
¿La soja está desplazando a otras actividades? Acercamiento a los modos de expansión de la frontera agrícola (Argentina, 1988-2000)

Mesa 30: Reflexiones sobre la producción material y simbólica de fronteras que construyen territorios en contextos neoliberales.

Germán Rosati

CONICET / IDAES-UNSAM / PIMSA

Resumen

Este trabajo forma parte de un proyecto más general que se propone analizar e identificar a nivel de la formación social argentina las principales relaciones entre los procesos de expansión de la frontera agrícola, los cambios en el uso del suelo y algunos procesos demográficos existentes en zonas rurales/agropecuarias en las últimas tres décadas. Se propone, específicamente, realizar una primera aproximación a las diferentes modalidades de expansión de la frontera agrícola. Particularmente, se abordan las siguientes preguntas: ¿la expansión de la frontera se da hacia zonas “nuevas” o existe un proceso de sustitución de producciones previas? ¿Bajo qué modalidades se produce dicha expansión? Se hará especial énfasis en el proceso de expansión sojera.

Se utilizaron fuentes secundarias (datos censales agropecuarios del año 1988 y 2002) a nivel departamental. A partir de las distribuciones marginales de superficie sembrada para cada

departamento se estimaron modelos de inferencia ecológica para abordar el problema de la sustitución. A su vez, se intentó construir una tipología de modalidades de expansión para cada uno de los departamentos del país. Finalmente, se analizaron las formas de expansión de la frontera agraria en las diferentes estructuras agrarias de Argentina.

Fue posible detectar zonas en las que este efecto es más alto: Norte de Buenos Aires y Sur de Santa Fé; Centro-Oeste de Chaco, Centro-Este de Santiago del Estero; Oeste de Tucumán y Este de Salta. Así, fue posible ver cómo en la zona núcleo pampeana y en otras zonas como el área algodonera de Chaco, aparecen caracterizadas por cambios importantes en el uso del suelo.

Por último, se presentan algunos resultados preliminares de un ejercicio basado en modelos de regresión regularizados que intentan aportar algunas evidencias empíricas sobre la relación entre el proceso de expansión sojero y ciertos movimientos en la estructura agraria (particularmente, la desaparición de explotaciones).

Palabras claves: frontera agraria - sojización - estructura agraria - inferencia ecológica - uso del suelo - sustitución de actividades

1. Introducción

Una de las transformaciones recientes en el sector agropecuario es la incorporación a la producción agrícola de zonas consideradas previamente como marginales o no productivas. Este proceso de expansión de la frontera agrícola parece iniciarse hacia mediados de la década del 70 y no incluye solamente a provincias “extrapampeanas”, sino que la frontera se

corre también hacia zonas con una menor intensidad de explotación como el interior de la misma región pampeana (Borrelo 1988).

Desde el punto de vista empírico, el problema de la expansión reciente de la frontera agrícola en Argentina y Latinoamérica ha sido estudiado en varias dimensiones. Particularmente relevante es la pregunta acerca de los impactos o efectos que la expansión de la frontera ha provocado. El presente trabajo se centrará en la expansión sojera en Argentina (probablemente, el principal motor de la expansión de frontera) entre 1988 y 2002. Este período comprende lo que se ha denominado la primera fase de expansión y su consolidación (Martínez Dougnac 2013). En efecto, la expansión sojera comienza hacia la década del '80, luego de un período de experimentación y hacia el año 2002 la cantidad de soja sembrada en Argentina se encontraba en pleno ascenso.

La discusión alrededor de estos impactos parece centrarse en algunos ejes principales. El primero de ellos se vincula a los impactos ambientales y ecológicos de dicha expansión. Existen numerosos trabajos que han tratado de aproximarse a este problema, particularmente, en relación a la expansión sojera. Al interior de las ciencias naturales, los estudios abordan los impactos que dicha expansión produce sobre el ambiente, como la deforestación y el desmonte en determinadas zonas (como el norte argentino, el Gran Chaco Latinoamericano, etc.), la pérdida de nutrientes del suelo, etc. (Viglizzo y Jobbágy 2010, Mena 2007, Kaprovicas 2009, LART/FAUBA 2004, Fehlenberg et al 2017, Volante et al 2015).

Un segundo eje se centra en los cambios en las estructuras agrarias. En el caso argentino, el concepto de “pampeanización” fue acuñado para intentar conceptualizar dichos cambios. Esta noción intentaría marcar la generalización de un tipo de estructura agraria y productiva (asociada a la gran escala, a la utilización de determinados paquetes productivos, etc.) tanto

en zonas de expansión de la frontera, como en zonas productivas antiguas, que comienzan a asemejarse a la estructura vigente en la región pampeana (Zarrili 2008). Así, la adopción de procesos tecnológicos intensivos en el uso de capital, los procesos de concentración de la superficie productiva bajo diversas modalidades serían constitutivas de dichas dinámicas.

Un tercer eje de análisis vinculados a estos impactos se vincula con cambios en el uso del suelo. Algunos de los interrogantes que pueden plantearse acerca de las modalidades predominantes de la expansión de la superficie con soja en el campo argentino son los siguientes: ¿cuál es la relación entre este proceso de expansión sojera y el resto de las actividades agropecuarias? ¿Se produce un desplazamiento de otras producciones y actividades? ¿O se trata más bien de una expansión cuyo único efecto será “adicionarse” al resto de las producciones¹?

Desde las ciencias sociales no existen demasiados estudios que se aproximen a los efectos de la expansión agrícola sobre los usos del suelo (y particularmente en relación a la expansión sojera) con una mirada general. La conexión de estos procesos con un nivel macro de análisis ha sido más esporádica, por ejemplo, la relación con la dinámica de la(s) estructura(s) sociales y agrarias. Existen dos grandes estrategias metodológicas en estos estudios: algunos intentan una aproximación de tipo general partir de la consideración de estadísticas agregadas a nivel nacional o regional (Ackhar et al 2011, Azcuy Ameghino y León 2005, Azcuy Ameghino y Ortega 2010). Otros trabajos se basan en estudios de caso, es decir, del estudio de estos

¹ Otro proceso que ha sido asociado a la expansión de la frontera agrícola (particularmente, en torno al incremento de la soja) se relaciona con los cambios producidos en la ganadería (sobre todo, vacuna). Tres de los principales cambios estarían dados por a) una reducción del stock ganadero total del país en los últimos 20 años; b) una redistribución geográfica de dicho stock (especialmente, hacia las provincias del Noreste); c) una intensificación en la explotación ganadera (a través del uso de feedlots). En este trabajo no será abordada esta cuestión.

procesos en determinados departamentos y/o provincias (León, Prudkin y Reboratti 1985, Rodríguez 2008, Steimbregger, Radnonich y Bendini 2003, Valenzuela 2014). La primera estrategia tiene la ventaja de brindar un panorama general del fenómeno; tiene, a su vez, la desventaja de perder en dicha agregación la especificidad de la información analizada. La segunda estrategia intenta aportar mediante un análisis más focalizado, ganando en una descripción detallada del fenómeno. La desventaja es que puede llevar a perder de vista procesos estructurales de mayor escala.

Este trabajo forma parte de un proyecto más general que se propone analizar e identificar a nivel de la formación social argentina las principales relaciones entre los procesos de expansión de la frontera agrícola, los cambios en el uso del suelo y algunos procesos demográficos existentes en zonas rurales/agropecuarias en las últimas tres décadas. Se propone realizar una primera aproximación a las diferentes modalidades de expansión de la frontera agrícola. Particularmente, se abordan las siguientes preguntas: ¿la expansión de la frontera se da hacia zonas “nuevas” o existe un proceso de sustitución de producciones previas? ¿Bajo qué modalidades se produce dicha expansión? Se hará especial énfasis en el proceso de expansión sojera.

En trabajos previos (Rosati 2013) se había estimado que entre 1987/88 y 2009/10, aproximadamente un 87% de la expansión de la superficie sojera total se había hecho en zonas en las que el incremento de la superficie sojera iba acompañado del descenso absoluto o relativo de otras producciones. Sin embargo, tal aproximación anterior tenía algunas deficiencias:

- la sustitución se había estimado mediante el cálculo de medidas agregadas (tasas de variación): la relación entre la tasa de variación de la soja y la del resto de los cultivos

-
- no era posible estimar cuántas hectáreas de otros cultivos eran reemplazadas por la soja.

