



Estimaciones de ingresos en censos. Mapas de ingreso y una aplicación a resultados electorales (2003 - 2023)

Tobías Yatche

Tesis de Maestría

Maestría en Economía

Universidad Nacional de La Plata

Directores de tesis: Leonardo Gasparini y Pablo Gluzmann

Fecha de defensa: 20/02/2025

Códigos JEL: J30, D72

Estimaciones de ingresos en censos. Mapas de ingreso y una aplicación a resultados electorales (2003 - 2023)

Octubre 2024

Tobías Yatche

Maestría en Economía - UNLP *

Resumen

En esta investigación se construye un índice de ingreso estimado por área geográfica utilizando la técnica de *small area estimation* (SAE). Esta técnica permite proyectar los estados socioeconómicos semestrales obtenidos de la Encuesta Permanente de Hogares (EPH) sobre las variables del Censo Nacional, generando un índice del ingreso estimado por radio censal durante el período 2003-2023. A partir de este índice, se realizan dos aplicaciones: primero, se elabora un mapa semestral con datos de ingresos por radio censal para toda Argentina, segundo, se analiza la relación entre las variaciones de ingresos y la performance electoral de los oficialismos donde se observa una predominancia del voto sociotrópico sobre el egotrópico^{**}.

JEL: J30, D72

*Director de tesis Leonardo Gasparini y co-director Pablo Gluzmann.

**El presente trabajo es una derivación de la monografía “Distribución geográfica del ingreso en Argentina” de Kenny, Walovnik e Yatche y el documento de trabajo de Abbate, Gasparini, Gluzmann, Montes-Rojas, Sznaider y Yatche (2023), “Ingreso Estimado Por Area Geográfica: una aplicación para Argentina”, presentado en la LVIII Reunión Anual de la Asociación Argentina de Economía Política.

1. Introducción

En esta investigación se construye un índice de ingreso estimado por área geográfica utilizando la técnica de *small area estimation* (SAE). Esta técnica permite proyectar los estados socioeconómicos semestrales obtenidos de la Encuesta Permanente de Hogares (EPH) sobre las variables del Censo Nacional, generando un índice del ingreso estimado por radio censal durante el período 2003-2023.

La recopilación de información sociodemográfica se lleva a cabo principalmente mediante encuestas a hogares. Estas herramientas se utilizan ampliamente a nivel mundial para medir ingresos, pobreza y otros indicadores de bienestar. Los cuestionarios necesarios para relevar ingresos o consumo son extensos y requieren una formación adecuada de los encuestadores, lo que resulta en muestras relativamente pequeñas en comparación con la población de referencia. Esto implica que, aunque las mediciones son fiables a nivel de grandes agregados (por ejemplo, a nivel nacional, regional o de aglomerados urbanos), carecen de representatividad cuando se busca un mayor nivel de desagregación geográfica. En Argentina, estos problemas se agravan: la Encuesta Permanente de Hogares (EPH), realizada trimestralmente, es la principal fuente de datos para medir ingresos, pobreza, desigualdad y otros indicadores socioeconómicos importantes y su cobertura se limita únicamente a los principales aglomerados urbanos.

Por otra parte, los censos, que se llevan a cabo aproximadamente cada diez años, recopilan información de prácticamente toda la población del país. Esto permite una desagregación granular: la unidad máxima de desagregación son los radios censales, que agrupan en promedio a 300 viviendas, permitiendo que en zonas de alta densidad poblacional la información se presente a nivel de manzana. No obstante, en Argentina los censos no preguntan sobre ingresos ni gastos, lo que impide la estimación directa de indicadores de pobreza monetaria. Con esta información, solo es posible calcular indicadores de pobreza multidimensional que no incluyen variables monetarias.

Dadas las limitaciones de los datos de las encuestas, la literatura ha propuesto diversas soluciones para abordar la falta de estimaciones directas con alta desagregación espacial. Estas técnicas se conocen comúnmente como estimaciones de área pequeña (*small area estimation*, SAE). Las técnicas de SAE proporcionan un marco analítico que permite mejorar la granularidad de la información sin necesidad de recolectar más datos, al incorporar fuentes auxiliares para aumentar la precisión.

En la primera sección, se implementa *sae* para construir un índice de ingresos esperados semestrales con alta desagregación geográfica. A partir de este índice, se realizan dos aplicaciones. Primero, se elabora un mapa semestral con datos de ingresos por radio censal para toda Argentina. Segundo, se analiza la relación entre las variaciones de ingresos y la performance electoral de los oficialismos. En esta última sección, se discuten los determinantes del voto económico, donde se observa una predominancia del voto sociotrópico sobre el egotrópico.

2. Estimación del ingreso en censos

2.1. Implementación de *small area estimation*

Uno de los métodos más utilizados de *small area estimation* (SAE) es el desarrollado por [Elbers et al. \(2003\)](#). Este método requiere la disponibilidad de microdatos del censo y de una encuesta contemporánea. Usando datos de encuestas, se construye un modelo que relaciona una medida del bienestar del hogar, como el gasto per cápita, con un conjunto de características, sus miembros y su ubicación. Utilizando los parámetros estimados del modelo, se aplican estas características a los datos del censo de la población del país para simular el gasto per cápita del hogar (o los ingresos) para cada hogar individual. A partir de los valores simulados del censo, podemos calcular cualquier indicador de bienestar. Esta metodología ha sido utilizada por organismos nacionales de estadísticas y organizaciones internacionales para estimar mapas de pobreza. Entre muchos ejemplos se pueden citar los mapas de pobreza de Brasil ([Ebers et al., 2008](#)), Vietnam ([Cuong et al., 2010](#)) y Sri Lanka ([Department of Census and Statistics and World Bank, 2015](#)). Desde la publicación de esta metodología, ha habido avances significativos en la literatura, como el estimador *Empirical Best* (EB), que además incorpora al modelo información adicional de la encuesta ([Molina and Rao, 2010](#)).

El objetivo en esta sección es aplicar los desarrollos en términos de SAE para el caso argentino, incorporando la dimensión estructural de las estimaciones espaciales del ingreso. Para ello, siguiendo la estrategia propuesta [Gasparini et al. \(2022\)](#) se utilizan todas las ondas disponibles de la Encuesta Permanente de Hogares en su versión continua (en lugar de utilizar únicamente los datos contemporáneos). De esta manera, se consigue el primer panel por radio censal con el ingreso correspondiente a cada onda de la EPH lo que permite evaluar el ingreso que hubiera tenido cada hogar en distintos estados de la naturaleza, y para estimar un ingreso promedio o estimado con alta desagregación geográfica.

2.2. Metodología

El presente trabajo se centra en la construcción de un índice de ingreso estimado por área geográfica, es decir, un indicador del ingreso espacial para todo el territorio argentino. La construcción del índice se realiza estimando el ingreso per cápita familiar para cada individuo del censo desde el 2003 hasta el 2023 y agregando estas estimaciones del ingreso por área geográfica. Se utilizan las EPH (Encuesta Permanente de Hogares) publicadas por el INDEC y se incluye la homogeneización del método de imputación de ingresos propuestos en [Tornarolli \(2018\)](#). Los ingresos per cápita familiares se expresan a precios 2011 para la consistencia intertemporal.

Para estimar el ingreso per cápita familiar en cada semestre se utiliza un modelo de regresión condicional que relaciona los ingresos con el capital humano de las personas y las características de los hogares e individuos. De esta manera es posible evaluar cómo las variables observadas explican el ingreso per

cápita familiar para los jefes de hogares, el indicador más importante de los ingresos de los hogares y las personas, en cada onda de la EPH.

Una vez construido el IPCF estimado para cada hogar en cada semestre de la EPH, se construye el ingreso estimado de cada hogar del censo como el promedio de las predicciones del ingreso para el período 2003-2023. Por último, se estima el ingreso estimado por radio censal como la media del ingreso estimado individual para cada radio censal para los últimos dos censos disponibles.

La estimación del ingreso se realizó mediante una variación del método desarrollado por [Elbers et al. \(2003\)](#). En el trabajo original, los autores plantean una estimación de área pequeña que consiste en: (i) seleccionar un indicador de interés que se encuentre disponible en una encuesta pero no en el censo, (ii) identificar covariables que se relacionen con el indicador y que se encuentren tanto en el censo como en la encuesta, (iii) estimar un modelo lineal de mínimos cuadrados generalizados (GLS) sobre la encuesta, que relacione el indicador de interés con las covariables seleccionadas, y (iv) generar predicciones del indicador en el censo, utilizando las covariables censales disponibles y las estimaciones de los parámetros del modelo del punto previo. Estas predicciones tienen la ventaja de estar geográficamente localizadas, ya que típicamente los censos permiten una desagregación espacial mucho mayor que las encuestas de hogares. En general, luego de esta predicción se agregan los indicadores al mínimo nivel espacial disponible, lo que permite construir mapas con las estimaciones correspondientes.

