

Estudios de Economía Aplicada
Nº 12, 1999. Págs. 17-33

Análisis de la Función de Producción Agraria para distintos niveles de Agregación

CEPAS LÓPEZ, S.
DIOS PALOMARES, R.
Departamento de Estadística

Esta versión incluye todas las correcciones sugeridas por el evaluador, las cuales nos han parecido oportunas y por las que les quedamos muy agradecidos.

RESUMEN

El principal objetivo de este estudio es estimar la función de producción para el sector agrario español (FPA). También se pretende comprobar si el planteamiento de diferentes niveles de agregación de los datos afecta a las repercusiones de la multicolinealidad, realizar un análisis descriptivo de las variables que van a entrar en los modelos, y evaluar la incidencia de la multicolinealidad sobre la FPA. Se utilizan cuatro niveles de agregación de las variables explicativas, y se estiman las correspondientes funciones de producción por MCO tras introducir dichas variables linealizadas en una especificación Cobb-Douglas. Igualmente, se detecta la intensidad de la multicolinealidad mediante el cálculo de los índices de condición y de la descomposición de la varianza. Los resultados obtenidos indican que la estimación de la FPA se ve afectada negativamente por la multicolinealidad, y que el uso de agregaciones no consigue mejorar esa situación. Sin embargo, los modelos presentaron excelentes propiedades desde el punto de vista predictivo.

Palabras clave: multicolinealidad, función de producción agregada, descomposición de la varianza

ABSTRACT

The main aim of this research is to estimate the aggregate production function (APF) for Spain's agricultural sector. Other objectives are to carry out a descriptive analysis of the production inputs, to assess the effects of multicollinearity on the APF, and to find out whether such effects are affected when using several aggregative levels of the inputs. Specifically, four aggregative levels are used, and the four outcoming models are estimated by Ordinary Least Squares, employing a Cobb-Douglas specification. Equally, multicollinearity's severity is detected using variance decomposition and condition indexes. It is

concluded that multicollinearity seems to affect negatively the APF estimation, with no improvement provided by the aggregation. However, the models estimated showed favourable properties in terms of predictive capacity.

Key words: multicollinearity, aggregated production function, variance decomposition.

Clasificación A.M.S.: 62P20

Artículo recibido en abril de 1998. Revisado en enero de 1999.

1. Introducción

En el presente artículo se va a analizar la función de producción agregada para el sector agrario español. El estudio de la función de producción tiene dos utilidades fundamentales: por un lado, nos permite conocer como afectan los diferentes factores productivos a la producción total, y por el otro, nos es posible realizar predicciones sobre la evolución futura de la producción.

Podemos definir la función de producción a nivel microeconómico como "la relación entre las cantidades de recursos empleadas por una empresa y la cantidad de bienes y servicios que ésta produce por unidad de tiempo". Si las funciones de producción correspondientes a todas las unidades productivas de un sector o economía se agregan, obtenemos lo que se conoce como función de producción agregada. El concepto de la función de producción se ha aplicado empíricamente a numerosas economías nacionales y sectoriales, empleando diferentes especificaciones. La más común es la Cobb-Douglas, que se ha impuesto debido a su facilidad de cálculo, pese a que impone una serie de restricciones sobre los parámetros a estimar. Ejemplos de estimaciones de funciones de producción agregadas basadas en la especificación Cobb-Douglas serían los de Solow (1957), Griliches y Ringstad (1971), y Woodfield (1972).

En lo referente a la agricultura, es muy abundante la elaboración de funciones de producción a nivel micro, como en Cañas y otros (1994) o Sebastián y Rodríguez (1978), pero no tanto la de funciones correspondientes al sector agrícola en su conjunto. Un interesante análisis comparativo de funciones de producción agrarias agregadas entre países puede encontrarse en Yamada y Ruttan (1990). También Walters (1963) recopiló numerosos estudios de funciones de producción siendo un gran número de ellos relativos al sector agrícola.

El estudio de cualquier fenómeno económico en el contexto del modelo econométrico, se puede realizar desde dos perspectivas generales que son el análisis estructural y la predicción. Concretando al entorno económico que ocupa nuestro interés, son muchos los problemas que se presentan al analista para llegar a conseguir una estructura aceptable y con buenas propiedades predictivas de la función de producción agraria a nivel nacional. Dichos problemas se refieren tanto al aspecto de la especificación del modelo como al de la estimación. Debido a estas razones son realmente escasos los trabajos aplicados que hasta la fecha han desarrollado modelos fiables de dicha función.

En cuanto a la especificación, resulta problemático no solo establecer los factores de producción sino medirlos de forma adecuada, siendo una de las dificultades principales la valoración de la tierra. En la definición de los factores, hay que decir que es fundamental el nivel de agregación que se establezca y en este sentido entra en juego la finalidad del estudio. Si el objetivo del mismo es la predicción, resulta totalmente irrelevante el número y nivel de agregación de los factores, siendo deseable conseguir

un modelo estable y con buena fiabilidad predictiva. En cambio, si tenemos un interés de tipo estructural, sería conveniente incluir aquellos factores cuyas elasticidades queramos estimar. Así, cabría pensar que la mejor especificación desde este último punto de vista, sería aquella que contara con el máximo número de factores, ya que nos permitiría extraer gran cantidad de información con respecto al problema en estudio.