En este trabajo se utilizarán fuentes secundarias (datos censales agropecuarios del año 1988 y 2002) y se intentará abordar estas limitaciones. A partir de las distribuciones marginales de superficie sembrada a nivel departamental se estimarán modelos de inferencia ecológica para abordar el problema de la sustitución. A su vez, se intentará construir una tipología de modalidades de expansión para cada uno de los departamentos del país.

Ahora bien, los datos de los Censos Nacionales Agropecuarios muestran una generalizada tendencia a la desaparición de explotaciones agropecuarias en toda la formación social argentina: entre 1988 y 2002 desaparecieron unas 87.000 explotaciones, es decir, un descenso del 20%. Si bien este proceso tiene variaciones regionales, parece relativamente claro que las explotaciones pequeñas y medianas (hasta 100 hectáreas) fueron las que mayores tasas de desaparición presentan.

Muchos estudios marcan la concomitancia entre el proceso de expansión sojera y la desaparición de las explotaciones (Martínez Dougnac, G. 2013, Rodríguez 2008). Sin embargo, los mecanismos causales y las explicaciones de este fenómeno no resultan tan claras. Algunos enfoques atribuyen a la expansión sojera un rol fundamental: “El avance de la soja en el interior del país desplaza con violencia a campesinos de sus tierras amparados en las leyes veinteañales, y a comunidades indígenas que tienen derechos ancestrales sobre esos terrenos.” (Giarracca y Teubal 2013: 30). Otros autores presentan una visión diferente. Azcuy Ameghino y León (2005), afirman que el proceso de desaparición de las explotaciones no se debe a la expansión sojera, sino a que “el proceso actual de agriculturización está concentrado en ese cultivo”. Es decir que su visión es más abarcadora y parecen considerarlo más un efecto del proceso de agriculturización.

Por ello, este trabajo buscará aportar evidencias sobre la influencia que el proceso de expansión sojera tiene sobre uno de los indicadores de la llamada “pampeanización”: la desaparición de explotaciones agropecuarias en el período. La pregunta podría formularse de la siguiente forma: ¿cuál es la importancia que los procesos de sojización tienen en la disminución del número de explotaciones? ¿Qué modalidades asume en las diferentes estructuras agrarias de Argentina en este período?

2. Estimando el peso de la “sustitución de actividades”

El objetivo de este primer ejercicio es realizar una primera aproximación a la estimación del llamado "efecto sustitución", es decir, la cantidad de hectáreas que han sido plantadas en una campaña t0 con otros cultivos y que en una campaña posterior t1 fueron reemplazadas por soja. Si conociéramos qué cultivo habían sido plantadas en cada momento del tiempo podríamos construir una tabla 2x2 como la siguiente:

Tabla 1. Relación entre los parámetros a estimar

	Soja 2001 2002	Resto 2001 2002	Total
Soja 1987 1988	β_s^i	$1 - \beta_s^i$	$S_{87,88}^s$
Resto 1987 1988	β_r^i	$1 - \beta_r^i$	$1 - S_{87,88}^i$
Total	$S_{01,02}^s$	$1 - S_{01,02}^i$	1

La variable independiente X será la superficie sembrada total en t0 (por ejemplo, la campaña 1987/88). La variable dependiente T será la superficie sembrada total en t1 (campaña 2001/02).

$i_{s}^{t_0}$ representa la proporción de hectáreas que fueron plantadas con soja en ambos momentos del tiempo, respecto al total de hectáreas en 1987/88. En cambio, $i_{n}^{t_0}$, cuantifica la proporción de superficie que habían sido plantadas con algún otro cultivo en t0 pero fueron plantadas con soja en t1. Es decir que $i_{n}^{t_0}$ es el parámetro de mayor interés en este trabajo. Sería posible construir una tabla de 2x2 para cada uno de los departamentos, es decir, la unidad de agregación más pequeña disponible en la fuente analizada, en los que existe información. El subíndice i expresa este nivel de agregación².

El problema será estimar cuántas hectáreas fueron ocupadas por otros cultivos en t0 y fueron transformadas en soja en t1 para cada departamento. Podemos plantearlo como un problema de inferencia ecológica. Diversos métodos se han desarrollado para proveer alguna respuesta al problema. Este trabajo no se propone realizar una historia de los métodos de inferencia ecológica (Kousser 2001). Sin embargo, es necesario realizar algunos comentarios al respecto. Si se observa la tabla 1 nuevamente, es posible definir una relación clara entre las marginales:

² La mejor fuente de análisis para este tipo de análisis son las imágenes satelitales. Las mismas proveen un menor nivel de desagregación para estimar dichos valores. No obstante, dado que este trabajo intenta una primera aproximación más general a este problema, se propone un ejercicio que logre identificar áreas a nivel agregado -departamental- en las que la soja esté probablemente desplazando otras actividades. Con estas áreas localizadas, podría hacerse un uso más efectivo de las imágenes satelitales. Incluso, podría intentarse utilizar las técnicas de inferencia ecológica presentadas aquí para realizar inferencias de sustitución a nivel de pixel.

$$S_{01,02}^i = S_{87,88}^i \times \beta_s^i + (1 - S_{87,88}^i) \times \beta_r^i$$

$$S_{01,02}^i = \beta_r^i + (\beta_s^i - \beta_r^i) \times S_{87,88}^i$$

Es decir, en el i -ésimo departamento la proporción de hectáreas totales de soja cosechadas en 2001-02 ($S_{01,02}^i$) es igual a la proporción total cosechada con soja en 1987-88 ($S_{87,88}^i$) multiplicada la proporción de superficie sembrada con soja en los dos períodos (β_s^i) sumado a la proporción de superficie total que no había sido sembrada con soja en 1987-88 ($1 - S_{87,88}^i$) multiplicada la proporción de hectáreas que habían sido sembrada con soja en 1987-88 pero no en 2001-02 (β_r^i).

Es importante notar que esto es una identidad contable de la tabla (King 1997) y no un supuesto de linealidad de la relación entre ambas variables o algún modelo estadístico. Al reordenar los términos de la ecuación anterior podemos definir (ahora sí) una relación lineal entre los dos parámetros de interés (β_s^i y β_r^i).

$$\frac{S_{01,02}^i}{1 - S_{87,88}^i} - \frac{S_{87,88}^i}{1 - S_{87,88}^i} \beta_s^i = \beta_r^i$$

La ecuación anterior habilita a estimar los rangos posibles de valores que podrían asumir ambos parámetros: los valores posibles deben caer en la línea definida por dicha ecuación. De esta forma, para cada departamento podemos definir el rango de valores posibles, simplemente introduciendo los valores marginales de la tabla. Luego, podemos generar un gráfico -llamado en inglés "tomography plot"- en el que cada rango de posibles valores

aparece representado como una línea. De esta forma podemos resumir toda la información presente en los datos.

Estas ideas están en el centro de lo que se ha llamado "método de los límites" (Duncan and Davis 1953). La idea es poder extraer toda la información determinística de los datos para poder inferir los parámetros de ella. Goodman (1959) propone atacar el problema de la inferencia ecológica desde un marco estadístico: propone una regresión lineal para estimar las cantidades de la ecuación 2, agregando un término de error aleatorio,

$$S_{01,02}^i = \beta_r^i + (\beta_s^i - \beta_r^i) \times S_{87,88}^i$$

El primer término sería el intercepto y el segundo, la pendiente. El método de estimación es simplemente Mínimos Cuadrados Ordinarios. El enfoque de Goodman ha sido criticado por diversos motivos: 1) asume homogeneidad espacial: los coeficientes del modelo no varían a lo largo de los departamentos; 2) los distritos no están ponderados según su tamaño: en caso de existir diferencias importantes en su tamaño podrían introducirse sesgos y 3) las estimaciones del modelo pueden producir valores por fuera del intervalo [0,1] propio de una proporción.

Uno de los más recientes avances en el problema es el de Gary King (1997). Desarrolló diversos métodos que tratan de resolver varias de las limitaciones previas. En el presente trabajo se usará el método de King apropiado para tablas de 2x2. La propuesta de King (1997) utiliza tanto la información proveniente del método de límites como de la regresión de Goodman para fortalecer las estimaciones de los parámetros a nivel de precintos. Luego, a partir de ambas estimaciones a nivel de precintos estima los parámetros agregados.