En concreto, se estimó un modelo lineal sobre los microdatos de cada uno de los semestres de la Encuesta Permanente de Hogares entre 2003 y 2023, estimando la relación entre el ingreso per cápita familiar de cada hogar i del aglomerado j (y_{ij}^t) y un vector de variables observables (X_{ij}^t). Las variables incorporadas incluyen características del hogar y la vivienda (disponibilidad de agua potable, acceso a baño, condición de propiedad, cantidad de miembros) y del jefe del hogar (género, edad, nivel educativo). En el modelo se incluyen efectos fijos por aglomerado δ_j^t . Por lo tanto, para el semestre t de la Encuesta Permanente de Hogares, se estimó un modelo correspondiente es de la forma:

$$y_{ij}^t = \beta^t X_{ij}^t + \delta_j^t + \varepsilon_{ij}^t \quad (1)$$

Una vez estimado el modelo, se guardan las estimaciones del vector de parámetros de cada una de las ondas β^t y δ_j^t . Utilizando estas estimaciones, se generan predicciones del ingreso para cada uno de los hogares del censo, utilizando las covariables de los últimos dos censos disponibles X_{icj} . En otras palabras, se utilizan la relación estimada en la EPH para predecir el ingreso utilizando las variables observables del censo. A partir de esta predicción, se agrupa el ingreso estimado:

$$\hat{y}_{icj}^{ts} = \beta^t X_{icj}^s + \delta_j^t \quad (2)$$

Esta relación, estimada para cada semestre de la EPH, puede interpretarse como el ingreso que el

hogar i del radio censal c del aglomerado j encuestado en cada uno de los censos s tendría bajo la situación económica correspondiente al semestre t . En este sentido, las variaciones de los rendimientos de cada una de las características observables estimadas según la ecuación 1 a lo largo del tiempo se entienden como cambios en la situación económica. Para los hogares rurales se imputó el efecto fijo del aglomerado más cercano, ya que no existen datos de ingresos de personas que viven en la ruralidad en la EPH.

Finalmente, para construir el indicador de *ingreso estimado* por radio censal se calcula la media del ingreso estimado individual para cada uno de los radios censales. Por lo tanto, definiendo n_c como el número de hogares contenidos en el radio censal c , el *ingreso estimado* del radio censal c (\bar{Y}_{cj}) queda definido como:

$$\bar{Y}_{cjt}^s = \sum_{i \in c} \frac{\bar{y}_{icjt}^s}{n_{jt}} \quad (3)$$

Con esta estimación, resulta posible generar mapas del ingreso estimado por radio censal para cada onda de la EPH y, de ese modo, tener una estimación del ingreso geográfico semestral para todos los radios censales disponibles en el Censo.

2.3. Ingreso estimado con alta desagregación geográfica y variabilidad intertemporal

2.3.1. Precisión de las estimaciones sobre la EPH

En esta sección nos concentramos en evaluar la performance de las predicciones del modelo de ingreso estimado (IE) en comparación con el ingreso per cápita familiar (IPCF) observado en la encuesta. Antes de analizar los resultados resulta útil establecer ciertas hipótesis respecto de dicho desempeño. Podemos separar las discrepancias entre IE e IPCF en “deseables” e “indeseables”.

Las discrepancias deseables están relacionadas con el carácter estructural del IE y el carácter corriente del IPCF. En términos de variabilidad, se espera una menor dispersión (y desigualdad) del IE, ya que se trata de una estimación que, como se mostrará más adelante en la Figura ??, tiende a subestimar los outliers. En términos de sesgo, podrían presentarse diferencias si el ingreso corriente presentara mayores o menores niveles en algún grupo no aleatorio de hogares, pero en este caso el sesgo no tiene a priori un signo definido. Otra diferencia deseable se relaciona con la cobertura, mientras que la EPH se realiza en grandes aglomerados urbanos, las estimaciones IE permiten al extender los resultados a zonas urbanas más pequeñas y zonas rurales.

Las discrepancias “indeseables” se relacionan con la falta de información (i.e inobservables de carácter

Cuadro 2.1: Estadísticas descriptivas del Ingreso estimado y el ingreso per cápita familiar.

	Ingreso Estimado (predicho)	Ingreso per cápita familiar (observado)
Media	1.447,0	1.847,5
Mediana	1.190,1	1.374,7
Media por quintiles	1	520,7
	2	859,3
	3	1.199,7
	4	1.705,0
	5	2.950,2
Coefficiente de gini	0,336	0,436

Fuentes: IPCF calculado utilizando la EPH del 2do semestre de 2010, Ingreso Estimado calculado utilizando características de la EPH 2do semestre de 2010 y modelos estimados para todas las ondas de EPH 2003-2022.

estructural potencialmente relevantes), así como la incorrecta especificación del modelo econométrico. En ambos casos las discrepancias pueden provocar inconsistencias tanto en términos de sesgo como de varianza. El carácter lineal del método de estimación (MCO) elegido puede provocar sesgos por error de especificación en la medida que se presenten no linealidades entre el ingreso y las variables explicativas. Adicionalmente, las covariables elegidas ad-hoc pueden no ser suficientes para predecir correctamente la relación. La razón por la cual se aplica esta estrategia es simplemente operativa, ya que en futuras versiones se buscará evaluar métodos más sofisticados de estimación (métodos semi paramétricos), así como incorporar de distintos métodos de selección de variables relevantes, como por ejemplo la inclusión más variables disponibles tanto en la EPH como en el censo, la incorporación de interacciones, la selección de variables mediante *lasso* (Tibshirani, 1996), Global Search Regression (Gluzmann and Panigo, 2015; Panigo et al., 2020, 2023), entre otros. El modelo elegido debería subestimar la variabilidad, y el método carácter lineal probablemente genere sesgos en las colas de la distribución. En términos de desigualdad, el IE presenta un coeficiente de Gini sustancialmente menor, aunque es difícil a priori establecer en qué medida esto se debe a discrepancias deseables o indeseables.

El Cuadro 2.1, muestra un menor valor del IE para medidas centrales (media y mediana). Una posible hipótesis que explique este resultado se relaciona con que el año 2010 fue un año relativamente “bueno” para Argentina. Al observar las diferencias por niveles de ingreso, se evidencia un mayor nivel respecto del IPCF para el quintil más bajo, y menores niveles para el resto de los quintiles. Sin embargo, esta diferencia es creciente a medida que se incrementa el nivel de ingresos, especialmente en el 5to quintil. Si bien parte de estas discrepancias pueden ser de carácter “deseable”, es probable que éstas sean principalmente producto de diferencias indeseables, especialmente en el tramo de mayores ingresos.

La Figura 1 permite visualizar las diferencias antes mencionadas a lo largo de toda la distribución. El panel (a) muestra estimaciones de kernel del logaritmo del IE y el IPCF. Se observa que existen diferencias a favor del IPCF en el centro de la distribución, pero las curvas difieren en las colas, con mayor concentración del IPCF en estos tramos. El panel (b) compara, a nivel de individuo, el IPCF con su IE. Puede observarse que existe una mayor dispersión de los puntos en los niveles más bajos

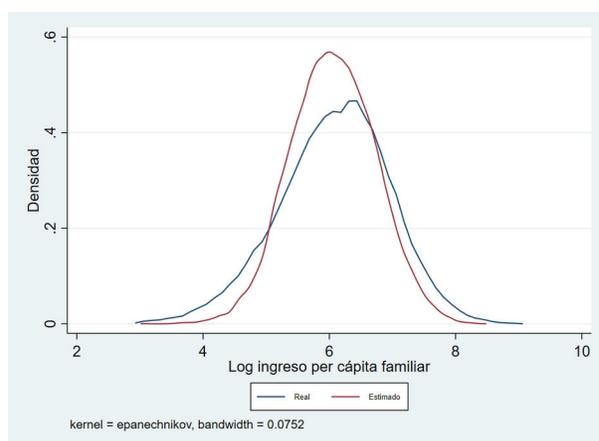
Cuadro 2.2: Ingreso estimado e IPCF, por regiones

	Ingreso Estimado (predicho)	Ingreso per cápita familiar (observado)
Gran Buenos Aires	1.586,8	2.063,9
Pampeana	1.424,5	1.788,0
Cuyo	1.175,3	1.349,6
Noroeste Argentino	1.026,9	1.271,1
Noreste Argentino	977,0	1.166,3
Patagonia	1.918,8	2.606,5
Total	1.447,0	1.847,5

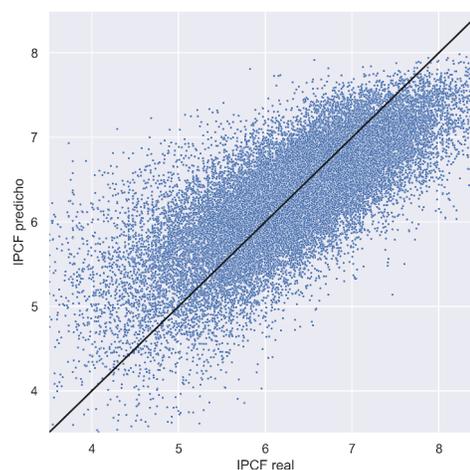
Fuentes: IPCF calculado utilizando la EPH del 2do semestre de 2010, Ingreso Estimado calculado utilizando características de la EPH 2do semestre de 2010 y modelos estimados para todas las ondas de EPH 2003-2022.

de la distribución del ingreso y una mayor masa a la derecha de la curva de 45 grados en el tramo de mayores ingresos.