Sin embargo, hay que tener en cuenta la incidencia casi segura de un problema añadido a los hasta aquí comentados, que es la presencia de multicolinealidad en la matriz de datos de los factores de producción. Sabido es que cuando existe colinealidad entre las variables exógenas del modelo econométrico, las estimaciones de los parámetros sufren una gran inestabilidad, siendo extremadamente sensibles a pequeños cambios en las observaciones.

En cuanto a la repercusión en la predicción, la multicolinealidad no debe afectar siempre que se mantenga la misma relación entre las variables para el periodo de interés. Dado que la gravedad del problema en el caso de la función de producción es importante debido a las propias características de la relación existente entre los factores, resulta imprescindible realizar una valoración de la repercusión de la multicolinealidad en la estimación del modelo.

La multicolinealidad es un problema inherente a la matriz de datos, y por tanto, es de esperar que si especificamos distintos modelos para explicar la producción, también sea distinto el efecto pernicioso provocado por dicho problema. Teniendo en cuenta que la especificación está en principio abierta al criterio del analista, ante la imposibilidad de encontrar otros remedios mejores, es aceptable la idea de minimizar la repercusión de la multicolinealidad, adoptando aquella especificación que menos problema presente, siempre y cuando sea aceptable la estimación conseguida.

Basándonos en todo lo anterior, hemos desarrollado el análisis econométrico de la función de producción agraria a nivel nacional, para cuatro niveles de agregación distintos, de modo que el más agregado contiene solo los factores Capital y Trabajo, y el menos agregado incluye Amortizaciones, Conservación, Ganadería, Carburantes, Electricidad, Fertilizantes, Fitosanitarios, Piensos, Semillas, Regadío, Secano, Trabajo Asalariado y Trabajo Familiar.

Con el fin de llegar a la estimación del modelo que mejor nos permita conocer el proceso de producción agraria, se estudian las cuatro especificaciones teniendo en cuenta que mientras menor es la agregación más información obtenemos. Sin embargo, podría ocurrir que para el modelo menos agregado la inestabilidad fuera mayor. Esta problemática es la que afrontamos, sobre todo desde el punto de vista estructural de modo que analizaremos la gravedad de la multicolinealidad para cada modelo y las propiedades de los mismos en las dos finalidades comentadas.

En definitiva, nos planteamos como objetivo principal estimar la función de producción para el sector agrario español siendo objetivos secundarios los siguientes:

- Evaluar la incidencia de la multicolinealidad sobre la función de producción agregada agraria, para distintos niveles de agregación.
- Comprobar si el planteamiento de diferentes niveles de agregación de los datos afecta a las repercusiones de la multicolinealidad.
- Contrastar la hipótesis de existencia de rendimientos constantes a escala en el sector agrario español.

2. Metodología

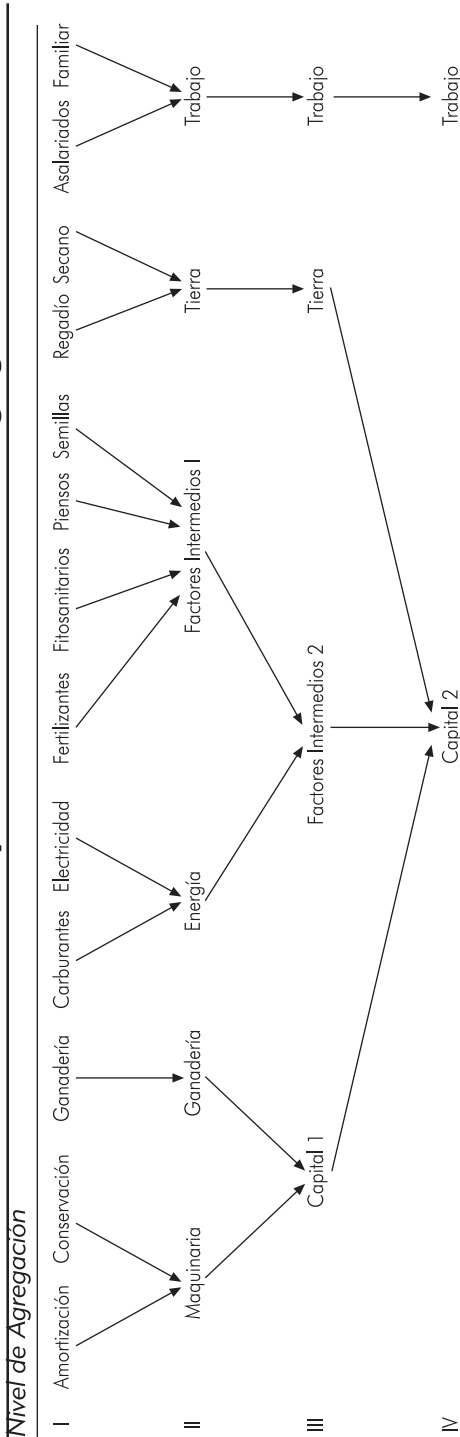
Para plantear las cuatro especificaciones de la función de producción con diferentes niveles de agregación, hemos considerado un total de 13 factores que hemos ido agrupando en diferentes variables a medida que hemos ido ascendiendo en el nivel de agregación. En la tabla 1 pueden verse cuales van a ser las variables que vamos a utilizar según la agregación deseada, y en el Anexo los datos empleados en el trabajo.