Usando el "tomography plot" se determina el rango de valores posibles de los parámetros en los precintos de interés. Luego, se modela la distribución de probabilidad conjunta de ambos parámetros. Así, en lugar de asumir (como en el modelo de Goodman) que los parámetros son "constantes" a lo largo de los distritos, el modelo de King asume que existe cierto nivel de correlación. Al mismo tiempo, el modelo es forzado a producir valores dentro del rango [0,1]. Ambos requisitos se logran asumiendo que (β_r^i, β_s^i) son extraídos de una distribución bivariada normal truncada (TBN).

La segunda etapa es estimar el vector de los cinco parámetros de la TBN: medias, desvíos estándar y correlaciones. Estos parámetros son estimados mediante el uso de regresiones lineales. Una vez que los parámetros de la TBN son estimados, el modelo produce estimaciones de la distribución de probabilidad y la plotea sobre el tomography plot.

La etapa final del modelo es la producción de estimaciones de las cantidades de interés usando la información determinística sobre los límites de los parámetros en base a variantes de la ecuación 3. Esto se logra mediante procesos de simulaciones iterativas dado que la derivación analítica del modelo es compleja.

Los resultados de la aplicación del modelo son los siguientes: una estimación de las cantidades agregadas de interés, n estimaciones (donde n es la cantidad de distritos en el dataset) de los parámetros a nivel distrito y otros productos como los límites de los parámetros para cada distrito, desvíos estándar de cada parámetro estimado, etc.

Sin embargo, existen algunos supuestos que tiene el enfoque de King que es necesario explicitar. En primer lugar, asume la inexistencia de sesgos de agregación, es decir: la forma en que la data ha sido agregada (generalmente, en forma de agrupamientos administrativos como provincias, departamentos, etc.) puede causar la existencia de correlación entre los parámetros estimados y las variables independientes. En este caso, la estimación de los parámetros podría estar sesgada.

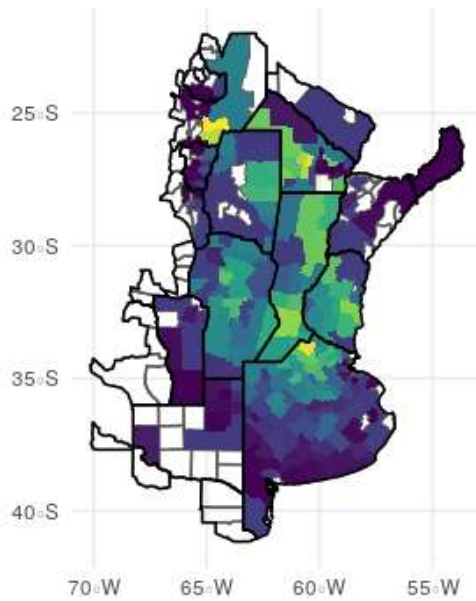
En segundo lugar, el modelo asume que no existe autocorrelación espacial. De esta forma, la distribución de los parámetros y las cantidades agregadas se asumen como aleatorias.

Finalmente, existe un tercer problema potencial: si las hipótesis acerca de la distribución de los parámetros (es decir, que los datos no provengan de una distribución normal bivariado truncada). King afirma que el modelo es bastante robusto a este problema.

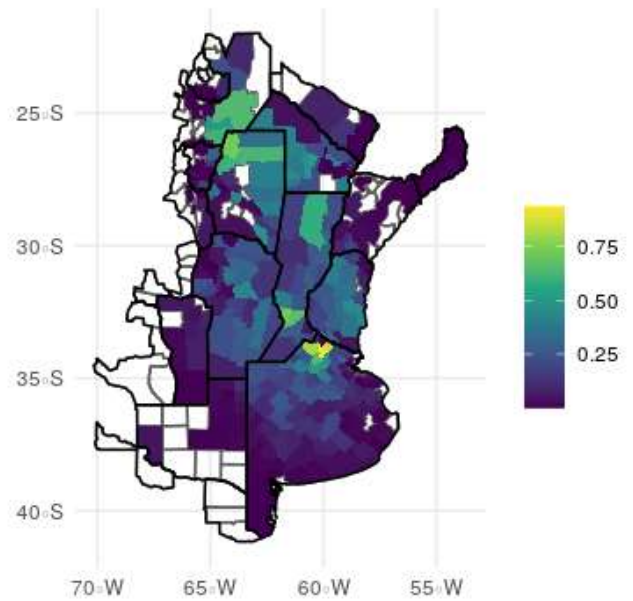
Se han estimado otros dos métodos con diferentes mecanismos de estimación y supuestos al de King que intentan abordar algunas de las limitaciones mencionadas. El método de Wakefield (Haneuse y Wakefield 2004) busca incorporar la dependencia espacial, intentando así morigerar los efectos de los sesgos de agregación. A su vez, el método de Quinn (2004) intenta incorporar el efecto temporal en este tipo de modelos. De esta forma, se intentará evaluar si diferentes especificaciones producen resultados concordantes.

Mapa 1. Estimación del efecto sustitución soja (β_{11}) según diferentes modelos. Argentina 1988-2002.

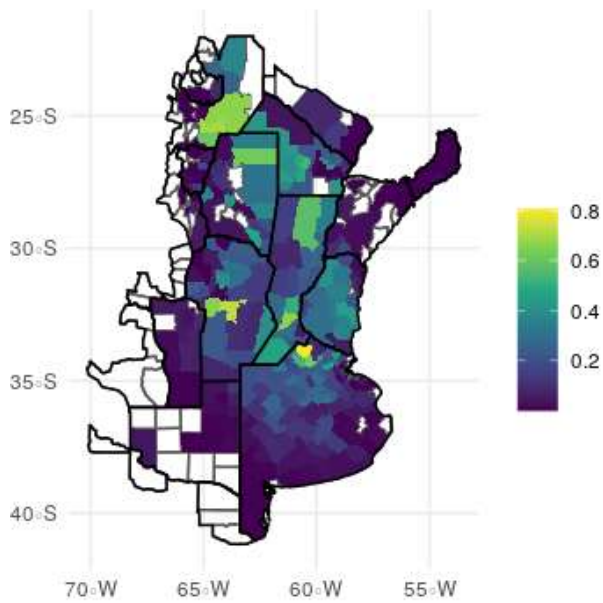
Método de King



Método de Wake



Método de Quinn



Fuente: elaboración propia sobre datos de CNA 1988 y 2002

Más allá de las diferencias en las estimaciones del efecto de sustitución en cada uno de los métodos analizados, puede verse que existen ciertos *hot spots* que son coincidentes. Es decir, aquellos lugares con mayores valores en el coeficiente β_{11} tienden a solaparse en los tres mapas: 1) Norte de Buenos Aires y Sur de Santa Fe; 2) Centro-Oeste de Chaco, Centro-Este de Santiago del Estero; 3) Oeste de Tucumán y Este de Salta. Esto puede notarse en las correlaciones de Spearman (no paramétricas) entre las tres estimaciones (por encima de 0.84 en todos los casos).

3. Identificando tipos de expansión entre 1988 y 2002

Dados los resultados relativamente consistentes de los tres métodos, se avanzará en este trabajo en un análisis en profundidad del método de Gary King, dado que parece mostrar una mayor variabilidad en sus resultados. Ahora bien, puede notarse la existencia de evidencias de un sesgo de agregación. Aquellos departamentos en los que la proporción sojera es mayor en 1988 parecen presentar intervalos para las estimaciones del efecto de conservación más grandes.

No obstante, y más allá de los problemas para la estimación de los valores y sus intervalos de confianza, los resultados obtenidos parecen lograr detectar zonas que se han caracterizado por procesos de cambios en el uso del suelo relativamente intensos en el período.