Figura 1: Distribución IPCF e IE



(a) Comparación del IPCF predicho y el real



(b) Distribución real y estimada del IPCF

El Cuadro 2.2 muestra las diferencias a nivel de regiones del país. En todas las regiones se observan mayores ingresos en términos del IPCF, siendo las regiones de GBA y La Pampa (regiones con ingresos mayores al promedio nacional) las que presentan mayores diferencias.

Finalmente, el Cuadro 2.3 muestra distintas variables asociadas al bienestar. En la medida en que el IE represente mejor niveles de ingreso de mediano/largo plazo, deberíamos esperar que los individuos de menores ingresos presenten peores desempeños en términos de vivienda e infraestructura, años educación y necesidades básicas insatisfechas. En el cuadro siguiente se observan efectivamente estos patrones para el quintil de menores ingresos. En la última fila se muestra el porcentaje de desocupados respecto de la población económicamente activa, una variable predominantemente de carácter coyun-

Cuadro 2.3: Ingreso estimado (IE) e Ingreso per capital familiar (IPC) en relación a características observables de los hogares

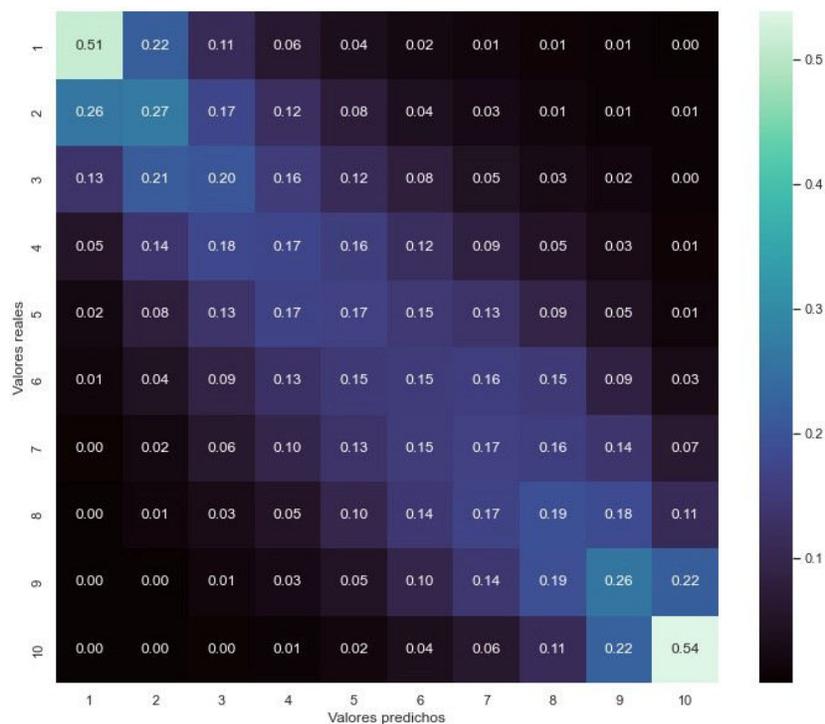
	Quintiles de ingreso estimado y observado									
	1		2		3		4		5	
	IE	IPC	IE	IPC	IE	IPC	IE	IPC	IE	IPC
Vivienda en lugar precario (%)	9.0	6.5	3.3	3.4	2.2	2.2	0.9	1.1	0.6	0.6
Vivienda con mat. precarios (%)	7.3	5.3	2.4	2.6	1.3	1.3	0.6	0.9	0.3	0.2
Vivienda con acceso a agua (%)	95.9	97.3	99.3	98.9	99.2	99.0	99.6	99.7	100	99.9
Vivienda con baño (%)	54.5	71.0	87.7	85.2	93.1	91.9	96.7	96.5	99.8	99.0
Vivienda con acceso a cloaca (%)	21.7	37.1	40.6	50.0	55.6	56.4	70.3	68.9	87.6	84.6
Hogares con NBI (%)	69.7	51.9	32.1	31.5	18.4	22.1	12.5	11.8	4.5	4.1
Miembros por habitación	1.61	1.24	0.98	0.94	0.75	0.73	0.56	0.58	0.38	0.44
Años de educación promedio	5.67	6.18	7.36	7.34	8.60	8.32	9.75	9.55	12.1	12.1
Desempleo (% personas activas)	11.0	17.6	8.9	10.8	7.9	7.8	6.9	5.5	4.3	2.1

Fuentes: IPCF calculado utilizando la EPH del 2do semestre de 2010, Ingreso Estimado calculado utilizando características de la EPH 2do semestre de 2010 y modelos estimados para todas las ondas de EPH 2003-2022.

tural. Se observa que la diferencia se revierte respecto de las variables anteriores, lo que es “deseable” para un indicador que busca representar niveles de ingreso de mediano/largo plazo.

La Figura 2 muestra la correlación por deciles entre el IPCF y el IE. Se observa que, en general, el modelo capta mejor los extremos de la distribución, mientras que tiene más problemas para captar los deciles medios de manera correcta. Sin embargo, se observa que cuando el modelo predice “incorrectamente” el decil, lo hace por un valor cercano al real.

Figura 2: Correlación por deciles entre el IPCF y el IE



Nota: Valores normalizados respecto a los deciles reales

3. Ingreso estimado por Área Geográfica

A partir de la metodología descrita en las secciones anteriores es posible construir mapas del Ingreso estimado por radio censal utilizando los dos últimos censos disponibles.

En la Figura 3 se presenta la distribución relativa del Ingreso Estimado por Área Geográfica (IEPAG) para la Argentina utilizando las características observables del censo 2001 y 2010. En ambos casos se muestra el decil del IEPAG correspondiente a cada radio censal. En términos generales, los mapas presentan un patrón similar: el sur de la Argentina tiene un ingreso medio-alto, el centro un ingreso medio que decrece hacia el este, y el norte presenta un ingreso bajo que decrece en la región del NOA. En términos temporales, hubo un empobrecimiento relativo en el interior de la Provincia de Buenos Aires, San Luis, Córdoba y Santa Fe entre el 2001 y 2010. Sin embargo, en el resto del país la pobreza relativa se mantiene constante entre ambos períodos. Esto se explica por el crecimiento relativo del ingreso en los centros urbanos.

Al comparar la distribución del ingreso estimado, se observa un incremento en 2010 respecto a 2001, lo que indica una mejora en el ingreso por radio censal. Como muestra la Figura 4, el desplazamiento de la distribución hacia la derecha refleja una mejora en las características observables de los hogares, como

mayor nivel educativo, mejores condiciones de vivienda y mejoras generales en las variables utilizadas para la predicción.

Figura 3: Ingreso estimado por Área Geográfica, evolución de la distribución relativa