Para medir la producción, se considera la Producción Final Agraria (PFA), expresada en pesetas constantes de 1986, a lo largo de todo el período de análisis 1964-1994. Dicha serie se obtiene de los Anuarios de Estadística Agrícola publicados por el Ministerio de Agricultura. Sin embargo, a finales de los años ochenta se produjo un cambio en la metodología de medición, pasando de la *metodología nacional* que se había venido utilizando, a la *metodología europea*. Debido a las complejidades propias del cambio metodológico, se estimó la relación existente entre ambas metodologías mediante un modelo de regresión lineal ($R^2=0,998$), y se transformaron todos los datos a metodología nacional.

En lo que se refiere a los datos de los factores productivos, se recogieron en magnitudes desagregadas, considerándose las siguientes series:

- Maquinaria: está recogida mediante la variable conservación que incluye los gastos en conservación y reparación de maquinaria, y otra variable en concepto de flujo del stock de maquinaria (es decir, amortización de la maquinaria). Para calcular esta segunda variable, estimamos la tasa de amortización dividiendo la amortización del stock de maquinaria que aparece en los Anuarios de Estadística Agraria (AEA) por el stock de maquinaria publicado en las Cuentas del Sector Agrario de los años 1972 y 1976, obteniendo de este modo una aproximación al porcentaje de amortización de la maquinaria, resultando un 14% en ambos años. El input amortización que incluimos en nuestro modelo se obtiene después de sumar a la depreciación de maquinaria obtenida al aplicar el 14% de amortización a los datos de stock de maquinaria publicados en los AEA, el 4% del valor total de la maquinaria. Este porcentaje se considera una aproxima-

Tabla 1. Variables para Diferentes Niveles de Agregación



ción a la tasa de interés a largo plazo¹. Se usan pesetas constantes de 1986 tras deflactar mediante el índice de precios pagados por los agricultores.

- Input Ganadero: este factor productivo se estima como un porcentaje (10%) del stock ganadero. Para calcular el stock ganadero nos valemos de la valoración del stock de capital publicada en 1976. Dividimos el valor de cada cabaña ganadera (distinguiendo ganado bovino, ovino, caprino y porcino) por el número de cabezas, obteniendo de este modo una aproximación al valor de cada individuo. Posteriormente deflactamos dicho precio por el índice de precios pagados por los agricultores correspondiente a cada especie ganadera hasta tener precios de 1986; finalmente, multiplicamos las series históricas de censo de animales por dichos precios para obtener nuestra serie de stock ganadero.
- Factores Intermedios: carburantes, electricidad (al agregar estas dos series obtenemos el gasto en energía), fertilizantes, fitosanitarios, piensos, semillas y plantones constituyen el gasto en factores intermedios. Las series anteriores se expresan en pesetas constantes de 1986 que se obtienen al deflactar los datos en pesetas corrientes mediante el índice de precios pagados por los agricultores.

1. Véase Fernández (1995), págs. 109-110 para una detallada explicación tanto del método de cálculo del input de flujo de maquinaria, como de la justificación del porcentaje aplicado.

- Tierra: el factor tierra lo incluimos desagregándolo en seco y regadío. Se consideran los servicios de la tierra como un porcentaje (7%) del valor total de ésta, en concepto de aproximación a la renta de la tierra. Se utiliza el precio medio de la tierra tanto de seco como de regadío en 1986² para estimar el valor de la tierra. Dada la gran variación existente en los precios de la tierra según el tipo de suelo y de cultivo, es necesario tomar dichos datos con cierta provisionalidad. También sería conveniente destacar que tanto las plantaciones como las infraestructuras, que no se incluyen dentro del factor capital por la dificultad de obtención de dichos datos, vienen, en cierto modo, reflejados dentro del factor tierra.
- Trabajo: debido a la dificultad de obtener datos de trabajo expresados en Unidades de Trabajo Agrícolas para todos los años de la serie, se utiliza el número total de trabajadores empleados en el sector agrario, distinguiendo entre trabajo asalariado y familiar (incluyendo aquí a los trabajadores no remunerados).

La totalidad de los datos correspondientes a los factores productivos se obtuvo a partir de los Anuarios de Estadística Agrícola y de las Cuentas del Sector Agrario.

La especificación elegida ha sido la Cobb-Douglas con una variable de tiempo como indicador de la incidencia del progreso técnico sobre la producción agraria. Una vez linealizada mediante una transformación logarítmica, y añadido un término de error, dicha función, en su especificación más agregada, toma la forma:

$$\ln Q = \ln A + qt + a \ln K + b \ln L + e \quad (1)$$

Aun siendo conscientes de las numerosas limitaciones que suponen el adoptar una especificación Cobb-Douglas, su facilidad de manejo, y la dificultad de estimar otras especificaciones como la translog y la función de elasticidad de sustitución constante cuando hay tantas variables explicativas como en el caso de la agregación primera, ha hecho que nos inclinemos finalmente por ella.