El siguiente paso será la identificación de las diferentes modalidades en que se produce la expansión de la frontera en los departamentos analizados. Para ello, se calculó, a partir de los parámetros estimados mediante el modelo de inferencia ecológica, la cantidad (y la

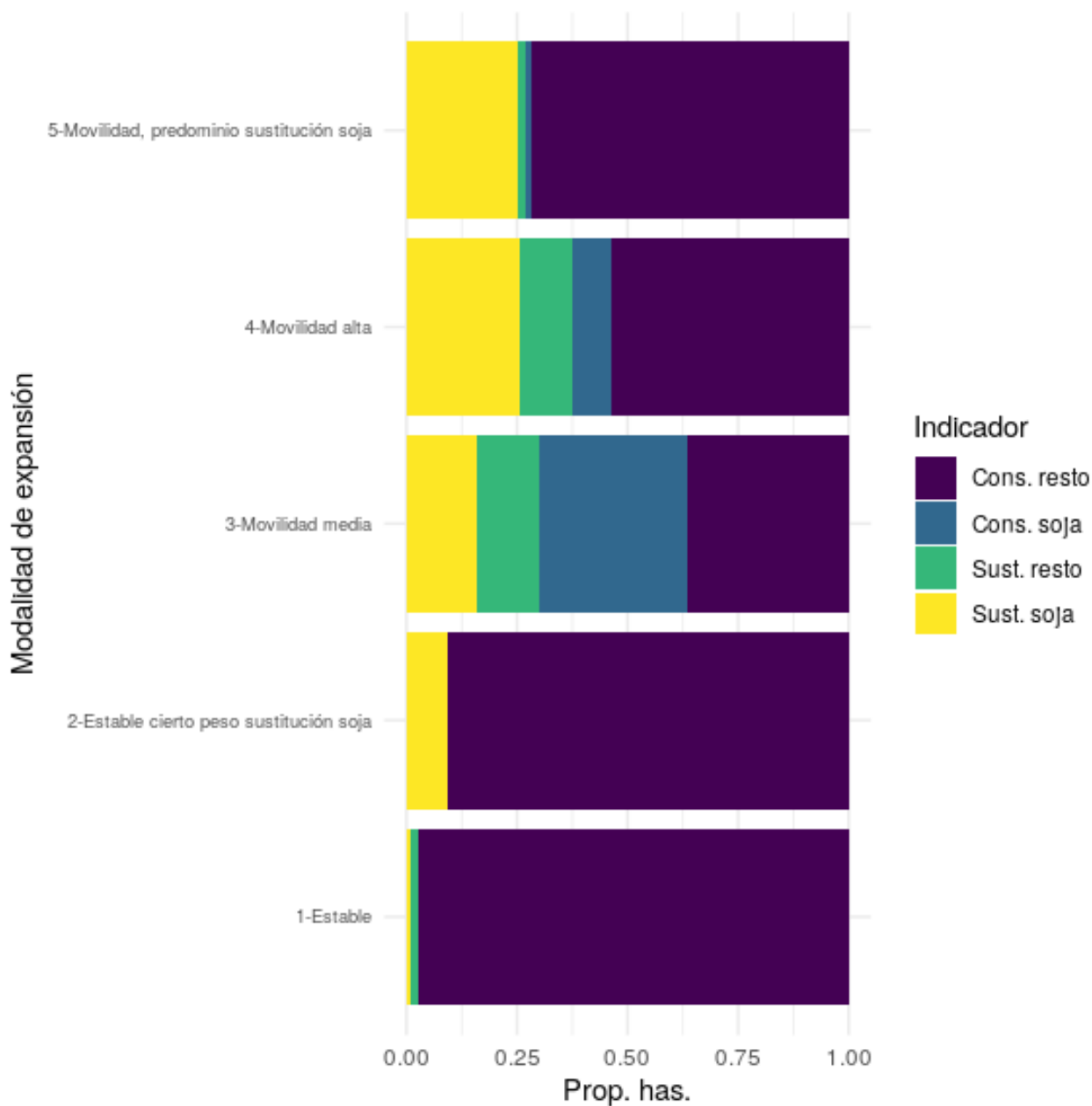
proporción) de hectáreas que en cada departamento se siembra bajo cada una de las cuatro modalidades.

Dado que lo que se busca es poder identificar patrones comunes de expansión, se decidió realizar un análisis de clusters con dichas variables como inputs. El análisis de clusters busca construir grupos de objetos (en este caso departamentos) intentando que cada grupo tenga la máxima homogeneidad al interior y la máxima diferencia entre los grupos en función de las variables consideradas. (James, Witten, Hastie y Tibshirani 2017). La métrica utilizada puede ser una distancia o una similitud. En la primera, los grupos estarán formados por objetos con distancias pequeñas entre ellos. En el segundo caso, los grupos formados contendrán objetos similares entre sí (por ejemplo, el coeficiente de correlación).

En este trabajo se utilizó el método de clustering jerárquico. La matriz de distancias fue calculada utilizando distancias de tipo manhattan y el método de aglomeración utilizado fue el de Ward. Ahora bien, el problema habitual a la hora de trabajar con análisis de conglomerados es la selección de la cantidad de grupos o clusters a formar. Existen diversos criterios para realizar la selección de este hiperparámetro, en este trabajo se utilizó el llamado *gap statistic* (Tibshirani, Walther y Hastie 2001).

¿Qué características tienen los clusters conformados? Para poder caracterizar e interpretar cada uno de los clusters (que se corresponde con un tipo de modalidad de expansión) se calcula la mediana de cada una de las variables originales (no de los componentes utilizados para la estimación de los clusters) para cada uno de los clusters:

Gráfico 1. Medianas de cada una de las variables originales según cluster identificado.



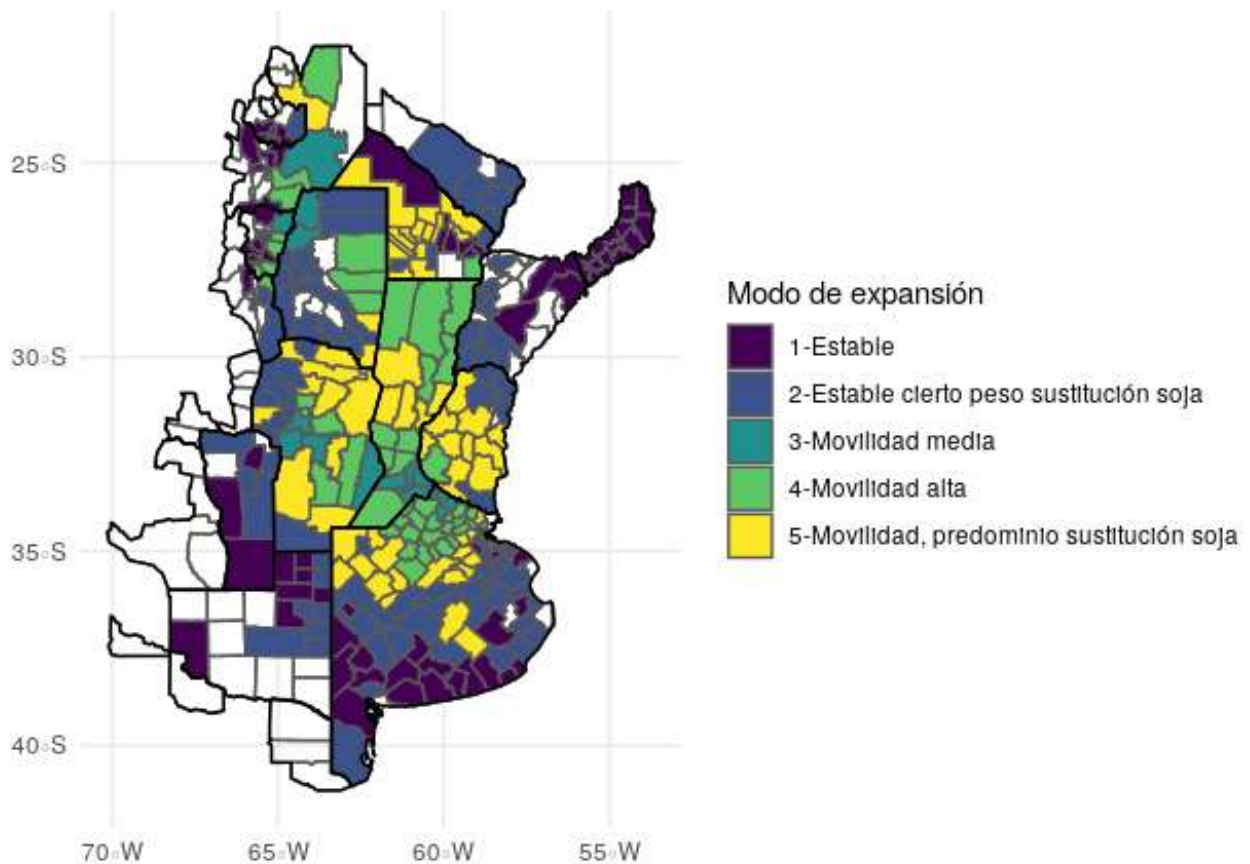
Fuente: elaboración propia sobre datos de CNA 1988 y 2002

Puede verse, entonces, que los clusters 3 a 5 se caracterizan por valores elevados de hectáreas que son sustituidas (ya sea, por soja o por otros). Particularmente, presentan proporciones medias superiores al 12% de hectáreas que han sido reemplazadas con soja.