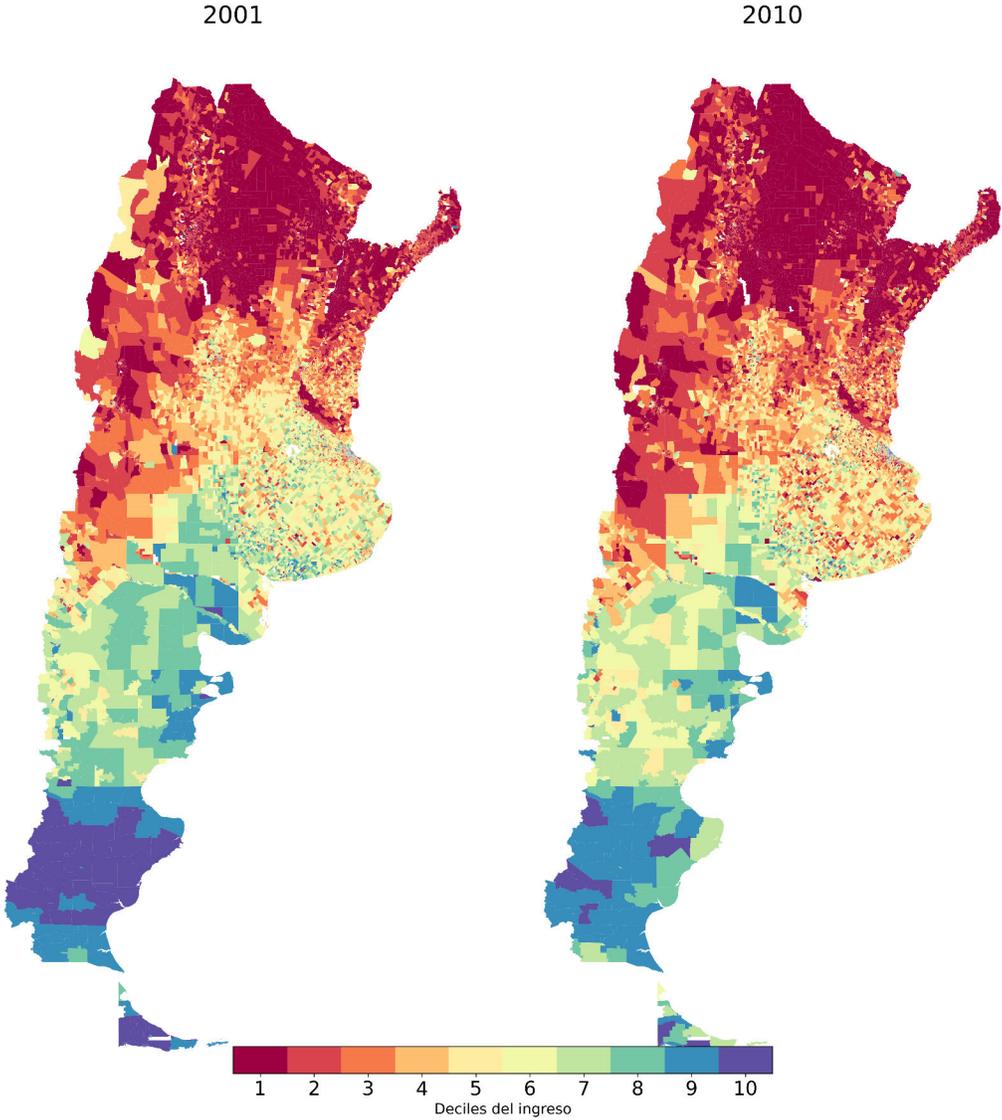
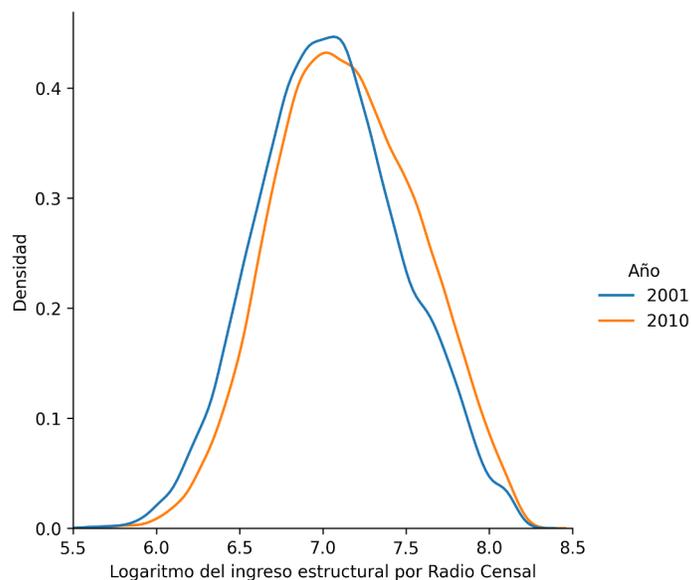


Figura 4: Distribución Logaritmo Ingreso estimado 2001 y 2010



Además de estudiar lo que sucede con la distribución relativa del ingreso estimado, resulta posible, a partir de la metodología aplicada en el presente trabajo, analizar la distribución absoluta del ingreso estimado.

En la Figura 5 se observa que, en las principales zonas urbanas del país, el ingreso crece en términos absolutos entre los años 2001 y 2010. Al igual que nivel nacional, este fenómeno se da por una mejora en las características observables de los individuos. Este crecimiento refleja un cambio real en las condiciones estructurales de los hogares, ya que los coeficientes de los modelos se mantienen constantes.

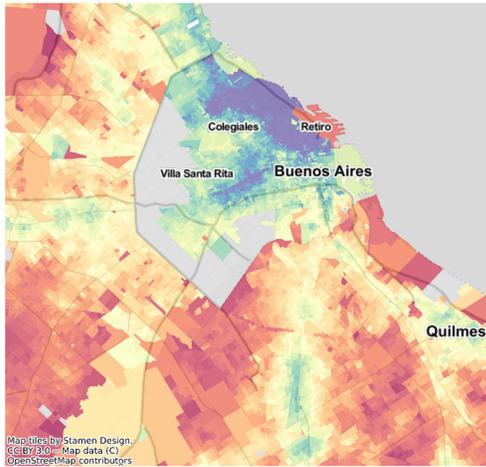
La Tabla 3.1 se presenta un correlograma entre el IE y otros estimadores del bienestar. Se observa que, en términos generales, la correlación es alta, particularmente entre el ingreso estimado y los indicadores de bienestar que, a su vez, no se correlacionan entre sí. Lo anterior permite interpretar al IE como una variable latente —el bienestar— que expresan los restantes indicadores.

	IEPAG	ICV	Pm^2	%NBI	MRI
IEPAG	1.00	0.82	0.79	0.74	0.32
ICV	0.82	1.00	0.56	0.71	0.24
Pm^2	0.79	0.56	1.00	0.36	0.25
%NBI	0.74	0.71	0.36	1.00	0.18
MRI	0.32	0.24	0.25	0.18	1.00

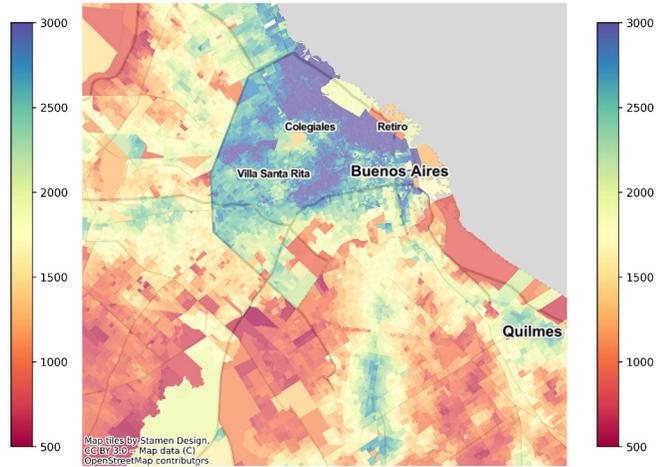
Cuadro 3.1: Tabla de correlaciones de Spearman entre variables.

Nota: IEPAG: Ingreso Estimado por Área Geográfica, ICV: Índice de Calidad de Vida, Pm^2 : Precio Promedio del Metro Cuadrado, %NBI: Porcentaje de Hogares sin Necesidades Básicas Insatisfechas, MRI: Meta Relative Income Indicator.

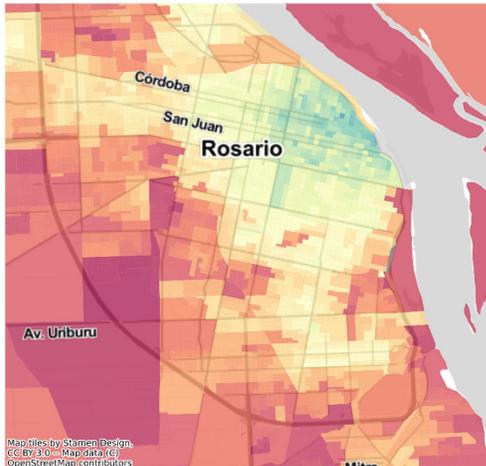
Figura 5: Ingreso estimado por Área Geográfica, evolución de la distribución absoluta



