓMICAS

MODELACIÓN DE LA PRODUCCIÓN CON INNOVACIONES
TECNOLÓGICAS VERDES

UNA ESPECIFICACIÓN PROBABILÍSTICA CON CÓPULA

TESIS

QUE PARA OPTAR POR EL GRADO DE:

DOCTOR EN ECONOMÍA

PRESENTA:

JOSÉ MANUEL MÁRQUEZ ESTRADA

TUTOR PRINCIPAL: DR. ARMANDO SÁNCHEZ VARGAS

INSTITUTO DE INVESTIGACIONES ECONÓMICAS DE LA UNAM

Ciudad Universitaria, CD.MX.

Noviembre del 2018.



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

INDICE

INTRODUCCIÓN	1
CAPITULO 1: MARCO TEÓRICO: FUNCIÓN DE PRODUCCIÓN, INNOVACIÓN TECNOLÓGICA E INNOVACIONES VERDES	11
1.1 MOTIVACIÓN	11
1.2 LA FUNCIÓN DE PRODUCCIÓN CON INNOVACIONES TECNOLÓGICAS	13
1.4 LA FUNCIÓN DE PRODUCCIÓN CON INNOVACIONES VERDES	17
1.3 EL ENFOQUE ESTÁNDAR PARA LA MODELACIÓN EMPÍRICA DE LA FUNCIÓN DE PRODUCCIÓN	18
1.5 COMENTARIOS GENERALES	20
CAPITULO 2: METODOLOGIA EMPÍRICA, ESPECIFICACIÓN PROBABILÍSTICA Y TEORÍA DE CÓPULAS	23
2.1 MOTIVACIÓN	23
2.2 EL ENFOQUE DE REDUCCIÓN PROBABILÍSTICA	25
2.3 INTRODUCCIÓN A LA TEORÍA DE CÓPULAS	36
2.4 EL ENFOQUE DE REDUCCIÓN PROBABILÍSTICA CON TEORÍA DE CÓPULAS	47
2.6 COMENTARIOS GENERALES	50
CAPITULO 3: UNA PROPUESTA DE MODELACION DE LA FUNCION DE PRODUCCION BASADO EN EL ENFOQUE DE REDUCCION PROBABILISTICA CON CÓPULA	51
3.1 INTRODUCCIÓN	51
3.2 ENFOQUE DE REDUCCIÓN PROBABILÍSTICA DEL MODELO GENERAL	52
3.3 ESPECIFICACIÓN DE UN MODELO DE PRODUCCIÓN CON INNOVACIONES PARETO Y CÓPULA PRODUCTO-CLAYTON	63
3.4 COMENTARIOS GENERALES	75
CAPITULO 4: MODELO EMPÍRICO DE PRODUCCIÓN CON INNOVACIONES TECNOLÓGICAS VERDES	77
4.1 INTRODUCCIÓN	77
4.2 DESCRIPCIÓN DE LOS DATOS	78
4.3 CARACTERÍSTICAS PROBABILÍSTICAS DE LOS DATOS	80
4.5 ESPECIFICACIÓN DEL MODELO	83
4.6 ESTIMACIÓN DEL MODELO Y PRUEBA DE CORRECTA ESPECIFICACIÓN	85
4.4 IDENTIFICACIÓN DEL MODELO DE PRODUCCIÓN CON INNOVACIONES TECNOLÓGICAS VERDES	87
4.8 HALLAZGOS DEL MODELO	89
APÉNDICE A: PRUEBAS DEL MODELO	95

A.1 PRUEBAS DEL MODELO GENERAL	95
A.2 PRUEBAS DEL MODELO DE PRODUCCIÓN CES	97
APÉNDICE B: RUTA DE BÚSQUEDA DE LA BIBLIOGRAFÍA	105
BIBLIOGRAFÍA	109

INTRODUCCIÓN

La modelación de la producción con cambio tecnológico en economía es una herramienta importante pues permite el estudio de las diferencias en productividad y de empleo de recursos en distintas economías, ayuda a explicar cambios y ciclos de la producción y la evolución de la innovación tecnológica, además de que constituye una de las fuentes de estudio del crecimiento económico [Aghion y Howitt (1992, 2009), Barro y Sala-i-Martin (2004)]. La evidencia empírica sugiere que hay un efecto de escala en la producción basada en la investigación y desarrollo de tecnología en economías industrializadas, lo cual explica en gran parte las brechas entre economías industrializadas y las no industrializadas y determina sinergias en el comercio internacional y su dinámica [Romer (1990), Barro y Sala-i-Martin (2004)]. Este tipo de modelos ha sido estudiado desde diferentes perspectivas de la economía como con el modelo insumo-producto, las teorías de la función de producción, las nuevas teorías de producción y la frontera tecnológica. Además, los modelos de producción con innovaciones han sido una herramienta recurrente para modelar a las empresas que utilizan tecnologías verdes, las cuales contribuyen a comprender la dinámica y el impacto de la innovación en la mitigación o reducción de la contaminación o del consumo excesivo de los recursos naturales.

Los primeros modelos que se hicieron de la producción con innovación tecnológica fueron durante los años 1950's. Houthakker (1955) elaboró un modelo para obtener el valor esperado de la producción en una economía suponiendo que las innovaciones se daban de manera independiente y tenían una distribución marginal del tipo Pareto. Habakkuk (1962) desde una perspectiva histórica, plantea a la carencia relativa de factores de capital o de trabajo para la producción como condicionante del tipo de innovaciones que se desarrollaron en EEUU e Inglaterra. Continuando con la línea de investigación de Houthakker, Johansen (1972) y Hildenbrand (1981) hicieron una formalización del modelo de producción a través de análisis convexo y la estadística aplicada, estableciendo los conceptos y resultados básicos de la producción de corto y largo plazo, así como los elementos de la dinámica en la producción con innovación, diferenciando el progreso

tecnológico que aumenta la capacidad de producir del que ahorra insumos y explicando el comportamiento macroeconómico a través del modelo de crecimiento a máxima capacidad y proponiendo estudios de sección cruzada para el análisis de la producción a nivel microeconómico y de series de tiempo a nivel macroeconómico.

Posteriormente surgieron modelos, microeconómicos como en Kortum (1997) y Acemoglu (2003), que combinan el uso de la estadística y métodos de optimización restringida bajo el supuesto de crecimiento a máxima capacidad, para poder incorporar al modelo un conjunto de tecnologías disponibles para estimar a la función de producción, sin embargo estos modelos consideran al cambio tecnológico como variable exógena al modelo, por lo que se tiene en cuenta como un factor del proceso tecnológico con carácter no rival y no como una variable a determinar. Un estudio importante para el análisis de la innovación fue hecho por Caselli y Coleman (2006), pues en este formalizan la idea de frontera de producción, la cual usamos como función de control para resolver el problema de elección de la tecnología, y obtiene un modelo consistente para dicha función basado la tecnología CES.

Pero esta metodología para el análisis de la producción con innovación fue incorporando de manera ex ante a la modelación empírica supuestos de la microeconomía como los rendimientos decrecientes, la homogeneidad de la producción, el impacto de orden lineal de las innovaciones, el Teorema de Uzawa (1961), la racional de los agentes productivos, entre otras, además de supuestos probabilísticos de independencia entre los insumos, por lo que en lugar de proponer ciertas definiciones con base en la evidencia empírica para orientar el análisis, construyen modelos de estimación de la producción probabilísticos basados en las definiciones y supuestos del modelo de producción con innovaciones neoclásico. Así, algunos autores como Lucas (1998, 2009), Jones (2005) y Caselli y Coleman (2006), modelaron la producción con cambio tecnológico tomando las innovaciones como variables aleatorias explicativas y asumiendo funciones de producción de naturaleza determinista de la forma Cobb-Douglas y hacen su análisis asumiendo que se cumplen condiciones de regularidad estadística.

Sin embargo, la evidencia empírica muestra que estos supuestos resultan inadecuados, ya que se obtiene el modelo empírico a partir de estos supuestos económicos sin tomar en cuenta el comportamiento de los datos. Acemoglu (2003) y Caselli y Coleman (2006) mostraron que la función de producción Cobb-Douglas no es adecuada para la modelación de economías como la de los EEUU, dado que no se sostiene el supuesto de que la participación de los factores en la producción es constante a través del tiempo y que tampoco es factible al no sostenerse de manera empírica la hipótesis de que elasticidad de sustitución entre los factores de la producción debe ser igual a uno. Además, como se demuestra en Habakkuk (1962), Caselli y Coleman (2006) y Sánchez y Márquez (2013, 2015), cada economía, e incluso cada rama productiva en ella, se comporta de diferente manera de acuerdo, entre otras cosas, a su nivel de especialización, a su acceso a la tecnología y a la capacitación de sus trabajadores, por lo que es sensato pensar que cada uno de ellos tendrá una función de producción distinta de acuerdo a su nivel de especialización y acceso a la tecnología, y una elasticidad de sustitución distinta entre sus insumos de la producción de acuerdo a la escasez de insumos adecuados para la producción que enfrente cada economía. Además, Growiec (2006, 2013, 2017) generalizó estos modelos eliminando el supuesto de independencia entre sus insumos y suponiendo una función de producción CES de tipo determinista, razón por lo cual sólo logra encontrar formas analíticas cerradas de la función de producción global en casos especiales con probabilidad cero de ocurrir.

Así, aunque estos últimos modelos basados en la teoría neoclásica de la producción arrojan luz sobre la forma en que podemos modelar la producción con innovación y abren el debate teórico a cerca de las características de este tipo de modelación, estos tienen varias carencias tanto metodológicas como de origen, puesto que asumen formas funcionales predeterminadas de la función de producción y de la frontera tecnológica sin tomar en cuenta si estas son compatibles con la evidencia empírica, contrario a lo que se demuestra en Spanos (1986, 1990) y Spanos y McGuirk (2001) que se debería de hacer, dada la naturaleza observacional de los datos en economía.

Finalmente, Aoyama (2010) hizo una modelación empírica de la función de producción, la cual utiliza datos financieros de empresas japonesas y construye su distribución conjunta utilizando la teoría de cópulas, llamándole función de producción copula, y encontrando una fuerte correlación entre capital, trabajo y valor agregado.

Por otra parte, es importante destacar que existe un desfase de la modelación empírica de la producción con innovaciones verdes con la teórica, dado que la mayoría de los artículos que usan datos para analizar este tema ocupan formas funcionales estándar (principalmente de la forma Cobb-Douglas) para describir la producción, obligando a suponer que las variables aleatorias que modelan la producción son independientes, lo que resulta incompatible con la evidencia empírica como se demuestra en Acemoglu (2003) y Sánchez y Márquez (2013, 2015). Por lo que los trabajos teóricos y aplicados que se basan en este tipo de modelación, han llevado a conclusiones parciales y de corto alcance para la economía, validas sólo para ciertos casos especiales (Barro y Sala-i-Martin, 2004), en lugar de partir de la evidencia empírica, identificar los parámetros del mismo y validar este modelo probabilístico con los datos reales de la economía. Ejemplos de esto son los modelos econométricos de regresión basados en funciones de producción ambientales, la frontera de transformación, las funciones direccionales ambientales de distancia o el análisis insumo-producto.

Los modelos econométricos se han planteado en la mayoría de los trabajos empíricos presuponiendo una función de producción ambiental en la cual el producto de la empresa depende de un conjunto de insumos (tomados estocásticamente como variables aleatorias independientes) que se relacionan a través de una forma funcional dada, reduciendo el modelo a un análisis residual de los errores en la función. El modelo de frontera de transformación plantea una especificación parecida al anterior, pero donde el error estadístico está constituido por dos tipos de componentes independientes, el ruido estadístico v_j , que se distribuye normal con media cero, y el error unilateral $u_j \geq 0$, que representa la ineficiencia técnica de la producción. Así, este modelo supone que los componentes del error tienen una relación lineal de la forma $\varepsilon_j = v_j - u_j$, por lo que basta

con calcular los momentos de u_j para estimar la ineficiencia técnica de la producción en la economía, como se hace en Jondrow (1981). Las funciones direccionales ambientales son modelos útiles para medir los efectos de regulaciones ambientales en la productividad de las empresas comparándola cuando se está bajo este régimen y cuando no, vea por ejemplo Färe, Grosskopf y Pasurka (2006). Por último, el modelo de insumo-producto analiza la producción en una economía a través del análisis de su estructura y sus capacidades de producción, pudiendo incorporar en el mismo diversos estímulos, como los de innovaciones verdes, con el fin de estimar sus efectos, o estimar impactos ambientales del uso de materiales o productos en diferentes etapas de la producción; vea por ejemplo Hendrickson (1998) y Sudhakar (2012).

Así, dada la necesidad de plantear una nueva forma de construir modelos de producción con innovaciones verdes basados en la aproximación probabilística de sus parámetros, en este trabajo analizaremos la forma en que podemos modelar la producción con innovaciones tecnológicas verdes en una economía competitiva de manera empírica con un enfoque de reducción probabilística (Spanos, 1986 y 1999) utilizando datos de patentes verdes como variables proxy, las cuales incluyen según la clasificación de la USTPO, las que contribuyen a mejorar la calidad del medio ambiente, la conservación de la energía, el desarrollo de recursos de energía renovable o la reducción de los gases de efecto invernadero. En este sentido, esta sección es una aplicación directa de nuestro modelo teórico para ejemplificar con datos reales los usos que puede tener esta forma de modelación estadística.

El enfoque de reducción probabilística expuesto en Spanos (1989, 1990), Spanos y McGuirk (2001), permite construir un modelo probabilístico y muestral que describe consistentemente el comportamiento de los datos y establecer una vinculación adecuada con elementos teóricos que ayudan a explicar el problema que analizamos con esta herramienta. Este método consiste primeramente en *especificar* este modelo de acuerdo al comportamiento de los datos de nuestro modelo, para lo cual en nuestro caso utilizamos la teoría de cópulas y en particular al Teorema de Sklar para construir una función de

probabilidad conjunta que capture su naturaleza estocástica y especificar su modelo muestral, después se prosigue a *estimar* los parámetros de nuestro modelo a través de algún método estadístico adecuado para el modelo que ocupamos, seguido de una *prueba de correcta especificación*.

Si en este punto el modelo no pasa la prueba de correcta especificación, entonces se procede a hacer una nueva y se repiten los pasos anteriores. Pero si las regularidades de nuestro modelo son capturadas de manera adecuada por la especificación de nuestro modelo probabilístico (cuando no se detectan problemas de especificación en él), entonces continuamos con el proceso de *identificación* de nuestro modelo probabilístico con el modelo económico en cuestión, interpretando las formas funcionales y los parámetros estimados en términos de este modelo.

En este caso, el enfoque de reducción probabilística que proponemos identificará a la distribución conjunta de las variables de nuestro modelo con el proceso de producción para nuestra economía y a su especificación muestral, lo cual refleja al conjunto de planes de producción factibles y nos permitirá identificar los parámetros probabilísticos como la dependencia y parámetros marginales con los del modelo teórico-económico. Este enfoque nos permitirá deducir a su vez, a partir de los parámetros estimados de manera estadística, una función de producción local y frontera tecnológica que permiten caracterizar a la tecnología usada en esta economía. Además, dado que la teoría de cópulas permite separar los efectos individuales de las variables aleatorias de su efecto conjunto, nuestro modelo nos permite analizar la dinámica interna de la producción y sus variaciones debido a impactos en sus parámetros.

Nuestro esquema estadístico resulta ser lo suficientemente general para permitir incluir diversas variables endógenas y exógenas y discutir sus efectos en un modelo de producción para una economía determinada, así como validar supuestos más realistas de la economía como el crecimiento logístico de la población (Brida, 2008) o supuestos específicos de la teoría de los recursos naturales como la curva de Kuznets (Stern, 1996) utilizando evidencia

empírica. Además de que permite la construcción de modelos estadísticos para el análisis de otros temas importantes de la teoría económica como el problema del consumidor, la economía de la información y el modelo de equilibrio general.

Esto último nos ayuda en la modelación y el análisis de las variables fundamentales y contrastar los resultados de nuestra modelación con los supuestos y principios que presuponen los modelos teóricos, fungiendo así como una herramienta de validación empírica de los supuestos de la teoría económica en un problema específico y constituyéndose además como un puente entre la teoría económica y la modelación estadística. Note además que esta metodología puede usarse para modelar otros problemas de interés en economía como la demanda, las brechas económicas o los efectos específicos de ciertas variables económicas.

Sin embargo, en este proceso nos enfrentamos a tres tipos de problemas que tiene que resolver para constituirse como una herramienta adecuada. Por un lado está el *problema de especificación*, el cual consiste en escoger de manera adecuada un modelo estadístico que aproxime de manera adecuada la distribución poblacional y la muestral de nuestro fenómeno. El siguiente es el *problema de estimación*, el cual es el elegir los métodos adecuados para la estimación de los parámetros de nuestra distribución poblacional. Por último están los *problemas de distribución*, los cuales se presentan en las distintas etapas de nuestro proceso y consisten en la deducción matemática de la naturaleza de la distribución de las muestras aleatorias en nuestras estimaciones de los parámetros y de los estadísticos para validar la especificación de nuestro modelo (pruebas de correcta especificación) (Spanos y McGuirk, 2001). Así, durante el desarrollo de este trabajo y en el modelo empírico que desarrollamos estaremos haciendo énfasis en estos problemas y sus soluciones.

En el presente trabajo mostramos que se puede construir un modelo de producción con innovaciones basados en un modelo probabilístico especificado a partir de los datos para una economía. En particular identificamos de manera teórica y un modelo con función

global de producción CES a partir de un modelo probabilístico basado en una cópula producto-Clayton, identificando al parámetro de dependencia estadístico con uno que determina la elasticidad de sustitución entre insumos del modelo económico, así como los parámetros marginales del modelo con los indicadores de escasez en la economía.

Por otra parte, con el fin de ilustrar la utilidad de éste método, se lleva a cabo la modelación empírica para la producción con innovaciones verdes para la economía de alta especialización de Estados Unidos, para lo cual se ocuparon usando datos de patentes verdes del año 2014. Tomando en cuenta las restricciones del modelo, se ocuparon los datos del núcleo de empresas que tenían concedidas patentes verdes de ambos tipos en ese año, por lo que observando las gráficas de los mismos y suponiendo que la muestra aleatoria era independiente e idénticamente distribuida, se propuso que las patentes se podían modelar con una distribución Pareto y la modelación de su distribución conjunta por una cópula producto-Clayton, por lo que usando el método máxima verosimilitud completa se obtuvieron los valores para los parámetros marginales y de dependencia.

Después el modelo fue sometido a la realización de pruebas de correcta especificación, las cuales pasó, por lo que se continuó con el proceso de identificación del modelo estadístico con el teórico-económico de la función de producción, obteniendo una función de producción global de la familia CES con parámetro de sustitución $\theta = 1.724$, que se corresponde con una elasticidad de sustitución $\sigma = 0.367$ (los insumos son complementarios dados los valores de los parámetros de la función CES).

En el primer capítulo de nuestro trabajo se hará un resumen y análisis de las teorías que se han empleado para la modelación estocástica de la producción con innovaciones en la economía, así como con innovaciones verdes. Además, haremos un recuento de los modelos empíricos estándar usados en la modelación de la función de producción.

En el segundo capítulo describimos el enfoque de reducción probabilística propuesto por Aris Spanos, el cual constituye la base teórica de nuestra modelación. Además hacemos una

introducción a la teoría de cópulas y enlazamos el enfoque anterior con esta teoría para describir el enfoque que usaremos en nuestro análisis.

En el tercer capítulo describimos el modelo de función de producción general basado en el enfoque de reducción probabilística con cópulas tomando en cuenta los factores capital y trabajo de la producción e innovaciones de estas dos variables. Además, desarrollamos la aplicación de este modelo utilizando distribuciones marginales Pareto y una cópula producto-Clayton, así como suponiendo independencia y homogeneidad en la muestra de datos para la especificación del modelo, y desarrollaremos la identificación del modelo.

En el cuarto capítulo desarrollamos el modelo empírico de la producción donde incluimos innovaciones verdes como factores que mejoran la productividad en la economía altamente innovadora, estimamos y validamos la especificación del modelo con los datos de patentes verdes como valores proxy de las innovaciones, además, interpretamos los resultados de nuestro modelo y destacamos sus propiedades. Por último concluimos en el último capítulo y platicamos desde nuestra visión sobre los alcances y tareas futuras desarrollables con la metodología aquí planteada.

CAPITULO 1: MARCO TEÓRICO: FUNCIÓN DE PRODUCCIÓN, INNOVACIÓN TECNOLÓGICA E INNOVACIONES VERDES

1.1 Motivación

En este capítulo vamos a describir el modelo económico de producción con innovaciones tecnológicas y con innovaciones verdes que vamos a emplear para el proceso de identificación de nuestro modelo estadístico una vez que se haya especificado éste a partir de la evidencia empírica que nos darán los datos que emplearemos en nuestro trabajo.

La modelación del cambio tecnológico resulta de vital importancia para el estudio de varias de las ramas de la economía, como la macroeconomía, el crecimiento y desarrollo económico y la economía de los recursos naturales. Los primeros modelos surgieron en los años 1950's como una necesidad para la rama del crecimiento económico, ya que los modelos propuestos hasta ese momento no podían explicar el crecimiento de la acumulación del capital per cápita ni el crecimiento per cápita observado en varias economías que se estudiaban en ese entonces (Barro y Sala-i-Martin, 2004).

Además, ayudaron para explicar lo que la evidencia empírica sistemáticamente ha señalado sobre el efecto de escala en la producción que tienen la investigación y el desarrollo de tecnología en las distintas economías del mundo, en particular en las industrializadas, las cuales crecen más rápidamente y les permite apropiarse del valor agregado de la producción, lo que explica en gran parte las brechas entre economías industrializadas y las no industrializadas.

Por otra parte, la distribución de la inversión mundial en investigación y desarrollo de tecnología determina sinergias en el comercio internacional y su dinámica, ya que por un lado es un factor importante en el intercambio entre países y juega un papel importante en la asignación de tareas en la producción de bienes en las cadenas internacionales de producción (Habbakkuk (1962), Romer (1990), Barro y Sala-i-Martin (2004)).

Los modelos para explicar el cambio tecnológico de Houthakker (1955), Johansen (1972) y Hildenbrand (1981) pretendieron explicar su conformación basados únicamente en las características estocásticas de los factores que intervienen en la producción. Habakkuk (1962) hizo un ensayo histórico de la innovación en EEUU e Inglaterra y postula que la diferencia del tipo de innovaciones entre ambos países se explica por la carencia de capital o de trabajo relativo en sus economías.

Más recientemente, los modelos de Acemoglu (2003), Jones (2005) y Growiec (2006), incorporaron supuestos de la microeconomía para la construcción de los modelos estocásticos de producción como los rendimientos constantes a escala, la convexidad de los conjuntos y las formas funcionales de la función de producción y la frontera tecnológica.

Estos trabajos han dado pauta para incluir en la discusión de la teoría de la producción la participación de otros tipos especiales de bienes, como la contaminación ambiental, e innovaciones tecnológicas verdes, los cuales han sido considerados en la literatura tanto como causa de desarrollo potencial de la economía, como una causa de distorsión del desarrollo óptimo de una economía capitalista (Zhang, 2006).

Además, vamos a discutir el enfoque estándar del modelo de función de producción con innovaciones tecnológicas e innovaciones verdes con los hechos estilizados a cerca del estado del arte de este tipo de modelos, analizando las principales implicaciones de los artículos en los cuales nos basamos para nuestro trabajo y su importancia en la teoría económica de la producción.

Por último haremos un recuento de la ruta de búsqueda que se hizo sobre la revisión bibliográfica para este trabajo, en la cual describimos a groso modo las referencias que tomamos de acuerdo al nivel de relevancia con nuestro trabajo y la diferencia de estos con el trabajo actual. La ruta de búsqueda se puede consultar en el Apéndice B.

1.2 La Función de Producción con Innovaciones Tecnológicas

Nuestro modelo de producción con innovaciones se inspiró en los trabajos de Johansen (1972) y Hildenbrand (1981), los cuales establecieron una formalización matemática de este concepto empleando el análisis convexo. Posteriormente, el análisis de la tecnología a partir de la función de producción local y la frontera tecnológica se retomó de los trabajos de Caselli y Coleman (2006), Houtakker (1955), Kortum (1997) y Jones (2005).

El modelo pionero de Houtakker consiste en construir una función de probabilidad conjunta de los insumos productivos para obtener el valor esperado de la producción., la cual resulta tener la forma funcional de una tecnología Cobb-Douglas de la forma

$$X_0 = CX_1^\alpha X_2^\beta X_3^\gamma, \text{ donde } \alpha + \beta + \gamma = 1.$$

Años más tarde, Johansen (1972) derivó un modelo de una función de producción de corto plazo macroeconómica, la cual se deriva, en cualquier momento del tiempo, de la distribución de la capacidad de producción existente en un punto del tiempo, es decir, de las técnicas disponibles para producir. En su modelo hay sustituibilidad de insumos antes de la elección de la técnica, pero una vez que se elige ya no se permite este intercambio entre ellos.

Hildenbrand (1981) generaliza esta propuesta suponiendo mecanismos de producción que usan los insumos de manera proporcional y requieren de convergencia en las medidas de distribución, dada la naturaleza del modelo, el cual se basa en conjuntos convexos. Este trabajo propone una aproximación no paramétrica basada en la distribución de la información tecnológica, la cual consiste en calcular la función de producción media de corto plazo de manera estocástica.

Pero esta idea es abandonada y comienza a ponderar principalmente supuestos de la teoría microeconómica para su hechura. Kortum (1997) construye un modelo basado en patentes para explicar el comportamiento de la investigación en innovaciones tecnológicas,

demostrando que si tenemos innovaciones independientes con distribución Pareto, la función de producción resultante debe ser del tipo Cobb-Douglas.

Jones (2005) retoma esta idea y propone un esquema en el que una empresa representativa resuelve el problema de elección de la tecnología con la que va a producir, el cual consiste en maximizar una función de producción con innovaciones sujeta a una frontera tecnológica, la cual está constituida por los pares tecnológicos eficientes disponibles para la empresa, es decir, propone resolver

$$\max_{a,b} Y = f(a, L, b, K) \text{ sujeto a } H(a, b) = N,$$

donde $f(a, L, b, K)$ es una función de producción, H es la frontera tecnológica en un nivel N . Jones concluye que la función de producción que obtiene de resolver el problema de elección de la tecnología depende exclusivamente del comportamiento conjunto de las innovaciones tecnológicas en el corto plazo, y que en el largo plazo el crecimiento de la economía es explicado exclusivamente por las innovaciones ahorradoras de trabajo. Así, en este modelo, el comportamiento conjunto de las innovaciones tecnológicas asumidas en el modelo, al cual llama menú tecnológico, determina la forma de la función de producción global y la dirección del cambio tecnológico en el largo plazo.