Particularmente, el cluster 3, a su vez, tiene como rasgo distintivo una significativa proporción media de hectáreas que se han conservado con soja y una proporción también elevada de hectáreas de soja reemplazadas por otros cultivos, por ello lo hemos caracterizado como de "movilidad media". El cluster 4, no muestra ese peso de las hectáreas conservadas por soja y muestra altas proporciones de sustitución por soja y por otros cultivos, por ello lo caracterizamos como de "movilidad alta". El cluster 5, en cambio, muestra como rasgo principal el peso que las hectáreas reemplazadas por soja adquieren. Estos dos clusters parecen caracterizarse por una gran movilidad de la frontera.

Finalmente, los clusters 1 y 2 aparecen caracterizados sobre todo por la preponderancia que adquiere la proporción media de hectáreas que se mantienen sembradas con cultivos que no son soja. Estos dos clusters parecen estar caracterizados por cierta inmovilidad en la frontera, al menos en lo que se refiere a la expansión sojera.

Mapa 2. Modos de expansión de la frontera agraria, Argentina, 1988-2002.



Fuente: elaboración propia sobre datos de CNA 1988 y 2002

De esta forma, puede notarse en el mapa anterior que la modalidad "estable" parece predominar en los departamentos más externos del área analizada: la provincia de Misiones, el sur de Buenos Aires, norte de Chaco, norte de La Pampa y algunas zonas de Tucumán. Incluso, el cluster que hemos denominado "estable con cierto peso de sustitución soja" presenta este rasgo.

En cambio, las zonas de movilidad en todas sus formas parecen tener mayor peso en las zonas núcleo. Así, el sur y norte de Santa Fe, el norte de Buenos Aires aparece como zonas caracterizadas por una movilidad media y alta. A su vez, la zona algodonera chaqueña aparece caracterizada por patrones de movilidad y cambio en el uso del suelo.

4. Los modos de expansión según estructuras agrarias

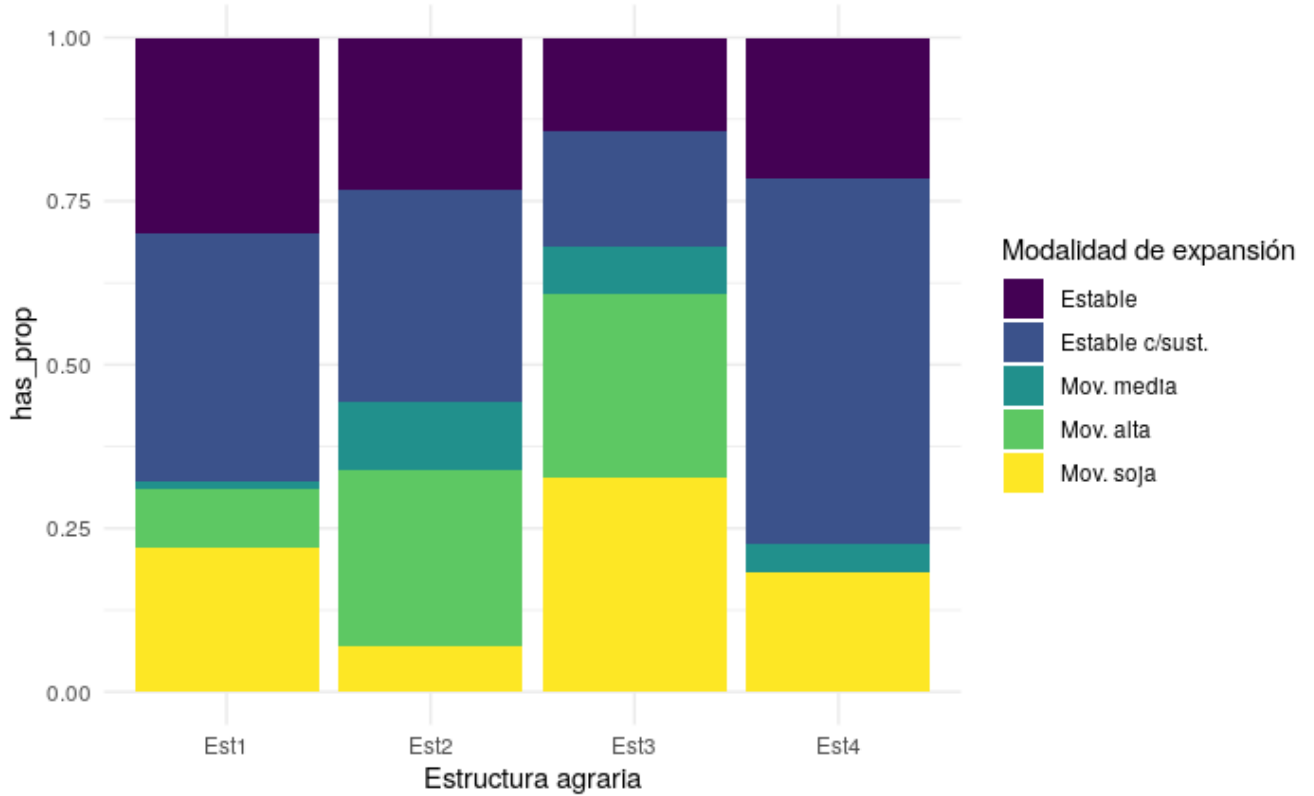
Ahora bien, ¿en qué estructuras agrarias se producen estos patrones de cambios en el uso del suelo? En un trabajo anterior (Rosati y Chazarreta 2020) se habían identificado cuatro estructuras agrarias que componen el campo argentino: 1. alto grado de desarrollo de relaciones salariales; 2. alto grado de desarrollo de relaciones salariales y existencia de superpoblación latente; 3. alto desarrollo de relaciones salariales articuladas con fracciones de pequeños propietarios simples y acomodados y 4. pequeña producción mercantil simple con elementos de descomposición. Cada una se caracteriza por diferentes proporciones de tipos de explotaciones y grupos sociales³.

Para intentar ponderar la relación entre las estructuras agrarias concretas y las modalidades de expansión de la frontera agraria se calculará la proporción de hectáreas cultivadas bajo las

³ Las estructuras agrarias fueron construidas a partir de la información de dos fuentes censales (Censo Nacional de Población y Vivienda 2001 y Censo Nacional Agropecuario 2002). Se utilizó información desagregada a nivel departamental, lo cual permite desagregar las distintas provincias, dado que, como es esperable, coexisten en su interior diferentes formas de organización de la producción agropecuaria. Se construyeron 4 clusters a partir de información de las explotaciones y de la población ocupada en la rama 1 de cada departamento. De esta forma, se logró una diferenciación de zonas o áreas no basada en criterios de cercanía geográfica, como la categorización en regiones frecuentemente utilizada (noroeste, noreste, patagonia, pampeana y cuyo), sino en la homogeneidad respecto a las estructuras sociales agrarias. Para mayor detalle puede consultarse Rosati y Chazarreta (2020).

diferentes modalidades de expansión (clusters) en cada una de las estructuras agrarias construidas.