Análisis de la Multicolinealidad

Basándonos en el modelo econométrico lineal general $\mathbf{y} = \mathbf{Xb} + \mathbf{e}$, para detectar la multicolinealidad partimos del sistema de ecuaciones normales $\mathbf{X'Xb} = \mathbf{X'y}$, calculamos los índices de condición de cada uno de los valores propios λ_k de la matriz

2. Ministerio de Agricultura Pesca y Alimentación (1987) Boletín Mensual de Estadística. Junio.

$X'X$ como

$$\mathbf{h}_k = (\mathbf{I}_{max} / \mathbf{I}_k)^{1/2}$$

así como el número de condición de X ,

$$\mathbf{k}(X) = (\mathbf{I}_{max} / \mathbf{I}_{min})^{1/2}$$

que es una medida de la elasticidad de \mathbf{b} ante cambios en $X\mathbf{y}$ o en $X'X$. Cuando las variables que entran en el modelo se normalizan para tener longitud unidad, números de condición mayores que 30 suelen considerarse problemáticos. Belsley (1980) y Estellares y Prieto (1996) son algunos de los autores que recomiendan esta metodología.

Una vez detectada la presencia de la multicolinealidad surge la pregunta de cómo afectará a los parámetros estimados. La varianza de un solo coeficiente b puede expresarse de la forma:

$$\text{var}(b_k) = \mathbf{s}^2 \sum_{j=1}^K \frac{p_{kj}^2}{\mathbf{I}_j} \quad (2)$$

donde p_{kj} es la j -ésima componente del k -ésimo vector propio de $X'X$, y \mathbf{I}_j el j -ésimo valor propio de dicha matriz. Vemos que la proporción de $\text{var}(b_k)$ asociada con una raíz característica \mathbf{I}_j será:

$$\mathbf{f}_{kj} = \frac{p_{kj}^2 / \mathbf{I}_j}{\sum_{j=1}^K p_{kj}^2 / \mathbf{I}_j} \quad (3)$$

donde puede verse que una raíz característica pequeña no implica de por sí que una variable se vaya a ver envuelta en la multicolinealidad. Para una mejor interpretación, tabulamos los valores de (3) en la tabla 2.

Tabla 2. Descomposición de la Varianza

Índice de Condición	$\text{var}(b_1)$	$\text{var}(b_2)$...	$\text{var}(b_K)$
\mathbf{h}_1	\mathbf{f}_{11}	\mathbf{f}_{21}	...	\mathbf{f}_{K1}
\mathbf{h}_2	\mathbf{f}_{12}	\mathbf{f}_{22}	...	\mathbf{f}_{K2}
...
\mathbf{h}_K	\mathbf{f}_{1K}	\mathbf{f}_{2K}	...	\mathbf{f}_{KK}

Los elementos de cada columna de la tabla anterior suman 1. La presencia de dos o más valores de grandes en una misma fila indicarían que la multicolinealidad, y en particular que la dependencia lineal asociada a valores altos del índice de condición, está afectando negativamente a la precisión de la estimación de los coeficientes. Valores de ϕ_{kj} mayores de 0,50 se consideran grandes (Judge (1985)). No interesa usar datos desviados con respecto a la media porque de este modo enmascaramos la existencia de una posible dependencia lineal que afecte al término constante.

Para evaluar los efectos que haya podido tener la multicolinealidad sobre la bondad de la estimación y sobre la capacidad predictiva usaremos los conocidos estadísticos R^2 y U de Theil respectivamente.

Para contrastar la existencia de rendimientos constantes a escala en la función de producción agregada de la agricultura española, efectuamos sobre los modelos una prueba F, que nos indique si la hipótesis de que la suma de los coeficientes de los factores productivos sea igual a uno es aceptable para el modelo estimado.

Tratamiento Informático

Para realizar este estudio nos hemos valido fundamentalmente del programa Eviews 2.0 de Quantitative Micro Software, así como de Microsoft Word 6.0 y Microsoft Excel 5.0 para Windows. Gran parte de los resultados, como el cálculo de los índices de condición, la descomposición de la varianza, o la evaluación de la capacidad predictiva, se han obtenido tras elaborar diversos macros en Eviews.

3. Resultados

Análisis de la presencia de multicolinealidad

Al calcular los índices de condición para la agregación 1 (tabla 3 en página siguiente), vemos que estos son anormalmente altos: el índice máximo alcanza el valor de 9002, y todos excepto dos alcanzan valores mayores que 30. En la primera fila puede apreciarse con cierta claridad una relación de dependencia entre *regadío* y *secano*, con porcentajes de la varianza asociados a ese número del 51% y 86% respectivamente.

La tercera fila muestra de igual modo una dependencia entre las variables *amortización* (0.544), *conservación* (0.614) y *trabajo familiar* (0.680). Sin embargo, por tratarse de una dependencia dominada por la de la fila 1, no podemos tener la seguridad de que esas y sólo esas son las variables que se ven envueltas.

Tabla 3. Descomposición de la Varianza. Agregación 1.