Así, el modelo de producción con innovaciones al cual llegan Kortum (1975) y Jones (2005) consiste en una función de producción de la forma:

$$Y_t = A(N_t)K_t^\beta L_t^\alpha \varepsilon_t,$$

donde el término ε_t es el término estocástico del modelo y está distribuido como una normal, una Pareto o una Frechét dependiendo de los supuestos.

Caselli y Coleman (2006) por su parte retoman la idea de Jones y pretenden la estimación de una frontera tecnológica para cada varios países, para la cual propone que las innovaciones no se comportan de manera independiente entre sí, sino que existe una relación inversa entre los avances tecnológicos que ahorran capital y trabajo, por lo que

emplean un modelo matemático del tipo CES para modelar estas fronteras, manteniendo el supuesto de que las innovaciones son variables aleatorias que siguen una distribución del tipo Pareto.

Por último, Growiec (2006) parte del modelo propuesto por Jones (2005), suponiendo que el productor quiere resolver el problema de elección de la tecnología, para lo cual asume que la empresa se enfrenta a una función de producción determinista del tipo CES de la forma

$$Y = A[\varphi L^{-\theta} + (1 - \varphi)K^{\theta}]^{\frac{1}{\theta}},$$

junto con una frontera tecnológica como la que modelan Caselli y Coleman (2006), de la forma

$$H(a, b) = H(a, b) = \left(\frac{a}{\gamma_a}\right)^{-\alpha\theta} + \left(\frac{b}{\gamma_b}\right)^{-\beta\theta}.$$

Así, Growiec demuestra que cuando el comportamiento de las innovaciones tecnológicas está descrito por marginales Pareto, las funciones resultantes constituyen una nueva clase de funciones de producción que sólo tienen forma funcional cerrada para ciertos valores de los parámetros de las funciones, por lo que concluye que la forma de la función de producción óptima de la economía depende tanto las relaciones entre innovaciones tecnológicas como de función de producción genérica de la economía. Concluye además que estas formas funcionales son importantes para determinar la dirección del cambio tecnológico, por lo que este último puede aumentar tanto por un incremento en la tasa de crecimiento de las innovaciones que ahorran trabajo como en las que ahorran capital, así como lo proponen Atkinson y Stiglitz (1969) y contrario a la conclusión de Jones (2005). Además, en Growiec (2017) establece una relación cercana entre las soluciones de los problemas de elección de la tecnología y de maximización del beneficio, estableciendo una dualidad entre la función de producción global y la frontera tecnológica y obteniendo esta caracterización para varios tipos de función de producción.

Así, Jones (1999) hace un recuento de los modelos de crecimiento analizando sus implicaciones de los efectos de escala en los modelos, concluyendo que la no rivalidad entre las innovaciones en los modelos de crecimiento basados en variables exponenciales implica que siempre hay un efecto a escala en el modelo de crecimiento. Acemoglu (2003) propone un modelo de equilibrio general asumiendo un modelo de producción con una función determinística del tipo CES en una economía competitiva donde se cumplen los supuestos del modelo de crecimiento neoclásico con cambio tecnológico. De esta forma, Acemoglu concluye que en la dinámica de transición del modelo, el crecimiento económico está motivado por innovaciones ahorradoras de capital y de trabajo, pero que en el largo plazo sólo pesan las ahorradoras de trabajo.

Por otra parte, otros trabajos como los de Jones (1999), Grossman (2017), León-Ledezma y Satchi (2018) y Sánchez y Márquez (2018) exploran el comportamiento del crecimiento basado en innovaciones en el largo plazo como una aplicación de este tipo de modelos.

Por último, Grossman (2017) extiende el Teorema de Uzawa, el cual caracteriza el crecimiento de largo plazo en un modelo de crecimiento neoclásico, encontrando un modelo de crecimiento balanceado consistente con la evidencia empírica para los EEUU, suponiendo que la variable educación es endógena y que el capital es más complementario con el nivel de escolaridad que con el pago a salarios en la producción. León-Ledezma y Satchi (2018), utilizando una función de producción del tipo CES y el Teorema de Uzawa, obtienen una caracterización de cómo la frontera tecnológica determina la elasticidad de sustitución de largo plazo entre el capital y el trabajo. Finalmente, Sánchez y Márquez (2018) generalizan el Teorema de Uzawa, encontrando que el importan ambos tipos de innovación para el crecimiento de la producción con innovaciones en el largo plazo.

1.4 La Función de Producción con Innovaciones Verdes

A parte de los trabajos que mencionamos en la sección anterior, otros autores han desarrollado modelos de producción considerando otro tipo de variables como las de energía, recursos naturales y contaminación ambiental, las cuales se consideran componentes de la producción en la teoría de la innovación y el crecimiento económico, con el objetivo de estudiar su interacción con las variables clásicas de capital (físico y humano) y trabajo, así como su impacto individual en el proceso productivo (Zhang, 2006). Más aún, la nueva teoría del crecimiento desarrollada en la última década del siglo pasado, propone tomar a estas variables de contaminación ambiental como factores que permiten la producción, es decir, como variables que tienen inferencia directa sobre el nivel de producción para una tecnología dada (Boztkurt y Akan, 2014).

En este sentido, se han encontrado una relación negativa entre la producción y la emisión de contaminantes como el dióxido de carbono (CO_2) [Andersson (2008), Boopen (2011), Boztkurt y Akan (2014)], probando además que la elasticidad de las emisiones respecto al ingreso se ha ido incrementando en el tiempo. Además, otros factores que intervienen en este proceso productivo según los autores son el consumo energético, la combinación de combustibles utilizada para la producción y la población. En este trabajo consideraremos a las innovaciones verdes como factores endógenos del crecimiento.

Un modelo para explicar la dinámica del medio ambiente desarrollado por Plourde (1972) y Forster (1975) toma en cuenta una función de producción de la forma:

$$Y_t = AL_t^\beta K_t^\alpha \exp(-h_p E_t),$$

donde el término $\exp(-h_p E_t)$ indica que la productividad está negativamente relacionada con el nivel de contaminación ambiental pero el nivel de innovación descrito en el término tecnológico A influye de manera positiva en la producción. Así, el modelo propone que la dinámica de la acumulación de innovación tecnológica A_t está descrita por una ecuación diferencial que depende de manera positiva del nivel del consumo y de la producción, así

como de un proceso natural de purificación del ambiente y de acciones humanas pro-ambientales (como el desarrollo de tecnologías limpias o procesos de purificación, reforestación, entre otros).

Por otra parte, basados en la escasez de los recursos naturales que se pueden emplear en la producción, Stern (1996) propone que hay una relación en forma de U-invertida entre la degradación ambiental y el producto interno per-cápita, la cual se conoce como la Curva de Kuznets, de tal forma que el crecimiento reduce el impacto ambiental de la actividad económica. Finalmente, Brock y Taylor (2004) construyeron un modelo de crecimiento basado en el de Solow pero con innovaciones verdes, con el cual muestra que la curva de Kuznets y el modelo original de Solow están íntimamente relacionados, dado que el segundo es necesario para la convergencia de la senda de crecimiento del primero. Además, este modelo propone que los modelos empíricos de la curva de Kuznets fallan por la heterogeneidad de las condiciones iniciales a lo largo de un país.

1.3 El Enfoque Estándar para la Modelación Empírica de la Función de Producción

Dada la revisión anterior de la literatura podemos ver que el enfoque más usado para la modelación empírica de la producción con innovaciones está descrito en los trabajos de Kortum (1997), Jones (2005), Acemoglu (2003), Growiec (2006, 2009, 2017), Plourde (1972) y Forster (1975), entre otros.

Estos autores retomaron el trabajo seminal de Houtakker (1955), en el cual propone que los insumos intermedios para la producción siguen una distribución probabilística Pareto independientes y demuestra que entonces, el modelo adecuado para la función de producción de la economía es la tecnología Cobb-Douglas dada por la función

$$Y_t = A(N_t)K_t^\beta L_t^\alpha \varepsilon_t.$$

Así, los autores por lo general proponen que el modelo de producción con innovaciones, donde éstas se distribuyen de acuerdo a la distribución Pareto y pueden ser aproximadas por variables de patentes, corresponde a esta especificación sin tomar en cuenta el comportamiento estocástico de los datos. Por lo que lo que el paso siguiente para el proceso es estimar el valor de los parámetros de éste modelo usando algún método estadístico, suponiendo además que el modelo muestral está constituido por muestras independientes e idénticamente distribuidas, por lo que se opta principalmente por estimar un modelo de la forma:

$$\sum_{t=1}^m \log(Y_t) = n \log(A) + \sum_{t=1}^m \log(N_t) + \beta \sum_{t=1}^m \log(K_t) + \alpha \sum_{t=1}^m \log(L_t) + \sum_{t=1}^m \log(\varepsilon_t),$$

donde los parámetros del modelo se pueden obtener e interpretar directamente gracias a que partimos de un modelo de producción definido inicialmente.

Así entonces, este enfoque estándar consiste en partir del modelo de función de producción que queremos estimar, suponiendo que conocemos la forma funcional de nuestro modelo y que además ésta cumple con propiedades que impone la teoría neoclásica de la producción tales como convexidad, homogeneidad y rendimientos a escala, entre otras. El modelo empírico estándar consiste en estimar los parámetros de este modelo a partir de aplicar algún método de estimación probabilístico a los datos, suponiendo que el conocemos el mecanismo generador de información en los mismos, es decir, suponiendo que el modelo de función de producción que estimamos es el que produce los datos de innovación de los cuales partimos, lo cual es muy conveniente para el momento de identificar el modelo pues ya parte de un modelo con una forma específica y solamente utiliza la estimación probabilística de los parámetros del modelo teorico-económico que presupuso desde un inicio.

Sin embargo, por esta forma en que se procede, es muy común que el modelo no pase las pruebas de correcta especificación, pues al no tomar en cuenta el comportamiento de los datos de entrada, la probabilidad de que estos se comporten de acuerdo al modelo que

presuponemos es muy baja, además, como podemos ver de la evidencia empírica, ni los insumos de la producción ni las innovaciones tecnológicas (en general y tampoco las verdes) suelen ser independientes y la forma funcional Cobb-Douglas no es un buen modelo por las restricciones adicionales que impone para su modelación [Acemoglu (2013), Sánchez y Márquez (2013, 2015)].

Por lo que, en este punto, los modeladores en lugar de regresar a re-especificar el modelo para encontrar el más adecuado de acuerdo a la evidencia empírica, lo que hacen normalmente es utilizar métodos alternativos para tratar que los datos se ajusten a este modelo o a una transformación de él, empleando por ejemplo variables indicadoras para quitar datos que consideran atípicos del modelo y por lo tanto descartables o considerando que hay cambio estructural en los mismos. El problema principal con esta parte de la modelación es que pocas veces estos pasos se justifican técnicamente y más bien se aplican de manera discrecional para que su modelo pase las pruebas de correcta especificación [Spanos (1986, 1999), Spanos y McGuirk (2001)].

Aunque estos modelos basados en el enfoque estándar constituyen la base para nuestro análisis de la producción, presentan graves carencias en la modelación que ya discutimos, por lo que en la mayoría de los casos nos son modelos correctamente especificados, o consideran parámetros desconocidos que tienen una repercusión importante en el modelo pero que no se pueden aproximar estadísticamente como en el caso de Growiec (2006, 2013 y 2017) o que recurren previamente a un tratamiento incorrecto de los datos para que estos ajusten al modelo previamente especificado para los mismos.

1.5 Comentarios Generales

Esta revisión de la literatura nos permite ver que no hay una forma única de modelar la producción con innovaciones e innovaciones verdes con ayuda de la estadística en la economía, sino que esto se hace de acuerdo al enfoque teórico que se adopte y al

entendimiento de cómo se debe modelar el proceso de producción. Así, cada forma va a tener un alcance distinto, de acuerdo al marco conceptual en el que se inscriba, y por lo tanto, implicaciones distintas tanto en el análisis del corto y largo plazo de la producción con innovaciones, es decir, en el modelo de función de producción con innovaciones, la producción con innovaciones verdes y la dirección del cambio tecnológico.

Sin embargo los modelos basados en el enfoque estándar tienen en general muchas carencias al no tomar en cuenta el comportamiento real de los datos sino imponerles un mecanismo generador de los mismos previamente establecido y suponer ciertos comportamientos estadísticos que, a pesar de que algunos están basados en la observación empírica, se apegan más al modelo teórico-estadístico que nos asegura un resultado final que coincide con la forma funcional que presuponemos al inicio. Por estas razones, queremos construir un modelo que permitan una correcta especificación estadística basada en el comportamiento observado de los datos y adoptando los modelos probabilísticos que mejor ajusten a los datos.

Así, vamos a usar los patrones probabilísticos y muestrales de los datos como un medio para obtener el mecanismo generador de información de los mismos y construir un modelo probabilístico que explique coherentemente el comportamiento de los datos y que permita identificar los parámetros del mismo con los de la teoría microeconómica [siguiendo el enfoque de Spanos (1999) sobre la modelación de fenómenos observacionales]. Este tipo de modelación tiene la ventaja de que deja abierta la posibilidad de que las formas funcionales y propiedades de las funciones del modelo se comporten de acuerdo a las características individuales de los datos y a su funcionamiento conjunto. Además, este modelo puede incorporar diversas especificaciones adicionales dadas por la evidencia empírica como la curva de Kuznets o el crecimiento poblacional diferente al exponencial.

Es importante mencionar que se estuvieron haciendo búsquedas sistematizadas de artículos y libros sobre el tema, basadas en las palabras clave relacionadas con nuestros temas de investigación así como de artículos relacionados, con lo cual pudimos constatar que esta tesis tomó en cuenta todos los artículos referentes al tema y que ninguno de los trabajos

anteriores se ha dado a la tarea de conjuntar el enfoque de reducción probabilística con la teoría de cópulas ni de hacer la especificación de un modelo de producción con innovaciones tecnológicas o innovaciones verdes como el que realizamos, por lo que los resultados del proceso de identificación son novedosos y tienen un gran alcance.

CAPITULO 2: METODOLOGIA EMPÍRICA, ESPECIFICACIÓN PROBABILÍSTICA Y TEORÍA DE CÓPULAS

2.1 Motivación

La modelación empírica de la teoría de la producción con innovaciones es una herramienta importante en economía ya que nos permite hacer una aproximación al comportamiento de los agentes de la producción en temas importantes como el aprovechamiento de los recursos naturales y la incorporación de innovaciones verdes en el proceso de producción, además de que al hacerse de forma correcta puede usarse como una herramienta para contrastar la evidencia empírica con los modelos económicos teóricos.

El objetivo de la descripción estadística de diferentes aspectos de la economía es el aproximar de la mejor manera posible el mecanismo generador de información de los datos económicos basados en el modelo axiomático de probabilidad y la especificación del modelo muestral, los cuales que nos va a permitir hacer inferencia sobre los modelos económicos una vez que se prueba que la especificación probabilística es adecuada y se realiza un proceso de identificación de los parámetros estimados (Spanos, 1986 y 1999).

Así, la estadística inferencial aplicada a datos observacionales (que no se pueden reproducir en un ambiente controlado sino sólo se recogen al observar el fenómenos, como los que se generan en las ciencias económicas, las ciencias sociales o la biología) debería basar su modelación en métodos estadísticos que tomen en cuenta esta metodología probabilística, en lugar de sólo adoptar los métodos de análisis estadístico de las ciencias experimentales, los cuales hacen supuestos simplificadores que rara vez cumplen los datos ocupados en estas ciencias [Spanos (1999) y Spanos y McGuirk (2001)].

Empíricamente, nuestro modelo toma en cuenta para el proceso de identificación los hechos estilizados propuestos en los trabajos que ya citamos anteriormente, donde se observa que la distribución de los datos de innovación son modelados de manera consistente usando distribuciones Pareto [Caselli y Coleman (2006), Jones, (2005); Growiec (2006, 2017)] y

que el modelo más adecuado de función de producción es aquel que permita una participación de los insumos en el producto cambiante en el corto plazo pero con tendencia de largo plazo (Acemoglu, 2003).

Nuestro esquema estadístico resulta ser lo suficientemente general para permitir incluir diversas variables endógenas y exógenas así como modelos más realistas de la economía como el crecimiento logístico de la población (Brida, 2008) o supuestos específicos de la teoría de los recursos naturales como la curva de Kuznets (Stern, 1996) utilizando evidencia empírica.

También tiene la ventaja de que, al mezclar este enfoque con la teoría de cópulas, no nos tenemos que restringir a un menú de distribuciones conjuntas del cual podamos especificar nuestro modelo probabilístico, sino que podemos construir la distribución que mejor se adapte a nuestros datos, tomando en cuenta los hechos estilizados y la distribución empírica de los mismos.

Así, un modelador con este enfoque puede hacer inferencia estadística utilizando argumentos inductivos que permitan plantear conclusiones inferidas, y establecer generalizaciones y señalamientos a cerca de futuras observaciones del mismo tipo, con base en los datos económicos observados. Además de que permite la construcción de modelos estadísticos para el análisis de otros temas importantes de la teoría económica como el problema del consumidor, la economía de la información, el modelo de equilibrio general, las brechas económicas o los efectos específicos de ciertas variables económicas.

2.2 El Enfoque de Reducción Probabilística

El enfoque estadístico que consideramos en este trabajo es el de *Reducción Probabilística*, el cual se expone en Spanos (1986, 1999) y Spanos y McGuik (2001), el cual por un lado nos permite construir un modelo consistente con la naturaleza observacional de los datos en economía, por lo que resulta ser bastante útil al poder obtener un mecanismo generador de información y estimar los parámetros que modelan de manera consistente los datos observados en nuestro fenómeno de estudio. A este proceso le llamamos *especificación* del modelo empírico.

Por otra parte, nos permite *identificar* los parámetros estadísticos de nuestro modelo estocástico y posteriormente identificarlos con los del fenómeno que estamos estudiando, lo cual nos sirve para hacer inferencia y permite contraponer los resultados y supuestos teóricos con los hallazgos empíricos, por lo que nos ayuda a mejorar los modelos o tomar en cuenta aspectos que escapan a la intuición teórica inicial (Sánchez y Márquez, 2015).

Así, nuestro enfoque consiste inicialmente en *especificar* el modelo estadístico a partir de la información de los datos de los que partimos, para lo cual identificamos patrones de comportamiento de los mismos especificando un modelo probabilístico de los mismos y un modelo muestral, valiéndonos de las distribuciones empíricas, la teoría de cópulas y la observación del comportamiento de la muestra de datos.

Una vez que concluimos el proceso de especificación de nuestro modelo, se prosigue a estimar los parámetros del mismo a través de algún método estadístico adecuado para las formas funciones a las que llegamos. En nuestro caso, basados en la especificación de nuestro modelo estadístico, construimos su función de verosimilitud y la maximizamos para obtener los estimadores de los parámetros del mismo. Entonces, con especificación de nuestro modelo y con los parámetros estimados, construimos pruebas de correcta especificación que nos ayude a comprobar si ésta ha sido adecuada.

Si nuestro modelo no pasa las pruebas, se procede a hacer una re-especificación del modelo y a repetir los pasos anteriores hasta volver a hacer una análisis de correcta especificación, pero si las pasa, quiere decir que describe de manera apropiada el comportamiento de nuestros datos, por lo que las regularidades de nuestro fenómeno son capturadas por la especificación de nuestro modelo estadístico y podemos proceder a *identificar* nuestro modelo con el económico al que queremos llegar, interpretando las formas funcionales y los parámetros estimados en términos de los parámetros y las formas funcionales que nos ayudan a explicar el fenómeno que estamos estudiando en este caso.

Desde el punto de vista empírico, es importante resaltar que este tipo de modelación consiste en realizar la descripción de un fenómeno estocástico, es decir, la descripción de un fenómeno cuyos datos presentan un patrón de ocurrencia regular, usando métodos de inferencia basados en los patrones de distribución, dependencia y homogeneidad de los datos, por lo que para hacer esto se debe verificar la naturaleza estocástica del fenómeno de estudio, la existencia de datos observables para dicho fenómeno, así como la correcta especificación del modelo estadístico como lo describe (Spanos, 1986 y 1999).

Este enfoque, el cual tomaremos como base para hacer inferencia, parte de identificar de acuerdo al comportamiento de nuestros datos dos modelos:

- i) Un *modelo probabilístico* $E = \{T(\mathbf{x}; \theta), \theta \in \Theta, \mathbf{x} \in \text{Ran}(\mathbf{X})\}$, compuesto por una familia de funciones $\{T(\mathbf{x}; \theta), \theta \in \Theta\}$ llamadas densidades conjuntas, las cuales miden probabilidades de ocurrencia de vectores aleatorios \mathbf{X} sobre eventos de interés en su rango $\text{Ran}(\mathbf{X})$, y un conjunto de parámetros Θ sobre el cual toman valores estas densidades y están determinados empíricamente por los valores de nuestros datos.
- ii) Un *modelo de muestreo* $X = \{X_1, X_2, \dots, X_n\}$, constituido por muestras de vectores aleatorios con cierta distribución, los cuales pueden presentar dependencia y heterogeneidad.

Comenzamos construyendo el *Modelo Probabilístico* para variables observacionales, Apoyándonos en Spanos (1986) y Rincón (2007), el cual le proporciona sustento a nuestro enfoque. Spanos (1986) construye este modelo a partir de un distinguir la distribución empírica de los datos a partir de catálogos conocidos de distribuciones conjuntas, basado en trabajos de Ken Hendricks y Trygve Haavelmo. En este trabajo se utiliza la teoría de cópulas y el Teorema de Sklar para especificar una función de probabilidad conjunta que capture su naturaleza estocástica basada en los datos de nuestro modelo, como veremos en la siguiente sección.

Dado un *experimento aleatorio* (aquel que aun cuando se repita bajo las mismas condiciones iniciales, no podemos saber a priori el resultado del mismo), definimos el *espacio muestral* como el conjunto de todos los posibles resultados del experimento aleatorio, y lo denotamos por S . Así, el espacio muestral puede tener cardinalidad numerables o no dependiendo de si el número de posibles resultados es numerable o continuo de acuerdo a la naturaleza del experimento aleatorio.

Decimos que un *evento de interés* A es un conjunto formado por resultados posibles de nuestro experimento aleatorio, es decir, un subconjunto del espacio muestral S , sobre el cual nos interesa establecer una medida de probabilidad para poder hacer inferencia sobre él. En general estaremos interesados en establecer de un sistema axiomático probabilístico sobre un conjunto eventos de interés, así como de los eventos derivados de estos a través de las operaciones de unión, intersección, resta y complemento.

Con este propósito, supondremos que el conjunto de eventos de interés y demás eventos obtenidos de las intersecciones, uniones, complementos y restas de estos, forman un conjunto \mathcal{F} que es subconjunto del potencia de S y cumple con los siguientes axiomas:

- i. $S, \emptyset \in \mathcal{F}$.
- ii. Si $A \in \mathcal{F}$, entonces $A^c \in \mathcal{F}$ (donde A^c es el complemento de A).
- iii. Si $A_1, A_2, A_3, \dots \in \mathcal{F}$, entonces $\bigcup_{k=1}^{\infty} A_k \in \mathcal{F}$.

Así, diremos que el conjunto \mathcal{F} cumple ser una *sigma álgebra* de subconjuntos de S y podemos establecer una medida de probabilidad sobre él.

Un ejemplo de sigma-álgebra es el conjunto de subconjuntos del espacio de los números reales llamada la *Sigma-álgebra de Borel* dada por:

$$\mathcal{B}(\mathbb{R}) = \sigma\{(a, b) \subseteq \mathbb{R}: a \leq b\}$$

donde $\sigma\{\cdot\}$ es el conjunto generado por los abiertos de la forma (a, b) con $a \leq b$, es decir el conjunto de todas las uniones, intersecciones, diferencias, complementos y diferencias simétricas que podemos hacer con intervalos abiertos, a los cuales llamaremos *borelianos*.

La sigma-álgebra de Borel contiene intervalos cerrados, semicerrados, puntos, intervalos infinitos, así como uniones de estos, entre muchos otros elementos.

DEFINICIÓN 1: Una medida de probabilidad es una función $P: \mathcal{F} \rightarrow [0,1]$, que satisface lo siguiente:

- i. $P(S) = 1$.
- ii. $P(A) \geq 0$, para cualquier $A \in \mathcal{F}$.
- iii. Si $A_1, A_2, A_3, \dots \in \mathcal{F}$, son eventos ajenos dos a dos, es decir, $A_n \cap A_m = \emptyset$ para todo $n \neq m$, entonces

$$P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k)$$

Note que esta medida cumple varias propiedades matemáticas como que $P(\emptyset) = 0$ o que $P(A^c) = 1 - P(A)$. También está bien definida la medida de probabilidad de uniones finitas, intersecciones, diferencias y diferencias simétricas de eventos, además de propiedades de convergencia y asintóticas. Por lo que tenemos una medida de probabilidad bien definida para todos los eventos de interés.

Sin embargo, para poder hacer inferencia estadística, necesitamos definir una medida de probabilidad real que nos permita asignarle un valor de probabilidad a conjuntos reales.

DEFINICIÓN 2: Definimos una variable aleatoria como una función $X: S \rightarrow \mathbb{R}$, tal que para cualquier conjunto boreliano $B \in \mathcal{B}(\mathbb{R})$, se cumple que la imagen inversa $X^{-1}(B)$ es un elemento de \mathcal{F} , es decir, la imagen inversa bajo X de cualquier boreliano es un elemento de la sigma-álgebra \mathcal{F} .

Note que si consideráramos a los elementos borelianos $B \in \mathcal{B}(\mathbb{R})$ y a la sigma-álgebra \mathcal{F} como dos espacios de conjuntos, donde construimos un sistema de pre-vecindades abiertas con los conjuntos de cada respectiva sigma-álgebra del espacio, entonces una variable aleatoria sería una *función continua* del conjunto \mathcal{F} a $\mathcal{B}(\mathbb{R})$.

Ahora, dada una variable aleatoria X , podemos asociarle una función que mide la probabilidad de eventos descritos por esta variable, de la siguiente forma.

DEFINICIÓN 3: La función de distribución de la variable aleatoria X es una función $T: \mathbb{R} \rightarrow [0,1]$, definida por:

$$T(x) = P(X \leq x).$$

Esta función mide la probabilidad de cualquier boreliano de la forma $(-\infty, x]$, los cuales también forman una base sobre la cual se puede obtener la sigma-álgebra de Borel. Así, cualquier evento aleatorio de interés se puede asociar a un elemento de la sigma-álgebra de Borel, por lo que se le puede asignar a su vez una medida de probabilidad no negativa.

Se puede probar fácilmente que la función de distribución conjunta cumple con las siguientes propiedades:

- i. $\lim_{x \rightarrow \infty} T(x) = 1.$
- ii. $\lim_{x \rightarrow -\infty} T(x) = 0.$
- iii. Si $x_1 \leq x_2$, entonces $T(x_1) \leq T(x_2).$
- iv. $T(x)$ es continua por la derecha, es decir $\lim_{z \rightarrow x_+} T(z) = T(x).$

Además, recíprocamente una función $T(x): \mathbb{R} \rightarrow [0,1]$ se llama *función de distribución* si cumple las cuatro propiedades anteriores.