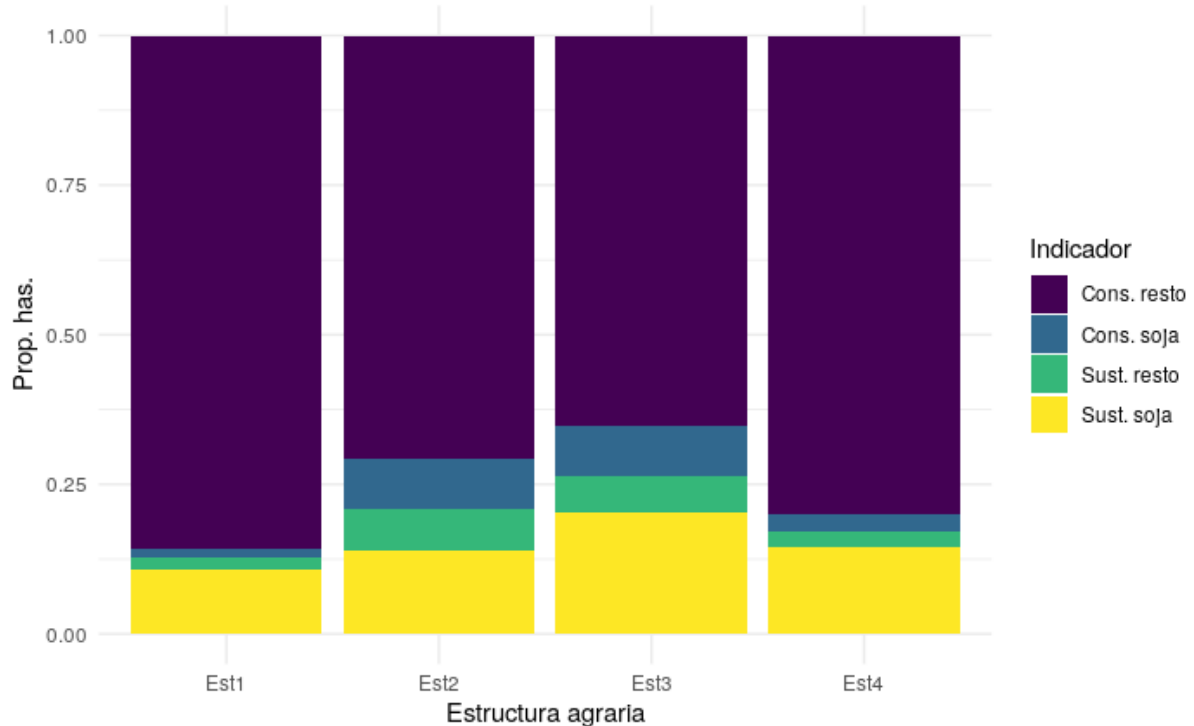
Gráfico 2. Proporción de hectáreas en cada modalidad de expansión según tipo de estructura agraria, Argentina, 1988-2002



Fuente: elaboración propia sobre datos de CNA 1988 y 2002

Al mismo tiempo, para cada estructura agraria fueron agregadas las hectáreas estimadas en el modelo de inferencia ecológica, es decir, las celdas interiores de la tabla.

Gráfico 3. Proporción de hectáreas estimadas por modelo de inferencia ecológica según tipo de estructura agraria, Argentina, 1988-2002



Fuente: elaboración propia sobre datos de CNA 1988 y 2002

Utilizando cualquiera de los dos indicadores, puede notarse que en aquellos departamentos comprendidos en las estructuras 1 y 4 (es decir, aquellas de alto desarrollo capitalista y de pequeña producción mercantil en descomposición) se nota un predominio claro de superficie en departamentos bajo la forma de estabilidad o de conservación de los cultivos ya existentes. Es decir, que los mayores movimientos de la frontera se concentran en las estructuras agrarias caracterizadas por un alto grado de desarrollo capitalista articuladas con formas de superpoblación latente o con fracciones de pequeños propietarios.

5. Un intento de cuantificación de algunos efectos en la expansión de la frontera

Ahora bien, ¿cómo se vincula la expansión sojera con ciertos procesos más generales en la estructura agraria argentina? Particularmente, ¿cuál es su relación con uno de los indicadores más analizados: la desaparición de explotaciones entre 1988 y 2002? Este apartado intentará aportar algunos elementos empíricos a estas preguntas. Particularmente, intentará evaluar qué modalidades tal desaparición tiene en las diferentes estructuras agrarias construidas y su relación con los modos de expansión sojera. Se presentan algunos resultados preliminares de un trabajo en curso.

Para ello se construyeron una serie de modelos de regresión lineal a nivel departamental cuya variable dependiente será la tasa de variación de las explotaciones entre 1988 y 2002 y como variables dependientes cuantitativas indicadores referidos al proceso de sojización (la proporción de hectáreas reemplazadas por soja y la proporción de hectáreas que se mantienen de otros cultivos) y como variable dependiente cualitativa la estructura agraria a la que pertenece el departamento. También se incluyeron todas sus interacciones.

Ahora bien, existe un número relativamente elevado de variables: las 2 variables cuantitativas, las cuatro *dummies* correspondientes a la estructura agraria y todas las interacciones en sus diferentes órdenes. Dado que esto constituye un modelo complejo de interpretar, se utilizará un proceso de regularización basado en el método LASSO. El mismo permite hacer una selección automática de aquellos coeficientes que resultan significativos para el ajuste del modelo, incluyendo las interacciones entre los diversos predictores. De esta forma, solamente quedan incluidos en el modelo las variables (y/o sus interacciones de diferente orden) que tiene una importancia significativa.

La precisión predictiva de un modelo de regresión puede incrementarse a través del encogimiento de los valores de los coeficientes o, incluso, haciéndolos cero. Haciendo esto, se introduce algún sesgo, pero se reduce la varianza de los valores predichos y, por lo tanto, se incrementa la precisión predictiva total.

En muchos casos, cuando existen muchos predictores puede ser necesario identificar un subconjunto más pequeño de estos predictores que muestren los efectos más grandes. Es por ello que resulta útil imponer restricciones en el proceso de estimación. A esta operación se la define como “regularización”. Existen varios métodos de regularización de modelos lineales: non negative garrotte (Breiman, 1995); ridge regression (Hoerl y Kennard, 1970). Este trabajo se centrará en el LASSO. Este método utiliza la norma l1 como medida de penalización para definir las restricciones al modelo lineal. LASSO busca minimizar la siguiente expresión (en lugar de la suma de los cuadrados de los residuos -RSS-):

$$RSS + \lambda \sum_1^P |\beta_p|$$

Particularmente, en este trabajo se utiliza una versión del LASSO para incluir interacciones (Lim y Hastie 2014).

En el primer modelo (M1) se ajustó un modelo sobre el total de los departamentos considerados (n=482) y se introdujeron como variables explicativas: el tipo de estructura agraria, la proporción de hectáreas que son sustituidas por soja y la proporción de hectáreas que se conservan de otros cultivos.

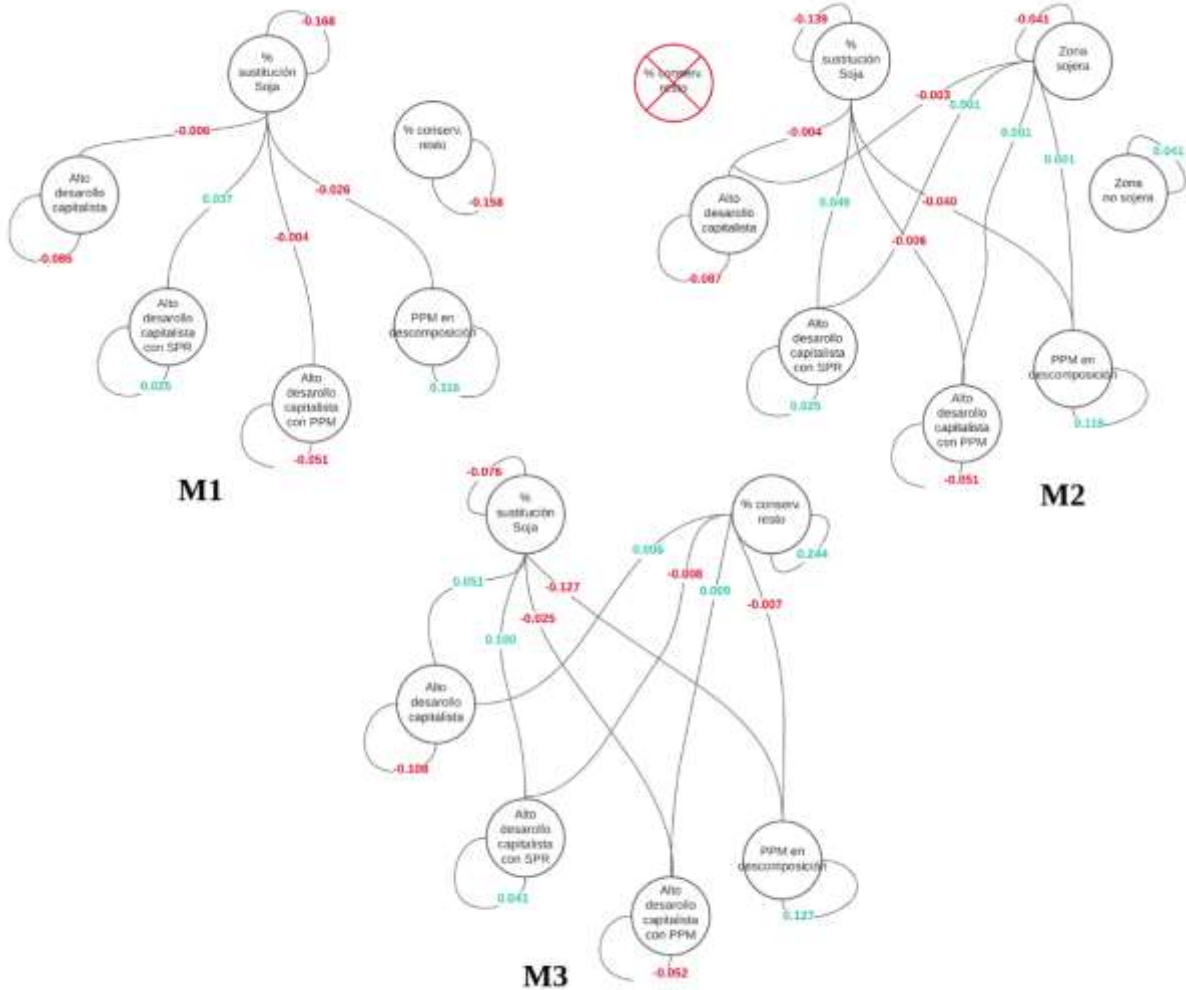
En el segundo modelo (M2) se ajustó un modelo sobre el mismo dataset anterior, pero se agregó una variable que diferencia a los departamentos en "sojeros" y "no sojeros"(aquellos