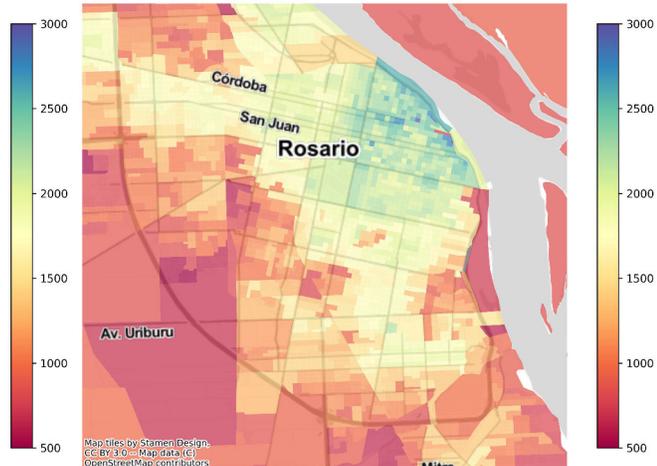
(a) AMBA, 2001



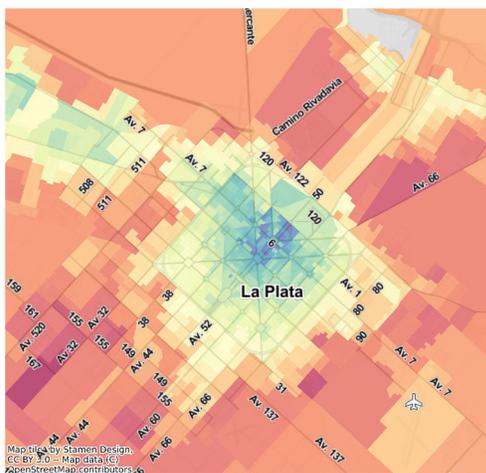
(b) AMBA, 2010



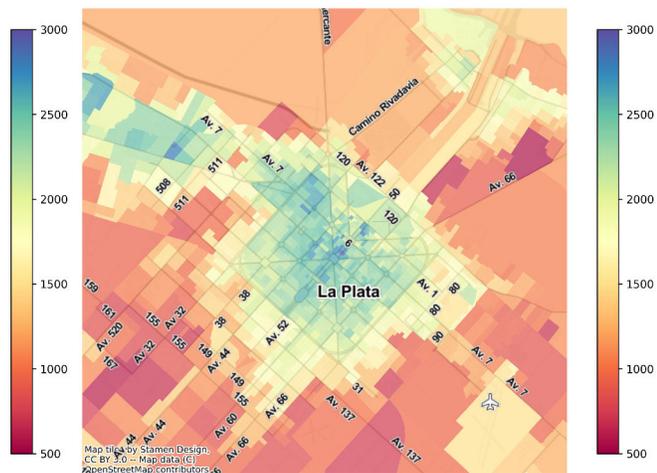
(c) Rosario, 2001



(d) Rosario, 2010



(e) La Plata, 2001



(f) La Plata, 2010

4. Relación ingreso y resultados electorales

4.1. Voto económico: sociotrópico vs egotrópico

En el estudio de los factores que determinan el comportamiento electoral, la teoría del **voto económico** se enfoca en analizar la influencia de las percepciones económicas de los votantes en sus decisiones electorales. Existe un amplio consenso en la literatura de ciencia política (Lewis-Beck, 1986; Duch, 2001; Singer and Carlin, 2013) que, en sociedades democráticas, una mejora en el bienestar económico generalmente favorece la reelección de los oficialismos. En el debate sobre el voto económico, lo esencial es determinar si predomina la economía general o el bienestar económico personal, una distinción que da origen a dos enfoques: el sociotrópico y el egotrópico. Mientras que el **voto sociotrópico** se basa en la percepción del bienestar económico general de la sociedad, el **voto egotrópico** se enfoca en las circunstancias económicas personales del votante.

El voto sociotrópico se refiere a la tendencia de los votantes a basar sus decisiones electorales en la evaluación de la economía en un contexto más amplio que su situación personal. Este enfoque sugiere que las personas votan considerando el bienestar económico general de la nación o la comunidad. Por ejemplo, en el análisis de Anderson (2006), se observa cómo la gobernanza multinivel y la claridad de responsabilidad afectan la percepción de la economía nacional y, por ende, el apoyo al partido en el poder. Similarmente, Singer and Carlin (2013) encuentran que en América Latina, el contexto electoral y el nivel de desarrollo influyen en la naturaleza del voto económico, destacando la importancia de la percepción de la economía nacional sobre las finanzas personales.

Por otro lado, el voto egotrópico se centra en cómo las experiencias económicas personales afectan las preferencias electorales. Aquí, la decisión de voto se basa en cómo la situación económica del individuo ha cambiado, por ejemplo, en términos de empleo o ingresos. Un estudio relevante de Gomez and Wilson (2001) en América Latina encontró que las dificultades económicas personales estaban significativamente relacionadas con un menor apoyo a los partidos gobernantes. De manera similar, Duch (2001) destaca cómo las percepciones personales del rendimiento económico influyen en las decisiones electorales en nuevas democracias.

Los estudios revisados abordan el voto económico con diversas metodologías. Por ejemplo, Pacek and Radcliff (1995) usan análisis de series temporales para examinar la relación entre el desarrollo del estado de bienestar y el voto económico en democracias industriales. Por otro lado, Lewis-Beck (1986) enfocan su análisis en evaluaciones retrospectivas, prospectivas y afectivas del desempeño económico del gobierno y su impacto en la decisión de voto. Estas diferencias metodológicas subrayan cómo se interpretan las decisiones de voto bajo distintas perspectivas, resaltando la complejidad y la multidimensionalidad del voto económico.

Distribución del voto según el ingreso estimado

En este análisis, voy a estudiar cómo la distribución del ingreso se relaciona con la orientación política en Argentina. Específicamente, me pregunto si los niveles de ingresos influyen en las preferencias políticas de los votantes.

La figura 6 presenta seis filas correspondientes a cada elección y tres columnas para los partidos políticos. Cada celda muestra un gráfico de dispersión de circuitos electorales donde el eje horizontal representa el ingreso estimado y el eje vertical el porcentaje de votos. En la primera columna (en celeste), se observa que los partidos peronistas tienen una pendiente negativa, indicando que obtuvieron más votos en circuitos con menores ingresos. En la columna del medio, la pendiente es positiva, mostrando que estos partidos recibieron más votos en circuitos con mayores ingresos. En la tercera columna, la tendencia es transversal, excepto para Carlos Menem (Frente Leal, 2003) y Javier Milei (La Libertad Avanza, 2023), donde también se observa una pendiente negativa. Si bien esta figura no permite ver variaciones en el ingreso que den cuenta del voto económico, se observa que en Argentina existe una clara relación entre el nivel de ingresos y la orientación política.

La figura 7 muestra la transferencia de votos en las elecciones presidenciales generales desde 2003 hasta 2023, incluyendo el traspaso de votos entre las generales de 2023 y el balotaje. Se observa que algunos partidos acumulan votos en ciertos períodos y luego los pierden. ¿Está esta fluctuación influenciada por la performance económica del oficialismo? ¿En qué medida? El Anexo A profundiza en el análisis de estas transferencias y detalla la metodología de inferencia ecológica utilizada.