Así, a partir de su distribución, podemos clasificar a las variables aleatorias como *discretas*, si se corresponde a una función de distribución constante por pedazos, es decir, que puede tomar a lo más un número numerable de valores; o *continúas*, si su función de distribución correspondiente es continua. Otras variables recurrentes se comportan como discretas en unos intervalos y continuas en otros, las cuales se llaman variables aleatorias *mixtas*.

Esta función de distribución probabilística captura la medida de probabilidad de una familia de eventos de interés a partir de la construcción de una variable aleatoria que los mapea en conjuntos borelianos manteniendo la estructura de la sigma-álgebra de S en la sigma-álgebra de Borel. Por lo que, podemos hacer inferencia a cerca de este tipo de experimentos aleatorios al trabajar en un espacio real valuado y poder contrastar este modelo axiomático con los datos.

En nuestro análisis nos interesará en especial un tipo de variables aleatorias llamadas *absolutamente continuas*, las cuales se nombran así si existe una función no negativa e integrable f tal que para cualquier valor de x se cumple

$$T(x) = \int_{-\infty}^x f(u)du.$$

En este caso, a la función $f(x)$ la llamamos *función de densidad* de la variable aleatoria X .

Note que cualquier función de densidad marginal asociada a una variable aleatoria se comporta como una variable aleatoria distribuida uniformemente en el intervalo $[0,1]$, como en el siguiente lema, lo cual será importante al observar más adelante las características de las cópulas. La demostración de este lema la podemos encontrar en Spanos (1999).

LEMA 1: Sea $U = T(x)$ la función de densidad acumulada de una variable aleatoria continua X . Entonces U es una variable aleatoria que se distribuye uniformemente en el intervalo $I = [0,1]$.

Por otro lado, normalmente los fenómenos aleatorios en economía son multifactoriales, es decir, varios de los eventos que queremos modelar dependen de varios factores, por lo que nos gustaría ampliar nuestro modelo de probabilidad una sólo variable a uno multivariado, donde participen varias variables aleatorias para explicar nuestro fenómeno.

Comenzaremos por definir lo que es una función de distribución bivariada, lo cual va a servir de base para generalizar nuestro modelo.

DEFINICIÓN 4: Decimos que la función T es una función de distribución bivariada con dominio $\bar{\mathbb{R}}^2$ (donde $\bar{\mathbb{R}} = [-\infty, \infty]$) si cumple que:

1. T es 2-cresciente, es decir, dados cualesquiera dos puntos (x_1, x_2) y (y_1, y_2) en $\bar{\mathbb{R}}^2$, se tiene que

$$T(x_2, y_2) - T(x_2, y_1) - T(x_1, y_2) + T(x_1, y_1) \geq 0.$$

2. $T(x, -\infty) = T(-\infty, y) = 0$ para cualquier $x, y \in \bar{\mathbb{R}}$ (la función está fija).
3. $T(\infty, \infty) = 1$.

En este caso, se dice que la función de distribución cumple medir la probabilidad:

$$T(x, y) = P(X \leq x, Y \leq y).$$

Además, diremos que estas variables son independientes si cumplen que

$$T(x, y) = P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y).$$

Note que la función de distribución bivariada actúa sobre un vector aleatorio $\mathbf{X} = (X_1, X_2)$, el cual está formado por dos variables aleatorias X_1 y X_2 que se comportan probabilísticamente de manera conjunta, es decir, es un vector formado por dos variables aleatorias con una distribución conjunta que describe su comportamiento probabilístico.

Se puede probar además que, dados dos vectores aleatorios \mathbf{X} y \mathbf{Y} , entonces su función de distribución conjunta T cumple ser no decreciente, es decir, $T(\mathbf{X}) \leq T(\mathbf{Y})$, para todo $\mathbf{X}, \mathbf{Y} \in \text{Dom}(F)$ con $\mathbf{X} \leq \mathbf{Y}$,¹ vea Durante y Sempi (2010).

DEFINICIÓN 5: Dado un vector aleatorio (X_1, X_2) , con función de densidad conjunta $T(x_1, x_2)$, llamaremos función de distribución marginal de X_1 a la función definida como

$$T_1(x_1) = \lim_{x_2 \rightarrow \infty} T(x_1, x_2).$$

Similarmente, llamaremos función de distribución marginal de X_2 a la función

$$T_2(x_2) = \lim_{x_1 \rightarrow \infty} T(x_1, x_2).$$

Estas funciones marginales cumplen las propiedades de las funciones de distribución univariadas. Para el caso de vectores multidimensionales, se pueden obtener las marginales con un procedimiento similar, dejando una variable fija y tomando el límite cuando las demás tienden a infinito.

De manera natural se generalizan toda esta construcción a un espacio multidimensional donde los eventos de interés los podamos describir a través de vectores aleatorios de la forma (X_1, X_2, \dots, X_n) , los cuales tomarán eventos en un espacio de probabilidad multivariado y nos regresará conjuntos borelianos en \mathbb{R}^n . Así, un *vector aleatorio n-dimensional* es un arreglo formado por n -variables aleatorias, las cuales tienen un comportamiento estocástico conjunto.

DEFINICIÓN 6: Definimos la función de distribución conjunta del vector (X_1, X_2, \dots, X_n) como la probabilidad dada por:

$$T(x_1, x_2, \dots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n).$$

¹ En general, si \mathbf{X} y \mathbf{Y} son vectores aleatorios, entonces $\mathbf{X} \leq \mathbf{Y}$ sí y sólo si $X_i \leq Y_i$, para todo $i = 1, 2, \dots, n$.

Además, diremos que estas variables son independientes si cumplen que

$$T(x_1, x_2, \dots, x_n) = P(X_1 \leq x_1)P(X_2 \leq x_2) \cdots P(X_n \leq x_n).$$

Se puede probar fácilmente que la función de distribución conjunta cumple con las siguientes propiedades:

- i. $\lim_{x_1, x_2, \dots, x_n \rightarrow \infty} T(x_1, x_2, \dots, x_n) = 1.$
- ii. $\lim_{x_1, x_2, \dots, x_n \rightarrow -\infty} T(x_1, x_2, \dots, x_n) = 0.$
- iii. $T(x_1, x_2, \dots, x_n)$ es no decreciente y continua por la derecha en cada variable.
- iv. Si $x_i < y_i$, para $i = 1, 2, \dots, n$, entonces se cumple que

$$\sum_{z_i \in \{x_i, y_i\}} (-1)^{\#x} T(z_1, z_2, \dots, z_n) \geq 0$$

en donde $\#x$ es el número de veces que alguna de las variables z_i toman el valor x_i cuando evaluamos la función T .

Hasta aquí hemos descrito al modelo probabilístico para darle sustento a nuestro enfoque de reducción probabilística, por lo que procedemos ahora a construir el *Modelo de Muestreo* a partir del cual completamos la especificación de nuestro modelo estadístico, el cual nos permite estimar el mecanismo generador de información de nuestros datos, establecer un método de estimación de los parámetros y construir las pruebas de correcta especificación para asegurarnos de que tenemos un modelo consistente con nuestros datos.

Spanos (1986, 1999) muestra que se puede capturar la esencia del comportamiento estocástico los datos observacionales describiendo sus características de distribución, independencia y homogeneidad a través de este planteamiento metodológico de reducción probabilística, las cuales describimos a continuación:

1. La *distribución* establece que al tener observaciones de diversas repeticiones del fenómeno aleatorio, si estos datos tienen distribución formarán patrones estables.

2. La *independencia* caracteriza a las realizaciones que no tiene repercusión ni son repercutidas por alguna otra realización de los datos.
3. La *homogeneidad* se da si las probabilidades asociadas a varias realizaciones permanecen idénticas para todas las observaciones.

En nuestro caso, la distribución la describimos a partir de las definiciones y resultados anteriores, los cuales constituyen el modelo axiomático de probabilidad. Pero la dependencia y la homogeneidad corresponden al *modelo muestral*, el cual está dado por los vectores aleatorios X_1, X_2, \dots, X_m de nuestras variables y su comportamiento conjunto.

Por un lado, puede ser que la ocurrencia de los vectores se de manera *dependiente* o *independiente*, y que además, las distribuciones de los mismos varíen de un vector a otro, en cuyo caso diremos que son *heterogéneos*, pero si esto no ocurre diremos que los vectores son *homogéneos*. Ejemplos de este tipo de muestras se dan en los diferentes trabajos de series de tiempo, datos panel o datos que dependen de alguna característica como los de la econometría regional.

Dado una muestra de vectores aleatorios X_1, X_2, \dots, X_m , compuesto de m -vectores aleatorios, definimos la función de *distribución conjunta* de la muestra como su probabilidad acumulada hasta un punto $(x_1, x_2, \dots, x_m) \in \bar{\mathbb{R}}^m$, y está dada por:

$$T(x_1, x_2, \dots, x_m; \Theta) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_m \leq x_m; \Theta),$$

donde Θ es el conjunto de parámetros asociado a la distribución de este vector. Si esta muestra la constituyen vectores aleatorios independientes, su distribución conjunta es el producto de sus distribuciones marginales, es decir:

$$T(x_1, x_2, \dots, x_m; \Theta) = T_1(x_1)T_2(x_2) \cdots T_m(x_m).$$

Por cuestiones de simplificación del modelo, vamos a suponer que el conjunto de vectores de variables aleatorias de nuestra muestra cumplen homogeneidad y son independientes entre sí.

DEFINICIÓN 7: *Una muestra aleatoria es una colección de variables o vectores aleatorios X_1, X_2, \dots, X_m , que cumplen ser independientes y de tener cada uno de ellos la misma distribución. Al número m se le llama tamaño de la muestra aleatoria.*

Así, para hacer la identificación de los parámetros de nuestro modelo, vamos a suponer que nuestros datos provienen de una muestra aleatoria, es decir que los vectores aleatorios que la conforman son independientes e idénticamente distribuidos, la cual se constituye de m vectores aleatorios n -dimensionales X_1, X_2, \dots, X_m . Cada uno de estos vectores tienen cierta distribución conjunta $T(x_k)$ determinada por la relación estocástica entre las variables aleatorias que los componen, las cuales supondremos absolutamente continuas.

La *función de verosimilitud* en este caso está dada por

$$\mathcal{L}(x_1, x_2, \dots, x_m; \Theta) = f(x_1)f(x_2) \cdots f(x_m),$$

donde la función $f(x_k)$, para $k = 1, 2, \dots, n$, es la *densidad del vector aleatorio* X_k dada por

$$f(x_k) = \frac{\partial^n}{\partial x_k^1 \partial x_k^2 \cdots \partial x_k^n} T(x_k).$$

En nuestro caso, la distribución de nuestros vectores aleatorios se identificará a partir de los datos, determinando el tipo de distribución marginal de las variables aleatorias y la cópula que mejor captura la estructura de los datos de manera empírica como veremos más adelante. Además, supondremos que estos vectores conforman una muestra aleatoria. Así, a continuación vamos a dar el concepto de cópula y algunos resultados útiles para el propósito de este trabajo.

2.3 Introducción a la Teoría de Cópulas

La importancia de las cópulas en este trabajo es que a partir de ellas construimos la distribución conjunta de las variables que conforman los vectores aleatorios de nuestro modelo, por lo que nos permite realizar la especificación de nuestro modelo probabilístico. Pero además, al poder descomponer los efectos individuales de los colectivos en un vector aleatorio, la cópula nos permitirá estudiar la estructura del modelo de distribución de nuestros datos y del modelo muestral, lo cual permite nos facilitará realizar la identificación del modelo estadístico con el modelo económico de producción con innovaciones verdes.

Con la finalidad de dotar al lector con una introducción a la teoría de cópulas y sus aplicaciones en nuestro modelo, a continuación enunciamos algunos de sus principales conceptos y resultados más importantes de esta teoría. Comenzaremos con el estudio del caso bivariado para tener mayor claridad en la exposición y posteriormente generalizaremos al caso multivariado n -dimensional siguiendo los trabajos de Durante y Sempi (2010) y el de Nelsen (2006).

Un problema importante de la probabilidad multidimensional que estudiaba Fréchet (1951) era el poder describir al conjunto de todas las posibles funciones de densidad conjunta para un conjunto de variables aleatorias definidas en el mismo espacio de probabilidad con distribuciones marginales dadas. Una respuesta a este cuestionamiento la dio Sklar en su artículo de 1969, donde no solamente asegura que existen familias de funciones llamadas cópulas, las cuales al componerse funcionalmente con las distribuciones marginales de estas variables aleatorias generan distribuciones conjuntas, sino que de manera recíproca, dada una distribución conjunta de un conjunto de variables aleatorias y las distribuciones marginales de las mismas, existe una función cópula tal que al componerla con las distribuciones marginales nos dan la distribución conjunta de las mismas (Durante y Sempi, 2010).

Por otro lado observamos que al ser la cópula una función que conjunta a las densidades marginales de las variables aleatorias que participan en el modelo, esta determina la forma

en que las variables interaccionan entre ellas, es decir, le da una estructura matemática a la densidad conjunta con la cual determina la dependencia entre ellas. Esta es la razón por la que las medidas de dependencia más importantes como la τ de Kendall y la ρ de Spearman dependen únicamente de la cópula y no de la forma funcional ni de los parámetros de las marginales en muchos casos.

Así, el mayor número de aplicaciones de la teorías de cópulas se han ocupado de modelar dependencia estocástica y eventos extremos (dada la revisión bibliográfica que hicimos, la cual está resumida en la ruta de búsqueda de artículos en el apéndice B), por lo que han sido herramientas muy útiles para aplicaciones en finanzas, hidrología y ciencias ambientales, cálculo de riesgos, entre otras disciplinas.

En este sentido, en este trabajo utilizamos la teoría de cópulas para el proceso de especificación de nuestro modelo probabilístico, el cual nos permitirá, una vez correctamente especificado y estimados los parámetros, identificar sus componentes y sus parámetros estocásticos con los de la teoría económica, siguiendo la idea de dependencia probabilística entre las innovaciones que ahorran capital y las que ahorran trabajo planteada por Habakkuk (1962) y Growiec (2006).

De manera intuitiva, podemos definir a las *cópulas* como funciones que mezclan distribuciones marginales de variables aleatorias para obtener una función de distribución conjunta. Así, la cópula dota de una estructura matemática a la distribución conjunta de las variables, lo cual junto con la forma funcional de las marginales de las variables aleatorias, nos permitirá la identificación del modelo probabilístico que describe al fenómeno económico que pretendemos modelar. Así, estas funciones cópula deben cumplir con propiedades adecuadas tanto en sus límites como en su forma de mezclar variables.

Comenzamos definiendo formalmente la noción de volumen de una función. Lo cual captura la propiedad (1) de la definición de distribución conjunta.

DEFINICIÓN 8: Sean $S_1, S_2 \subset \overline{\mathbb{R}}$ conjuntos no vacíos y H una función cuyo dominio cumple $\text{Dom}H = S_1 \times S_2$; sea $B = [x_1, x_2] \times [y_1, y_2]$ un rectángulo cuyos vértices cumplen $x_1, x_2 \in S_1$ y $y_1, y_2 \in S_2$. Entonces el T -volumen de B se define como: $V_T(B) = T(x_2, y_2) - T(x_2, y_1) - T(x_1, y_2) + T(x_1, y_1)$

Así, el T -volumen es una función definida sobre un rectángulo, lo cual, recordando la definición de distribución bivariada de la definición 2, es importante para definir una función de distribución conjunta pues note que si T es una función de este tipo entonces debe cumplir que el T -volumen sobre cualquier rectángulo debe ser mayor o igual a cero.

Por otra parte, note que para obtener las demás características de las funciones de distribuciones bivariadas necesitamos que al componer esta función T con distribuciones marginales y la evaluemos obtengamos cero cuando alguna de las marginales valgan cero (su menor valor). Con esto mente, hacemos la siguiente definición.

DEFINICIÓN 9: Una función T es 2-creciente si $V_T(B) \geq 0$ para cualquier rectángulo B cuyos vértices están en el conjunto $\text{Dom}T$. Además, se dirá que T está fija si dados los valores $a_1 = \inf(S_1) \in S_1$ y $a_2 = \inf(S_2) \in S_2$, se cumple que, para todo $(x, y) \in S_1 \times S_2$, $T(x, a_2) = 0 = T(a_1, y)$.

Las definiciones 2 y 3 nos permiten definir una función que conecta a las distribuciones marginales para obtener una distribución bivariada. Por un lado, pedir que una función sea 2-creciente nos sirve para garantizar que el valor de la distribución sobre cualquier rectángulo es no negativo, y por otro, pedir que esté fija nos ayudará para asegurar que la probabilidad acumulada en los extremos inferiores del dominio de la función sea cero.

Por otro lado, trabajando en sentido inverso, dada una distribución bivariada conjunta, es importante poder obtener las distribuciones marginales de sus variables aleatorias así como la función que las combine para obtener dicha distribución bivariada. La siguiente definición nos ayuda en esta tarea ya que permite precisar el concepto de función marginal de una distribución conjunta.

DEFINICIÓN 10: Sean $a_1 = \inf(S_1) \in S_1$ y $a_2 = \inf(S_2) \in S_2$. Definimos las distribuciones marginales de la función $T: S_1 \times S_2 \rightarrow \mathbb{R}$ como las funciones F y G , con dominios S_1 y S_2 respectivamente, definidas como:

$$F(x) = T(x, a_2), \text{ para toda } x \in S_1.$$

$$G(y) = T(a_1, y), \text{ para toda } y \in S_2.$$

Así, se prueba utilizando la desigualdad del triángulo (vea Nelsen (2006)) que:

$$|T(x_2, y_2) - T(x_1, y_1)| \leq |F(x_2) - F(x_1)| + |G(x_2) - G(x_1)|.$$

Esta última desigualdad es del tipo “desigualdad de Lipschitz”, la cual nos importa pues nos indica que las distribuciones bivariadas resultan ser uniformemente continuas y cumplen con condiciones de convexidad que se analizarán más adelante a través de las cópulas.

Ahora formalizamos el concepto matemático de cópula bidimensional, el cual es uno de los pilares de nuestro trabajo. El concepto de cópula está ligado totalmente a la definición de función de distribución conjunta, de forma tal que al combinarla con las marginales de las variables ésta cumpla con todas las propiedades descritas en la definición 1.

DEFINICIÓN 11: Una cópula bidimensional o 2-cópula es una función $C: I^2 \rightarrow I$ con las siguientes propiedades:

- i. C está fija y es 2-creciente, es decir, para todo $u, v \in I$ y todo $u_1, u_2, v_1, v_2 \in I$ tales que $u_1 \leq u_2$ y $v_1 \leq v_2$, se cumple que

$$C(u, 0) = 0 = C(0, v)$$

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0$$

- ii. Para todo $u, v \in I$, se tiene que $C(u, 1) = u$ y $C(1, v) = v$.

Note que (i) asegura las propiedades de distribución conjunta que habíamos planteado anteriormente, mientras que (ii) tiene que ver con la marginalización de las distribuciones univariadas de las variables aleatorias del modelo de distribución, el cual equivale a sumar (en el caso discreto) o integrar (en el caso continuo) sobre todos los posibles valores de una de las variables aleatorias para obtener la distribución marginal de la otra.

Algunos ejemplos de cópulas cotidianamente utilizados son los siguientes:

a) La cópula producto $C(u, v) = uv$, la cual denota independencia entre sus variables.

b) La cópula normal dada por

$$C_R(u, v) = \int_{-\infty}^{\theta^{-1}(v)} \int_{-\infty}^{\theta^{-1}(u)} \frac{1}{2\pi(1-\rho^2)} \exp\left[-\frac{s^2 - 2\rho st + t^2}{2(1-\rho^2)}\right] ds dt$$

donde θ^{-1} es la función inversa de la normal estándar.

c) Cópula Clayton definida como

$$C(u, v) = \max\left\{[u^{-\theta} + v^{-\theta} - 1]^{-\frac{1}{\theta}}; 0\right\}; \theta \in [-1, \infty] \setminus \{0\}.$$

d) Cópula Gumbel-Hougaard dada por

$$C(u, v) = \exp\{-[(-\ln u)^\alpha + (-\ln v)^\alpha]^{1/\alpha}\}, \quad \alpha \geq 1.$$

Esta definición de función cópula se puede generalizar para una función de varias variables, es decir, una cópula n -dimensional o n -cópula es una función $C: I^n \rightarrow I$ que cumple las siguientes propiedades:

- i. Para cada $i = 1, 2, \dots, n$, $C(u) = u_i$ cuando cada componente de u es igual a 1, excepto la componente i -ésima que vale $u_i \in I$.

ii. C es isotónica, es decir, para todo $u, v \in I$ tales que $u \geq v$ se tiene que $C(u) \geq C(v)$.

iii. C es n -creciente, es decir el C -volumen de cualquier caja en I^n es positivo

$$\sum_{z_i \in \{u_i, v_i\}} (-1)^{\#u} C(z_1, z_2, \dots, z_n) \geq 0,$$

en donde $\#u$ es el número de veces que alguna de las variables z_i toman el valor u_i cuando evaluamos la cópula C .

Un ejemplo de este tipo de cópulas está dado por la *Cópula Clayton Trivariada Anidada* definida como

$$C(u, v, w) = \max \left\{ \left[u^{-\theta} + [v^{-\mu} + w^{-\mu} - 1]^{\frac{\theta}{\mu}} - 1 \right]^{-\frac{1}{\theta}} ; 0 \right\}; \theta, \mu \in [-1, \infty] / \{0\}.$$

Así, una característica importante de las cópulas es que están acotadas por arriba y por abajo por funciones llamadas Cotas de Fréchet-Hoeffding, las cuales resultan ser también cópulas en el caso bidimensional. Esta propiedad es necesaria y permite poder construir una función de distribución multivariada a partir de una cópula.

En el siguiente lema se establece la desigualdad de Fréchet-Hoeffding, la cual da cuenta de esto para el caso bidimensional.

LEMA 2 (Desigualdad de Fréchet-Hoeffding): Sea C una cópula bivariada, entonces para cada $(u, v) \in I \times I$, se cumple la desigualdad

$$\max\{u + v - 1, 0\} \leq C(u, v) \leq \min\{u, v\}.$$

La desigualdad sigue siendo válida para el caso n -dimensional. Sea C una cópula n -dimensional y dados u_1, u_2, \dots, u_n en I^n , se cumple la siguiente desigualdad

$$\max\{u_1 + u_2 + \dots + u_n - n + 1, 0\} \leq C(u_1, u_2, \dots, u_n) \leq \min\{u_1, u_2, \dots, u_n\}.$$

Así entonces, estas cotas inferior $W(u_1, u_2, \dots, u_n) = \max\{u_1 + u_2 + \dots + u_n - n + 1, 0\}$ y superior $M(u_1, u_2, \dots, u_n) = \min\{u_1, u_2, \dots, u_n\}$ de Frèchet-Hoeffding acotan cualquier cópula n-dimensional, pero sólo $M(u_1, u_2, \dots, u_n)$ sigue siendo una cópula para el caso de más de dos variables [vea Nelsen (2006)]. Esto también ocurre con otras cópulas, por lo que debemos ser cuidadosos al generalizar nuestros conceptos.

El siguiente resultado nos indica una propiedad importante de las distribuciones bivariadas, la cual se conoce como Desigualdad de Lipschitz. Así, las cópulas cumplen ser uniformemente continuas en su dominio, ya que sus imágenes están más cercanas entre sí que los puntos de su dominio.

LEMA 3: *Sea C una cópula, entonces para cada par (u_1, u_2) y (v_1, v_2) en su dominio tenemos que*

$$|C(u_2, v_2) - C(u_1, v_1)| \leq |u_2 - u_1| + |v_2 - v_1|.$$

Esta propiedad también se cumple para las n-cópulas, es decir, $\forall u, v \in I^n$, se tiene que

$$|C(u) - C(v)| \leq \sum_{i=1}^n |u_i - v_i|.$$

Esta desigualdad permite obtener varias propiedades topológicas tales como la convexidad y el poder acotar a las cópulas a través de funciones con características muy específicas como las distribuciones marginales, lo cual resulta útil para la descripción y análisis de las propiedades de las mismas.

Nuestro siguiente resultado, demostrado en Sklar (1959), es uno de los teoremas más importantes de la teoría de cópulas y nos ayuda a identificar el modelo probabilístico y hacer la identificación con el modelo económico suponiendo independencia y

homogeneidad en la muestra aleatoria y centrándonos en su distribución, ya que nos garantiza dos hechos fundamentales.

Por un lado nos permite construir funciones de distribución multivariadas dada una cópula y las distribuciones marginales de las variables aleatorias de nuestro modelo, debido a que la función que se obtiene al componer las marginales con la cópula es una función de distribución conjunta. Además, de manera inversa, nos asegura que dada una función de distribución multivariada conjunta y sus marginales, existe una cópula tal que al componerla con estas marginales obtenemos dicha función de distribución multivariada conjunta.

Igual que anteriormente, comenzaremos la versión bivariada del teorema, para identificarlo con el modelo de dos distribuciones conjuntas, y después generalizaremos al caso multivariado.

TEOREMA DE SKLAR (Sklar, 1959): *Sea T una función de distribución bivariada conjunta con marginales F y G . Entonces existe una cópula C tal que, para todo $x, y \in \overline{\mathbb{R}}$, tenemos que*

$$T(x, y) = C(F(x), G(y)).$$

Si F y G son continuos, entonces C es única; en caso contrario, C esta determinada de manera única en $\text{Ran}F \times \text{Ran}G$. De manera inversa, si C es una cópula y F y G son funciones de distribución, entonces la función $H(x, y) = C(F(x), G(x))$ es una función de distribución conjunta con marginales F y G .

Así, por ejemplo, dada a cópula $W(u, v) = \max\{u + v - 1, 0\}$ y las siguientes distribuciones marginales:

$$F(x) = P(X \leq x) = 1 - e^{-\varphi x}, \text{ para } x \geq 0 \quad (\text{Exponencial})$$

$$G(y) = P(Y \leq y) = 1 - \left(\frac{y}{\gamma y}\right)^{-\alpha}, \text{ para } y \geq \alpha \quad (\text{Pareto})$$

Entonces la función que resulta de la composición de estas marginales con la cópula, dada por

$$W(F(x), G(y)) = \max \left\{ 1 - e^{-\varphi x} - \left(\frac{y}{\gamma y} \right)^{-\alpha}, 0 \right\}, \text{ para } x \geq 0, y \geq \alpha,$$

resulta ser una función de densidad bivariada conjunta.

El caso multivariado, suponiendo que se tienen n variables que conforman mi vector aleatorio, es el siguiente.