departamentos que no tienen ninguna hectárea de soja sembrada en ninguno de los dos censos relevados).

El último modelo (M3) consiste en las mismas variables que las utilizadas para el modelo 1 pero fue entrenado para los departamentos sojeros exclusivamente.

Se presentan a continuación los coeficientes de los tres modelos estimados. Se representan como una red en la cual cada nodo es una variable. Los vínculos entre los diferentes nodos y los valores asociados representan los coeficientes de los efectos seleccionados por el LASSO: tanto de los efectos principales (los bucles) como de las interacciones (los vínculos entre los nodos).

Figura 3. Coeficientes de regresión para los tres modelos de regresión lineal regularizados mediante LASSO



Fuente: elaboración propia sobre datos de CNA 1988 y 2002

En los tres modelos, las estructuras agrarias en las que la desaparición de EAPs es más intensa (efectos principales de los tres modelos) son las estructuras 1 y 3, caracterizadas por un alto grado de desarrollo de relaciones salariales y por su articulación con explotaciones "grandes" (estructura 1) y con fracciones de pequeña producción mercantil simple y capitalista (estructura 3). A su vez, aquellos departamentos que presentan un alto grado de desarrollo

de relaciones salariales articuladas con la presencia de superpoblación relativa (estructura 2) y alta prevalencia de pequeña producción mercantil en crisis (estructura 4) presentan un efecto positivo sobre la tasa de variación de las EAPs entre 1988 y 2002.

En cambio, el efecto de las hectáreas que se "conservan", es decir que no son sustituidas por soja, presentan efectos diferentes en cada modelo:

- negativo y de una magnitud similar al de las hectáreas sustituidas por soja en el modelo que considera a los 482 departamentos (sin discriminar por su carácter sojero-no sojero)
- no significativo en el modelo 2 (al introducir el la variable dummy sojero-no sojero)
- positivo y de magnitud muy relevante al considerar solamente los departamentos sojeros

Al observar las interacciones puede notarse que al analizar el total de los departamentos e independientemente de la introducción de una variable dummy sojero-no sojero (modelos 1 y 2) se observa que la interacción entre la proporción de hectáreas sustituidas por soja y las variables de estructuras agrarias presentan efectos positivos en las estructuras 1, 3 y 4: es decir, estructuras de alto grado de desarrollo de relaciones salariales, con peso de explotaciones "grandes", articulación con elementos de pequeña producción capitalista (1 y 3). En cambio, en aquellas estructuras en las que la proporción de hectáreas sustituidas por soja es alta y se trata de estructuras en las que pesan fracciones de superpoblación relativa el efecto es positivo. Al analizar los departamentos sojeros, esta interacción cambia: se hace negativa en las estructuras 3 y 4 y positiva en la 1 y 2.

De esta forma, puede notarse que en los tres modelos el "efecto estructura" tiene un peso relevante y casi equivalente al efecto sustitución. Es decir, que la desaparición de EAPs se

debe aproximadamente en la misma magnitud al proceso de sustitución de cultivos por la soja y al tipo de estructura agraria en el que tales procesos se desarrollan.

Esto parece mostrar que los procesos de sustitución de soja presentan efectos diferentes según el tipo de estructura agraria en el que se desarrollan. Así, en las estructuras de mayor grado de desarrollo de relaciones capitalistas (1 y 3) el efecto es negativo. En cambio, en las estructuras que articulan superpoblación relativa y pequeña producción en crisis, este signo es positivo. ¿Hasta qué punto esto confirma el rasgo que estas estructuras presentan en tanto reservorios de superpoblación en el campo (bajo la forma de pequeña producción descapitalizada)? Un análisis de flujos migratorios y saldos migratorios podría arrojar luz sobre estos aspectos.

6. Comentarios finales y nuevos problemas

En el presente trabajo se buscó abordar tres aspectos de la llamada primera fase de expansión sojera reciente (1988-2002): 1) estimar cuál es el peso de la sustitución de la superficie ocupada por otros cultivos a causa del avance sojero; 2) identificar algunos patrones generales en la expansión de la frontera agraria y 3) aportar información (aún preliminar) sobre la relación entre estos procesos y las diferentes estructuras agrarias.

Fue posible detectar zonas en las que este efecto es más alto: Norte de Buenos Aires y Sur de Santa Fé; Centro-Oeste de Chaco, Centro-Este de Santiago del Estero; Oeste de Tucumán y Este de Salta. A su vez, fue posible observar diversas formas en las que la expansión de la frontera se produce: en la zona núcleo pampeana y en otras zonas como el área algodonera de Chaco, aparecen caracterizadas por cambios importantes en el uso del suelo.

Al mismo tiempo, es necesario remarcar algunos procesos que quedan fuera del análisis de este trabajo: los fenómenos de deforestación y los aparentes cambios en la distribución de la ganadería. Ambos son difíciles de abordar con fuentes censales. En el primer caso, los Censos Agropecuarios solamente relevan superficie de monte ubicadas dentro de los límites de las explotaciones. En el segundo, los Censos contabilizan la cantidad de cabezas y no la superficie ocupada por ganado, por lo cual se hace imposible incluirlos en los modelos de inferencia ecológica sin tener algún tipo de coeficiente que permita traducir la cantidad de cabezas a superficie ocupada.

Las existencias de algunas evidencias de sesgos de agregación en los modelos de inferencia ecológica hacen necesario intentar una mayor apertura de las diferentes clases de cultivo en la estimación. En lugar de utilizar tablas de 2x2 (soja-resto) para cada departamento, utilizar una desagregación de la categoría “resto”, quizás en términos de los requerimientos de fuerza de trabajo de cada cultivo.

En ese sentido, un problema que se abre es la necesidad de replicar ambos ejercicios (la estimación de las diferentes formas de avance de la frontera agrícola y los impactos de la sojización sobre las estructuras agrarias) con datos más recientes para detectar si existen cambios en estas modalidades de expansión. Será necesario extender el ejercicio para el período correspondiente al Censo Nacional Agropecuario de 2018, cuyos datos preliminares recién se estaban publicando al momento de escribir el presente artículo. Si bien existe un relevamiento censal agropecuario previo al de 2018, (el Censo Nacional Agropecuario de 2008), el operativo de campo se desarrolló en una coyuntura signada por el conflicto por las retenciones móviles, por lo cual parece haber presentado una serie de problemas que es necesario indagar antes de utilizar dicha fuente. También las series de estimaciones agrícolas del Ministerio de Agricultura, Ganadería y Pesca son una fuente con desagregación espacial y temporal útiles para este tipo de análisis.