I. Condición	C	Amort	Conser	Ganadería	Carbur	Elect	Fert	Filo	Piensos	Semillas	Regadío	Secano	Asalariados	Familiar	Tiempo
9002	0,122	0,010	0,033	0,000	0,011	0,060	0,006	0,002	0,002	0,017	0,508	0,864	0,042	0,019	0,100
6193	0,674	0,278	0,221	0,453	0,094	0,065	0,006	0,100	0,000	0,000	0,361	0,116	0,005	0,125	0,173
5314	0,134	0,544	0,614	0,118	0,125	0,055	0,133	0,120	0,336	0,102	0,125	0,017	0,029	0,680	0,156
2870	0,041	0,075	0,100	0,405	0,107	0,000	0,041	0,082	0,009	0,016	0,005	0,003	0,038	0,001	0,207
2108	0,004	0,000	0,003	0,018	0,348	0,089	0,005	0,480	0,092	0,161	0,000	0,000	0,466	0,089	0,039
1714	0,000	0,065	0,002	0,002	0,046	0,164	0,071	0,025	0,471	0,126	0,001	0,000	0,018	0,027	0,074
1488	0,003	0,008	0,000	0,002	0,097	0,003	0,347	0,031	0,009	0,436	0,000	0,000	0,187	0,015	0,097
1214	0,006	0,000	0,014	0,001	0,019	0,016	0,322	0,083	0,011	0,055	0,000	0,000	0,125	0,011	0,011
979	0,007	0,000	0,008	0,001	0,072	0,103	0,039	0,005	0,058	0,008	0,000	0,000	0,038	0,030	0,057
770	0,002	0,014	0,000	0,000	0,001	0,208	0,024	0,041	0,000	0,044	0,000	0,000	0,027	0,002	0,064
644	0,000	0,004	0,002	0,000	0,064	0,197	0,000	0,025	0,006	0,002	0,000	0,000	0,000	0,000	0,006
411	0,002	0,000	0,001	0,000	0,013	0,037	0,000	0,005	0,000	0,024	0,000	0,000	0,023	0,000	0,001
341	0,005	0,002	0,002	0,000	0,004	0,003	0,005	0,001	0,006	0,008	0,000	0,000	0,002	0,000	0,014
16	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001
1	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000

Fuente: Elaboración propia

Valores superiores a 0,5 en negrita

Para formalizar la relación exacta existente entre las variables implicadas, llevamos a cabo regresiones auxiliares entre esas variables. Al efectuar una regresión entre *regadío*, una constante y *secano*, obtuvimos una relación negativa, lo que tiene su lógica ya que la superficie destinada a regadío ha venido creciendo en el período de estudio a expensas de una disminución en la superficie de secano, aunque con un R^2 de tan solo 0.4, lo que nos hace pensar que tal relación no está del todo clara. Al elaborar una segunda regresión auxiliar del factor conservación en función de una constante, *amortización* y *trabajo familiar*, ambas variables explicativas aparecieron como significativas al 5%, con un valor de R^2 del 0.988. El coeficiente del factor *amortización* fue de 1 (un aumento en la amortización de maquinaria vino acompañado por un aumento en la conservación y reparación de ésta en la misma proporción), mientras que el de *trabajo familiar* fue de -0.26, indicando un posible proceso de sustitución de factores productivos entre el trabajo y la maquinaria. Los resultados referentes a la agregación 2 están expuestos en la tabla 4.

Aquí aparecen tres claras relaciones, cada una correspondiente a un nivel de dominancia. La primera, implicaría al término independiente (0.899) y al factor *tierra* (0.996), la segunda a *ganadería* (0.839), *trabajo* (0.702) y *tiempo* (0.844), y la última a maquinaria (0.684) y factores intermedios 1 (0.561). También aquí todos los índices de condición menos dos son mayores que 30, siendo el valor máximo de 6763.

Tabla 4. Descomposición de la Varianza. Agregación 2

I.Condición	C	Maquinaria	Ganadería	Energía	F.Intermedios1	Tierra	Trabajo	Tiempo
→ 6763	0.899	0.060	0.076	0.046	0.202	0.996	0.003	0.048
→ 3133	0.049	0.012	0.839	0.273	0.179	0.002	0.702	0.844
→ 1962	0.012	0.684	0.064	0.057	0.561	0.002	0.085	0.000
1059	0.030	0.119	0.014	0.071	0.017	0.000	0.197	0.042
883	0.002	0.061	0.006	0.509	0.024	0.000	0.013	0.026
474	0.008	0.064	0.001	0.044	0.017	0.000	0.000	0.038
15	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Fuente: Elaboración propia

Al llevar a cabo las correspondientes regresiones auxiliares, puede destacarse la alta correlación positiva existente entre *maquinaria* y *factores intermedios*. El proceso de capitalización de la agricultura en España no ha estado fundamentado sólo en la maquinaria, sino también en el aumento del uso de factores intermedios. Igualmente, la relación de dependencia entre la tierra y la constante del modelo se deben a que el factor tierra como agregación del secano y el regadío no ha variado significativamente en el período 1964-88.

También es clara la presencia de multicolinealidad en la agregación 3 (tabla 5). El índice de condición sigue siendo del mismo orden de magnitud, y solamente dos filas aparecen con un índice inferior a 30. Además, se van haciendo más evidentes si cabe las relaciones entre las variables que estaban involucradas en la agregación anterior. Únicamente puede apreciarse una nueva relación, entre las variables trabajo y tiempo.

Tabla 5. Descomposición de la Varianza. Agregación 3

<i>I.Condicion</i>	<i>C</i>	<i>Capital1</i>	<i>Factint2</i>	<i>Tierra</i>	<i>Trabajo</i>	<i>Tiempo</i>
→ 6317	0.942	0.059	0.289	0.997	0.026	0.001
→ 2178	0.021	0.818	0.577	0.003	0.334	0.087
→ 1181	0.007	0.093	0.002	0.000	0.619	0.832
705	0.029	0.031	0.133	0.000	0.021	0.070
15	0.000	0.000	0.000	0.000	0.000	0.009
1	0.000	0.000	0.000	0.000	0.000	0.000

Fuente: Elaboración propia

Al llevar a cabo la oportuna regresión auxiliar obtuvimos un valor negativo del coeficiente del trabajo, lo que puede explicarse fácilmente en base a la disminución del número de trabajadores empleados en el sector agrario en las últimas décadas.