TEOREMA DE SKLAR (Caso multivariado): *Sea C una n -cópula y sean U_1, U_2, \dots, U_n distribuciones marginales. Entonces existe una función de distribución multivariada T con marginales U_1, U_2, \dots, U_n , tal que para todo $x_1, x_2, \dots, x_n \in \overline{\mathbb{R}}$, se tiene que:*

$$T(x_1, x_2, \dots, x_n) = C(U_1(x_1), U_2(x_2), \dots, U_n(x_n); \theta).$$

De manera inversa, si T es una función de distribución conjunta con marginales U_1, U_2, \dots, U_n . Entonces, para todo vector aleatorio (x_1, x_2, \dots, x_n) in $\overline{\mathbb{R}}^n$ existe una n -cópula C que cumple la ecuación anterior. Si las marginales U_1, U_2, \dots, U_n son todas continuas, entonces C es única; de otra forma, C esta determinada de manera única en su rango $\text{Ran}(U_1) \times \text{Ran}(U_2) \times \dots \times \text{Ran}(U_n)$.

Existen familias importantes de cópulas por sus propiedades y uso, como las elípticas, dentro de las cuales ésta la Cópula Normal, o las EFMG construidas a partir de marginales y sus marginales de supervivencia. Para nuestro trabajo nos concentramos en la familia de *Cópulas Arquimedianas*, las cuales se definen a través de su función llamada generador de la cópula como sigue.

DEFINICIÓN 12: *Una n -cópula C se llama arquimediana si admite la representación*

$$C(u) = \psi(\psi^{-1}(u_1) + \psi^{-1}(u_2) + \dots + \psi^{-1}(u_n)),$$

para todo $u \in I^n$, donde la función convexa ψ se conoce como generador de C y cumple con $\psi(1) = 0$ y admitir pseudoinversa.

Ejemplos de este tipo de cópulas con la Gumbel y la Clayton con generadores dados respectivamente por

$$\psi(t) = [-\ln(t)]^\omega \quad \text{y} \quad \psi(t) = \frac{t^\omega - 1}{\omega}.$$

Finalmente abordaremos el tema de las medidas de dependencia entre las variables de un modelo probabilístico, entre las cuales destacan la correlación de Pearson, la Rho Spearman y la Tao de Kendall.

Note que al ser *medidas de dependencia estandarizadas*, mientras mayor dependencia positiva tengan estas variables, sus medidas de dependencia alcanzarán valores positivos más altos, y entre más dependencia negativa haya entre estas, su coeficiente negativo de dependencia será también más grande (en valor absoluto). Además, cuando las variables aleatorias sean independientes, el valor de su coeficiente de dependencia será cero.

Una buena medida de dependencia cuando las variables aleatorias presentan una relación lineal es el *coeficiente de correlación de Pearson*, el cual se define para un vector de variables aleatorias $(X, Y)^T$ distinto de cero y con varianzas finitas como:

$$\rho(X, Y) = \frac{E(XY) - E(X)E(Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}.$$

Para casos más generales usamos la noción de concordancia, la cual es una medida de afinidad entre las variables aleatorias. Formalmente, dadas dos observaciones (x_i, y_i) y (x_k, y_k) de un vector aleatorio (X, Y) decimos que estas son *concordantes* si cumplen que $(x_i - x_k)(y_i - y_k) > 0$ y que son *discordantes* si $(x_i - x_k)(y_i - y_k) < 0$.

Una medida de concordancia es a *Tao de Kendall*, la cual se define como la probabilidad de que concordancia menos la probabilidad de discordancia:

$$\tau_c(X, Y) = P((X_1 - X_2)(Y_1 - Y_2) > 0) - P((X_1 - X_2)(Y_1 - Y_2) < 0).$$

Luego, su medida muestral está dada por

$$t_c = \frac{c - d}{c + d} = \frac{c - d}{\binom{n}{2}}$$

donde c y d son el número de pares concordantes y discordantes respectivamente y n es el número total de observaciones, por lo que $\binom{n}{2}$ es el total de pares.

Se prueba que la *versión poblacional* de Tao de Kendall está dada por:

$$\tau_c(X, Y) = 4 \iint_{[0,1]^2} C(u, v) dC(u, v) - 1 = 4E[C(U, V)] - 1$$

donde $U, V \sim U(0,1)$ con distribución conjunta descrita por la cópula C .

Además, en particular, para el caso de que la cópula sea del tipo Arquimediana con generador $\psi(t)$, se tiene que la Tao de Kendall está dada por

$$\tau_c = 1 + 4 \int_0^1 \frac{\psi(t)}{\psi'(t)},$$

donde t es una variable auxiliar que representa la distribución marginal.

Por otra parte, la *Rho de Spearman* también es una medida de dependencia basada en la concordancia, la cual tiene por rango al intervalo $[-1,1]$ y su interpretación es similar al del coeficiente de correlación de Pearson. La versión muestral de la Rho de Spearman es:

$$\rho_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)};$$

donde d_i representa la diferencia entre ordenamientos y n es el número de casos totales de la muestra. Formalmente, sean $(X_1, Y_1), (X_2, Y_2)$ y (X_3, Y_3) vectores aleatorios

independientes con una función de distribución conjunta común H y cópula C , la *versión poblacional* de la *Rho Spearman* tiene la siguiente expresión:

$$\rho_S(X, Y) = 3(P\{(X_1 - X_2)(Y_1 - Y_3) > 0\} - P\{(X_1 - X_2)(Y_1 - Y_3) < 0\})$$

Luego entonces, se prueba que se puede calcular mediante la formula

$$\rho_S(X, Y) = 12 \iint_{I^2} uv dC(u, v) - 3 = 12 \iint_{I^2} C(u, v) dudv - 3.$$

Note que las medidas de dependencia descritas anteriormente dependen de la cópula que se usa y sus parámetros y además, al cálculo de una u otra se hace más sencillo o más complicado de acuerdo a las formas funcionales de las cópulas que se emplean, por lo que podemos ocupar una media u otra dependiendo de qué nos facilite más los cálculos y ayude en el proceso de identificación de nuestro modelo.

2.4 El Enfoque de Reducción Probabilística con Teoría de Cópulas

En esta sección, finalmente vamos a definir el enfoque de reducción probabilístico probabilística de Spanos (1986, 1999), Spanos y McGuirk (2001), usando la teoría de cópulas para la especificación del modelo muestral y hablar de las implicaciones que tiene esto ya que esta nueva especificación generaliza el enfoque anterior y nos va a permitir implementar dos cambios respecto al enfoque sin cópulas. Por un lado, nos permite descomponer a la distribución probabilística de nuestro modelo en dos partes, una que captura el efecto marginal de cada una de las variables aleatorias que componen el modelo y la otra conjunta que dota lo de estructura y captura el efecto conjunto de las mismas.

Por otra parte, la teoría de cópulas amplía el espectro de distribuciones que se pueden ocupar para la modelación, pudiendo incluso construir una nueva distribución que capture de mejor manera el comportamiento estocástico de los datos, combinando distribuciones marginales de las variables aleatorias del modelo a través de una función cópula que las

agrupa en una distribución conjunta. Todo esto es posible gracias al teorema de Sklar, el cual va a resultar crucial para nuestro proceso de especificación y por consiguiente para el de identificación.

Formalmente, la especificación de nuestro modelo estadístico con cópula se compone de los siguientes elementos:

- i. Un *modelo probabilístico* para cada vector aleatorio X^k dado por:

$$\Phi \equiv \{T(x_1^k, x_2^k, \dots, x_n^k) = C(U_1(x_1^k), U_2(x_2^k), \dots, U_n(x_n^k); \theta); x_i^k \in \mathbb{R}, \forall i, C \in \mathcal{H}\},$$

donde \mathcal{H} es una familia de cópulas Arquimedianas y las funciones U_i son marginales en general de transformaciones de las variables aleatorias x_i^k , para todo $i = 1, 2, \dots, n$.

- ii. Un *modelo muestral* dado por una muestra aleatoria $\mathbf{X} := \{X^1, X^2, \dots, X^m\}$, la cual es independiente e idénticamente distribuida, por lo que la distribución individual para cada vector aleatorio $X^k = (X_1^k, X_2^k, \dots, X_n^k)$ está definida por la función

$$T(x_1^k, x_2^k, \dots, x_n^k) = C(U_1(x_1^k), U_2(x_2^k), \dots, U_n(x_n^k); \theta), \forall k = 1, 2, \dots, m.$$

Es decir, para hacer la identificación de los parámetros de nuestro modelo, vamos a suponer que nuestros datos provienen de una muestra aleatoria \mathbf{X} compuesta de vectores aleatorios independientes e idénticamente distribuidos $\{X^1, X^2, \dots, X^m\}$, cada uno de los cuales tienen cierta distribución conjunta $T(x_1^k, x_2^k, \dots, x_n^k)$ determinada por la relación estocástica entre las variables aleatorias X_1^k que los componen. Vamos a suponer que estas variables aleatorias son absolutamente continuas.

A partir de esta especificación obtenemos la función de verosimilitud $\mathcal{L}(\theta)$ para estimar los parámetros de nuestro modelo, que en este caso está dada por

$$\mathcal{L}(\theta) = \prod_{k=1}^m \left[\frac{\partial^n}{\partial x_1^k \partial x_2^k \dots \partial x_n^k} C(U_1(x_1^k), U_2(x_2^k), \dots, U_n(x_n^k); \theta) \right].$$

En nuestro caso, la distribución de nuestros vectores aleatorios se identificará a partir de los datos, determinando el tipo de distribución marginal de las variables aleatorias y la cópula que mejor captura la estructura de los datos de manera empírica como veremos más adelante. Además, supondremos que estos vectores conforman una muestra aleatoria

Una vez que tenemos la especificación del modelo y la estimación de los datos, se procede a realizar una prueba de correcta especificación, que al pasarla nuestro modelo nos permitirá estar seguros de que nuestro mecanismo generador de información es congruente con los datos. En caso de que nuestro modelo no pase la prueba de correcta especificación, se procede a hacer una re-especificación del mismo y a seguir los pasos anteriores nuevamente. Pero si pasa las pruebas, el siguiente paso será continuar con el proceso de identificación de los parámetros y formas funcionales obtenidas de nuestro modelo probabilístico con el modelo económico que estamos especificando.

Por último, una vez que tenemos bien especificado el modelo estadístico, procedemos a identificar el modelo de la función de producción con cambio tecnológico de nuestra economía. En nuestro caso queremos obtener la *función de producción con innovaciones verdes* a partir de este proceso de identificación, para posteriormente deducir las características y poder hacer inferencia con este modelo [Hildenbrand (1981), Jones (1999, 2005), Growiec (2006, 2017)].

Para el proceso de identificación, las medidas de dependencia enunciadas anteriormente, junto con los parámetros de las cópulas y sus características, serán muy importantes, ya que haciendo uso del teorema de Sklar y la especificación muestral se va a construir un puente con el modelo económico de producción, para el cual la cópula le dotará de una estructura matemática.

2.6 Comentarios Generales

Note que el enfoque de reducción probabilística es una herramienta indispensable para la modelación de datos observacionales, ya que nos permite construir nuestro modelo partiendo completamente del comportamiento estocástico de los datos y llegando a un modelo mediante el proceso de identificación, por lo que los modelos construidos a partir de este enfoque son congruentes con la naturaleza observacional de los datos y con la teoría de probabilidad y estadística. Además que este enfoque no requiere de hacer manipulación de los datos, permitiendo la construcción de pruebas de correcta especificación y el uso de métodos de estimación de parámetros.

Por otro lado, el enfoque de reducción probabilística con cópulas generaliza al modelo original propuesto por Spanos (1986,1999) ya que las cópulas nos permiten obtener diversas especificaciones del modelo probabilístico que podemos elegir entre muchas familias de cópulas y combinarlas con cualquier distribución marginal que modele el comportamiento de nuestros datos, incluso podemos elegir variables con distintos tipos de distribución y obtener una distribución conjunta, por lo que podemos incluso crear alguna especificación no usada anteriormente para este tipo de modelos económicos.

Por último, es importante resaltar que se hizo una búsqueda sistematizada de libros y artículos de investigación centrados en los temas de crecimiento económico endógenos, modelos de producción con innovaciones tecnológicas e innovaciones verdes, la modelación estocástica de la producción, las aplicaciones de la teoría de cópulas y la combinación de estos tópicos anteriores con el fin de conocer el estado del arte en estos temas. Esta búsqueda ayudó a complementar la bibliografía sugerida por el comité tutorial, a partir de la cual también se siguió el camino de sus aplicaciones en otros trabajos. Así, se hizo la búsqueda tanto por palabras clave del tema de investigación como de términos o frases para hacer la búsqueda más especializada.

CAPITULO 3: UNA PROPUESTA DE MODELACION DE LA FUNCION DE PRODUCCION BASADO EN EL ENFOQUE DE REDUCCION PROBABILISTICA CON CÓPULA

3.1 Introducción

Una de las grandes causas del cambio tecnológico y de las desigualdades en la tecnificación de la industria entre los países en el tiempo es la disponibilidad de los activos de mano de obra y maquinaria, por lo que al potenciar a los insumos de la producción, se puede considerar al progreso tecnológico como motor de cambio en los procesos productivos, el cual surge por la acción deliberada de agentes económicos, públicos o privados, a través de la investigación para la innovación y el desarrollo, y es respaldada por un interés de carácter económico (Habakkuk, 1962). Así, en esta tesis construiremos la *especificación estadística* de un modelo de función de producción cuyo progreso tecnológico está dado por innovaciones verdes descritas a partir de la información empírica relevante de este proceso y haremos la *identificación* de los parámetros y las formas funcionales entre estos dos modelos.

Una vez que especificamos el modelo estadístico a partir de las características de los datos, vamos a estimar los parámetros del mismo mediante algún método adecuado, y posteriormente a comprobar la correcta especificación del mismo mediante pruebas estadísticas del modelo. Finalmente, procedemos al *proceso de identificación* con el modelo teórico-económico que pretendemos modelar, por lo que en este punto debemos hacer algunas suposiciones teóricas y definir un modelo marco con el cual vamos a identificar a nuestro modelo estadístico.

Nuestro supuesto de fundamental para esta parte es que los datos observacionales de los cuales partimos, compuestos por las innovaciones de capital b_i y trabajo a_i y el producto Y_i , nos describirán por un lado las combinaciones de innovaciones que producen determinada cantidad de producto para valores fijos dados de los insumos capital K y trabajo L , así como por otro lado, todas las posibles combinaciones de pares tecnológicos (a_i, b_i)

disponibles para una empresa. Así entonces, fijando los valores de la probabilidad podemos obtener los componentes que caracterizan a la producción sujeta al proceso de innovación en una economía [Aghion y Howitt (2009), Basu y Weil (1998), Lucas (1988), Romer (1986, 1990), Growiec (2006), Sánchez y Márquez (2013)].

3.2 Enfoque de Reducción Probabilística del Modelo General

En este apartado desarrollaremos en primer lugar la *especificación estadística* de nuestro modelo basados en los datos del mismo, después haremos la *estimación* de sus parámetros y formas funcionales, seguido de la descripción de las *pruebas de correcta especificación* para concluir con el proceso de *identificación* para un modelo de producción con innovaciones tecnológicas.

Así, vamos a comenzar especificando el *modelo probabilístico* de los datos, para lo cual inicialmente supondremos que las innovaciones tecnológicas y la producción se comportan como variables aleatorias con ciertas funciones de distribución marginal caracterizadas de acuerdo a la evidencia empírica.

SUPUESTO 1: *Existe un espacio de probabilidad (Ω, \mathcal{F}, P) , donde están definidos los vectores aleatorios (a_i, b_i, Y_i) para la i -ésima técnica de producción en una economía competitiva, el cual incluye al producto Y_i , y a las innovaciones tecnológicas (a_i, b_i) asociadas a ciertos niveles dados de trabajo L y capital K .*

Note que este supuesto nos obliga a asumir también que existen funciones de distribución marginales asociadas a cada una de las variables aleatorias a_i , b_i y Y_i . Así, dado un vector de variables aleatorias (a_i, b_i, Y_i) con distribuciones marginales conocidas (o marginales de alguna transformación monótona de ellas) dadas por $U(a; \theta_a)$, $V(b; \theta_b)$ y $W(Y; \theta_Y)$, nos preguntamos cuál es la función de distribución conjunta $T(a, b, Y)$ que permita capturar el comportamiento estocástico de nuestros datos.

Sabemos por un lado que si podemos determinar una cópula trivariada que refleje la relación de dependencia entre las marginales de nuestras variables, el Teorema de Sklar (1959) nos garantiza que existe una función de distribución conjunta T con marginales U, V y W , definida por

$$(1) \quad T(a, b, Y) = C(U(a; \theta_a), V(b; \theta_b), W(Y; \theta_Y), \theta),$$

la cual describirá el comportamiento estocástico de nuestros datos. Además, este mismo teorema nos garantiza que si podemos estimar una función de distribución conjunta $T(a, b, Y)$ que describa el comportamiento estocástico de nuestros datos, entonces podemos encontrar una cópula trivariada C , definida como en la ecuación anterior, que permite asociar a las marginales U, V y W con dicha distribución.

Así, de la composición de la probabilidad conjunta, siempre es posible descomponer a la distribución conjunta para capturar dos tipos de efectos, el efecto individual dado por las funciones marginales U, V y W , y el efecto conjunto dado por la cópula C (Nelsen, 2006). Siguiendo el desarrollo del modelo probabilístico del capítulo anterior y basados en la evidencia empírica en Aoyama (2010) y los trabajos de Growiec (2006, 2017) y Sánchez y Márquez (2015), sean $Z_1 = Z_1(a)$, $Z_2 = Z_2(b)$ y $Z_3 = Z_3(Y)$, transformaciones monótonas de las innovaciones y la producción, entonces calculamos mediante algún método adecuado² las distribuciones de estas variables dadas por

$$U_{Z_1}(a) = U(Z_1(a); \theta_a), V_{Z_2}(b) = V(Z_2(b); \theta_b), \text{ y } W_{Z_3}(Y) = W(Z_3(Y); \theta_Y).$$

Así la para obtener la densidad conjunta de nuestras nuevas variables podemos definir la cópula que modelo su dependencia con esta expresión la podemos reescribir como la distribución conjunta de las variables aleatorias auxiliares de la siguiente forma

$$T(a, b, Y) = C[U(Z_1(a); \theta_a), V(Z_2(b); \theta_b), W(Z_3(Y); \theta_Y), \theta] = C[U_{Z_1}(a), V_{Z_2}(b), W_{Z_3}(Y)].$$

² Podemos emplear al método de las funciones de distribución, al de las transformaciones o la de la función generadora de momentos.

Así, la densidad conjunta escrita en términos de las variables auxiliares anteriores está dada por la expresión:

$$f(z_{1i}, z_{2i}, z_{3i}) = \frac{\partial^3 [C[U_{Z_1}(a), V_{Z_2}(b), W_{Z_3}(Y)]]}{\partial U_{Z_1} \partial V_{Z_2} \partial W_{Z_3}} \frac{\partial U_{Z_1}}{\partial Z_1} \frac{\partial V_{Z_2}}{\partial Z_2} \frac{\partial W_{Z_3}}{\partial Z_3},$$

Note además que las diferentes distribuciones conjuntas que pueden ajustar para la correcta especificación del modelo probabilístico de nuestras variables no sólo depende de las formas funcionales de la cópula C y sus marginales U, V y W , sino también de los valores de la familia de parámetros de estas funciones $\theta, \theta_a, \theta_b$ y θ_Y , por lo que la estimación de los parámetros resulta crucial para el proceso de identificación de nuestro modelo estadístico con el teórico-económico. Esta última expresión es la especificación del modelo probabilístico basado en los datos.

Luego, vamos ahora a especificar el *modelo muestral* de nuestros datos basados en la evidencia empírica sobre la distribución de probabilidad de las innovaciones tecnológicas siguiendo el trabajo de Aoyama (2010) y Spanos y McGuirk (2001). Note que en general podemos suponer que estos vectores aleatorios que constituyen la muestra de tamaño m dada por

$$\{Z_k\}_{k=1}^m = \{(Z_{1k}, Z_{2k}, Z_{3k})\}_{k=1}^m$$

es una colección de vectores aleatorios que pueden tener cierta dependencia entre sí, es decir, que se les puede especificar una que los relacione de acuerdo a la forma de la dependencia entre ellos. Además, la muestra también puede en general presentar heterogeneidad entre sus componentes, por lo que entre un vector aleatorio y otro los parámetros y las formas funcionales pueden cambiar.

El caso más usado, por cuestiones de simplicidad, pero también porque que nace a partir de observar el comportamiento de los vectores aleatorios de la muestra, es que haya independencia e idéntica distribución y es el que vamos a adoptar.

SUPUESTO 2: La muestra de tamaño m que describe nuestro modelo muestral dada por $\{Z_k\}_{k=1}^m = \{(Z_{1k}, Z_{2k}, Z_{3k})\}_{k=1}^m$ es una colección de vectores aleatorios independientes e idénticamente distribuidos.

Así, basados en el supuesto anterior, la función de verosimilitud de nuestro modelo está dada por

$$\mathcal{L}(\theta | Z_k = z_k) = \prod_{k=1}^m P(Z_k = z_k | \theta) = \prod_{k=1}^m f(z_k | \theta),$$

donde para cada valor k , el vector $z_k = (z_{1k}, z_{2k}, z_{3k})$ es una realización del vector aleatorio $Z_k = (Z_{1k}, Z_{2k}, Z_{3k})$. Así, tomando logaritmos tenemos

$$\log[\mathcal{L}(\theta | Z_k = z_k)] = \sum_{i=1}^m \log \left[\frac{\partial^3 [C[U_{Z_1}(a), V_{Z_2}(b), W_{Z_3}(Y)]]}{\partial U_{Z_1} \partial V_{Z_2} \partial W_{Z_3}} \frac{\partial U_{Z_1}}{\partial Z_1} \frac{\partial V_{Z_2}}{\partial Z_2} \frac{\partial W_{Z_3}}{\partial Z_3} \right].$$

Una vez que hemos completado el proceso de especificación de nuestro enfoque de reducción probabilística, podemos proceder a hacer la *estimación de los parámetros* de nuestro modelo maximizando la función anterior, así como a plantear las *pruebas de correcta especificación* del mismo, las cuales dependen de las formas funcionales de las distribuciones y de los parámetros obtenidos de los datos de nuestro modelo. Si estas pruebas no se superan, quiere decir que nuestro modelo no se especificó correctamente y no captura el comportamiento estocástico de nuestros datos de manera adecuada, por lo que es necesario regresar al proceso de especificación y re-especificar el modelo.

Sin embargo, si nuestro modelo pasa las pruebas de correcta especificación, entonces se continua con el *proceso de identificación* de nuestro modelo estadístico con el modelo teorico-económico a través de establecer ciertas relaciones entre las formas funciones y los parámetros de ambos modelos. Para este fin, definimos primero el marco de referencia en el que nos vamos basar para identificar nuestro modelo.

SUPUESTO 3: *Formalmente, suponemos que en nuestra economía hay un conjunto J de empresas, indexadas por $j \in J \subseteq \mathbb{R}$, las cuales tienen a su disposición un conjunto I de técnicas para la producción, indexadas por $i \in I \subseteq \mathbb{R}$, las cuales dependen de las innovaciones tecnológicas que ahorran capital b_i y trabajo a_i en un punto dado del tiempo $t \in [0, \tau]$.*

Por lo que una vez que elegimos una empresa para su modelación, vamos a fijar los valores de los insumos de capital K y trabajo L respectivamente, puesto que queremos modelar solamente la relación de la producción con las innovaciones para después describir la toma de decisiones sobre la tecnología para la producción.

Note que, dado que nuestra especificación muestral la constituye un conjunto de vectores aleatorios independientes entre sí y homogéneos respecto a su distribución, basta con identificar el modelo estadístico con el teórico-económico para una sola empresa, ya que todas las demás tendrían la misma identificación.

Así, partimos de la *especificación probabilística usando cópulas* dada por la ecuación

$$T(a, b, Y) = C(U(a; \theta_a), V(b; \theta_b), W(Y; \theta_Y), \theta),$$

la cual nos indica las combinaciones de innovaciones que producen determinada cantidad de producto para valores dados de insumos, por lo que si dejamos fijos el nivel de probabilidad $T(a, b, Y)$ a un valor dado, la especificación probabilística de nuestro modelo nos permite describir todas las interacciones entre las innovaciones y la producción, para niveles fijos de insumos L y K respectivamente. Podemos entonces obtener las características de la producción sujeta a la innovación en una economía, es decir, a partir de ésta podemos describir la función de producción para nuestra empresa j -ésima, la cual llamaremos *función de producción local* (FPL) descrita a partir de la relación entre los datos de las variables aleatorias consideradas en nuestro problema.

Además, siguiendo el trabajo de Caselli y Cólman (2006) y Sánchez y Márquez (2013), al provenir nuestros datos de la producción de una economía, si fijamos un nivel de

producción Y , esta especificación contiene la información de todos los pares tecnológicos disponibles para la j -ésima empresa, al cual le llamamos *frontera tecnológica* (FT). La siguiente definición de Caselli y Coleman (2006) es muy útil para la derivación de la FT a partir de la especificación probabilística anterior.

DEFINICIÓN 13: *La Frontera Tecnológica es una curva en el espacio (a, b) tal que la probabilidad $P(a_i > a, b_i > b)$ es constante a lo largo de la curva, para niveles dados de Y, K y L .*

Basados en estas características de la función de probabilidad obtenida en la especificación del modelo probabilístico, tenemos la siguiente proposición.

PROPOSICIÓN 1: *La ecuación (1) obtenida a partir de la cópula C describe al conjunto de planes factibles de producción, por lo que podemos derivar la FPL y la FT como sigue:*

- a) *La FPL se obtiene eligiendo valores específicos para el conjunto de parámetros Θ de C y fijando un valor P_0 de la probabilidad $T(a, b, Y)$ a un nivel de producto dado.*
- b) *La FT se obtiene eligiendo valores específicos de los parámetros Θ de C , fijando el valor de la densidad marginal del producto en $W=1$ y fijando la probabilidad P_1 de la distribución asociada $T(a, b) = C(U(a), V(b), 1)$ a un nivel dado de eficiencia en la producción.*

PRUEBA: Vea apéndice A.1.

Note que aunque la FPL y la FT son derivadas de la misma especificación probabilística de nuestro modelo, es decir, suponemos que los datos de ambas funciones tienen en principio un mismo mecanismo generador de información, esta relación se puede modificar si se cambia la especificación de las innovaciones tecnológicas con la variable producto, las distribuciones marginales o la relación entre las innovaciones mismas con los insumos de la producción.

Una justificación alternativa de este resultado se da en Sánchez y Márquez (2015), en el cual se comprueba la dualidad entre la función de densidad conjunta de las variables utilizadas para la modelación empírica y la función de transformación de la empresa, la cual describe al conjunto factible de producción y de la cual se pueden obtener la función de producción local y la frontera tecnológica.