A su vez, un campo de interrogantes se abre acerca de la influencia que el proceso de expansión sojera tiene sobre las estructuras agrarias en Argentina. Particularmente, en términos de la llamada “pampeanización”: ¿cuál es la importancia que los procesos de sojización tienen en la homogeneización de las diferentes estructuras agrarias y muchos de sus efectos asociados (disminución del número de explotaciones, centralización de la producción, incremento de fracciones de rentistas, cambios en la propiedad del suelo, etc.)? ¿Qué modalidades asume en las diferentes estructuras agrarias de Argentina en este período?

En este punto, se presentan algunos resultados preliminares de un ejercicio basado en modelos de regresión regularizados que intentan aportar algunas evidencias empíricas sobre la relación entre el proceso de expansión sojero y ciertos movimientos en la estructura agraria (particularmente, la desaparición de explotaciones). El ejercicio parece aportar elementos que permiten esbozar algunas hipótesis y preguntas: los procesos de sustitución de soja presentan efectos diferentes según el tipo de estructura agraria en el que se desarrollan. En las estructuras de mayor grado de desarrollo de relaciones capitalistas (el efecto es negativo; en las que articulan superpoblación relativa y pequeña producción en crisis, este signo es positivo. ¿Hasta qué punto esto confirma el rasgo que estas estructuras presentan en tanto reservorios de superpoblación en el campo (bajo la forma de pequeña producción descapitalizada)? Un análisis de flujos migratorios y saldos migratorios podría arrojar luz sobre estos aspectos.

7. Bibliografía citada

Ackhar, Marcel, Domínguez, Ana, Díaz, Ismael y Pesce, Fernando (2011). La intensificación del uso agrícola del suelo en el litoral oeste del Uruguay en la última década. *Revista Pampa*, 7, 143-157.

Azcuy Ameghino, Eduardo y León, Carlos (2005). La sojización: contradicciones, intereses y debates". *Revista Interdisciplinaria de Estudios Agrarios*, 23, 133-157.

Azcuy Ameghino, Eduardo y Ortega, Lucía (2010). Sojización y expansión de la frontera agropecuaria en el NEA y NOA: transformaciones, problemas y debates. *Documentos del PIEA*, 5, 141-159.

Borrello, José (1988). El Desplazamiento de la Frontera Agrícola en la Argentina Contemporánea. *Revista Geográfica*, 108, 5-18.

Breiman, Leo (1995). Better subset selection using the nonnegative garrote. *Technometrics*, 37(4), 738-754.

Duncan, Otis Dudley and Beverly Davis (1953). An Alternative to Ecological Correlation. *American Sociological Review*, 18(6), 665–666.

Fehlenberg, Verena, Baumann, Matthias, Gasparri, Nestor, Piquer-Rodríguez, María, Gavier-Pizarro, Gregorio y Kuemmerle, Tobias (2017). The role of soybean production

as an underlying driver of deforestation in the South American Chaco. *Global Environmental Change*, 45, 24-34.

Giarracca, Norma y Teubal, Miguel (2013). Las actividades extractivas en la Argentina. En Giarracca, Norma y Teubal, Miguel (coords.), *Actividades extractivas en expansión ¿Reprimerización de la economía argentina?* (pp. 19-44). Buenos Aires: Antropofagia.

Goodman, Leo (1959). Some Alternatives to Ecological Correlation. *American Journal of Sociology*, 64(6), 610–624.

Hoerl, Arthur y Kennard, Robert (1970). Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, 12(1), 55-67

James, Gareth, Witten, Daniela, Hastie, Trevor y Tibshirani, Robert (2017). *An Introduction to Statistical Learning with Applications in R*. Berlín: Springer.

Krapovickas, Julieta (2009). *Cambio socio-ambiental en el Chaco Argentino y su relación con la expansión de soja en la década de 1990* [Tesis de Maestría en Estudios Territoriales y de la Población], Universidad Autónoma de Barcelona.

King, Gary (1997). *A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data*. Princeton: Princeton University Press.

LART/FAUBA (2004). *Patrones espaciales y temporales de la expansión de Soja en Argentina. Relación con factores socio-económicos y ambientales*, Buenos Aires: Laboratorio de Análisis Regional y Teledetección (LART)/FAUBA.

León, Carlos, Prudkin, Nora y Reboratti, Carlos (1985). El conflicto entre producción, sociedad y medio ambiente: la expansión agrícola en el sur de Salta. *Desarrollo Económico*, 25(99), 399-420.

Lim, Michael y Hastie, Trevor (2014). Learning Interactions via Hierarchical Group-Lasso Regularization. *Journal of Computational and Graphical Statistics*, 24(3), 627-654.

Martínez Dougnac, Gabriela (2013). De los márgenes al boom. Apuntes para una historia de la sojización. En Martínez Dougnac, Gabriela (comp.), *De especie exótica a monocultivo. Estudios sobre la expansión de la soja en Argentina* (pp. 1-38), Buenos Aires: Imago Mundi.

Mena, Carlos (2007). *Characterizing and modeling agricultural and forest trajectories in the northern ecuadorian amazon: spatial heterogeneity, socioeconomic drivers, and spatial simulations* [Tesis Doctoral University of North Carolina], Carolina Campus Repository <https://cdr.lib.unc.edu/concern/dissertations/gx41mj94k>.

Ortega, Lucía (2011). ¿Qué es la expansión de la frontera agropecuaria? Aproximación al caso de Chaco. *Documentos del PIEA*, 6, 87-109.

Quinn, Kevin (2004). Ecological Inference in the Presence of Temporal Dependence. En King, Gary, Rosen, Ori y Tanner, Martin (coords.), *Ecological Inference: New Methodological Strategies* (pp. 207-232). New York: Cambridge University Press.

Rodríguez, Javier (2008). *Consecuencias económicas de la soja transgénica. Argentina 1996-2006*. Buenos Aires, CLACSO-Ediciones Cooperativas.

Rosati, Germán (2013). Patrones espaciales de expansión de la frontera agrícola: la expansión sojera en Argentina (1987/88-2009/10). En Gras, Carla y Hernández,

Valeria (comps.), *El agro como negocio: producción, sociedad y territorios en la globalización* (pp- 97-122), Buenos Aires: Biblos.

Rosati, Germán y Chazarreta, Adriana (2020). Tipos de estructuras sociales agrarias en la formación social argentina. Un análisis a nivel departamental: 2001-2002. *Mundo Agrario*, en prensa.

Steimbregger, Norma, Radonich, Marta y Bendini, Mónica (2003). Expansiones de frontera agrícola y transformaciones territoriales: procesos sociales diferenciales. En

Bendini, Mónica y Steimbregger, Norma (comps.), *Territorios y organización social de la agricultura* (pp. 17-39). Buenos Aires, La Colmena.

Tibshirani, Robert, Walther, Guenter y Hastie, Trevor (2001). Estimating the number of data clusters via the Gap statistic. *Journal of the Royal Statistical Society*, 63(2), 411-423.

Valenzuela, Cristina (2014). Implicancias del avance de la 'frontera' agropecuaria en el Nordeste Argentino en las últimas dos décadas. *Estudios Socioterritoriales*, 2(16): 95-109.

Viglizzo, Ernesto, Jobbágy, Esteban (eds.) (2010). *Expansión de la Frontera Agropecuaria en Argentina y su Impacto Ecológico-Ambiental*, Buenos Aires, Ediciones INTA.

Volante, José Norberto, Mosciaro, María Jesús, Morales Poclava, María Cecilia, Vale, Laura Marina, Castrillo, Silvana Alejandra, Sawchik, Jorge, Tiscornia, Guadalupe, Fuente, Marcel, Maldonado Ibarra, Isaac, Vega, Trujillo, Richard, Cortéz, Paruelo, José (2015). Expansión agrícola en Argentina, Bolivia, Paraguay, Uruguay y Chile entre 2000-2010: caracterización espacial mediante series temporales de índices de vegetación. *Revista de Investigaciones Agropecuarias*, 41(2), 179-191.

Haneuse, Sebastien y Wakefield, Jonathan (2004). Ecological Inference Incorporating Spatial Dependence. En King, Gary, Rosen, Ori y Tanner, Martin (coords.), *Ecological Inference: New Methodological Strategies* (pp. 266-302). New York: Cambridge University Press.

Zarrilli, Adrián (2008). El proceso de agriculturización en las regiones extrapampeanas argentinas: insostenibilidad y límites de un modelo de transformación. La provincia del Chaco (1980-2006). *XII Congreso de Historia Agraria*, Córdoba: Argentina.