Finalmente, y al realizar el análisis de descomposición de la varianza correspondiente al último nivel de agregación (tabla 6), obtuvimos de igual modo altos valores de los números de condición que indican que la presencia de la multicolinealidad es severa. Al realizar la correspondiente regresión auxiliar entre *capital2* y *tiempo*, se observó el resultado esperado: una correlación positiva entre ambos factores.

Tabla 6. Descomposición de la Varianza. Agregación 4

<i>I.Condicion</i>	<i>C</i>	<i>Capital2</i>	<i>Trabajo</i>	<i>Tiempo</i>
→ 2418	0.679	0.996	0.114	0.585
→ 1098	0.321	0.004	0.886	0.407
15	0.000	0.000	0.000	0.008
1	0.000	0.000	0.000	0.000

Fuente: Elaboración propia

También aparece otra relación entre trabajo y tiempo (aunque el valor de este último no llegue a alcanzar 0.5). La regresión auxiliar indicó una dependencia ligeramente negativa entre ambas, lo que puede explicarse por la disminución de trabajadores agrícolas en los años analizados. El hecho de que apareciera una misma variable

(tiempo) involucrada en dos relaciones, nos hizo pensar que quizás existiera también una relación entre *capital2* y *trabajo*. Una vez más, se realizó una regresión auxiliar entre esas variables, obteniéndose un coeficiente negativo para *capital2*, con un alto valor del coeficiente de determinación (0.937). Este resultado indicaría la condición de factores substitutivos existentes entre *capital2* y *trabajo*.

Estimación de los modelos de producción

Al llevar a cabo la estimación de la función de producción agraria para las cuatro agregaciones (tabla 7), obtuvimos altos valores de R^2 , aunque ninguna de las variables, con la excepción de *capital 2* en MCO4, puede considerarse como estadísticamente significativa.

Es reseñable también el alto valor de las desviaciones típicas de los coeficientes. Estos resultados son los que pueden esperarse cuando la multicolinealidad se encuentra presente de una forma tan importante.

Como dijimos previamente, hemos prestado atención a la capacidad predictiva del modelo. Al calcular la capacidad de predicción, vemos la *U* de Theil, descomponiéndola en sesgo, varianza y error, que se obtiene al predecir los valores de la variable dependiente (PFA) mediante los valores muestrales de las variables dependientes. Los resultados obtenidos están resumidos en la Tabla 8; en general se trata de unos valores de *U* de Theil bastante aceptables, ya que la proporción debida al error recoge la mayor parte del error de predicción.

Tabla 8. Evaluación de la Capacidad Predictiva Muestral

<i>Modelo</i>	<i>U Theil</i>	<i>Sesgo</i>	<i>Varianza</i>	<i>Error</i>
MCO1	0.35791	0.10922	0.00018	0.89060
MCO2	0.45380	0.15765	0.07491	0.76744
MCO3	0.45781	0.14784	0.15308	0.69909
MCO4	0.48769	0.16514	0.12859	0.70627

Fuente: Elaboración propia

Como cabría esperarse, los modelos estimados que obtuvieron un mayor valor del coeficiente de determinación (R^2), muestran la mejor capacidad predictiva muestral, con una proporción debida al sesgo y a la varianza muy pequeñas.

En cuanto a la interpretación económica de los modelos estimados, hay que comentar que debido a la baja estabilidad de los estimadores, hay que analizar con cierta cautela, tanto sus valores como sus signos.

Más interesante que el estudio individual de los coeficientes estimados, es el contrastar si se cumple la hipótesis de rendimientos constantes a escala. Con fines compa-

Tabla 7. Resultados de la Estimación

Agregación 1	C	Amort	Conser	Ganad	Carbu	Elec	Fert	Filo	Piensos	Semillas	Regadio	Secano	Asal	Familiar	Tiempo	R2
	5,207 (9,693)	-0,256 (0,268)	0,327 (0,302)	-0,285 (0,334)	0,027 (0,130)	-0,155 (0,094)	0,005 (0,121)	0,205 (0,163)	-0,021 (0,139)	0,017 (0,127)	0,431 (0,565)	0,146 (0,645)	0,092 (0,218)	0,336 (0,427)	0,039 (0,027)	0,99
Agregación 2	C	Maquinaria	Ganad	Energía	Facint1	Tierra	Trabajo	Tiempo	R2							
	9,308 (8,051)	-0,013 (0,112)	-0,285 (0,270)	-0,096 (0,098)	0,262 (0,153)	0,407 (0,537)	0,044 (0,263)	0,03 (0,020)	0,984							
Agregación 3	C	Capital1	Facint2	Tierra	Trabajo	Tiempo	R2									
	2,765 (7,031)	0,051 (0,163)	0,235 (0,133)	0,638 (0,517)	-0,094 (0,194)	0,009 (0,010)	0,982									
Agregación 4	C	Capital2	Capital2	Trabajo	Tiempo	R2										
	2,134 (2,501)	0,874** (0,180)	-0,02 (0,160)	0,006 (0,010)	0,984											

Fuente: Elaboración propia

Las desviaciones típicas de los coeficientes estimados se incluyen entre paréntesis.