Figura 6: Relación ingreso estimado y votos por partido político

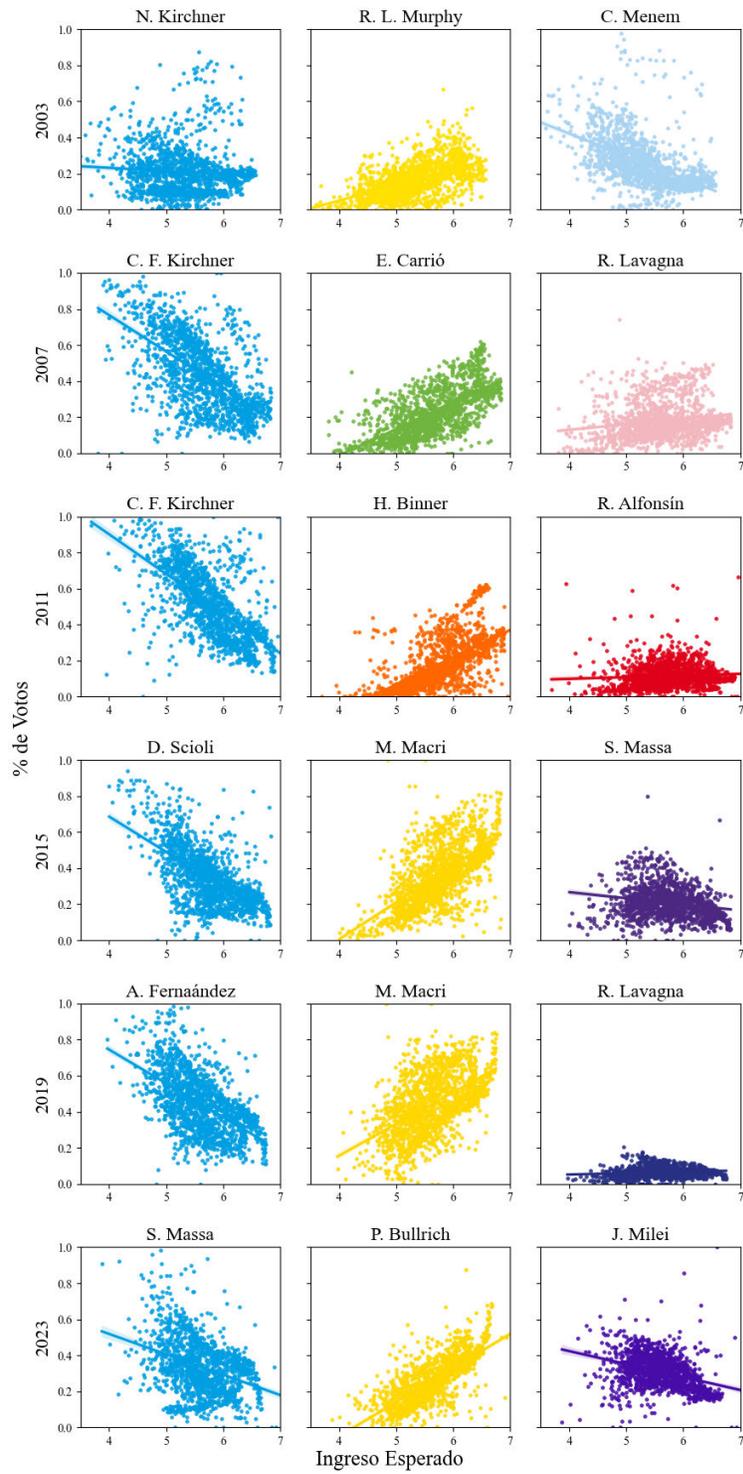
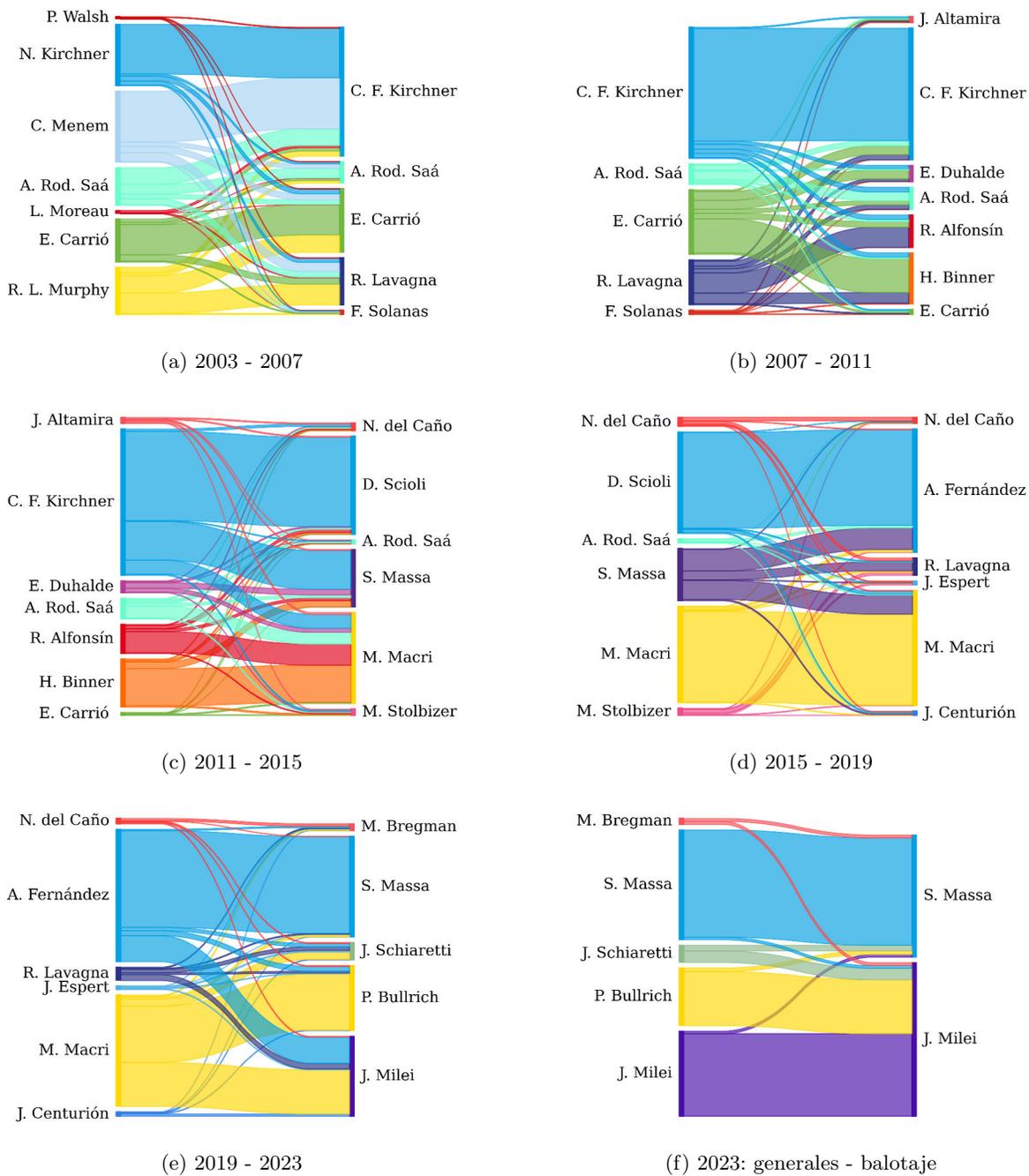


Figura 7: Transferencia de votos



4.2. Metodología: regresión con efectos fijos

Para analizar los determinantes de los votos al oficialismo teniendo en cuenta los efectos fijos por circuito electoral, consideramos la siguiente especificación de modelo:

$$Votos_{t,i} = \beta_1 IngresoEstimado_{t,i} + \beta_2 Controles_t + \alpha_i + \epsilon_{t,i} \quad (4)$$

donde:

$Votos_{t,i}$: Votos al oficialismo en el circuito electoral i en el tiempo t .

$IngresoEstimado_{t,i}$: Ingreso estimado en el circuito electoral i en el tiempo t .

$Controles_t$: es el vector de variables de control que varían solo en el tiempo t .

α_i : Efectos fijos por circuito electoral, capturando características invariables en el tiempo.

β_1 : Coeficiente para el ingreso estimado, indicando cómo este afecta los votos al oficialismo.

β_2 : Vector de coeficientes para las variables de control (PBI e IPC, nacionales en logaritmos).

$\epsilon_{t,i}$: es el término de error estocástico.

Este modelo permite analizar la influencia del ingreso estimado y otras variables de control en los votos al oficialismo.

4.3. Predominancia del voto sociotrópico frente al egotrópico

Cuadro 4.1: Resultados de la Regresión sobre el Porcentaje de Votos al Oficialismo

	(1)	(2)	(3)
Ingreso Esperado (en log)	0.197*** (0.0342)	-0.0407 (0.0486)	-0.0604 (0.0401)
año = 2007		0.241*** (0.0388)	
año = 2011		0.321*** (0.0367)	
año = 2015		0.154*** (0.0337)	
año = 2019		0.245*** (0.0474)	
año = 2023		0.136*** (0.0343)	
PBI (en log)			0.586*** (0.0769)
IPC (en log)			-0.0332*** (0.00792)
Constante	-0.727*** (0.192)	0.426 (0.257)	-6.873*** (0.860)
Observaciones	9,082	9,082	9,082
R-cuadrado	0.067	0.326	0.233
Circuitos electorales	1,594	1,594	1,594

Errores estándar robustos en paréntesis

*** p<0.01, ** p<0.05, * p<0.1

Los resultados del cuadro 4.1 muestran que el coeficiente para la variable Ingreso Estimado a nivel de circuito electoral es positivo y significativo en la primera regresión, lo que indica que, ceteris paribus, un aumento en el ingreso de un circuito electoral se asocia con un incremento en los votos al oficialismo. Sin embargo, en la segunda regresión, cuando se controla por elección (considerando que estos controles agregan variables sociales tanto económicas como no económicas), el coeficiente del ingreso estimado pierde significatividad. De manera similar, en la tercera regresión, donde se controla por las principales variables económicas (Producto Bruto Interno e Índice de Precios al Consumidor), el coeficiente del ingreso estimado para cada circuito electoral también pierde significatividad. El coeficiente del PBI es positivo y significativo, lo que sugiere que a mayor crecimiento económico, el oficialismo obtiene más votos. Un aumento del 10 % en el PBI durante los cuatro años de gestión se traduce en un incremento de 5,8 puntos porcentuales en los votos del oficialismo. Por el contrario, el coeficiente del IPC es negativo y significativo, indicando que una mayor inflación durante el período de gobierno reduce el apoyo al oficialismo. En concreto, una inflación acumulada del 100 % en todo el período implica una caída de 3,3 puntos porcentuales en los votos del oficialismo.

Este cuadro 4.1 evidencia que, para el periodo analizado, las variables generales del país predominan

sobre el ingreso por circuito electoral, apoyando la corriente del voto sociotrópico como uno de los principales determinantes del voto. Esto sugiere que los votantes tienden a basar sus decisiones en la percepción de la economía nacional más que en sus circunstancias económicas personales, predominando así el voto sociotrópico sobre el egotrópico.

