

23- ALGUNOS FACTORES QUE EXPLICAN LAS BRECHAS DE RENDIMIENTO DE ARROZ A NIVEL COMERCIAL

M.N. Sanguinetti¹, B. Ferraro², B. Lanfranco³

PALABRAS CLAVE: Variedad, Zona de producción, Rendimiento, Zafra

INTRODUCCIÓN

Bajo el liderazgo de la Asociación de Cultivadores de Arroz (ACA) y junto al Instituto Nacional de Investigación Agropecuaria (INIA) y los molinos integrantes de la Gremial de Molinos Arroceros (GMA), este año dio inicio la ejecución del proyecto “Fortalecimiento de las estrategias de transferencia para reducir las brechas de rendimientos en el sector arrocerero”, con financiamiento del programa Bienes Públicos Sectoriales de la Agencia Nacional de Desarrollo (ANDE).

El componente 1 del proyecto plantea la cuantificación detallada de las brechas de rendimiento entre productores. En particular, se propone:

1. analizar los potenciales de rendimiento y las brechas existentes entre productores de acuerdo con la diferente escala, variedades de arroz y tecnologías disponibles, pudiendo cuantificar el impacto económico y las principales causales de estas brechas.
2. definir el conjunto de tecnologías a aplicar en los campos de los productores participantes y los criterios de ajustes regionales y por distintos públicos objetivo.
3. crear una estructura sectorial interinstitucional de transferencia, para reducir las brechas de rendimiento e incrementar la competitividad de los productores

En este artículo se presentan los resultados de un primer análisis llevado a cabo para

identificar algunos factores (variedad utilizada y localización geográfica del cultivo) que explican los rendimientos de arroz a nivel de chacra. Constituye un primer paso para la definición de potenciales de rendimiento comercial alcanzables, en el marco de un proceso de intensificación sostenible.

MATERIALES Y MÉTODOS

Los datos utilizados para el análisis provienen de la Comisión Sectorial del Arroz (CSA). Los datos recabados anualmente por la CSA incluyen información del productor (nombre o razón social), variedad sembrada, área cosechada, producción, molino receptor y departamento. Los datos de las cuatro zafras consideradas (2014-2015, 2015-2016, 2016-2017 y 2017-2018) se agruparon en una sola base de datos, agregándose una variable identificadora de la zafra.

Como variable dependiente se tomó el logaritmo natural del rendimiento (en kg/ha), calculado a partir de los datos de área y producción registrados por la CSA. Debido a la inconsistencia de muchos registros, con importantes errores de digitación, se optó por trabajar con observaciones cuyo dato de rendimiento cayera dentro de un rango de rendimientos razonables definido en forma arbitraria. Aunque posibles, rendimientos por debajo de los siete mil kilos podrían considerarse como deprimidos para las condiciones tecnológicas promedio en el país. De la misma forma, rendimientos por encima de los 14.000 serían muy excepcionales. Buscando un balance razonable que permitiera eliminar observaciones con valores atípicos (*outliers*) sin incurrir en

¹ Ec., Gerente General, Asociación de Cultivadores de Arroz. msanguinetti@aca.com.uy

² Cr., Economía Aplicada, Instituto Nacional de Investigación Agropecuaria. bferraro@inia.org.uy

³ Ing. Agr., Ph.D., Economía Aplicada, Instituto Nacional de Investigación Agropecuaria. blanfranco@inia.org.uy

una pérdida de información demasiado alta, se descartaron las observaciones con rendimientos por debajo de 6.000 o superiores a 15.000 kg/ha.

Como regresoras se incluyeron tres variables cualitativas o categóricas: *zafra*, *variedad* y *zona*. Al utilizarse un modelo de análisis transversal que integra datos de cinco años diferentes es conveniente incluir una variable para controlar el efecto *zafra*. Inicialmente, el análisis incluía también la *zafra* 2013-2014. Debido a que los registros contenían muchos errores, lo que llevaba al descarte de un gran número de observaciones, se optó por trabajar con las cuatro correspondientes a 2014-2015, 2015-2016, 2016-2017 y 2017-2018.