** Significativo para una prueba al 1%

rativos, hemos calculado (tabla 9) la suma de los coeficientes estimados correspondientes a los factores de producción (excluimos por lo tanto los coeficientes del término independiente y del tiempo), para todas las agregaciones, así como el estadístico F y la probabilidad de que la variable F de Senedecor tome un valor superior al mismo, obtenidos ambos al contrastar la hipótesis nula de existencia de rendimientos constantes a escala.

Tabla 9. Contraste de Rendimientos Constantes a Escala

<i>Modelo</i>	<i>Suma</i>	<i>F</i>	<i>Probabilidad</i>
MCO1	0.869	0.044	0.836
MCO2	0.319	0.212	0.650
MCO3	0.83	0.648	0.428
MCO4	0.854	0.152	0.700

Fuente: Elaboración propia

Todas las estimaciones realizadas dan lugar a rendimientos a escala decrecientes, aunque se acepta la hipótesis de existencia de rendimientos constantes a escala para el sector agrario español, sin olvidar las oportunas reservas que hay que tener en la adopción de conclusiones, debido al problema que nos ocupa.

4. Conclusiones

Podemos concluir diciendo que parece evidente que la estimación de forma directa por MCO de una función de producción agregada para la agricultura española se ve adversamente afectada por la presencia de la multicolinealidad. Al calcular los índices de condición de los datos y realizar un análisis de la descomposición de la varianza obtuvimos resultados concluyentes. Como consecuencia directa de lo anterior, la estimación por mínimos cuadrados dio lugar a coeficientes muy inestables.

El análisis de las variables nos permitió poner de manifiesto tendencias que han tenido lugar en la agricultura española en las tres últimas décadas: la creciente mecanización del sector agrario, dentro de un proceso de sustitución de factores productivos de trabajo por maquinaria, y la mayor importancia que está tomando el sector ganadero, así como el papel neutral que ha jugado la tierra en los procesos anteriormente descritos.

El uso de variables más agregadas no consiguió generar modelos más precisos al estimar por MCO. Aunque el número de condición va disminuyendo cuando vamos agregando variables, éste sigue alcanzando valores muy superiores a los mínimos

requeridos para poder descartar la peligrosidad de la multicolinealidad. Si el investigador está interesado en efectuar un análisis estructural y por lo tanto en estudiar los coeficientes de variables desagregadas en una estimación de la función de producción por MCO, no parece aconsejable agregar variables, perdiendo entonces información, sólo con la finalidad de reducir los efectos de la multicolinealidad.

Los cuatro modelos presentan propiedades excelentes desde el punto de vista predictivo, por lo que son útiles para establecer entre que valores estará la producción final agraria a nivel nacional en el futuro, en base a la medición o extrapolación de los factores de producción.

Bibliografía

- BELSLEY, D. A. et al (1980): *Regression Diagnostics. Identifying Influential Data and Sources of Collinearity*. John Wiley & Sons.
- CAÑAS, J.A., FRESNO, R. y DIOS, R. (1994): "Funciones de Producción Lineales de Variedades de Maíz en Andalucía" en *Investigación Agraria, Economía*. vol 9(2), pp. 215-229.
- ESTALLERES, R. y PRIETO, J. (1996): "Sensibilidad del Análisis de la 'Estructura Propia' en la Detección de Colinealidad" en *Qüestió*, vol 20, pp 549-562.
- FERNÁNDEZ, M^a C. (1993): "La Contribución de la Investigación Agraria al Avance de la Productividad en el Sector Agrario Español durante el Período 1962-1989". Tesis doctoral. Universidad de Córdoba.
- GRILICHES, Z y RINGSTAD, V. (1971): *Economies of Scale and the Form of the Production Function*. North-Holland Publishing Company.
- JUDGE, G.G et al (1985): *The Theory and Practice of Econometrics*. John Wiley & Sons.
- MAPA (1973-1995): *Anuario de Estadística Agraria*.
- MAPA(1974-1986): *Las Cuentas del Sector Agrario*. No. 1-12.
- MAPA (1995): *Boletín Mensual de Estadística*. Junio.
- SEBASTIÁN, R.A. y RODRÍGUEZ BARRIO, J.E. (1978): *Funciones de Producción en Agricultura*, ETSIA, Universidad Politécnica de Madrid.
- SOLOW, R (1957): "El cambio técnico y la función de producción agregada", en Rosemberg, N (ed.): *Economía del Cambio Tecnológico*. Fondo de Cultura Económica.
- WALTERS, A.A. (1977): *Introducción a la Econometría*. Oikos-tau.
- WOODFIELD, A. (1972): "Estimates of Hicks-neutral Technical Progress, Returns to Scale and the Elasticity of Substitution in New Zealand Manufacturing 1926-68". *New Zealand Economic Papers*, vol VI, pp. 73-92.
- YAMADA, S y RUTAN, V.W. (1990): "Comparaciones internacionales de la productividad agraria" en San Juan, Carlos (ed.): *La Modernización de la agricultura española (1956-1986)*, pp. 73-159. Ministerio de Agricultura, Pesca y Alimentación.