Note también que los correspondientes valores θ_{LPF} y θ_{TF} de los parámetros de estas funciones pueden no coincidir (Basu y Weil, 1998) dado que dependen de los valores obtenidos de los parámetros de sus distribuciones $T(a, b, Y; \theta_{LPF})$ y $T(a, b; \theta_{TF})$. Para nuestro modelo inicial haremos el supuesto adicional de que estos valores coinciden, es decir, $\theta_{LPF} = \theta_{TF} = \theta$, aunque, este supuesto se puede omitir.

Ahora, en el caso especial en el que la cópula y la distribución marginal del producto sean *funciones invertibles*, dada la proposición 1, podemos identificar a la función de producción local y a la frontera tecnológica a partir de la distribución conjunta.

Entonces, podemos obtener a la *función de producción local* como:

$$(2) \quad Y = W_Y^{-1}(C_W^{-1}(U(a, L; \theta_a), V(b, K; \theta_b), \Theta), P_0; \theta_Y) ;$$

donde C_W^{-1} y W_Y^{-1} son las transformaciones inversas de la cópula C respecto a W y Y respectivamente y cuya correcta especificación se obtiene si C_W y W_Y son estrictamente monótonas.

Por otro lado, tomando $W \equiv 1$ y fijando la probabilidad $P(a_i > a, b_i > b) = P_1$, obtenemos la *frontera tecnológica* dada por:

$$(3) \quad H(a, b) = g(C_{ab}(U(a; \theta_a), V(b; \theta_b), \Theta), P_1).$$

Así, vamos a suponer que el nivel de producción de la empresa depende de las funciones $U(a; \theta_a)$ y $V(b; \theta_b)$, las cuales modelan la escasez de las innovaciones y exhiben un crecimiento exponencial como en Kortum (1997); Jones (2005); y Growiec (2006 y 2013),

por lo que las familias de parámetros θ_a y θ_b están asociados a la *disponibilidad* de las innovaciones tecnológicas (su escasez) y la familia de parámetros θ que determinan la forma de la función de producción y la frontera tecnológica, así está relacionada con la *sustituibilidad o las posibilidades de complementariedad* entre ellas.

Hasta aquí hemos *concluido con el proceso de identificación* de nuestro modelo estadístico con el teórico-económico y hemos descrito la relación de los parámetros y formas funcionales con el proceso de producción de una empresa en nuestra economía.

Ahora vamos por último a obtener el modelo de *función de producción global* (FPG) siguiendo los planteamientos en Jones (2005), Caselli y Coleman (2006), Growiec (2006 y 2017) y Sánchez y Márquez (2015), asumiendo homogeneidad entre las empresas y auxiliándonos del Teorema de la Envolvente. Vamos a suponer que nuestra empresa es competitiva y que se comporta como maximizadora del beneficio, por lo que está enfocada en resolver el problema del productor, el cual consiste en maximizar, respecto a las variables de las innovaciones, la producción dada por la ecuación (2) sujeto a la condición (3) que delimita al conjunto sobre el cual podemos decidir, como en Jones y Manuelli (1990).

Así, podemos obtener la FPG que caracteriza a la producción en nuestra economía simplemente resolviendo este problema y sustituyendo las soluciones en la función de producción local. Por lo que, formalmente definimos el *problema de elección de la tecnología óptima de la empresa*, que consiste en elegir de entre las innovaciones disponibles el par tecnológico (a^*, b^*) para la j -ésima empresa, que nos permite maximizar la cantidad de producto Y^* que puede elaborar dados fijos los valores de los insumos trabajo L y capital K .

DEFINICIÓN 14: Si la cópula empleada en la especificación probabilística del modelo y la distribución marginal del producto son funciones invertibles, entonces el *problema de elección de la tecnología de una empresa* que nos permite obtener la FPG está definido por:

$$(4) \quad \max_{a,b} Y = W_Y^{-1}(C_W^{-1}(U(a, L), V(b, K), \Theta), P_0),$$

$$\text{sujeto a } g(C_{ab}(U(a), V(b), \Theta), P_1) = N.$$

Las soluciones al sistema (4) nos dan diversas formas cerradas de las funciones de producción global, dependiendo de la cópula y las distribuciones marginales que usemos en nuestra modelación, así como del valor del conjunto de parámetros Θ que obtengamos para el modelo.

PROPOSICIÓN 2: *Dados los supuestos 1 y 2, el problema de optimización del productor (5) tiene una única solución si y sólo si la FPL y la FT son tangentes en el punto $[a^*(U(L), V(K), \Theta), b^*(U(L), V(K), \Theta)]$, lo cual puede ocurrir sólo si la FPL es convexa y la FT cóncava o si la FPL es más convexa que la FT.*

Así, obtenemos la FPG dada por $Y^ = Y^*(L, K, N, \theta, U(a^*, L), V(b^*, K))$, cuya forma funcional y elasticidad de sustitución $\sigma = \sigma(\theta, U(a^*, L), V(b^*, K))$ están determinadas por el conjunto de parámetros de la cópula Θ y las distribuciones marginales para valores dados de L y K .*

Finalmente, el conjunto de parámetros Θ determina también el grado de dependencia estocástica entre a_i, b_i y Y_i ; el cual puede ser representado mediante una medida de asociación $\mu = \mu(\theta)$.

PRUEBA: Vea Apéndice A.1.

Una de las implicaciones de la proposición 2 es que el conjunto de parámetros θ y las formas funcionales de las distribuciones marginales constituyen el vínculo entre los modelos económico y estadístico de la función de producción.

En el siguiente corolario, se describe el caso donde la cópula usada para la construcción del modelo de producción depende únicamente de un sólo parámetro θ . Así, en este modelo particular, dadas las distribuciones marginales de las variables aleatorias consideradas, la

medida de dependencia $\mu = \mu(\theta)$ es únicamente determinada por el valor del parámetro θ , el cual desde el punto de vista microeconómico, determina el valor de la elasticidad de sustitución σ de la función de producción de la economía.

COROLARIO 1: *Basados en la proposición 2, sea C un a cópula que depende exclusivamente de un único parámetro de dependencia θ . Entonces θ determina al parámetro de sustituibilidad de la FPG que se obtiene de resolver el problema del productor y constituye el vínculo entre la elasticidad de sustitución $\sigma = \sigma(\theta)$ y la medida de dependencia estocástica $\mu = \mu(\theta)$ como sigue:*

- a) $\sigma = \sigma(m^{-1}(\mu))$, donde $\mu = m(\theta)$;
- b) $\mu = \mu(n^{-1}(\sigma))$, donde $\sigma = n(\theta)$.

PRUEBA: Vea Apéndice A.1.

Note que es este caso, si la función de producción tiene una forma funcional determinada por la cópula C con las características del Corolario 1, entonces las ecuaciones dadas en (a) y (b) representan la identificación de nuestros dos modelos, es decir, estas funciones son el puente entre el modelo estadístico y el teórico-económico.

Ahora, siguiendo la agregación clásica en la economía e interpretando a la función de producción local Y_i como la producción individual para una j -ésima empresa dada la i -ésima técnica de producción, la cual quiere maximizar su función de beneficio asociado π_i , entonces la siguiente proposición garantiza que si cada una de estas maximizaciones es realizada de manera simultánea, entonces se maximiza el beneficio global, el cual es el objetivo final del proceso de optimización.

PROPOSICIÓN 3: *Sea (Y, π) un par de producción y beneficio para la j -ésima empresa en una economía competitiva, dada la i -ésima técnica. Entonces, si cada empresa de esta economía con función de producción local $Y = W_Y^{-1}(C_W^{-1}(U(a, L), V(b, K), \Theta), P_0)$ y frontera tecnológica $H(a, b) = g(C_{ab}(U(a), V(b), \Theta), P_1)$ maximiza su producción*

eligiendo al par tecnológico óptimo (a^, b^*) y si J es un conjunto a lo sumo numerable, entonces el beneficio óptimo máximo en la economía se obtiene agregando los valores óptimos de los beneficios de las J empresas.*

PRUEBA: Vea Apéndice A.1.

Note entonces que, la interpretación de este resultado sería que la solución a este problema, conocida como la *función de producción global*, caracteriza en general la producción de cualquier empresa maximizadora del beneficio representativa de nuestro modelo, por lo que basta con caracterizar a esta función para caracterizar la producción en toda la economía.

3.3 Especificación de un Modelo de Producción con Innovaciones Pareto y Cópula Producto-Clayton

Ahora vamos a hacer nuestra *especificación del modelo estadístico* siguiendo los pasos de la sección anterior y basados en el enfoque de reducción probabilística de Spanos (1986, 1999), por lo que especificaremos primero el modelo probabilístico de nuestros vectores aleatorios y después el modelo muestral.

Así, para especificar el *modelo probabilístico* tomando en cuenta la evidencia empírica en Kortum (1997), Acemoglu (2003), Jones (2005) y Caselli y Coleman (2006) y Sánchez y Márquez (2015), vamos a suponer que las variables aleatorias de innovación de capital y trabajo y la del producto tienen distribuciones Pareto, por lo que suponemos que sus distribuciones marginales están dadas por $F_{a_i}(a) = P(a_i \leq a) = 1 - \left(\frac{a}{\gamma_a}\right)^{-\alpha}$, $F_{b_i}(b) = P(b_i \leq b) = 1 - \left(\frac{b}{\gamma_b}\right)^{-\beta}$ y $F_{Y_i}(Y) = P(Y_i \leq Y) = 1 - \left(\frac{Y}{\gamma_Y}\right)^{-\eta}$, donde los parámetros cumplen que $0 < \gamma_a \leq a$, $0 < \gamma_b \leq b$, $0 < \gamma_Y \leq Y$, $\alpha > 0$, $\beta > 0$, y $\eta > 0$.

Ahora, basados en la evidencia empírica expuesta en el capítulo anterior en Caselli y Coleman (2006) y Aoyama (2010), junto con los trabajos de Growiec (2006, 2017) y Sánchez y Márquez (2013, 2015), vamos a suponer que la estructura probabilística de nuestro modelo está dado por una cópula Producto-Clayton, es decir, suponemos que las innovaciones tecnológicas que ahorran capital y trabajo están correlacionadas y que su comportamiento probabilístico, junto con el de la producción, está descrito por una distribución construida a partir de la cópula Producto-Clayton trivariada y dada por

$$(5) \quad C(U, V, W) = W \max \left\{ [U^{-\theta} + V^{-\theta} - 1]^{-\frac{1}{\theta}}, 0 \right\}$$

donde el parámetro de dependencia θ toma valores en el intervalo $[-1, \infty]$.³ Considerando sólo el soporte de la cópula (4) esta puede ser reescrita como

$$(6) \quad C(U, V, W) = W[U^{-\theta} + V^{-\theta} - 1]^{-\frac{1}{\theta}}.$$

La variable U de la cópula no es más que la distribución marginal de una transformación monótona de la variable a_i , la variable V de una transformación de b_i y W de una transformación de Y_i .

Además, siguiendo los pasos de la sección anterior, marginalizamos respecto a W en la ecuación (6), obteniendo la distribución que describe el comportamiento conjunto de a_i y b_i dada por la cópula Clayton

$$(7) \quad C_{ab}(U, V) = C(U, V, 1) = [U^{-\theta} + V^{-\theta} - 1]^{-\frac{1}{\theta}}.$$

Así, el comportamiento estocástico conjunto de estas variables no sólo está descrito a través de la cópula trivariada (6), de su cópula marginalizada (7) y su parámetro θ , sino de su comportamiento estocástico individual dada por sus distribuciones marginales U, V y W .

Introduciendo las distribuciones marginales de las transformaciones $Z_1 = \frac{1}{a_i}$, $Z_2 = \frac{1}{b_i}$ y $Z_3 = -Y_i$, en la cópula (6) (como en el lema 1 de la sección anterior) obtenemos la probabilidad conjunta que va a identificar la relación entre las innovaciones y el producto y que está dada por:

$$P\left(a_i > \frac{1}{a}, b_i > \frac{1}{b}, Y_i > Y\right) = \left(\frac{Y}{\gamma_Y}\right)^{-\eta} \left[(a\gamma_a)^{-\alpha\theta} + (b\gamma_b)^{-\beta\theta} - 1\right]^{-\frac{1}{\theta}}.$$

³ Para obtener esta cópula trivariada, introducimos las cópulas producto $C_{aY}(U, W) = UW$ y $C_{bY}(V, W) = VW$, y la cópula Clayton, $C_{ab}(U, V) = \max\left\{[U^{-\theta} + V^{-\theta} - 1]^{-\frac{1}{\theta}}, 0\right\}$ en la fórmula de truncamiento

$$C(U, V, W) = WC_{ab}\left(\frac{C_{aY}(U, W)}{W}, \frac{C_{bY}(V, W)}{W}\right),$$

la cual da como resultado una nueva cópula trivariada a la que llamaremos Producto-Clayton.

De igual manera, siguiendo los pasos de la sección anterior, tomando ahora las transformaciones $Z_1 = -a_i$ y $Z_2 = -b_i$, y sustituyendo sus distribuciones en la copula Clayton (7), obtenemos la probabilidad que describe el comportamiento de las innovaciones asequibles para la empresa, la cual está dada por:

$$P(a_i > a, b_i > b) = \left(\frac{Y}{\gamma_Y}\right)^{-\eta} \left[\left(\frac{a}{\gamma_a}\right)^{-\alpha\theta} + \left(\frac{b}{\gamma_b}\right)^{-\beta\theta} - 1 \right]^{-\frac{1}{\theta}}.$$

Formalmente, el siguiente lema que especifica este resultado.

LEMA 4: *Basados en el supuesto 2 y la proposición 2, suponiendo que la función de producción y la frontera tecnológica de nuestra economía se pueden describir mediante las cópulas de las ecuaciones 6 y 7, usando el Teorema de Sklar se obtiene la probabilidad asociada a la FPL*

$$(8) \quad P\left(Y_i > Y, a_i > \frac{1}{a}, b_i > \frac{1}{b}\right) = \left(\frac{Y}{\gamma_Y}\right)^{-\eta} \left[(\gamma_a a)^{-\alpha\theta} + (\gamma_b b)^{-\beta\theta} - 1 \right]^{-1/\theta},$$

y la probabilidad asociada a la FT

$$(9) \quad P(a_i > a, b_i > b) = \left[\left(\frac{a}{\gamma_a}\right)^{\alpha\theta} + \left(\frac{b}{\gamma_b}\right)^{\beta\theta} - 1 \right]^{-1/\theta},$$

donde $-1 \leq \theta < \infty$.

PRUEBA: Vea Apéndice A.2.

La ecuación (8) se usa para la obtención de la FPL ya que describe el comportamiento de dependencia conjunta entre las innovaciones tecnológicas y el producto y exhibe una relación positiva entre las innovaciones y el producto, en tanto que la (9) nos describe únicamente las posibilidades de combinación entre los dos tipos de innovación tecnológica,

mostrando a la vez una relación de complementariedad entre ellas, por lo que se ocupa para obtener la FT.

Note que para deducir la densidad conjunta de nuestras variables, esta expresión la podemos reescribir como la distribución conjunta de las variables aleatorias auxiliares $Z_1 = \frac{1}{a_i}$, $Z_2 = \frac{1}{b_i}$ y $Z_3 = -Y_i$, de la siguiente forma

$$P\left(a_i > \frac{1}{a}, b_i > \frac{1}{b}, Y_i > Y\right) = F\left(\frac{1}{a}, \frac{1}{b}; -Y\right) = C[U_{Z_1}(a), V_{Z_2}(b), W_{Z_3}(Y)].$$

Así, la densidad conjunta escrita en términos de las variables auxiliares anteriores está dada por la expresión:

$$f(z_{1i}, z_{2i}, z_{3i}) = f\left(\frac{1}{a_i}, \frac{1}{b_i}, -Y_i\right) = \frac{\partial^3 \left[W[U^{-\theta} + V^{-\theta} - 1]^{-\frac{1}{\theta}} \right]}{\partial U \partial V \partial W} \frac{\partial U_{Z_1}}{\partial Z_1} \frac{\partial V_{Z_2}}{\partial Z_2} \frac{\partial W_{Z_3}}{\partial Z_3},$$

la cual, definiendo $h(a, b) = [a^{-\alpha\theta} + b^{-\beta\theta} - 1]^{-\frac{1+\theta}{\theta}}$, se reduce a la función de densidad conjunta de las variables aleatorias del modelo, que es la forma más acabada de la especificación probabilística de nuestro modelo y está dada por

$$f(z_{1i}, z_{2i}, z_{3i}) = \alpha\beta\eta\theta^2(1 + \theta)h(a_i, b_i)a_i^{2\alpha\theta+1-\alpha}b_i^{2\beta\theta+1-\beta}\gamma_Y^\eta Y_i^{-\eta-1}.$$

Ahora, para la *especificación del modelo muestral*, dada la evidencia empírica siguiendo el trabajo de Aoyama (2010), Jones (2005), Growiec (2006) y Sánchez y Márquez (2013, 2015), vamos a suponer que estas variables constituyen una muestra independiente e idénticamente distribuida aleatoria $\{Z_k\}_{k=1}^m = \{(Z_{1k}, Z_{2k}, Z_{3k})\}_{k=1}^m$, es decir, son independientes entre sí y presentan homogeneidad entre ellos (las distribuciones y sus parámetros se mantienen al pasar de un vector a otro), por lo que la función de verosimilitud está dada por la ecuación:

$$\mathcal{L}(\theta | Z_k = z_k) = \prod_{k=1}^m P(Z_k = z_k | \theta) = \prod_{k=1}^m f(z_k | \theta),$$

donde para cada valor k , el vector $z_k = (z_{1k}, z_{2k}, z_{3k})$ es una realización del vector aleatorio $Z_k = (Z_{1k}, Z_{2k}, Z_{3k})$.

Así, tomando logaritmos tenemos

$$\log[\mathcal{L}(\theta | Z_k = z_k)] = \sum_{i=1}^m \log[f(z_k | \theta)] = \sum_{i=1}^m \log[f(z_{1k}, z_{2k}, z_{3k} | \theta)].$$

Finalmente, podemos obtener los valores de los parámetros del modelo maximizando la función de verosimilitud que resulta de sustituir la función de densidad $f(z_{1i}, z_{2i}, z_{3i})$ de nuestra especificación probabilística

$$\begin{aligned} \log[\mathcal{L}(\theta)] = \sum_{i=1}^m \left[\log[\alpha\beta\eta\theta^2(1+\theta)] - \frac{1+2\theta}{\theta} \log[(a\gamma_a)^{-\alpha\theta} + (b\gamma_b)^{-\beta\theta} - 1] \right. \\ \left. + \log[a^{2\alpha\theta+1-\alpha} b^{2\beta\theta+1-\beta} \gamma_Y^\eta Y^{-\eta-1}] \right] \end{aligned}$$

Por otra parte, una vez que estimamos los valores de los parámetros de nuestro modelo y establecemos una *prueba de correcta especificación*, la cual depende de la forma funcional de las distribuciones marginales y la conjunta, así como de los valores de los parámetros del modelo. Ejemplos de este tipo de pruebas son las pruebas de independencia, homogeneidad y distribución de nuestro modelo.

Así, si nuestra especificación estadística no pasa la prueba entonces regresamos a re-especificar nuestro modelo, ya sea cambiando las marginales de nuestro modelo, la cópula que captura su comportamiento conjunto, la dependencia de las muestras aleatorias, la homogeneidad entre las distribuciones de los vectores aleatorios o dos o más de estas cosas al mismo tiempo, de acuerdo a la evidencia empírica, siguiendo con una nueva estimación

de los parámetros y una prueba de correcta especificación. Se repiten estos pasos las veces que sea necesario.

En el caso de que la especificación estadística de nuestro modelos sea adecuada, procedemos a *identificar el modelo de producción* que queremos estimar, tanto con las formas funcionales como con la identificación de parámetros del mundo probabilístico como del económico. Así, siguiendo los pasos del apartado anterior, podemos determinar la FPL fijando el nivel de isoprobabilidad $F(a, b, Y) = P_0$ dado un nivel eficiente de producción en el espacio (a, b, Y) , y similarmente se determina la FT del modelo reducido fijando una curva de isoprobabilidad $F(a, b) = P_1$, en el espacio (a, b) dado un nivel eficiente de producción.

LEMA 5: *Basados en el lema 4 y fijando las probabilidades (8) y (9). Si sustituimos los valores $\gamma_a^{-\alpha\theta} \equiv \gamma_L^{-\alpha\theta}(1 + L^{-\theta})$ y $\gamma_b^{-\beta\theta} \equiv \gamma_K^{-\beta\theta}(1 + K^{-\theta})$ en la curva de isoprobabilidad asociada a (7) y ajustamos el tamaño de las variables y parámetros de tal forma que $(a\gamma_L)^{-\alpha\theta} + (b\gamma_K)^{-\beta\theta} = 1$, obtenemos la FPL dada por*

$$(10) \quad Y = A \left[(a\gamma_L)^{-\alpha\theta} L^{-\theta} + (b\gamma_K)^{-\beta\theta} K^{-\theta} \right]^{-\frac{\nu}{\theta}}.$$

Por otra parte, usando la curva de isoprobabilidad asociada a (8), la FT está descrita por

$$(11) \quad H(a, b) = \left(\frac{a}{\gamma_a} \right)^{\alpha\theta} + \left(\frac{b}{\gamma_b} \right)^{\beta\theta} = N,$$

donde $(a\gamma_L)^{-\alpha\theta} > 0$, $(b\gamma_K)^{-\beta\theta} > 0$, $A = \frac{\gamma_y}{P_0^\nu} > 0$, $\nu = \frac{1}{\eta} > 0$, $N = P_1^{-\theta} + 1 > 0$ y $-1 \leq \theta < \infty$.

PRUEBA: Vea Apéndice A.2.

Las definiciones para $\gamma_a^{-\alpha\theta}$ y $\gamma_b^{-\beta\theta}$ y la condición $(a\gamma_L)^{-\alpha\theta} + (b\gamma_K)^{-\beta\theta} = 1$ son identidades necesarias para mapear la función de probabilidad conjunta en la FPL CES, dotando así de microfundamentos al modelo. Los parámetros que dan forma al modelo son los de las

distribuciones marginales α y β , y el parámetro θ que determina la dependencia de la cópula.

Note que la FPL es una versión generalizada de la función CES, la cual permite definir retornos constantes a escala si $\nu = 1$, pero también obtenemos retornos crecientes si $\nu > 1$ y decrecientes en el caso $\nu < 1$ (las funciones con retornos decrecientes están asociadas a distribuciones de la producción con colas más pesadas para Y), además de incluir relaciones potenciales no lineales (en contraste con los modelos tradicionales de crecimiento) entre las innovaciones de trabajo y capital y los niveles de L y K respectivamente.

Por otro lado, observe que el parámetro de dependencia θ de la cópula se puede interpretar como el parámetro de sustituibilidad que determina la curvatura de la FPL CES, constituyéndose así como un enlace entre las cópulas (6) y (7), definidas en el espacio de probabilidad (Ω, \mathcal{F}, P) , y las funciones real valuadas de la teoría económica FPL y FT descritas en las ecuaciones (10) y (11) respectivamente.

En este sentido, una medida de dependencia estocástica está dada por la Tau de Kendall, cuyo valor para la cópula Producto-Clayton está dado por

$$(12) \quad \tau_c = \frac{\theta}{\theta + 2}, \quad \text{con } -1 \leq \tau_c \leq 1,$$

y la elasticidad de sustitución σ_{KL} (como función de θ) para nuestra función de producción generalizada CES está dada por

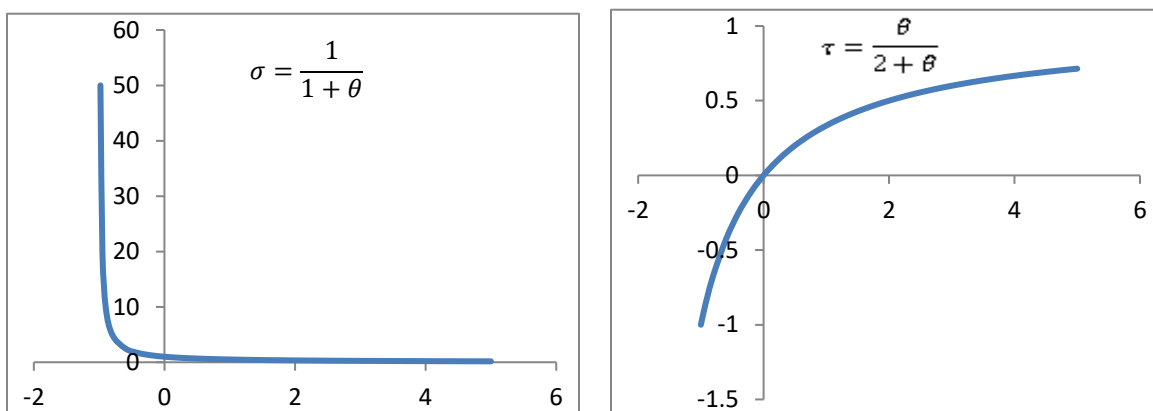
$$(13) \quad \sigma_{KL} = \frac{1}{1 + \theta}, \quad \text{con } 0 \leq \sigma_{KL} \leq \infty.^4$$

⁴ Estas fórmulas se deducen fácilmente usando las formulas $\tau_c = 1 + 4 \int_0^1 \frac{\psi(t)}{\psi'(t)}$, tomada de Nelsen (2006)

y de la definición de elasticidad de sustitución $\sigma = \left(\frac{d \ln[TMST_{L,K}]}{d \left(\frac{K}{L} \right)} \right)^{-1}$.

Estas dos ecuaciones nos muestran una relación inversa entre la elasticidad de sustitución de la FPL y la medida de dependencia estadística entre las variables aleatorias del modelo, dada en este caso por la Tao de Kendall. Cuando las innovaciones tecnológicas están positivamente correlacionadas, es decir, cuando a_i y b_i son complementos netos, un incremento en el nivel del factor de innovación del trabajo a_i está asociado a un incremento en el factor de innovación del capital b_i . En el caso contrario estos factores son sustitutos netos y su correlación se dan de manera inversa. Esta relación se describe gráficamente en la figura I, donde graficamos la relación entre estos dos parámetros.

Figura I: Elasticidad de sustitución y Tao de Kendall.



Graficas de la elasticidad de sustitución (σ) versus el parámetro de sustituibilidad (θ) a la izquierda y de la Tao de Kendall (τ_c) versus el parámetro de sustituibilidad (θ) a la derecha.

Note que las ecuaciones anteriores describen las relaciones entre los parámetros del modelo estadístico con los del modelo teorico-económico, lo cual si lo juntamos con la identificación de las funciones de producción local y frontera tecnológica, completa nuestro proceso de identificación entre estos modelos.

Entonces, a través del valor de la medida de dependencia τ_c entre las variables podemos determinar los valores del parámetro de sustituibilidad y de la elasticidad de sustitución asociados a la función de producción CES. Así, cambios en la dependencia se traducen de

manera directa en cambios en la sustituibilidad, y de manera inversa en cambios en la elasticidad de sustitución.

Además, note que para ciertos valores específicos del parámetro de dependencia podemos obtener los distintos modelos de producción local y fronteras tecnológicas planteadas en la literatura, lo cual se describe en el siguiente lema.