El análisis se focalizó exclusivamente en variedades de grano largo. De las 33 variedades plantadas en al menos una *zafra*, se seleccionaron 16, en base al área sembrada. De este número, 8 son variedades comunes (El Paso 144, INIA Olimar, INIA Tacuarí, CH Quebracho, Parao, INIA Merin, INIA Caraguatá y Urumati) y 8 son CL o Clearfield® (Inov CL, Gurí INTA CL, CL 212, CL 244, Titan CL, PL1-PL5, XP118 CL y XP102 CL). Para reducir el número de variables dicótomas, solo se consideraron como categorías individuales El Paso 144, INIA Olimar, INIA Tacuarí, Inov CL y Gurí INTA CL. Las otras 5 variedades comunes se agruparon en la categoría OCV (otras variedades comunes) mientras que las restantes 6 CL se agruparon con el nombre de OCL (otras CL). Con

esto la cantidad de variables dicótomas necesarias para representar la variedad se redujo a 7.

Finalmente, la localización original por departamento de cada observación se redujo a las tres zonas de producción arroceras del país: *Este* (Rocha, Treinta y Tres, Lavalleja y la parte este-noreste de Cerro Largo perteneciente a dicha cuenca), *Centro* (oeste y noroeste de Cerro Largo, Tacuarembó y Rivera) y *Norte* (Artigas, Salto y norte de Paysandú). Tras la eliminación de observaciones atípicas y la confección de las variables categóricas, el número definitivo de observaciones quedó definido como $N = 3.126$. En el cuadro 1 se informa acerca de algunas características descriptivas de los datos analizados.

El análisis empírico de los datos se realizó utilizando un modelo semi-logarítmico en el cual, la variable dependiente, cuantitativa y continua, se incluye mediante su transformación logarítmica, en tanto que las variables independientes son cualitativas y multicategóricas. Cuando una variable cualitativa admite solo dos categorías se denomina binaria o dicótoma. Expresa la presencia (1) o ausencia (0) del factor de interés y se incluye directamente en el modelo. Cuando el número de categorías es superior a dos, se necesita igual número de variables dicótomas. Las variables dicótomas (cualitativas) pueden utilizarse como variables independientes en los modelos de regresión, en las mismas condiciones en que se usan las cuantitativas.

Cuadro 1. Características de la base de datos analizada.

Zafra	Obs.	Total		Zona					
				Este		Centro		Norte	
		Área	Rend.	Área	Rend.	Área	Rend.	Área	Rend.
		ha	kg/ha	ha	kg/ha	ha	kg/ha	ha	kg/ha
2014-2015	763	143.432	8.606	89.976	8.754	20.616	8.162	32.840	8.618
2015-2016	706	140.964	8.420	96.245	8.702	17.609	7.951	27.110	7.809
2016-2017	855	154.705	8.621	100.104	8.665	21.252	8.489	33.349	8.592
2017-2018	802	136.725	8.337	88.121	8.385	18.800	8.128	29.804	8.396
Total	3.126	575.826	8.499	374.446	8.624	78.277	8.194	123.103	8.381

Un modelo cuyas variables explicativas son exclusivamente dicótomas, constituye una forma general de los llamados modelos de análisis de varianza (ANOVA).

Los modelos ANOVA se utilizan para medir la significancia estadística de la relación entre una variable dependiente (regresada) cuantitativa y un número de variables independientes (regresoras) cualitativas representadas, cada una de ellas, por dos o más variables binarias o dicótomas. Con la inclusión de un coeficiente de intersección, se debe descartar una variable dicótoma de cada uno de los factores incluidos en el modelo (*zafra*, *variedad* y *zona*). Esto evita un problema conocido como colinealidad perfecta o multicolinealidad. En ese caso, las variables regresoras ya no son independientes y el modelo no se puede estimar. Resuelto este problema, el modelo se estima mediante mínimos cuadrados ordinarios (MCO).

RESULTADOS

Los resultados de la estimación se observan en el Cuadro 2. La primera columna lista las variables independientes incluidas en forma explícita. La variable de base representa la

situación del cultivo de arroz en la zona este, sembrado con la variedad El Paso 144 en la zafra 2014-2015, siendo justamente las categorías representadas por las variables dicótomas omitidas en modelo empírico. De la segunda a la quinta columna, 2 a la 5 inclusive, se presentan los resultados de la estimación (coeficiente estimado, error estándar, valor del estadístico t, valor de probabilidad y significancia estadística).