ANEXO: Datos empleados en el trabajo

Año	Amortiz.	Conserv.	Ganad.	Carbur.	Eléctric.	Fertil.	Fitosanitar.	Plensos	Semillas	Regadío	Secano	Asalariados	Familiar	PFA
1964	39.256	25.391	70.008	22.748	3.386	66.911	15.351	1591.65	18.669	229.451	545.403	1.163.448	2865.452	1522515
1965	43.538	27.165	69.033	25.217	3.765	70.169	14.243	1616.17	21.150	235.009	537.319	1.071.819	2.789.881	1.453.546
1966	48.466	31.209	69.631	28.103	3.895	69.064	17.646	222.984	21.772	239.447	531.149	1.039.973	2.763.027	1.592.332
1967	57.636	36.728	72.766	31.403	5.102	73.824	18.528	243.629	24.773	246.022	524.341	1.020.065	2.740.445	1.687.524
1968	62.884	39.710	73.928	35.264	5.849	84.802	18.670	243.027	25.869	251.752	520.021	1.020.655	2.721.745	1.684.526
1969	72.575	47.453	71.225	38.843	6.038	94.254	21.124	271.040	25.303	255.274	517.416	1.029.367	2.617.533	1.754.647
1970	76.908	50.179	78.075	39.411	7.888	97.809	21.548	309.659	29.389	251.363	538.786	1.017.562	2.536.738	1.773.543
1971	82.855	53.873	76.126	41.705	9.525	100.908	23.475	306.582	32.424	277.094	551.870	1.020.879	2.423.721	1.925.380
1972	86.739	61.810	76.443	44.766	9.724	107.379	26.145	336.807	34.566	285.671	548.956	957.107	2.157.493	1.973.772
1973	88.105	65.481	80.691	45.194	10.148	110.452	28.175	419.789	35.854	290.131	542.315	917.807	2.106.993	2.103.665
1974	82.539	66.306	78.649	52.478	9.309	132.428	23.494	418.780	32.264	295.552	538.181	910.002	2.091.198	2.233.930
1975	90.612	78.227	77.073	65.030	10.781	134.081	23.768	382.820	36.472	299.257	535.719	806.835	1.890.765	2.244.367
1976	99.627	89.568	77.518	70.805	14.170	125.296	24.357	431.293	38.373	303.145	529.579	849.038	1.798.942	2.367.510
1977	116.917	100.892	79.441	71.692	13.020	125.608	26.404	498.442	35.355	307.754	526.767	785.167	1.684.632	2.362.287
1978	137.745	113.638	80.855	69.971	13.400	120.651	29.989	510.105	29.830	313.289	524.582	686.154	1.648.645	2.509.712
1979	145.893	115.690	78.226	67.807	13.628	143.571	32.352	509.387	32.654	319.476	521.509	655.775	1.572.884	2.500.980
1980	141.619	122.057	79.715	83.848	16.019	155.070	34.434	490.043	33.268	322.758	519.842	568.771	1.447.229	2.696.499
1981	137.910	117.195	79.604	102.352	23.033	148.920	32.745	503.043	29.178	326.166	518.618	559.690	1.431.209	2.503.673
1982	146.053	124.779	87.950	102.887	23.779	148.778	33.073	557.912	31.084	333.233	517.874	560.815	1.387.384	2.594.509
1983	145.944	125.931	89.613	105.068	21.984	129.043	38.414	583.185	32.377	334.274	517.130	522.008	1.392.292	2.702.018
1984	149.988	125.086	89.476	114.299	23.483	148.123	41.845	599.109	33.938	343.514	514.872	512.895	1.360.205	2.895.385
1985	163.167	125.606	89.268	117.811	29.150	154.057	47.078	591.254	36.784	343.811	511.963	548.176	1.281.523	2.959.609
1986	151.535	141.162	93.517	108.353	34.087	155.459	46.099	579.082	37.034	349.106	510.731	515.520	1.122.980	2.749.200
1987	163.778	152.255	104.356	107.700	33.882	161.383	50.379	584.935	42.085	355.179	508.284	507.298	1.108.601	3.005.198
1988	174.089	153.321	104.168	93.800	41.212	155.583	59.962	615.623	43.565	358.724	506.729	509.663	1.081.637	3.146.850
1989	172.733	161.569	105.529	92.563	43.928	147.919	57.612	611.133	46.967	362.429	504.488	458.664	1.037.635	2.975.648
1990	175.349	177.268	105.544	99.692	47.328	146.039	60.004	603.783	50.416	365.837	499.153	444.592	946.507	3.071.556
1991	154.684	186.281	106.007	107.049	50.803	146.070	58.989	599.990	54.893	365.197	496.857	433.640	818.259	3.069.915
1992	146.151	180.731	106.046	110.170	52.285	132.478	55.812	623.014	51.110	366.844	492.260	376.182	790.917	3.078.331
1993	162.121	173.794	105.626	114.580	47.973	113.031	52.844	641.065	43.123	370.457	482.797	345.838	761.961	2.965.991
1994	166.401	177.241	107.376	101.087	58.059	135.034	59.009	646.444	42.975	357.409	450.789	341.939	730.261	2.975.011

Todas las series están expresadas en millones de pesetas constantes de 1986, excepto Asalariados y Familiar que lo están en miles de trabajadores