5. Conclusiones

Este trabajo ha permitido la construcción de un índice de ingreso estimado con alta desagregación geográfica, utilizando la técnica de *small area estimation* (SAE) para el período 2003-2023. Este índice, calculado a nivel de radio censal, ofrece una representación detallada de los ingresos promedio a lo largo del tiempo. A partir de este mapa de ingreso, se pueden desarrollar diversas investigaciones tanto para mejorar la precisión del índice como para estudiar su relación con otros indicadores sociales y económicos. En este trabajo, específicamente, se analizó la correlación entre los niveles de ingreso estimado y los resultados electorales nacionales, lo que permitió explorar la relación entre las condiciones económicas y el comportamiento electoral.

El análisis de los resultados electorales arrojó evidencias claras sobre la relevancia del voto sociotrópico frente al egotrópico. Se observó que los votantes en Argentina tienden a basar sus decisiones en la evaluación del desempeño económico nacional, particularmente en lo referente al crecimiento del Producto Bruto Interno (PBI) y los niveles de inflación (IPC). A través de la implementación de un modelo de regresión con efectos fijos por circuito electoral, se concluyó que las percepciones económicas generales del país prevalecen sobre las condiciones económicas personales al momento de definir el voto.

Es importante resaltar que el período analizado (2003-2023) es inmediatamente posterior a una de las crisis económicas (2001) más significativas desde el retorno de la democracia en Argentina. En este contexto, la búsqueda de una gestión capaz de estabilizar la macroeconomía adquiere una relevancia especial lo que podría explicar la relevancia del voto sociotrópico frente al egotrópico.

Finalmente, el estudio plantea oportunidades para mejorar y ampliar la relación entre los ingresos y los resultados electorales. Una vía futura de investigación podría centrarse en las elecciones legislativas, en las cuales las dinámicas locales y regionales podrían ser más significativas que en las elecciones presidenciales. Asimismo, se podrían incluir nuevas especificaciones en los modelos que incorporen otras variables económicas, como el desempleo o las tasas de pobreza, y explorar su efecto en distintos tipos de elecciones y niveles de gobierno. Esto permitirá una comprensión más completa del impacto de las condiciones económicas sobre el comportamiento electoral, tanto a nivel local como nacional.

A. Comportamiento electoral intertemporal

En 2007, luego de cuatro años de crecimiento económico, ¿Cristina Fernández de Kirchner logró captar a los votantes que en la elección anterior habían apoyado a Carlos Menem? ¿Pudo Cristina, en 2011, retener ese respaldo obtenido en 2007? En 2015, ¿de dónde provinieron los votantes que apoyaron a Sergio Massa, y qué impacto tuvo su candidatura en la distribución del voto kirchnerista? Con el empeoramiento económico durante la gestión de Mauricio Macri, ¿hacia dónde migraron los votantes de Massa en 2019? En las elecciones generales de 2023, ¿de qué partidos llegaron los votantes que respaldaron a Javier Milei, y qué factores influyeron en esa transferencia de apoyo? ¿Logró Milei retener, en el balotaje, una porción significativa de los votantes de Juntos por el Cambio, tal como se observó en las generales?

Uno de los desafíos fundamentales en la ciencia política es comprender los determinantes del voto. ¿Por qué algunos votantes cambian su apoyo de un partido a otro en cada elección? Tradicionalmente, se ha atribuido este comportamiento a factores como la ideología, el liderazgo o las propuestas de campaña. Sin embargo, el contexto económico ha demostrado ser un factor determinante en las decisiones de los votantes. En muchas democracias, las fluctuaciones en los niveles de ingreso y bienestar económico de los ciudadanos influyen directamente en la continuidad de los gobiernos. En este sentido, estudiar la dinámica de transferencia de votos entre partidos a lo largo de las elecciones proporciona una base para analizar los factores que inciden en la permanencia o alternancia de los partidos en el poder.

En esta sección, utilizaremos la inferencia ecológica para analizar de manera descriptiva cómo se distribuyen los votos entre partidos en elecciones consecutivas, sin considerar aún factores económicos. La observación de estas transiciones nos ayudará a entender mejor los procesos de cambio y permanencia política, generando una base sólida para, en la próxima sección, analizar la influencia de los factores económicos, como el voto *sociotrópico* y *egotrópico*, en las decisiones electorales.

A.1. Contexto metodológico de inferencia ecológica

La inferencia ecológica es un método estadístico utilizado para inferir comportamientos individuales a partir de información agregada. Este enfoque es particularmente relevante en la ciencia política, donde los datos individuales generalmente no están disponibles debido a la naturaleza de los registros electorales.

El concepto de falacia ecológica fue formalizado por [Robinson and Comstock](#) en 1950, quien identificó que las correlaciones a nivel agregado no siempre son representativas de las relaciones a nivel individual, un desafío clave en la inferencia ecológica. Posteriormente, [Duncan and Davis \(1953\)](#) introdujo el método de límites determinísticos, enfocándose en establecer rangos inferiores y superiores para las cantidades de interés a partir de múltiples tablas. [Goodman \(1953, 1959\)](#) propusieron las regresiones ecológicas, aplicando técnicas de regresión a datos agregados. Sin embargo, ambos enfoques tenían

limitaciones significativas: los límites determinísticos de [Duncan and Davis](#) a menudo resultaban en intervalos demasiado amplios para ser informativos, mientras que las regresiones de [Goodman](#) dependían de suposiciones como la homogeneidad en las transiciones de votantes.

El trabajo de [King \(1997\)](#) representa un avance significativo en la inferencia ecológica al combinar las principales ideas de los enfoques previos. Su modelo se basa en coeficientes aleatorios con parámetros que varían y se distribuyen según una distribución normal bivariada truncada. Esta aproximación contrasta con la regresión ecológica tradicional al abordar la heterocedasticidad y proporcionar una forma más flexible de ajustar las estimaciones a un modelo lineal específico. Esta contribución sentó las bases para el desarrollo de modelos bayesianos en la inferencia ecológica, como el modelo propuesto por [King et al. \(1999\)](#), que utiliza un enfoque jerárquico bayesiano y ha sido adaptado para abordar casos más generales, como el RxC. Sin embargo, estos modelos bayesianos presentan desafíos computacionales significativos debido a la necesidad de realizar muestreos en la distribución posterior mediante métodos de Monte Carlo por cadenas de Markov (MCMC).

A raíz del trabajo de [King \(1999\)](#), se han propuesto otros modelos bayesianos y métodos alternativos para abordar la inferencia ecológica, como los modelos jerárquicos bayesianos de [Rosen et al. \(2001\)](#), que también requieren muestreos MCMC y comparten desafíos computacionales similares. Diversos investigadores han explorado enfoques alternativos, como el enfoque de dividir el problema RxC en múltiples problemas 2x2 secuenciales. Por ejemplo, ([Andreadis and Chadjipadelis, 2009](#)) se basa en un enfoque recursivo (de matrices 2x2) para estimar las transiciones de votantes en sistemas multipartidistas (RxC). Específicamente, selecciona pares de partidos políticos con fuertes correlaciones positivas en su desempeño electoral para estimar las tasas de transición de votantes, utilizando ciclos iterativos hasta cubrir una proporción deseada del electorado. Además, recientemente se ha presentado un enfoque basado en programación lineal por [Pavía and Romero \(2021\)](#), que introduce dos nuevos algoritmos (minimización del error cuadrático y minimización del error absoluto) mejorar la precisión de las estimaciones.

La siguiente tabla ilustra el problema de inferencia ecológica en un escenario general con R grupos en la primera elección y C grupos en la segunda. Los datos observados consisten en n_{ri} , el número de votantes en el grupo $r = 1, 2, \dots, R$ en la unidad i en la primera elección, y v_{ci} , el número de votantes en el grupo $c = 1, \dots, C$ en la unidad i en la segunda elección. Las variables desconocidas son p_{rci} para cada r, c, i , que indican la fracción de personas que votan por el partido r que también votan por el partido c en la unidad i . Cada entrada en el centro de la tabla en la fila r y columna c indica el número de votantes del grupo en la fila r que pertenecen al grupo en la columna c .

Para este análisis, se utilizó la biblioteca PyEI de Python, desarrollada por [Knudson et al. \(2021\)](#), que permite estimar el modelo jerárquico bayesiano RxC de [Rosen et al. \(2001\)](#). PyEI facilita la inferencia de patrones de votación a partir de datos agregados, permitiendo un análisis detallado de las transiciones de voto en diferentes grupos demográficos.