Aparte del intercepto, las únicas diferencias significativas se observaron para las zafras 2015-2016 y 2017-2018 (rendimientos menores en comparación con 2014-2015), el uso de INIA Olimar, Inov CL o de las restantes variedades comunes (no CL), con rendimiento promedio menor a El Paso 144 para la primera y superior en los dos casos restantes, así como los rendimientos de la zona centro, significativamente por debajo de la zona este. Dicho de otro modo, el efecto año de la zafra 2016-2017 no mostró diferencias significativas frente a la 2014-2015. Ni INIA Tacuarí, ni Gurí INTA CL, ni las restantes variedades CL mostraron superioridad a El Paso 144, mientras que los rendimientos de la zona norte no mostraron diferencias estadísticas frente a la zona este.

Cuadro 2. Resultados del modelo de regresión ANOVA.

Variables Independientes	Coefficiente estimado	Error estándar	Estadístico t	Valor-p y significancia	Variación porcentual	Mediana rendimiento
Intercepto ⁽¹⁾	9,0521	0,0067	1358,12	0,000**	---	8.536
Zafra 2015-2016	-0,0285	0,0078	-3,6584	0,000**	-2,81%	8.297
Zafra 2016-2017	-0,0031	0,0071	-0,4352	0,663	-0,31%	8.510
Zafra 2017-2018	-0,0408	0,0077	-5,2756	0,000**	-4,00%	8.195
INIA Olimar	-0,0225	0,0095	-2,3728	0,018*	-2,22%	8.346
INIA Tacuarí	-0,0016	0,0085	-0,1872	0,852	-0,16%	8.523
Inov CL	0,0783	0,0091	8,5787	0,000**	8,14%	9.231
Gurí INTA CL	0,0215	0,0126	1,7153	0,086	2,18%	8.722
Otras var. comunes	0,0438	0,0099	4,4150	0,000**	4,48%	8.918
Otras var. CL	-0,0032	0,0113	-0,2830	0,777	-0,32%	8.509
Zona centro	-0,0381	0,0081	-4,6877	0,000**	-3,73%	8.217
Zona norte	-0,0019	0,0094	-0,2046	0,838	-0,19%	8.520

Nota: El número de asteriscos refiere al nivel de significancia estadística: (**) 1% y (*) 5%

(1) La variable intercepto corresponde al valor de base (El Paso 144, zona este, zafra 2014-2015) frente al cual se comparan el resto de las variables.

Los valores de la penúltima columna expresan la variación porcentual de un cambio en los factores de base, por ejemplo la variedad, en el rendimiento. Así por ejemplo, sustituyendo la variedad El Paso 144 por Inov CL, todo lo demás constante (*ceteris paribus*), explicó un aumento en rendimiento de 8,14%. Por el contrario, manteniendo la variedad de base, al igual que todo lo demás, pero utilizando la zafra 2017-2018, se advierte una importante caída (-4%). La última columna representa la mediana de los rendimientos. El valor obtenido para la línea de base fue de 8.536 kg/ha. Si en vez de la zona este se considera un cultivo en la zona centro, todo lo demás constante, la mediana de los rendimientos se ubica en 8.217 kg/ha.

CONCLUSIONES

Los resultados presentados en este artículo constituyen un primer análisis en la determinación de los factores que afectan los rendimientos comerciales de arroz, utilizando

datos históricos de la CSA referidos a cuatro zafras (2014-2015 a 2017-2018). Se trata de un primer paso para establecer metas de rendimientos potenciales del cultivo de arroz. La reducción de las brechas entre los rendimientos actuales y los potencialmente alcanzables a nivel comercial es un objetivo clave de este proyecto. La evidencia obtenida sugiere que, aún cuando los insumos se apliquen en forma satisfactoria al cultivo para que el mismo pueda expresar al máximo su potencial, existen algunos factores como la variedad sembrada y la zona geográfica que afectan los rendimientos. El efecto año, que engloba las condiciones "agro-eco-climatológicas" prevalentes durante el desarrollo del cultivo, también es un factor relevante. Aunque los datos disponibles no incluían un factor importante como la época de siembra, la evidencia surgida a través de los trabajos de experimentación, ameritan la búsqueda de una forma de considerarla en forma explícita dentro de este modelo.