LEMA 6: *Bajo los supuestos del modelo microfundado CES de producción, se cumple:*

a) *Para el valor de la medida de dependencia $\tau = 0$, se tiene el valor del parámetro de sustituibilidad $\theta = 0$. La cópula que corresponde a este modelo es la producto, dada por $C(U, V, W) = UVW$. En este caso, la elasticidad de sustitución es $\sigma = 1$, la cual corresponde a la FPL que es del tipo Cobb-Douglas y tiene la forma*

$$Y = AL^{\alpha\nu} K^{\beta\nu}, \text{ donde } A = \frac{\gamma_Y}{P_0^\nu} [\alpha\gamma_L]^{\alpha\nu} [\beta\gamma_K]^{\beta\nu} \text{ y } \nu = \frac{1}{\eta}.$$

Además, la cópula marginalizada cuando $W = 1$ está dada por $C_{a,b}(U, V) = UV$, y su FT correspondiente es

$$H(a, b) = a^\alpha b^\beta = N, \text{ donde } N = \gamma_a^\alpha \gamma_b^\beta P_1.$$

b) *Si el valor de la medida de dependencia es $\tau = 1$, el valor del parámetro de sustituibilidad θ converge a infinito y la cópula que corresponde a este modelo es la producto-Cota Superior de Frechét-Hoeffding dada por $C(U, V, W) = W\min\{U, V\}$. En este caso, la elasticidad de sustitución es $\sigma = 0$, la cual corresponde con la FPL de Leontief de la forma*

$$Y = A[\min\{(\alpha\gamma_L)^\alpha L, (\beta\gamma_K)^\beta K\}]^\nu, \text{ donde } A = \frac{\gamma_Y}{P_0^\nu}.$$

Además, la cópula marginalizada cuando $W = 1$ esta dada por $C_{a,b}(U, V) = \min\{U, V\}$, cuya FT correspondiente es de la forma

$$H(a, b) = \min\{\gamma_a^\alpha a^{-\alpha}, \gamma_b^\beta b^{-\beta}\} = N \text{ donde } N = P_1.$$

c) Cuando la medida de dependencia es $\tau = -1$, el parámetro de sustituibilidad toma el valor $\theta = -1$ y la cópula que corresponde a este modelo es la producto-Cota Inferior de Frechét-Hoeffding dada por $C(U, V, W) = W \max\{U + V - 1, 0\}$. En este caso, la elasticidad de sustitución toma el valor $\sigma = \infty$, la cual corresponde a la FPL de sustitutos perfectos de la forma

$$Y = A[(a\gamma_L)^\alpha L + (b\gamma_K)^\beta K]^\nu \text{ donde } A = \frac{\gamma_Y}{P_0^\nu} > 0 \text{ y } \nu = \frac{1}{\eta}.$$

Además, la cópula marginalizada cuando $W = 1$ esta dada por $C_{a,b}(U, V) = \max\{U + V - 1, 0\}$, cuya FT correspondiente es de la forma

$$H(a, b) = \left[\left(\frac{a}{\gamma_a} \right)^{-\alpha} + \left(\frac{b}{\gamma_b} \right)^{-\beta} \right]^{-1} = N \text{ donde } N = [P_1 + 1]^{-1}.$$

PRUEBA: Vea Apéndice A.2.

Una vez que terminamos el proceso de identificación de nuestro modelo teorico-económico, podemos ahora utilizarlo para *hacer inferencia* sobre algunos aspectos que nos interesa ya en el campo de la teoría económica. En nuestro caso continuaremos con la modelación de una empresa competitiva que se comporte de manera maximizadora del beneficio. Así, siguiendo con los pasos del modelo base, el problema de elección tecnológica de la empresa en el caso general de la función CES está dado por

$$\max_{a,b} Y = A[(\gamma_a a)^{-\alpha\theta} L^{-\theta} + (\gamma_b b)^{-\beta\theta} K^{-\theta}]^{-\frac{\nu}{\theta}},$$

$$\text{sujeto a } \left(\frac{a}{\gamma_a} \right)^{\alpha\theta} + \left(\frac{b}{\gamma_b} \right)^{\beta\theta} = N.$$

Para resolverlo, planteamos el operador lagrangeano del problema y obtenemos el máximo de la función bivariada, cuyos resultados se plantean en la siguiente proposición.

PROPOSICIÓN 4: Los resultados de resolver el problema de elección de tecnología con función de producción CES son:

$$a^* = \left[\frac{\gamma_L^{\alpha\theta} \gamma_K^{\beta\theta} L^\theta K^\theta N}{\gamma_K^{\beta\theta} (1+L^\theta) K^\theta + \gamma_L^{\alpha\theta} L^\theta \sqrt{(1+L^\theta)(1+K^\theta)}} \right]^{\frac{1}{\alpha\theta}} \quad y \quad b^* = N \left[\frac{\gamma_L^{\alpha\theta} \gamma_K^{\beta\theta} L^\theta K^\theta N}{\gamma_K^{\beta\theta} \sqrt{(1+L^\theta)(1+K^\theta)} + \gamma_L^{\alpha\theta} L^\theta (1+K^\theta)} \right]^{\frac{1}{\beta\theta}}.$$

Luego, sustituyendo estos valores óptimos en la FPL, obtenemos la FPG dada por:

$$(13) \quad Y^* = AN^{\frac{\nu}{\theta}} \left[\gamma_L^{-\alpha\theta} \sqrt{1+L^\theta} L^{-\theta} + \gamma_K^{-\beta\theta} \sqrt{1+K^\theta} K^{-\theta} \right]^{-\frac{2\nu}{\theta}}.$$

PUEBA: Vea Apéndice A.2.

Note que esta función no es propiamente una función de producción global CES, aunque tiene una forma funcional y propiedades parecidas, además puede ser aproximada por la siguiente función de producción cuando $\theta > 0$ y para valores grandes de L^θ y K^θ

$$(14) \quad Y^* = AN^{\frac{\nu}{\theta}} \left[\gamma_L^{-\alpha\theta} L^{-\frac{\theta}{2}} + \gamma_K^{-\beta\theta} K^{-\frac{\theta}{2}} \right]^{-\frac{2\nu}{\theta}},$$

con valores respectivos $a^* = \left[\frac{\gamma_L^{\alpha\theta} \gamma_K^{\beta\theta} N L^\theta K^\theta}{\gamma_L^{\alpha\theta} L^{\frac{\theta}{2}} + \gamma_K^{\beta\theta} K^{\frac{\theta}{2}}} \right]^{\frac{1}{\alpha\theta}}$ y $b^* = \left[\frac{\gamma_L^{\alpha\theta} \gamma_K^{\beta\theta} N L^\theta K^\theta}{\gamma_L^{\alpha\theta} L^{\frac{\theta}{2}} + \gamma_K^{\beta\theta} K^{\frac{\theta}{2}}} \right]^{\frac{1}{\beta\theta}}.$

Del mismo modo, para el caso en que $-1 < \theta < 0$, se tiene que

$$(15) \quad Y^* = AN^{\frac{\nu}{\theta}} \left[\gamma_L^{-\alpha\theta} L^{-\theta} + \gamma_K^{-\beta\theta} K^{-\theta} \right]^{-\frac{2\nu}{\theta}},$$

con valores respectivos $a^* = \left[\frac{\gamma_L^{\alpha\theta} \gamma_K^{\beta\theta} N L^\theta K^\theta}{\gamma_K^{\beta\theta} K^\theta + \gamma_L^{\alpha\theta} L^\theta} \right]^{\frac{1}{\alpha\theta}}$ y $b^* = \left[\frac{\gamma_L^{\alpha\theta} \gamma_K^{\beta\theta} N L^\theta K^\theta}{\gamma_K^{\beta\theta} K^\theta + \gamma_L^{\alpha\theta} L^\theta} \right]^{\frac{1}{\beta\theta}}.$

En el siguiente lema, la forma de la FPG está determinada únicamente por los valores del parámetro de dependencia θ , los cuales se toman de acuerdo al lema 6.

PROPOSICIÓN 5: Bajo los supuestos del lema 6, asumiendo el planteamiento del problema de elección de la tecnología para la producción, dadas la FPL y la FT para los diferentes valores del parámetro θ , tenemos que;

a) Si $\theta = 0$, la solución al problema del productor es una FPG del tipo Cobb-Douglas y está dada por

$$Y^* = A^* L^{\alpha\nu} K^{\beta\nu}, \text{ donde } A^* = \frac{\gamma_Y}{P_0^\nu} \gamma_L^{\alpha\nu} \gamma_K^{\beta\nu} N^\nu.$$

b) Si $\theta = \infty$, la solución al problema del productor es una FPG del tipo Leontief dada por

$$Y^* = A^* [\min\{\gamma_L^\alpha L, \gamma_K^\beta K\}]^{2\nu} \text{ donde } A^* = \frac{\gamma_Y}{P_0^\nu} N^\nu.$$

c) Si $\theta = -1$, la solución al problema del productor es una FPG del tipo sustitutos perfectos y está dada por

$$Y^* = AN^\nu \left[\gamma_L^\alpha \sqrt{(1+L)L} + \gamma_K^\beta \sqrt{(1+K)K} \right]^{2\nu},$$

donde esta función se reduce, para valores grandes de L y K , a la forma

$$Y^* = AN^\nu \left[\gamma_L^\alpha L + \gamma_K^\beta K \right]^{2\nu},$$

la cual es una transformación monótona de una función de sustitutos perfectos.

PRUEBA: Vea Apéndice A.2.

Note que hemos obtenido la forma funcional general de la función de producción global de nuestro modelo, del cual pueden obtenerse los otros modelos considerados en la teoría económica cuando el parámetro de sustitución entre las innovaciones toma los valores cero, menos uno e infinito respectivamente.

3.4 Comentarios Generales

En este capítulo construimos el modelo general de la función de producción con innovaciones tecnológicas de una economía partiendo de las características de los datos de nuestro modelo. Además, ejemplificamos esta modelación con un modelo basado en la evidencia empírica y los trabajos anteriores, donde las innovaciones tienen distribución marginal Pareto y su distribución conjunta está descrita por la cópula Producto-Clayton, por lo que el comportamiento estocástico de nuestras variables permite identificar al modelo probabilístico que lo describe con el modelo de producción con innovaciones de nuestra economía, identificando además los parámetros probabilísticos de dependencia y de cola pesada de sus variables aleatorias con los parámetros de sustituibilidad y elasticidad de sustitución, así como de escases del modelo económico respectivamente.

Obtuvimos así un modelo con función de producción CES generalizada y frontera tecnológica como en Growiec (2006) y Sánchez y Márquez (2013), lo que nos llevó a un modelo de función de producción global del tipo CES. Además, suponiendo ciertos valores específicos del parámetro de sustituibilidad de nuestro modelo obtuvimos los demás modelos de producción de la literatura, el Cobb-Douglas, Leontief y Sustitutos Perfectos.

Así, la FPL permite definir retornos constantes a escala si $\nu = 1$, pero también obtenemos retornos crecientes si $\nu > 1$ y decrecientes en el caso $\nu < 1$ (las funciones con retornos decrecientes están asociadas a distribuciones de la producción con colas más pesadas para Y), además de incluir relaciones potenciales no lineales (en contraste con los modelos tradicionales de crecimiento) entre las innovaciones de trabajo y capital y los niveles de L y K respectivamente.

También, cambios en la dependencia se traducen de manera directa en cambios en la sustituibilidad, y de manera inversa en cambios en la elasticidad de sustitución. Así, este modelo resulta ser tan general que se pueden obtener los casos particulares de modelos en los que nos hemos basado y todos los modelos clásicos de la literatura económica a partir de él, permitiendo además innovar en la utilización y obtención de nuevas formas

funcionales de la producción de acuerdo a las características de los datos y el modelo estadístico que mejor se ajuste a la evidencia empírica.

Por un lado, nuestro modelo de origen completamente estocástico determina la forma de la función de producción y la frontera tecnológica, por lo que consecuentemente determina también la *función de producción global* agregando el supuesto de que nuestra empresa se comporta como maximizadora del beneficio y suponiendo homogeneidad e independencia en nuestra especificación muestral.

Por último, es importante mencionar que el modelo permite incluir varias variables que han sido omitidas u obviadas por la literatura, como el capital humano y las variables de innovaciones verdes, pero además incluso permite incluir diferentes tipos de dependencia entre ellas y formas funcionales de acuerdo a la naturaleza de los datos del modelo.

CAPITULO 4: MODELO EMPÍRICO DE PRODUCCIÓN CON INNOVACIONES TECNOLÓGICAS VERDES

4.1 Introducción

En esta sección utilizaremos de manera empírica las herramientas que hemos desarrollado para modelar el comportamiento de la producción en el sector de alta tecnología en la economía de EEUU que desarrolla y utiliza innovaciones tecnológicas verdes, las cuales incluyen a las que contribuyen a mejorar la calidad del medio ambiente, la conservación de la energía, el desarrollo de recursos de energía renovable o la reducción de los gases de efecto invernadero de acuerdo a la clasificación de la Oficina de Patentes y Marcas Registradas de los EEUU (USTPO por sus siglas en inglés).

Con el fin de ilustrar el alcance de nuestro modelo y ejemplificar los usos que puede tener esta forma de modelación estadística, después de la descripción de los datos que utilizaremos, haremos una modelación de nuestra teoría con datos reales de la economía y las patentes verdes de EEUU en 2014, finalizando con el proceso de identificación del mismo, lo que permite una discusión teórica a cerca de los resultados de la modelación.

La diferencia entre nuestro modelo con los desarrollados anteriormente en la literatura, es que partimos de la especificación del modelo estadístico construido a partir de la evidencia empírica que arrojan estos sin hacer supuestos adicionales del comportamiento de los mismos dado por la teoría económica.

Además, nuestro modelo captura directamente la interacción entre las innovaciones verdes que ahorran capital y las que ahorran trabajo, lo cual es importante para después, en el proceso de identificación, describir la dinámica del proceso de producción con innovaciones como lo señala la evidencia empírica, ya que permite estudiar su interacción y hacer inferencia estadística a cerca del comportamiento del mismo.

Otra de las ventajas de nuestro enfoque es que además se pueden obtener formas funcionales más generales de la función de producción (o incluso nuevas) que la función de

producción Cobb-Douglas, la cual, como vimos en el capítulo 1, es empleada comúnmente en este tipo de análisis y es resultado de pedir que haya independencia entre las variables analizadas (Sánchez y Márquez 2013).

Con esto en mente, tomamos en cuenta las características empíricas de nuestros datos y seguimos los pasos de nuestro enfoque para aplicar nuestra teoría y obtener finalmente la *función de producción global* que modele el comportamiento de la producción con innovaciones verdes en el sector de alta tecnología de la economía de Estados Unidos en 2014, todo esto partiendo de la especificación probabilística y muestral que caracteriza a nuestro modelo estadístico, suponiendo valores dados de las variables de capital y trabajo.

4.2 Descripción de los Datos

Dada la naturaleza de la modelación que proponemos en este trabajo, debemos especificar el tipo de datos que utilizaremos para la modelación empírica, lo cual va a ser determinante al tiempo en que tenemos que escoger el método y las distribuciones de nuestros datos. Así tenemos que, los datos que usamos al ser de patentes verdes otorgadas por la USTPO en 2014, estos son del tipo sección cruzada y las usaremos como variables proxy de las innovaciones tecnológicas verdes, tal y como lo hacen con las patentes en la literatura (Kortum (1956), Jones (2009), Acemoglu (2006)).

La clasificación que hace la USTPO de las patentes verdes está dada por las siguientes categorías:

- A) Producción Alternativa de Energía: la cual considera la energía producida a partir de residuos agrícolas, animales, humanos, de hospitales, industriales o químicos, biocombustibles, combustible derivado de la basura, sistemas domésticos de calentamiento de agua, calefacción pasiva por espacios, albercas, o celdas, gasificación, organismos genéticamente diseñados, geotermias, energía nuclear, energía solar, energía inercial o energía eólica.

- B) Conservación de la Energía: la cual considera vehículos de energía alternativa e híbridos, circuitos de rayos catódicos, reducción de arrastre, lámparas y vehículos eléctricos y dispositivos de carga, almacenamiento y distribución de la energía, motores y embarcaciones impulsados por el viento u ondas de agua, construcción de caminos, sistemas ópticos y sus elementos, ahorro de desplazamientos (como el trabajo en casa) y el comercio de emisiones.
- C) Agricultura Ambientalmente Amigable: considera innovaciones como técnicas de alternativas de riego, uso o reciclaje de desperdicios, fertilizantes alternativos, conservación del suelo y del agua, abatimiento de la contaminación y mejora del rendimiento.
- D) Purificación, Remediación o Protección Ambiental: considera innovaciones en destrucción o contención de virus o bacterias nocivas, productos biodegradables, remediación del suelo, plantas y plantaciones, recuperación del exceso de materiales de procesos o regeneración del flujo de residuos, reciclaje, limpieza de material tóxico, mejora de chimeneas, captura o secuestro de carbono, limpieza y remediación de desastres, confinamiento o destrucción de material contaminante y el uso de microbios o enzimas.

De la misma manera que el caso general, clasificamos a las patentes verdes en dos grupos, siguiendo a Acemoglu (2003) y Habakkuk (1967), aquellas que reducen los costos por el uso del insumo trabajo, como las patentes de herramientas, máquinas, aparatos y tipos específicos de procesos, las cuales clasificamos como ahorradoras de trabajo, y el resto de las patentes las clasificamos como ahorradoras de capital.

Sin embargo, aunque en la práctica una innovación en una empresa que ahorra capital muchas veces va acompañada de otras innovaciones de trabajo y de capital para su implementación y viceversa (Growiec, 2006) [piense por ejemplo en el caso de un cambio

de utilización de diésel a energía solar, en el cual se requiere encontrar nuevas formas de almacenar energía, de capacitación en el manejo de esta tecnología por el personal de la empresa, etc.], dada la complejidad y los costos para la obtención de patentes, esto no se ve reflejado directamente en la composición total de las patentes otorgadas por la USTPO, por lo que supondremos que este efecto si puede ser observado de manera adecuada en el núcleo de empresas que tiene innovaciones de ambos tipos.

Así, dado que queremos capturar el papel de la dependencia entre las innovaciones ahorradoras de capital y de trabajo en la producción, vamos a considerar únicamente los datos de las empresas que obtuvieron patentes verdes de ambos tipos en 2014 para hacer la aproximación de los parámetros de nuestro modelo de producción con cópulas y para su interpretación. Note además que por la forma de la copula Producto-Clayton que asumimos en nuestro modelo, los datos de innovación tecnológica son independientes de los de producción, por lo que no incluiremos estos últimos. Una ampliación de este trabajo sería el trabajar con funciones copula trivariadas donde todas las variables aleatorias del modelo estén correlacionadas entre sí, como por ejemplo, se podrían utilizar copulas anidadas.

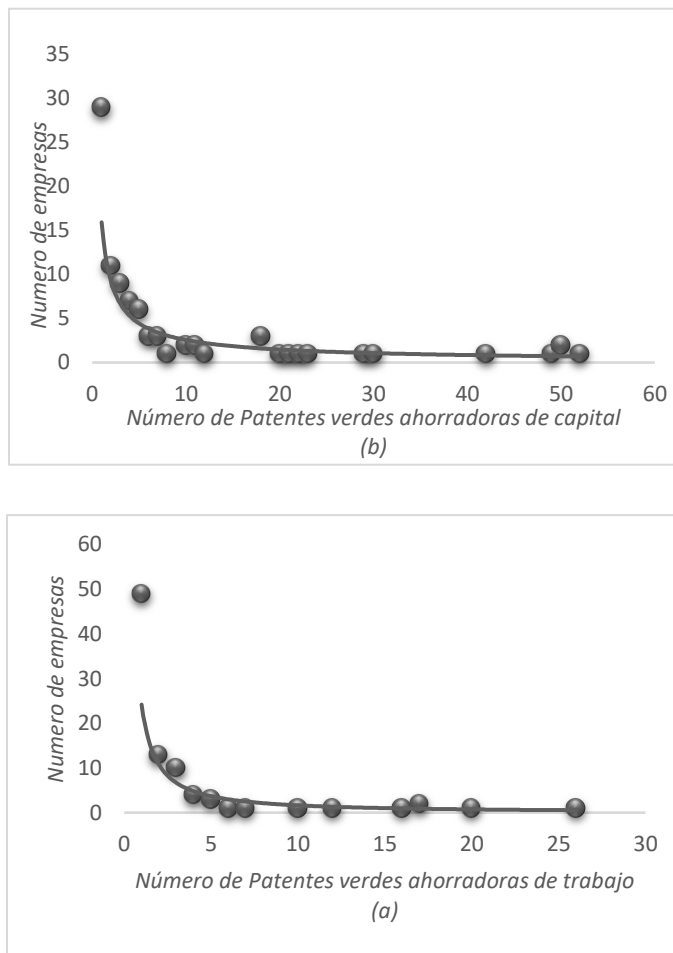
4.3 Características probabilísticas de los datos

Las patentes verdes otorgadas por la USTPO en 2014 ascendieron a 7302, de las cuales 6532 fueron ahorradoras de capital y 770 fueron ahorradores de trabajo. Estas patentes fueron repartidas entre 2876 empresas o inventores, de los cuales 88 obtuvieron patentes verdes de ambos tipos.

En las siguientes figuras graficamos el número de empresas o inventores de patentes verdes ahorradoras de capital o de trabajo contra las cantidades respectivas de patentes que obtuvieron (en este caso estamos graficando los datos para el núcleo de empresas que obtienen patentes de ambos tipos).

Note que los datos de este núcleo de estas patentes verdes muestran una distribución empírica de cola ancha, la cual se puede asociar, como lo hace la evidencia empírica, a la distribución Pareto con parámetro de escala igual a 1 (la línea de aproximación es del tipo Potencial).

FIGURA II: Densidades Kernel de las patentes verdes que ahorran capital y trabajo otorgadas por la Oficina de Patentes y Marcas Registradas de EEUU (USTPO) en 2014.



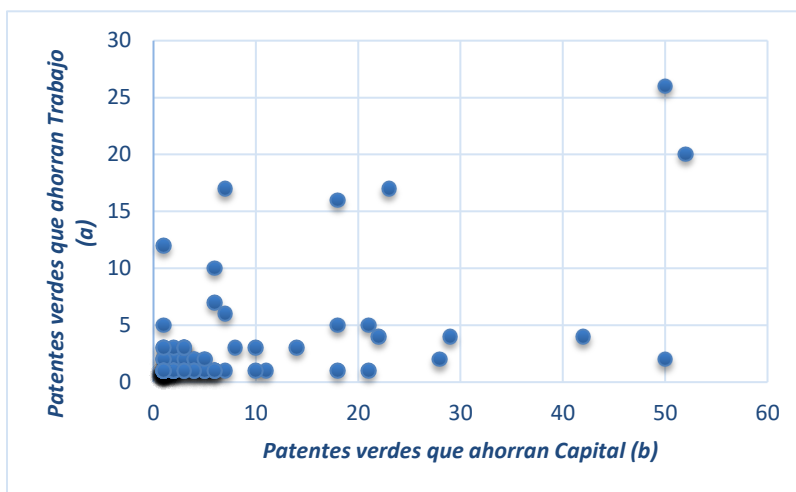
Fuente: Elaboración propia con datos de la USTPO

<http://www.uspto.gov/web/offices/ac/ido/oeip/taf/hcounts.htm>

La figura III muestra un diagrama de dispersión de las patentes de capital y trabajo de cada una de las empresas del núcleo de las que tienen ambos tipos de patentes, la cual sugiere

una relación de dependencia positiva entre estas patentes al tener varias empresas con niveles altos de ambos tipos, es decir, empresas que tienen un nivel más alto de algún tipo de patentes asociado a un nivel relativamente alto de patentes del otro tipo.

FIGURA III: Gráfico de dispersión entre las innovaciones verdes ahorradoras de capital y las que ahorran trabajo otorgadas por la USTPO en 2014.



Fuente: Elaboración propia con datos de la USTPO

<http://www.uspto.gov/web/offices/ac/ido/oeip/taf/hcounts.htm>

Desde el punto de vista empírico, estas gráficas sugieren que las patentes ahorradoras de capital y de trabajo respectivamente, pueden caracterizarse probabilísticamente de manera adecuada a través de una distribución conjunta con correlación positiva y con dos marginales del tipo Pareto en sus colas, igual que como se obtuvo en Sánchez y Márquez (2015), donde se graficaron las patentes del núcleo de empresas de alta innovación tecnológica en EEUU en 2014.

Así, las cópulas que pueden dar un buen ajuste de acuerdo a la distribución empírica de los datos son la Clayton y la Gumbel. En nuestro caso utilizaremos la cópula Clayton, la cual es la que tiene un ajuste más preciso y nos puede facilitar el proceso de identificación llegando a un modelo de producción conocido, como vimos en el capítulo anterior, de

acuerdo al procedimiento de Aoyama (2010) y siguiendo los trabajos de Growiec (2006) y Sánchez y Márquez (2013).

4.5 Especificación del Modelo

En esta sección se describe proceso de especificación del modelo basados en la evidencia empírica presentada en la sección anterior, así como en Acemoglu (2003), Caselli y Coleman (2006) y Aoyama (2010), el cual servirá de base para la estimación de los parámetros de la función de producción de empresas que emplea innovaciones verdes de capital y trabajo para mejorar su proceso de producción. Así, siguiendo los pasos de la especificación del capítulo anterior, para obtener los parámetros de nuestro modelo de producción con innovaciones verdes usando el *método de máxima verosimilitud completa*.

Vamos a comenzar especificando el *modelo probabilístico* de los datos, así de acuerdo a la evidencia empírica asumimos que las variables de innovación de capital y trabajo y la del producto tienen distribuciones Pareto, es decir, $F_{a_i}(a) = P(a_i \leq a) = 1 - \left(\frac{a}{\gamma_a}\right)^{-\alpha}$, $F_{b_i}(b) = P(b_i \leq b) = 1 - \left(\frac{b}{\gamma_b}\right)^{-\beta}$ y $F_{Y_i}(Y) = P(Y_i \leq Y) = 1 - \left(\frac{Y}{\gamma_Y}\right)^{-\eta}$, donde $0 < \gamma_a \leq a$, $0 < \gamma_b \leq b$, $0 < \gamma_Y \leq Y$, $\alpha > 0$, $\beta > 0$, y $\eta > 0$.