Una aplicación de la inferencia ecológica es estimar las matrices de transición de votantes, que nos

Cuadro A.1: Inferencia Ecológica - Caso RxC

1 ^{era} Elección	2 ^{da} Elección			
	Partido 1	Partido 2	...	Partido C
Partido 1	$n_{1i}p_{11i}$	$n_{1i}p_{12i}$...	$n_{1i}p_{1Ci}$
Partido 2	$n_{2i}p_{21i}$	$n_{2i}p_{22i}$...	$n_{2i}p_{2Ci}$
⋮	⋮	⋮	⋮	⋮
Partido R	$n_{Ri}p_{R1i}$	$n_{Ri}p_{R2i}$...	$n_{Ri}p_{RCi}$

indican cómo grupos de votantes cambian sus preferencias entre diferentes candidatos o opciones en elecciones sucesivas. En Argentina, esto podría implicar analizar cómo se movieron los votos entre, por ejemplo, las elecciones presidenciales de 2015 y 2019, identificando tendencias como el traslado de votantes de Daniel Scioli (Frente para la Victoria) a Alberto Fernández (Frente de Todos). Una matriz de transición revelaría la proporción de votantes de Scioli en 2015 que votaron por Fernández en 2019. Dado que los datos electorales se publican agregados por circuito electoral, las matrices de transición deben estimarse usando solo estos datos agregados.

Fuente de datos electorales Los resultados electorales se recopilaron a de información provista por la Dirección Nacional Electoral (DINE) facilitados por el proyecto *Política Argentina* de Juan Pablo Ruiz Nicolini. Con el objetivo de estudiar la relación entre el mapa de ingresos por área geográfica y los resultados electorales se elaboró un panel con datos de resultados electorales nacionales en Argentina desde 2003 al 2023. El panel cuanta con la mayor desagregación geográfica que el sistema electoral argentino provee que son circuitos electorales. Dentro de cada circuito electoral se distribuye en mesas al electorado que tiene domicilio dentro de ese circuito electoral. Dado que la distribución en mesas no se determina por un criterio de cercanía a la vivienda no es posible tener una mayor desagregación geográfica que dichos circuitos electorales.

A.2. Inferencia ecológica para transferencia de votos

La figura 7 muestra la transferencia de votos entre las elecciones presidenciales generales desde 2003 hasta 2023. También se presenta la transferencia de votos entre las generales de 2023 y el balotaje. Se observa que Cristina Fernández de Kirchner, en 2007, tras cuatro años de crecimiento económico durante la presidencia de Néstor Kirchner, atrae casi todos los votantes que lo eligieron en 2003. Además, capta una gran cantidad de votos de quienes optaron por Carlos Menem y Rodríguez Saá (incluso cuando este último también fue candidato en 2007).

Elisa Carrió, quien fue candidata en las primeras tres elecciones del periodo analizado, logró conservar sus votos entre 2003 y 2007, pero perdió casi todos sus votantes entre 2007 y 2011, quienes se inclinaron por Hermes Binner. Entre 2007 y 2011, con cuatro años de crecimiento económico, el Frente para la

Victoria mantiene a sus votantes y atrae votos de otros partidos, convirtiendo a Cristina Fernández en la presidenta más votada desde el regreso de la democracia, con un 54.11 % de los votos.

Para 2015, como se muestra en la figura (c), con un cierto estancamiento del crecimiento económico, Daniel Scioli (candidato del Frente para la Victoria) no logra retener a todos los votantes que habían elegido a Cristina en 2011 ni atraer nuevos votantes, perdiendo parte del electorado a manos de Sergio Massa. Con la profundización de la crisis económica entre 2015 y 2019 y la polarización del escenario político, Alberto Fernández atrae una gran cantidad de votantes de Massa, obteniendo una diferencia de 16 puntos porcentuales en las primarias y 8 en las generales.

Entre 2019 y 2023, Sergio Massa no consigue retener los votos de Alberto Fernández en las elecciones generales, que se transfieren a Javier Milei, lo que lleva a un balotaje. En la última figura, se observa cómo todos los votantes de Patricia Bullrich y gran parte de los votantes de Schiaretti optaron por Javier Milei.

Referencias

- Anderson, C. D. (2006), 'Economic voting and multilevel governance: A comparative individual-level analysis', *American Journal of Political Science* **50**(2), 449–463.
- Andreadis, I. and Chadjipadelis, T. (2009), 'A method for the estimation of voter transition rates', *Journal of Elections, Public Opinion and Parties* **19**(2), 203–218.
- Cuong, N. V., Truong, T. N. and van der Weide, R. (2010), 'Poverty and inequality maps in rural vietnam: An application of small area estimation', *Asian Economic Journal* **24**, 355–390.
- Department of Census and Statistics and World Bank (2015), 'The spatial distribution of poverty in sri lanka department of census and statistics-sri lanka poverty global practice, world bank group'.
URL: http://www.statistics.gov.lk/national_accounts/Press%20Release/2014ANNUAL.pdf
- Duch, R. M. (2001), 'A developmental model of heterogeneous economic voting in new democracies', *The American Political Science Review* **95**(4), 895–910.
- Duncan, O. D. and Davis, B. (1953), 'An alternative to ecological correlation.', *American sociological review* .
- Ebers, C., Lanjouw, P. and Leite, P. G. (2008), *Brazil Within Brazil: Testing The Poverty Map Methodology In Minas Gerais*, The World Bank.
- Elbers, C., Lanjouw, J. O. and Lanjouw, P. (2003), 'Micro-level estimation of poverty and inequality', *Econometrica* **71**, 355–364.
- Gasparini, L. C., Gluzmann, P. A. and Tornarolli, L. (2022), 'Caracterización de la población vulnerable: una propuesta con estimaciones para argentina', *Económica* **68**.
- Gluzmann, P. and Panigo, D. (2015), 'Global search regression: A new automatic model-selection technique for cross-section, time-series, and panel-data regressions', *The Stata Journal* **15**(2), 325–349.
- Gomez, B. T. and Wilson, J. M. (2001), 'Political sophistication and economic voting in the american electorate: A theory of heterogeneous attribution', *American Journal of Political Science* **45**(4), 899–914.
- Goodman, L. A. (1953), 'Ecological regressions and behavior of individuals', *American sociological review* **18**(6), 663.
- Goodman, L. A. (1959), 'Some alternatives to ecological correlation', *American Journal of Sociology* **64**(6), 610–625.
- King, G. (1997), 'A solution to the ecological inference problem: Reconstructing individual behavior from aggregate data', *Princeton University Press* .
- King, G. (1999), 'The future of ecological inference research: a comment on freedman et al.', *Journal of the American Statistical Association* **94**(445), 352–355.

- King, G., Rosen, O. and Tanner, M. A. (1999), ‘Binomial-beta hierarchical models for ecological inference’, *Sociological Methods & Research* **28**(1), 61–90.
- Knudson, K. C., Schoenbach, G. and Becker, A. (2021), ‘Pyei: A python package for ecological inference’, *Journal of Open Source Software* **6**(64), 3397.
URL: <https://doi.org/10.21105/joss.03397>
- Lewis-Beck, M. S. (1986), ‘Comparative economic voting: Britain, france, germany, italy’, *American Journal of Political Science* **30**(2), 315–346.
- Molina, I. and Rao, J. N. K. (2010), ‘Small area estimation of poverty indicators’, *Canadian Journal of Statistics* **38**, 369–385.
- Pacek, A. C. and Radcliff, B. (1995), ‘Economic voting and the welfare state: A cross-national analysis’, *The Journal of Politics* **57**(1), 44–61.
- Panigo, D., Glüzmann, P., Mocskos, E., Ungaro, A. M., Mari, V. and Monzón, N. (2020), ‘Globalsearchregression. jl: building bridges between machine learning and econometrics in fat-data scenarios’, **2**(13), 53.
- Panigo, D., Ungaro, A. M., Monzón, N., Tcach, A., Menduiña, J., Gluzmann, P., Barcellandi, A. R. and Panigo, N. (2023), ‘Accelerating economic research with julia’, *presentado en JuliaCon 2023* .
- Pavía, J. M. and Romero, R. (2021), ‘Improving estimates accuracy of voter transitions. two new algorithms for ecological inference based on linear programming’, *Sociological Methods & Research* p. 00491241221092725.
- Robinson, H. and Comstock, R. E. (1950), Genotypic and phenotypic correlations in corn and their implications in selection, Technical report, North Carolina State University. Dept. of Statistics.
- Rosen, O., Jiang, W., King, G. and Tanner, M. A. (2001), ‘Bayesian and frequentist inference for ecological inference: The $r \times c$ case’, *Statistica Neerlandica* **55**(2), 134–156.
- Singer, M. M. and Carlin, R. E. (2013), ‘Context counts: The election cycle, development, and the nature of economic voting’, *The Journal of Politics* .
- Tibshirani, R. (1996), ‘Regression shrinkage and selection via the lasso’, *Journal of the Royal Statistical Society. Series B (Methodological)* **58**(1), 267–288.
URL: <http://www.jstor.org/stable/2346178>
- Tornarolli, L. (2018), Series comparables de indigencia y pobreza: una propuesta metodológica, Technical report, Documento de Trabajo.