Así, siguiendo el desarrollo del modelo probabilístico del capítulo anterior y basados en la evidencia empírica del apartado 4.4, vamos a considerar que la densidad conjunta de las variables está descrita por la cópula

$$C(U, V, W) = W \max \left\{ [U^{-\theta} + V^{-\theta} - 1]^{-\frac{1}{\theta}}, 0 \right\}.$$

Así, introduciendo las distribuciones marginales de las transformaciones adecuadas de las variables aleatorias en la cópula (como en el lema 1 de la sección anterior) obtenemos la probabilidad de la forma:

$$P\left(a_i > \frac{1}{a}, b_i > \frac{1}{b}, Y_i > Y\right) = \left(\frac{Y}{\gamma_Y}\right)^{-\eta} \left[(a\gamma_a)^{-\alpha\theta} + (b\gamma_b)^{-\beta\theta} - 1 \right]^{-\frac{1}{\theta}}.$$

Note que para deducir la densidad conjunta de nuestras variables, esta expresión la podemos reescribir como la distribución conjunta de las variables aleatorias auxiliares $Z_1 = \frac{1}{a_i}$, $Z_2 = \frac{1}{b_i}$ y $Z_3 = -Y_i$, de la siguiente forma

$$P\left(a_i > \frac{1}{a}, b_i > \frac{1}{b}, Y_i > Y\right) = F\left(\frac{1}{a}, \frac{1}{b}; -Y\right) = C[U_{Z_1}(a), V_{Z_2}(b), W_{Z_3}(Y)].$$

Así, la densidad conjunta escrita en términos de las variables auxiliares anteriores está dada por la expresión:

$$f(z_{1i}, z_{2i}, z_{3i}) = f\left(\frac{1}{a_i}, \frac{1}{b_i}, -Y_i\right) = \frac{\partial^3 \left[W[U^{-\theta} + V^{-\theta} - 1]^{-\frac{1}{\theta}} \right]}{\partial U \partial V \partial W} \frac{\partial U_{Z_1}}{\partial Z_1} \frac{\partial V_{Z_2}}{\partial Z_2} \frac{\partial W_{Z_3}}{\partial Z_3},$$

la cual, definiendo $h(a, b) = [a^{-\alpha\theta} + b^{-\beta\theta} - 1]^{-\frac{1+\theta}{\theta}}$, se reduce a la densidad conjunta de las variables aleatorias del modelo, la cual es la *especificación probabilística de nuestro modelo* y está dada por:

$$f(z_{1i}, z_{2i}, z_{3i}) = \alpha\beta\eta\theta^2(1 + \theta)h(a_i, b_i)a_i^{2\alpha\theta+1-\alpha}b_i^{2\beta\theta+1-\beta}\gamma_Y^\eta Y_i^{-\eta-1}.$$

Luego, vamos ahora a especificar el *modelo muestral* de nuestros datos basados en la evidencia empírica sobre la distribución de probabilidad de las patentes verdes (que funcionan como nuestras variables proxy de las innovaciones verdes por lo que dada la evidencia empírica) siguiendo el trabajo de Aoyama (2010) e inspirados en Jones (2005), Growiec (2006) y Sánchez y Márquez (2013), vamos a suponer que estas variables constituyen una muestra aleatoria $\{Z_k\}_{k=1}^m = \{(Z_{1k}, Z_{2k}, Z_{3k})\}_{k=1}^m$ independiente e idénticamente distribuida, por lo que podemos obtener los valores de los parámetros del modelo maximizando la función de verosimilitud

$$\mathcal{L}(\theta | Z_k = z_k) = \prod_{k=1}^m P(Z_k = z_k | \theta) = \prod_{k=1}^m f(z_k | \theta),$$

donde para cada valor k , el vector $z_k = (z_{1k}, z_{2k}, z_{3k})$ es una realización del vector aleatorio $Z_k = (Z_{1k}, Z_{2k}, Z_{3k})$. Así, tomando logaritmos tenemos

$$\log[\mathcal{L}(\theta | Z_k = z_k)] = \sum_{i=1}^m \log[f(z_k|\theta)] = \sum_{i=1}^m \log[f(z_{1k}, z_{2k}, z_{3k}|\theta)].$$

Finalmente, substituyendo la función de densidad $f(z_{1i}, z_{2i}, z_{3i})$ de nuestra especificación probabilística obtenemos la verosimilitud

$$(16) \quad \log[\mathcal{L}(\theta)] = \sum_{i=1}^m \left[\log[\alpha\beta\eta\theta^2(1+\theta)] - \frac{1+2\theta}{\theta} \log[(a\gamma_a)^{-\alpha\theta} + (b\gamma_b)^{-\beta\theta} - 1] + \log[a^{2\alpha\theta+1-\alpha} b^{2\beta\theta+1-\beta} \gamma_Y^\eta Y^{-\eta-1}] \right].$$

4.6 Estimación del Modelo y Prueba de Correcta Especificación

Una vez planteada la función de verosimilitud dada en (16), estimaremos los parámetros del modelo utilizando el *Máxima Verosimilitud Completa (MVC)*, el cual estima todos los parámetros del modelo en un solo paso, por lo que tiene la ventaja de no necesitar más supuestos ni de aplicar algún tratamiento a los datos, por lo que provee de estimaciones robustas al poder capturar las interrelaciones entre las variables y minimizar la pérdida de información por manejo de los datos.

Note que esta función depende de los parámetros de las densidades marginales y del parámetro θ de la cópula con la que se obtuvo su distribución conjunta de nuestras variables aleatorias, por lo que se deben estimar todos estos parámetros en el proceso. Esta tarea es computacionalmente intensiva a medida que la muestra incrementa su tamaño, por lo que aunque tiene la ventaja de no necesitar más supuestos y de poder capturar las interrelaciones entre las variables y para el tamaño de nuestra muestra el método es computacionalmente eficiente.

Para hacer la *estimación empírica* del modelo se ocuparon los datos de la USTPO en 2014⁵ para patentes clasificadas como verdes por la misma organización, obteniendo los resultados resumidos en la siguiente tabla.

TABLA I: Estimación de los parámetros de la función de producción con innovaciones verdes con el método de máxima verosimilitud completa (MVC).

Parámetro	Valor estimado
α	1.642
β	0.806
θ	1.724
τ	0.463
σ	0.367

Fuente: Elaboración propia con datos de la USTPO
<http://www.uspto.gov/web/offices/ac/ido/oeip/taf/hcounts.htm>

Para validar estos parámetros obtenidos con el método de máxima verosimilitud total, se obtuvieron los valores de los parámetros de manera individual y aplicamos la prueba Kolmogorov-Smirnov suponiendo que las distribuciones que generan los datos son Pareto, encontrando los valores $\alpha = 1.64216$ y $\beta = 0.806145$ con un valor p igual a $1e-08$, con lo que con estos valores podemos concluir que nuestro método es lo suficientemente robusto para considerarlo una buena aproximación al modelo de probabilidad conjunta con los datos que se tiene. Note que como desde un principio estamos suponiendo independencia e idéntica distribución (supuesto 2) no es necesario realizar estas pruebas para comprobar la correcta especificación del modelo completo, sino que nos vamos a conformar con la prueba descrita anteriormente.

⁵⁵ Agradezco al Mtro. Mario Alejandro Mercado Mendoza por la ayuda que me brindó para la elaboración de esta tesis, tanto por sus comentarios, apoyándome con la descarga de los datos de patentes y la elaboración del código para la estimación de los parámetros de esta sección.

4.4 Identificación del Modelo de Producción con Innovaciones Tecnológicas Verdes

Siguiendo las observaciones de las características de los datos, observamos que al pasar nuestro modelo la prueba de correcta especificación, tenemos evidencia estadística para afirmar que este modelo nos da una descripción adecuada de los datos y sus parámetros, por lo que siguiendo con el proceso de identificación a partir de la probabilidad

$$P\left(a_i > \frac{1}{a}, b_i > \frac{1}{b}, Y_i > Y\right) = \left(\frac{Y}{\gamma_Y}\right)^{-\eta} \left[(a\gamma_a)^{-\alpha\theta} + (b\gamma_b)^{-\beta\theta} - 1\right]^{-\frac{1}{\theta}}.$$

Así, siguiendo los pasos del capítulo anterior obtenemos la *función de producción local para empresas con innovaciones verdes* dada por:

$$Y = A[\varphi L^{-\theta} + (1 - \varphi)K^{-\theta}]^{-\frac{\nu}{\theta}}.$$

Por otra parte, tomando en cuenta que $\gamma_a = \gamma_b = 1$, usando la curva de isoprobabilidad asociada a la FT, ésta relación viene dada por

$$H(a, b) = a^{\alpha\theta} + b^{\beta\theta} = N.$$

Hasta ahora hemos deducido los componentes de nuestro modelo de producción con patentes verdes, el cual tiene la virtud de capturar la dinámica del modelo puesta por el núcleo de empresas que hace innovaciones de ambos tipos, además de tener una forma funcional lo suficientemente general para capturar varios efectos de estas variables en el modelo. Luego entonces, suponemos que el primer objetivo de una empresa en esta economía competitiva es resolver el problema encontrar la mejor tecnología verde a su alcance, es decir, maximizar su función de producción sujeto a la restricción dada por la frontera tecnológica de nuestra empresa.

Los resultados de resolver el problema de elección de la tecnología para una economía con empresas con función de producción con innovaciones verdes del tipo CES son:

$$a^* = \left[\frac{\sqrt{1+K^\theta} N}{\sqrt{1+L^\theta} + \sqrt{1+K^\theta}} \right]^{\frac{1}{\alpha\theta}} \quad y \quad b^* = \left[\frac{\sqrt{1+L^\theta} N}{\sqrt{1+L^\theta} + \sqrt{1+K^\theta}} \right]^{\frac{1}{\beta\theta}}.$$

Sustituyendo estos valores óptimos en la FPL, obtenemos la FPG dada por:

$$Y^* = AN^{\frac{\nu}{\theta}}[(1 + L^\theta)^{-1/2} + (1 + K^\theta)^{-1/2}]^{-\frac{2\nu}{\theta}}.$$

Así, note que la Función de Producción Global con innovaciones verdes tiene una estructura parecida a la función CES. En particular cuando los valores de los insumos L y K son grandes, esta función se puede aproximar por la función CES

$$Y^* = AN^{\frac{\nu}{\theta}}[L^{-\theta/2} + K^{-\theta/2}]^{-\frac{2\nu}{\theta}}.$$

Con base en los resultados obtenidos, podemos obtener la identificación de las funciones que permiten caracterizar a esta economía, la cual tendría una Función de Producción Local dada por

$$Y = A[a^{-2.83}L^{-1.724} + b^{-1.389}K^{-1.724}]^{-\frac{1}{1.724}}.$$

También podemos obtener la función que representa la Frontera de Producción de las innovaciones verdes, la cual está dada por

$$H(a, b) = a^{2.83} + b^{1.389} = N.$$

Así, el modelo final que nos daría con los datos obtenidos de la estimación de la Función de Producción Global (para valores grandes de K y L) está dado por:

$$Y^* = AN^{1.160}[L^{-0.862} + K^{-0.862}]^{-1.160}.$$

Este es un modelo del tipo CES con elasticidad de sustitución positiva dada por $\sigma = 0.36$ lo cual nos indica que los insumos de nuestro modelo son sustitutos netos.

4.8 Hallazgos del Modelo

Observamos que las innovaciones en capital y trabajo están correlacionadas y que su comportamiento probabilístico, en conjunto con el de la producción, da lugar a una aproximación factible y robusta con la cópula Clayton. Además, como obtenemos una elasticidad de sustitución para nuestro modelo de $\sigma < 1$, se tiene que nuestra función de producción tiene curvas isocuantas estrictamente convexas. Por otro lado, note que la frontera de producción resultante tiene exponentes positivos en ambas variables, por lo que las curvas de nivel de esta función son estrictamente cóncavas, lo que nos habla de que las innovaciones se comportan como sustitutos netos, es decir, al aumentar la cantidad de uno disminuye la cantidad del otro y viceversa.

Este comportamiento es el esperado en este tipo de “insumos” pues al invertir una empresa en la innovación de un tipo, ya no puede invertir tanto en el otro tipo de innovación, aunque es importante señalar que hay un efecto de arrastre que le hace invertir en ambos tipos de innovaciones al mismo tiempo. Así, este tipo de comportamiento de las innovaciones repercute directamente en las relaciones de los insumos de capital y de trabajo en la función de producción global, ya que al ser estrictamente convexa, la solución al problema de maximización del producto dada la Frontera Tecnológica tiene un único valor, lo que nos da como resultado una FPG del tipo CES cóncava.

También, como $\theta > 0$, entonces las innovaciones pueden verse como sustitutos brutos, es decir, hay una compensación entre estas innovaciones y le heredan esta condición a los insumos de la producción de capital y trabajo a través del tiempo.

CONCLUSIONES

Las conclusiones de este trabajo se pueden resumir citando tres características importantes de este modelo. Por un lado, se demostró que siguiendo el método de reducción probabilística se puede obtener modelos más generales que los que se asumen en la teoría clásica del crecimiento con innovaciones y que modelan de manera consistente este tipo de fenómenos, ya que se especifica primero el modelo probabilístico basado completamente en el comportamiento estocástico de los datos, y luego validar la estimación con una prueba de correcta especificación, se procede a hacer el proceso de identificación con un la teoría económica, por lo que no se presuponen ciertos comportamientos y formas funcionales que presupone la teoría económica, ni tampoco se manipulan los datos para que ajusten a una especificación dada.

Por otro lado, al hacer de manera adecuada los pasos del enfoque de reducción probabilística, este enfoque arroja evidencia estadística sobre el comportamiento de los datos y el modelo de ajuste de los mismos, por lo cual podemos establecer un vínculo consistente entre el modelo probabilístico basado en la teoría de cópulas (suponiendo que las innovaciones tiene una distribución del tipo Pareto) y el modelo muestral con un modelo de producción con innovaciones del tipo CES, el cual generaliza a los demás modelos convencionales microfundados de la producción como los que plantean Acemoglu (2003), Jones (2005), Growiec (2006, 2017), Andersson (2008), Boopen (2011), Boztokurt y Akan (2014) y Sánchez y Márquez (2015). Así, podemos derivar todos estos modelos a partir de este modelo, e incluso encontrar nuevas formas funcionales si el comportamiento de los datos se especifica con distintas combinaciones de cópulas, distribuciones marginales, dependencia y homogeneidad muestral.

La tercera característica a destacar de nuestro modelo es que permite contrastar el comportamiento de los datos reales con los supuestos teóricos que se hace de estos, ya que al hacer la especificación del mismo basado en los datos, nuestro modelo no necesita hacer supuestos sobre el comportamiento de los datos ni presuponer relaciones funcionales para los mismos. Así entonces, se pueden especificar modelos que permiten responder a algunas

de las cuestiones que se hacen en los trabajos de Acemoglu (2003), Jones (2005), Growiec (2006), y en general en la literatura de la producción y el crecimiento basado en innovaciones tecnológicas y tecnologías verdes.

Por otra parte, una innovación que hacemos al enfoque de Spanos (1986,1999) es el usar cópulas para la especificación de la estructura de dependencia del modelo probabilístico, debido a lo cual podemos considerar relaciones funcionales más generales que en los otros trabajos similares desarrollados hasta el momento, además de que permite de manera natural incorporar nuevas variables y relaciones causales que se pueden plantear desde la empírea, así como combinar variables con diferentes distribuciones marginales.

Ejemplos de estas generalizaciones son el que el modelo considera diferentes tipos de dependencia entre las variables al contemplar distintas familias de cópulas como base para su especificación probabilística, permite obtener funciones de producción generalizadas e incluso obtener nuevas, dependiendo del mecanismo generador de información que revelen los datos.

Además, el modelo que desarrollamos aquí ejemplifica una forma de hacer inferencia estadística de procesos estocásticos con datos económicos observacionales, el cual por un lado, se puede extender para incluir nuevas variables en el proceso, nuevas relaciones funcionales u otros procesos que influyan en el análisis, y por otro lado se puede usar para modelar otros procesos de naturaleza similar a través de este esquema, como el de la demanda, la aversión al riesgo o el equilibrio general.

Los pasos para obtener un modelo de producción con este método son los siguientes:

1. Especificar el modelo estadístico que mejor describa el comportamiento de los datos en dos niveles. Por un lado se especifica la parte probabilística mediante la elección de las distribuciones marginales y la cópula que describa su dependencia. Por otra parte, se especifica el modelo muestral basados en el comportamiento de las muestras de datos que obtenemos. Así, consideraremos que esta especificación

captura el comportamiento estocástico de los datos si pasa las pruebas de correcta especificación del modelo.

2. Estimar los parámetros de nuestro modelo probabilístico mediante algún método de ajuste paramétrico, no paramétrico o semiparamétrico de acuerdo a las características de ajuste de los datos que utilizamos para este modelo.
3. Mediante pruebas de correcta especificación, verificamos que la especificación estadística de nuestro modelo ajusta de manera adecuada con los datos que ocupamos. Si nuestro modelo pasa estas pruebas, pasamos a la siguiente etapa, la de identificación de nuestro modelo con el modelo teórico-económico. En caso de que esto no ocurra, regresamos al paso 1, re-especificando nuestro modelo.
4. Una vez que verificamos la correcta especificación de nuestro modelo, se realiza la identificación de nuestro modelo probabilístico con el modelo teórico-económico que describe a nuestro fenómeno de estudio.

Así, por un lado vemos que se puede explorar la utilización de otros tipos de copulas y en aplicaciones para modelar otros tipos de problemas en la economía, así como profundizar en el estudio de la especificación muestral y las implicaciones de la dependencia y la heterogeneidad de las muestras, lo cual llena de contenido y aplicaciones a los modelos pero también representan un reto para la especificación e identificación de este tipo de modelos.

Se deben estudiar más a fondo los procesos estocásticos y las características de los modelos econométricos con copulas, plantear las pruebas de correcta especificación y los criterios que deben de satisfacer para ser buenos modelos econométricos, lo cual no se desarrolló en este trabajo pues no eran parte de sus objetivos.

Estos puntos imponen una agenda de investigación en sí que está en proceso e involucran herramientas más avanzadas como las series de tiempo, los datos panel y el problema de

causalidad con cópulas, pero requiere un análisis más detallado y el uso de otro tipo de herramientas estadísticas, distintas a las que aquí desarrollamos.

También debemos resaltar que estos modelos pueden incorporar variables endógenas y exógenas al análisis, las cuales pueden tener comportamientos marginales distintos y diferentes formas de dependencia; tomar en cuenta diferentes formas funcionales y efectos que modelan de manera más realista nuestro fenómeno de estudio como el crecimiento no exponencial de la población o la curva de Kuznets; además de incorporar restricciones de diferentes tipos como las institucionales, legales o por la disponibilidad de los insumos, modelos de monopolio, duopolio o cartel.

APÉNDICE A: PRUEBAS DEL MODELO

A.1 Pruebas del Modelo General

PRUEBA DE PROPOSICIÓN 1: *Basados en la especificación del modelo y usando el teorema de Sklar para el caso trivariado, podemos construir una distribución conjunta $T(a, b, Y) = C(U(a), V(b), W(Y), \theta)$ de nuestras variables aleatorias (a_i, b_i, Y_i) , con marginales U , V y W , y parámetro de dependencia θ .*

Así, $T(a, b, Y) = C(U(a), V(b), W(Y), \theta)$ se puede considerar como una forma reducida estocástica de la función de transformación aumentada de la empresa, por lo que fijando la probabilidad $T(a, b, Y) = P_0$, tenemos que $C(U(a), V(b), W(Y), \theta) = P_0$, por lo que podemos despejar (o deja expresado implícitamente) al producto como función de los insumos y las innovaciones tecnológicas como:

$$Y = W_Y^{-1}(C_W^{-1}(U(h_a(a), L; \theta), V(h_b(b), K; \theta); \theta), P_0),$$

donde C_W^{-1} y W_Y^{-1} son las transformaciones inversas de la cópula C respecto a W y Y respectivamente y cuya correcta especificación se obtiene si C_W y W_Y son estrictamente monótonas.

Además, marginalizando respecto de la producción Y (quitando su información de la distribución conjunta) obtenemos la probabilidad asociada $T(a, b) = C(U(a), V(b), 1) = P(a_i > a, b_i > b)$, luego tomando una curva de isoprobabilidad $P(a_i > a, b_i > b) = P_1$ y simplificándola obtenemos la frontera tecnológica de nuestro modelo dada por

$$H(a, b) = g(C_{ab}(U(a), V(b), \theta), P_1). \quad \text{Q.E.D.}$$

PRUEBA DE PROPOSICIÓN 2: *Dados los valores fijos L y K , la FPL y la FT, ambas están definidas sobre el plano (a, b) . Entonces como se maximiza Y sobre el conjunto de pares tecnológicos factibles $D = \{(a, b) | (a, b) \leq H(a, b)\}$ y la cópula es dos creciente*

(compuestas con funciones marginales no decrecientes de a y b), entonces, la FPL es también creciente en D , por lo que debe alcanzar su máximo sobre la FT.

Además, cuando la FPL y la FT son tangentes en el punto óptimo (a^*, b^*) el problema del productor (4) tiene una única solución. Esto solo puede ocurrir si la FPL tiene mayor convexidad que la FT o si la FPL es convexa y la FT es cóncava. Note que las probabilidades de donde provienen la FPL y la FT, dependen de no solo del conjunto de parámetros de la cópula Θ , sino también de los parámetros de las marginales $(\theta_a, \theta_b, \theta_Y)$, por lo que estas funciones dependen de estas dos familias de parámetros. Así, sustituyendo estos valores óptimos en la FPL obtenemos la FPG

$$Y^* = Y^*(L, K, N, U(a^*, L), V(b^*, K); \theta, \theta_a, \theta_b, \theta_Y).$$

Luego, tanto el conjunto de parámetros de la cópula y las distribuciones marginales (para valores dados de L y K) le dan forma a Y^* y a su elasticidad de sustitución

$$\sigma = \sigma(\theta, U(a^*, L), V(b^*, K)) = \left(\frac{d \ln [TMST_{L,K}(U(a^*, L), V(b^*, K)); \theta, \theta_a, \theta_b]}{d \left(\frac{K}{L} \right)} \right)^{-1},$$

donde $TMST_{L,K}(U(a^*, L), V(b^*, K)); \theta, \theta_a, \theta_b$) es la tasa marginal de sustitución técnica entre L y K .

Finalmente, el conjunto de parámetros elegido Θ determina también el grado de dependencia estocástica entre a_i, b_i y Y_i ; el cual puede ser representado mediante una medida de asociación $\mu = \mu(\theta)$ de la cópula (como vimos en el capítulo 1), siempre podemos definir una medida de asociación para la cópula usada en el modelo (Nelsen, 2006). Q.E.D.

PRUEBA DEL COROLARIO 1: Considere una cópula C determinada por un solo parámetro de dependencia θ . Si μ es una medida de dependencia para las variables aleatorias, entonces dicha medida en general depende del parámetro de la cópula, por lo

que se tiene $\mu = m(\theta)$ (Nelsen, 2006). Por otra parte, por el teorema 1, se tiene que la forma de la FPG depende del parámetro de sustituibilidad θ , por lo que su elasticidad de sustitución σ también depende en general de θ , a través de una relación, digamos $\sigma = n(\theta)$. Luego, si tomamos la inversa de ambas relaciones respecto a θ , podemos obtener:

- La elasticidad de sustitución de la FPG $\sigma = \sigma(\theta(\mu)) = \sigma(m^{-1}(\mu))$.
- La medida de dependencia $\mu = \mu(\theta(\sigma)) = \mu(n^{-1}(\sigma))$.

Así, en este caso particular el parámetro θ puede verse como un Puente que relaciona a la elasticidad de sustitución (un parámetro determinístico de la teoría económica) con la medida de dependencia (un concepto estadístico). Q.E.D.

PRUEBA DE LA PROPOSICIÓN 3: Sean a^* y b^* los niveles óptimos de a y b respectivamente para la j -ésima empresa de nuestra economía, los cuales se obtienen de resolver el problema (4). Entonces. El nivel óptimo de esta empresa está dado por

$$Y^* = W_Y^{-1}(C_W^{-1}(U(a^*, L), V(b^*, K), \Theta), P_0).$$

Luego, usando la FPG Y^* para definir la función de beneficio de la j -ésima empresa

$$\pi = W_Y^{-1}(C_W^{-1}(U(a^*, L), V(b^*, K^*), \Theta), P_0) - wL - rK.$$

Así, se obtiene el valor óptimo del beneficio usando el teorema de la envolvente, el cual está dado por $\pi = W_Y^{-1}(C_W^{-1}(U(a^*, L^*), V(b^*, K), \Theta), P_0) - wL^* - rK^*$. Finalmente, obtenemos el beneficio total de la economía sumando los beneficios individuales del total de las empresas de la economía. Q.E.D.

A.2 Pruebas del Modelo de Producción CES

En esta parte obtenemos las fórmulas de las distribuciones de las transformaciones de las variables aleatorias que vamos a ocupar más adelante.

Sea c_i una variable aleatoria con distribución Pareto, la cual tiene función de densidad dada por $f_{c_i}(c) = \rho \frac{\gamma_c^\rho}{c^{\rho+1}}$, donde $c \geq \gamma_c > 0$. Luego, si usamos la transformación $Z = -c_i$, su densidad está dada por

$$f_Z(z) = f_{c_i}(-z) \left| \frac{d(-z)}{dz} \right| = \rho \frac{\gamma_c^\rho}{(-z)^{\rho+1}};$$

por lo que tenemos que

$$P(c_i > c) = F_Z(-c) = \int_{-\infty}^{-c} f_Z(u) du = \int_{-\infty}^{-c} \rho \frac{\gamma_c^\rho}{(-u)^{\rho+1}} du = \left(\frac{c}{\gamma_c} \right)^{-\rho}.$$

Ahora, si para la transformación $R = \frac{1}{c_i}$, tenemos que su densidad está dada por

$$f_R(r) = f_{c_i}\left(\frac{1}{r}\right) \left| \frac{d\left(\frac{1}{r}\right)}{dr} \right| = \rho \frac{\gamma_c^\rho}{\left(\frac{1}{r}\right)^{\rho+1}} \left| -\frac{1}{r^2} \right| = \rho \gamma_c^\rho r^{\rho-1},$$

por lo que la probabilidad resultante es

$$P\left(c_i > \frac{1}{c}\right) = P\left(\frac{1}{R} > \frac{1}{c}\right) = P(R < c) = F_R(c) = \int_0^c \rho \gamma_c^\rho u^{\rho-1} du = \gamma_c^\rho c^\rho,$$

donde $0 < \gamma_c < \frac{1}{c}$.

Similarmente, si d_i una variable aleatoria con distribución Weibull, con rango $d \geq \gamma_d > 0$, su distribución está dada por $f_{d_i}(d) = \frac{\mu}{\gamma_d} \left(\frac{d}{\gamma_d}\right)^{\mu-1} \exp\left\{-\left(\frac{d}{\gamma_d}\right)^\mu\right\}$.

Así, usando la transformación $Z = -d_i$ se tiene que

$$f_Z(z) = f_{d_i}(-z) \left| \frac{d(-z)}{dz} \right| = \frac{\mu}{\gamma_d} \left(\frac{-z}{\gamma_d}\right)^{\mu-1} \exp\left\{-\left(\frac{-z}{\gamma_d}\right)^\mu\right\},$$

por lo que obtenemos la probabilidad

$$P(c_i > c) = F_Z(-c) = \int_{-\infty}^{-c} f_Z(u) du = \int_{-\infty}^{-c} \frac{\mu}{\gamma_d} \left(\frac{-u}{\gamma_d}\right)^{\mu-1} \exp\left\{-\left(\frac{-u}{\gamma_d}\right)^\mu\right\} du = e^{-\left(\frac{c}{\gamma_d}\right)^\mu}.$$

Y ahora, para la transformación $R = \frac{1}{d_i}$, donde $0 < \gamma_d < \frac{1}{d}$, tenemos que

$$f_R(r) = f_{d_i}\left(\frac{1}{r}\right) \left| \frac{d\left(\frac{1}{r}\right)}{dr} \right| = \frac{\mu}{\gamma_d} \left(\frac{1}{r\gamma_d}\right)^{\mu-1} \exp\left\{-\left(\frac{1}{r\gamma_d}\right)^\mu\right\} \left| -\frac{1}{r^2} \right| = \mu\gamma_d^{-\mu} r^{-\mu-1} e^{-(r\gamma_d)^{-\mu}},$$

por lo que la probabilidad resultante está dada por

$$P\left(d_i > \frac{1}{d}\right) = P\left(\frac{1}{R} > \frac{1}{d}\right) = P(R < d) = F_R(d) = \int_0^d \mu\gamma_d^\mu u^{-\mu-1} e^{-(u\gamma_d)^{-\mu}} du,$$

es decir, tenemos que

$$P\left(d_i > \frac{1}{d}\right) = F_R(d) = e^{-(d\gamma_d)^{-\mu}}. \quad Q.E.D.$$

PRUEBA DEL LEMA 4: Primero, para obtener la ecuación (7), usamos los resultados del apéndice B.1 para transformaciones de las variables aleatorias Pareto dadas por $a_i = \frac{1}{R}$, $b_i = \frac{1}{R}$ y $Y_i = -Z$, y los sustituimos en la cópula Producto-Clayton (5). Similarmente, para obtener la ecuación (8) usamos las transformaciones de las variables aleatorias con distribución Pareto dadas por $a_i = -Z$ y $b_i = -Z$ y la cópula Clayton (6). Q.E.D.

PRUEBA DEL LEMA 5: Fijando la probabilidad $P_0 = P\left(a_i > \frac{1}{a}, b_i > \frac{1}{b}, Y_i > Y\right)$ en la ecuación (9) y despejando $\left(\frac{Y}{\gamma_Y}\right)^\eta$, de tiene que $\left(\frac{Y}{\gamma_Y}\right)^\eta = \frac{1}{P_0} \left[(\alpha\gamma_a)^{-\alpha\theta} + (b\gamma_b)^{-\beta\theta} - 1 \right]^{-\frac{1}{\theta}}$.

Luego, defina $\gamma_a = \gamma_L(1 + L^{-\theta})^{-\frac{1}{\alpha\theta}}$ y $\gamma_b = \gamma_K(1 + K^{-\theta})^{-\frac{1}{\beta\theta}}$, y substituyendo en la ecuación anterior tenemos que:

$$Y = \frac{\gamma_Y}{P_0^{\frac{1}{\eta}}} \left[[a\gamma_L(1 + L^{-\theta})^{-\frac{1}{\alpha\theta}}]^{-\alpha\theta} + [b\gamma_K(1 + K^{-\theta})^{-\frac{1}{\beta\theta}}]^{-\beta\theta} - 1 \right]^{-\left(\frac{1}{\eta}\right)\left(\frac{1}{\theta}\right)}$$

$$= \frac{\gamma_Y}{P_0^v} \left[(a\gamma_L)^{-\alpha\theta} L^{-\theta} + (b\gamma_K)^{-\beta\theta} K^{-\theta} + (a\gamma_L)^{-\alpha\theta} + (b\gamma_K)^{-\beta\theta} - 1 \right]^{-\frac{v}{\theta}}, \text{ donde } v = \frac{1}{\eta}.$$

Finalmente, cambiando la escala de las variables para que $(a\gamma_L)^{-\alpha\theta} + (b\gamma_K)^{-\beta\theta} = 1$, tenemos que

$$Y = A \left[(a\gamma_L)^{-\alpha\theta} L^{-\theta} + (b\gamma_K)^{-\beta\theta} K^{-\theta} \right]^{-\frac{v}{\theta}},$$

donde se tienen los valores $A = \frac{\gamma_Y}{P_0^v}$, $\varphi = (a\gamma_L)^{-\alpha\theta}$, y $1 - \varphi = (b\gamma_K)^{-\beta\theta}$.

Ahora, para derivar la FT, primero fijamos la probabilidad $P_1 = P(a_i > a, b_i > b)$ en la ecuación (10). Así, simplificándola ecuación resultante tenemos que

$$H(a, b) = \left(\frac{a}{\gamma_a}\right)^{\alpha\theta} + \left(\frac{b}{\gamma_b}\right)^{\beta\theta} = P_1^{-\theta} + 1 = N. \quad \text{Q.E.D.}$$

PRUEBA DEL LEMA 6: *Tenemos los siguientes casos para algunos valores especiales del parámetro θ , en los cuales se calculan τ y σ simplemente sustituyendo los valores de θ en las formulas (11) y (12) respectivamente.*

a) *Caso Cobb-Douglas ($\theta = 0$): La cópula correspondiente para este caso es la Producto dada por $C(U, V, W) = UVW$. Así, tomando las transformaciones como en el Lema 4, tenemos la probabilidad*

$$P\left(Y_i > Y, a_i > \frac{1}{a}, b_i > \frac{1}{b}\right) = C(F_R(-Y), F_Z(a), F_Z(b)) = \left(\frac{Y}{\gamma_Y}\right)^{-\eta} (\gamma_a a)^\alpha (\gamma_b b)^\beta.$$

Luego, tomando la curva de isoprobabilidad de valor P_0 y asumiendo las identidades $\gamma_a = \gamma_L L$ y $\gamma_b = \gamma_K K$, tenemos que

$$P_0 = \left(\frac{Y}{\gamma_Y}\right)^{-\eta} (a^\alpha \gamma_L^\alpha L^\alpha) (b^\beta \gamma_K^\beta K^\beta),$$

y finalmente, definiendo $v = \frac{1}{\eta}$ y resolviendo para Y , obtenemos la FPL dada por

$$Y = \frac{\gamma_Y}{P_0^v} [a\gamma_L]^\alpha [b\gamma_K]^\beta L^\alpha K^\beta = AL^\alpha K^\beta.$$

Por otro lado, marginalizando respecto a la variable Y se obtiene la cópula para la FT dada por $C_{ab}(U, V) = C(U, V, 1) = UV$, por lo que sustituyendo como en el lema 4, tenemos la siguiente probabilidad;

$$P(a_i > a, b_i > b) = C_{a,b}(F_R(-a), F_R(-b)) = \left(\frac{a}{\gamma_a}\right)^{-\alpha} \left(\frac{b}{\gamma_b}\right)^{-\beta}.$$

Ahora, fijando esta probabilidad a un nivel P_1 y reduciendo esta ecuación, tenemos la FT

$$H(a, b) = a^\alpha b^\beta = N, \text{ donde } N = \frac{\gamma_a^\alpha \gamma_b^\beta}{P_1}.$$

b) Caso de Función de Producción Leontief ($\theta \rightarrow \infty$): La cópula (5) converge en este caso a la Producto-Cota Superior de Frechét-Hoefding $C(U, V, W) = W \min\{U, V\}$. Así, tomando las transformaciones de la variable aleatoria como en el lema 4, se tiene que:

$$\begin{aligned} P\left(Y_i > Y, a_i > \frac{1}{a}, b_i > \frac{1}{b}\right) &= C(F_R(-Y), F_R(a), F_R(b)) \\ &= \left(\frac{Y}{\gamma_Y}\right)^{-\eta} \min\{(\gamma_a a)^\alpha, (\gamma_b b)^\beta\} \end{aligned}$$

Luego, tomando una curva de isoprobabilidad a un valor P_0 , y asumiendo las identidades dadas por $\gamma_a^\alpha = \gamma_L^\alpha L$, $\gamma_b^\beta = \gamma_K^\beta K$, se tiene que

$$P_0 = \left(\frac{Y}{\gamma_Y}\right)^{-\eta} \min\{(a\gamma_L)^\alpha L, (b\gamma_K)^\beta K\}.$$

Finalmente, definiendo $v = \frac{1}{\eta}$ y despejando para Y se tiene que

$$Y = \frac{\gamma_Y}{P_0^v} [\min\{(a\gamma_L)^\alpha L, (b\gamma_K)^\beta K\}]^v = A [\min\{(a\gamma_L)^\alpha L, (b\gamma_K)^\beta K\}]^v.$$

Por otro lado, marginalizando respecto a la variable Y se obtiene la cópula para la FT dada por $C_{ab}(U, V) = C(U, V, 1) = \min\{U, V\}$, por lo que sustituyendo como en el lema 4, tenemos la siguiente probabilidad;

$$P(a_i > a, b_i > b) = C_{a,b}(F_R(-a), F_R(-b)) = \min \left\{ \left(\frac{a}{\gamma_a} \right)^{-\alpha}, \left(\frac{b}{\gamma_b} \right)^{-\beta} \right\}.$$

Ahora, fijando esta probabilidad a un nivel P_1 y reduciendo esta ecuación, tenemos la FT

$$H(a, b) = \max \left\{ \left(\frac{a}{\gamma_a} \right)^\alpha, \left(\frac{b}{\gamma_b} \right)^\beta \right\} = N, \text{ donde } N = P_1^{-1}.$$

c) Caso de Función de Producción de Sustituto Perfectos ($\theta = -1$): La cópula (5) converge en este caso a la Producto-Cota Inferior de Frechét-Hoeffding, la cual está dada por $C(U, V, W) = W \max\{U + V - 1, 0\}$. Así, tomando las transformaciones de la variable aleatoria como en el lema 4, se tiene que:

$$\begin{aligned} P\left(Y_i > Y, a_i > \frac{1}{a}, b_i > \frac{1}{b}\right) &= C(F_R(-Y), F_R(a), F_R(b)) \\ &= \left(\frac{Y}{\gamma_Y}\right)^{-\eta} \max\{(\gamma_a a)^\alpha + (\gamma_b b)^\beta - 1, 0\}. \end{aligned}$$

Luego, tomando una curva de isoprobabilidad a un valor P_0 , y asumiendo las identidades $\gamma_a = \gamma_L(1 + L)^{\frac{1}{\alpha}}$ y $\gamma_b = \gamma_K(1 + K)^{\frac{1}{\beta}}$, se tiene que

$$P_0 = \left(\frac{Y}{\gamma_Y}\right)^{-\eta} [(a\gamma_L)^\alpha L + (b\gamma_K)^\beta K + (a\gamma_L)^\alpha + (b\gamma_K)^\beta - 1].$$

Finalmente, ajustando la escala de los insumos tal que $(a\gamma_L)^\alpha + (b\gamma_K)^\beta = 1$, definiendo $v = \frac{1}{\eta}$ y despejando para Y se tiene que

$$Y = A[(a\gamma_L)^\alpha L + (b\gamma_K)^\beta K]^v.$$

Por otro lado, marginalizando respecto a la variable Y se obtiene la cópula para la FT dada por $C_{ab}(U, V) = C(U, V, 1) = W\max\{U + V - 1, 0\}$, por lo que sustituyendo como en el lema 4, tenemos la siguiente probabilidad;

$$P(a_i > a, b_i > b) = C_{a,b}(F_R(-a), F_R(-b)) = \max\left\{\left(\frac{a}{\gamma_a}\right)^{-\alpha} + \left(\frac{b}{\gamma_b}\right)^{-\beta} - 1, 0\right\}.$$

Ahora, fijando esta probabilidad a un nivel P_1 y recordando que como N indica el tamaño de la FT, debe haber una relación positiva de esta con las variables a y b , obtenemos la FT dada por la ecuación

$$H(a, b) = \left[\left(\frac{a}{\gamma_a}\right)^{-\alpha} + \left(\frac{b}{\gamma_b}\right)^{-\beta}\right]^{-1} = N, \text{ donde } N = [P_1 + 1]^{-1}. \quad Q.E.D.$$

PRUEBA DE LA PROPOSICIÓN 4: Para resolver el problema de la elección de la tecnología, definimos la función lagrangeana

$$L = A[(a\gamma_L)^{-\alpha\theta} L^{-\theta} + (b\gamma_K)^{-\beta\theta} K^{-\theta}]^{-\frac{v}{\theta}} - \lambda \left[\left(\frac{a}{\gamma_a}\right)^{\alpha\theta} + \left(\frac{b}{\gamma_b}\right)^{\beta\theta} - N \right].$$

Luego, combinando las condiciones de primer orden se tiene que

$$b^{2\beta\theta} = \frac{1 + L^\theta}{1 + K^\theta} a^{2\alpha\theta}.$$

Sustituyendo la condición anterior en la FT y resolviendo para $a^{\alpha\theta}$, tenemos que

$$a^{\alpha\theta} = \frac{\gamma_L^{\alpha\theta} \gamma_K^{\beta\theta} L^\theta K^\theta N}{\gamma_K^{\beta\theta} K^\theta (1 + L^\theta) + \gamma_L^{\alpha\theta} L^\theta \sqrt{(1 + L^\theta)(1 + K^\theta)}}$$

Similarmente, derivamos la expresión para el valor óptimo de $b^{\beta\theta}$ dada por

$$b^{\beta\theta} = \frac{\gamma_L^{\alpha\theta} \gamma_K^{\beta\theta} L^\theta K^\theta N}{\gamma_K^{\beta\theta} K^\theta \sqrt{(1+L^\theta)(1+K^\theta)} + \gamma_L^{\alpha\theta} L^\theta (1+K^\theta)}.$$

Finalmente, sustituyendo estas expresiones en la FPL, obtenemos la FPG dada por

$$Y^* = A' \left[\gamma_L^{-\alpha\theta} \sqrt{1+L^\theta} L^{-\theta} + \gamma_K^{-\beta\theta} \sqrt{1+K^\theta} K^{-\theta} \right]^{\frac{2\nu}{\theta}}, \text{ donde } A' = \frac{\gamma_y}{P_0^\nu} N^{\frac{\nu}{\theta}}. \quad Q.E.D.$$

PRUEBA DE LA PROPOSICIÓN 5: Para probar los casos de la FPL Cobb-Douglas y de Sustitutos Perfectos con su respectiva FT asociada, se siguen los mismos pasos que en la proposición 4, excepto en el caso de a función de producción del tipo Leontief con su FT, en la cual se dan argumentos de convexidad para su solución, de la misma forma que en la literatura convencional. Q.E.D.

APÉNDICE B: RUTA DE BUSQUEDA DE LA BIBLIOGRAFÍA

En estas tablas se ponen los resultados numéricos de la revisión bibliográfica que se hizo para esta tesis. Esta búsqueda ayudó a complementar la bibliografía sugerida por el comité tutorial y a conocer el estado del arte de esta investigación.

En la primera tabla, la búsqueda se hizo utilizando las palabras clave relacionadas con este trabajo, mientras que en la segunda tabla se utilizaron términos más específicos para la búsqueda sistematizada de artículos y libros que contuvieran estos temas. Así, en la columna izquierda se pusieron las palabras clave o términos de la búsqueda realizada, reportando en frente de ellas el número de resultados de libros y artículos académicos sobre este tema en el primer renglón de cada término, y el de libros y artículos académicos que contenían esta frase completa en el segundo renglón, para resultados en periodos desde un año específico hasta la actualidad.

Palabras Clave	Desde cualquier tiempo	Desde 2014	Desde 2017	2018
Modelos de crecimiento económico	966000	14100	13000	15300
	2020	438	150	69
Economic growth models	4220000	77400	20500	21500
	7320	2000	665	267
Modelos de crecimiento endógeno	50100	15400	10100	3670
	1990	431	131	55
Endogenous growth models	2400000	152000	32900	23100
	18800	4330	1330	530
Modelos de crecimiento económico endógeno	44700	15400	9030	3350
	95	21	8	7
Models of endogenous economic growth	780 000	18100	17100	17500
	306	72	28	7
Innovaciones tecnológicas	161000	13400	15400	9730
	53900	15400	5750	2010
Technological innovations	2510000	188000	32700	21700
	403000	28700	17600	10500
Teoría de cópulas	17900	6210	1790	638
	228	69	26	7
Copula theory	111000	17700	11800	4960
	4210	2030	870	395

Copulas theory	29000	11700	4730	2240
	381	160	54	24
Economía de los recursos naturales	630000	34800	16900	15500
	2490	526	150	61
Economy of natural resources	3850000	582000	98000	40400
	1270	534	185	71
Economía ambiental	766000	14200	14200	15500
	14500	3660	1180	462
Environmental economics	4420000	53000	17100	16200
	1430000	38000	21400	5100
Environmental economics natural resource	84600	18500	27200	19300
	747	157	53	19
Cópula Clayton	10400	4400	1710	733
	1060	530	213	94
Clayton copula	10400	4400	1710	733
	5560	2620	1020	437
Cópula Clayton en recursos naturales	67	19	9	5
	0	0	0	0
Clayton copula natural resources	2050	1020	444	219
	0	0	0	0
Cópula Clayton en economía ambiental	42	9	4	1
	0	0	0	0
Clayton copula environmental economics	2300	1160	509	256
	0	0	0	0

Búsqueda por términos	Cualquier tiempo	2014	2017	2018
Teoría de cópulas en modelos de crecimiento endógeno	1280	261	77	29
	0	0	0	0
Copula theory endogenous growth models	4860	1820	752	319
	0	0	0	0
Teoría de cópulas en modelos de crecimiento económico endógeno	1170	234	70	28
	0	0	0	0
Copula theory models of endogenous economic growth	3850	1490	620	268
	0	0	0	0
Teoría de cópulas crecimiento endógeno innovaciones tecnológicas	570	106	34	14
	0	0	0	0
Copula theory models of endogenous economic growth technological innovations	2650	927	393	157
	0	0	0	0
Teoría de cópulas crecimiento económico innovaciones tecnológicas	3210	826	248	97
	0	0	0	0
Copula theory economic growth technological innovations	12100	4090	1660	717
	0	0	0	0
Crecimiento endógeno innovaciones tecnológicas	19400	13900	4590	1810
	0	0	0	0
Endogenous growth technological innovations	135000	16500	16900	16800
	0	0	0	0
Modelos crecimiento económico endogenizado por innovaciones tecnológicas	439	107	42	17
	0	0	0	0
Models economic growth endogenous technological innovations	112000	16900	16800	17000
	0	0	0	0
Teoría de cópulas crecimiento endógeno innovaciones tecnológicas recursos naturales	516	96	30	13
	0	0	0	0
Copula theory endogenous growth technological innovations natural resources	1990	631	269	103
	0	0	0	0
Cópula Clayton recursos naturales	57	19	9	5
	0	0	0	0
Clayton copula natural resources	2050	1020	444	219
	0	0	0	0

BIBLIOGRAFÍA

ACEMOGLU, D. (2003): "Labor- and capital-augmenting technical change," *Journal of the European Economic Association*, 1, 1-37.

AGHION, P. y P. HOWITT (2009): *The economics of growth*. Cambridge MA: The MIT Press.

AGHION, P. y P. HOWITT (1992): "A Model of growth through creative destruction," *Econometrica*, 60, 323-351.

ANDERSSON, M. (2008): "Environmental and economic implications of taxing and trading carbon: some European experiences," en E. Milne (Ed). *The reality of carbon taxes in the 21st century*. Vermont, Environmental Tax Policy Institute: 61-87.

AOYAMA, H., H. IYETOMI, Y. FUJIWARA, Y. IKEDA y W. SOUMA (2010): "A paradigm shift from production function to production copula: statistical description of production activity firms," *Quantitative Finance*, 12-9, 1453-1466.

ATKINSON, A.B., AND J.E. STIGLITZ (1969): "A new view of technological change," *Economic Journal*, 79, 573-578.

BARRO, R. y X. SALA-I-MARTIN (2004): *Economic growth*. Cambridge MA: The MIT Press.

BASU, S., y D.N. WEIL (1998): "Appropriate technology and growth," *Quarterly Journal of Economics*, 113, 1025-1054.

BOOPEN, S. (2011): "On the relationship between CO₂ emissions and economic growth: the Mauritian experience," *Mauritius Environmental Outlook Report*, <http://www.csae.ox.ac.uk/conferences/2011-EDiA/papers/776-Seetanah.pdf>.

BOZTKURT, C., y Y. AKAN (2014): "Economic growth, CO₂ emissions and energy consumption: the Turkish case," *International Journal of Energy Economics and Policy*, Vol. 4, No. 3, pp.573-578.

BRIDA, G. (2008): "Población y crecimiento económico: una versión mejorada del modelo de Solow," *El trimestre económico*, Vol. LXXV, 5-22.

BROCK, W., y S. TAYLOR (2004): "The green Solow model," Working Paper 10557, *National Bureau of Economic Research*, Cambridge, MA.

CASELLI, F., y W.J. COLEMAN (2006): "The world technology frontier," *American Economic Review*, 96, 499-522.

DURANTE, F., y C. SEMPI (2010): "Copula theory: an introduction," en *Copula theory and its applications*, 96, 499-522.

FÄRE, R., S. GROSSKOPF y C. PASURKA (2006): "Multilateral Productivity Comparisons When Some Outputs Are Undesirable: A Nonparametric Approach," *The review of economics and statistics*, 1989, vol. 71, issue 1, 90-98.

FORSTER, B. (1975): "Optimal pollution control with a nonconstant exponential rate of decay," *Journal of Environmental Economics and Management*, Vol. 2, Issue 1, 1-6.

FRÉCHET, M. (1951): "Sur les tableaux de corrélation dont les marges sont données," *Annales of the University*, A14, 53-77.

- GROSSMAN, G., E. HELPMAN, E. OBERFIELD y T. SAMPSON (2017), “Balanced growth despite Uzawa,” *American Economic Review*, 107(4), 1293-1312.
- GROWIEC, J. (2006): “A new class of production functions and an argument against purely labor-augmenting technical change,” *International Journal of Economics Theory*, 4, 483-502.
- GROWIEC, J. (2013): “A microfundation for normalized CES production functions with factor-augmenting technical change,” *Journal of Economic Dynamics and Control*, 37, 2336-2350.
- GROWIEC, J. (2017): “Factor-specific technology choice,” *Narodowy Bank Polski Education & Publishing Department*, Working paper No. 265.
- HABAKKUK, H.J. (1962): *American and British technology in the nineteenth century: the search for labour-saving inventions*. Cambridge University Press.
- HENDRICKSON, C., A. HORVATH, S. JOSHI y L. LAVE (1998): “Economic input-output models for environmental life-cycle assessment,” *Environmental science and technology*, Abril, 184-191.
- HILDENBRAND, W. (1981): “Short-run production functions based on microdata,” *Econometrica*, 49-5, 1905-1125.
- HOUTHAKKER, H.S. (1955): “The Pareto distribution and the Cobb-Douglas production function in activity analysis,” *Review of Economic Studies*, 23, 27-31.
- JOHANSEN, L. (1972): *Production functions*. North-Holland, Amsterdam.

- JONDROW, J., C. KNOX LOVELL, I. MATEROV y P. SCHMIDT (1982): "On the estimation of technical inefficiency in the stochastic frontier production function model," *Journal of Econometrics*, 19, 233-238.
- JONES, C., (1999): "Growth: with or without scale effects", *American Economic Association*, vol. 89, No. 139-144.
- JONES, C., (2005): "The shape of production functions and the direction of technical change", *Quarterly Journal of Economics*, 120, 517-549.
- JONES, L.E., and R.E. MANUELLI (1990): "A convex model of equilibrium growth: theory and policy implications", *Journal of Political Economy*, 98, 1008-1038.
- KORTUM, S.S. (1997): "Research, patenting, and technological change," *Econometrica*, 65, 1389-1419.
- LEÓN-LEDEZMA, M. Y SATCHI, M. (2018). "Appropriate technology and balanced growth," *The review of economic studies*, rdy002, <https://doi.org/10.1093/restud/rdy002>
- LUCAS, E. R., Jr. (1988). "On the mechanics of economic development," *Journal of Monetary Economics*, 22, 3-42.
- LUCAS, E. R., Jr. (2009). "Ideas and growth," *Economica*, 76, 1-19.
- NELSEN, R. B. (2006): "An introduction to copulas," New York: Springer-Verlag.
- PLOURDE, C.G. (1972): "A model of waste accumulation and disposal," *The Canadian Journal of Economics*, Vol. 5, 119-125.
- RINCON, L. (2007): "Curso intermedio de probabilidad," UNAM, México.

ROMER, P.M. (1986): "Increasing Returns and Long-Run Growth," *Journal of Political Economy*, 94, 1002-1037.

ROMER, P.M. (1990): "Endogenous technological change," *Journal of Political Economy*, 98, S71-S102.

SÁNCHEZ, A., y J.M. MÁRQUEZ (2013): "What Drives the Shape of the Production Function and Technical Progress? A Copula Approach" Working paper, Economics Research Institute, UNAM, Mexico.

SÁNCHEZ, A., y J.M. MÁRQUEZ (2015): "The Complementarity between Scarce Technological Innovations" Working paper, Economics Research Institute, UNAM, Mexico.

SÁNCHEZ, A., y J.M. MÁRQUEZ (2018): "¿Debe Ser el Cambio Tecnológico Ahorrador Puro de Trabajo en un modelo Endógeno de Crecimiento de la Producción?" *Investigación Económica*, UNAM, México. Vol. 77, No. 303.

SKLAR, A. (1959): "Fonctions de Répartition à n Dimensions et Leurs Marges," *Publications de l'Institut Statistique de l'Université de Paris*, 8, 229-231.

SOLOW, R.M. (1956): "A contribution to the theory of economic growth," *Quarterly Journal of Economic*, 70, 65-94.

SPANOS, A. (1986): *Statistical foundations of econometric modeling*, Cambridge MA: Cambridge University Press.

SPANOS, A. (1999): *Probability theory and statistical inference: econometric modeling with observational data*, Cambridge MA: Cambridge University Press.

SPANOS, A. y A.MCGUIRK (2001): “The model specification problem from a probabilistic reduction perspective,” *American Journal of Agricultural Economics*, Volume 83, No. 5, 1168–1176.

STERN, D., M. COMMON y E. BARBIER (1996): “Economic Growth and Environmental Degradation: The Environmental Kuznets Curve and Sustainable Development,” *World Development*, Elsevier Science, Gran Bretaña, Vol. 24, No. 7, 1151-1160.

SUDHAKAR, J. (2012): “An input-output analysis to estimate embodied energy of goods”, *International Journal of Scientific and Research Publications*, Vol. 2, No. 11.

UZAWA, H. (1961): “Neutral inventions and the stability of growth equilibrium,” *Review of Economic Studies*, 28, 117-124.

ZHANG, W. B. (2006): “*Economic growth with income and wealth distribution*,” London: Palgrave McMillan.