



Estimación de la pobreza en Chile a nivel de comuna

SEMINARIO SOBRE DESAGREGACIONES DE DATOS
MEDIANTE TÉCNICAS DE ESTIMACIÓN EN ÁREAS PEQUEÑAS

MONTEVIDEO



Subsecretaría de Evaluación Social
Ministerio de Desarrollo Social y Familia
Noviembre 2022

Observatorio Social: Objetivos



General:

Caracterizar y analizar la realidad social, a través de la recopilación, procesamiento, análisis y difusión de datos, de modo de identificar las necesidades sociales de la población para el oportuno diseño de políticas sociales.

Específicos:

1. Analizar la realidad social nacional y regional, de modo de detectar las necesidades sociales de la población para el oportuno diseño de políticas sociales.
2. Estudiar y proponer metodologías para la recolección y procesamiento de información de encuestas sociales.
3. Sistematizar y analizar registros de datos, información y estadísticas que permitan caracterizar la población nacional, regional y comunal y sus necesidades sociales.



Proyectos Destacados



Encuesta de Caracterización Nacional (1987-2022)



Encuesta Nacional de Discapacidad y Dependencia



Nueva encuesta Bifásica de CASEN, medición Bienestar Social



Actualización mensual de CBA y Línea de la Pobreza en base a IPC



4 rondas encuesta. Busca medir consecuencias de la pandemia



Estimación de pobreza comunal, metodología SAE



Encuesta Nacional sobre Actividades de niños, niñas y adolescentes



Sistema de Indicadores en base a registros administrativos



Encuesta Longitudinal de Primera Infancia



Portal de visualización de datos. Casen, RRAA, BIPS, Donaciones Sociales, RSH



Rol de Coordinación Sub Comisión Social del Grupo Técnico de Indicadores

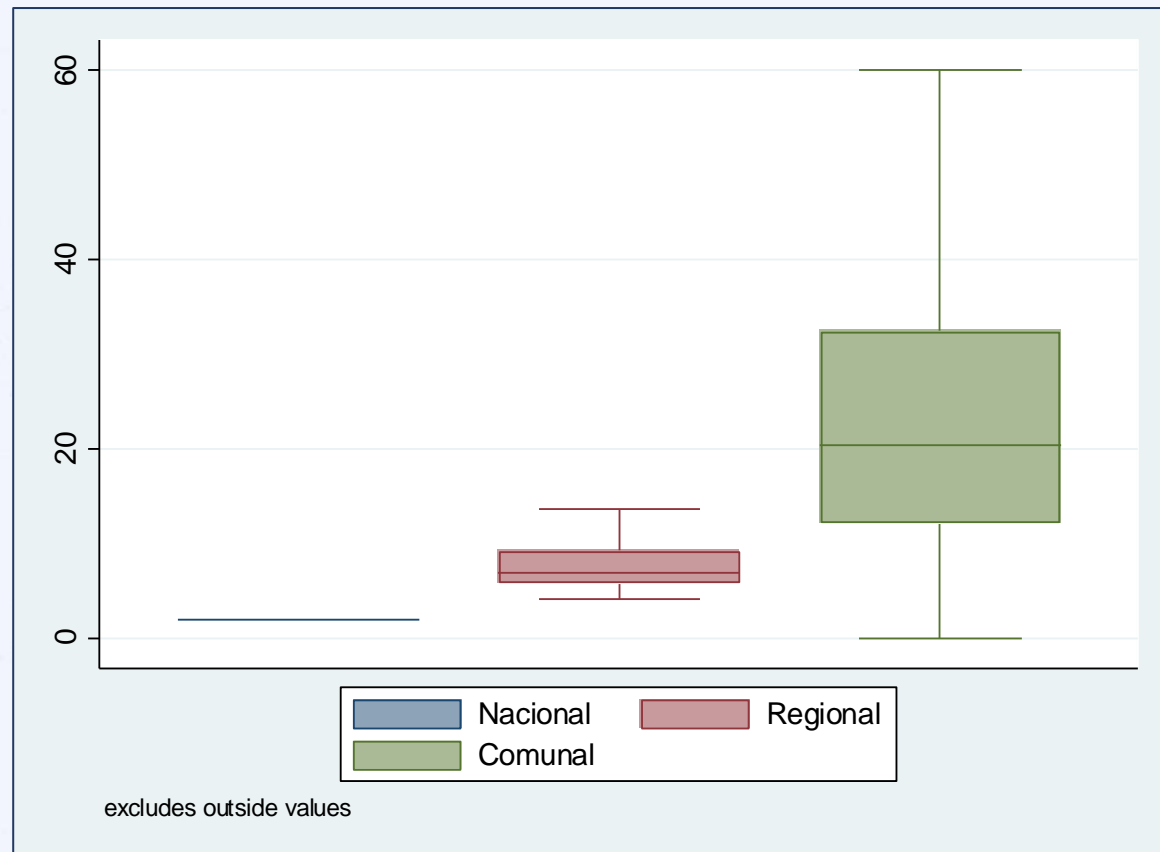


MOTIVACIÓN



- La encuesta Casen es la fuente oficial para la estimación de las cifras de pobreza a nivel nacional y regional. Esta encuesta se ha realizado cada 2 o 3 años desde 1987. En promedio se entrevista a cerca de 70.000 hogares en 324 comunas del país.
- La comuna no es un dominio de estudio de la encuesta, que tiene representatividad nacional, urbano/rural y regional, por lo que estimaciones realizadas a dicha escala de desagregación pueden tener niveles no adecuados de precisión.
- Actualmente, las estimaciones de la tasa de pobreza a nivel comunal son utilizadas para la asignación de recursos públicos, por lo que es fundamental su calidad y precisión.

Coeficiente de Variación Pobreza por Ingresos para distintos niveles de desagregación. Casen Pandemia 2020



MOTIVACIÓN



- El año 2010 se inició una consultoría internacional con el Dr. Partha Lahiri y PNUD, para estimar la pobreza por ingresos a nivel comunal. Estas cifras se publican desde Casen 2011.
- El Ministerio de Desarrollo Social aplica la metodología de Fay-Herriot (1979) para sus estimaciones de área pequeña. Esta metodología se eligió por la ventaja comparativa del Ministerio al tener acceso a registros administrativos y ser el encargado de levantar la encuesta Casen.
- Mediante un convenio de colaboración realizado con CEPAL, durante los años 2019 y 2020 se realizó una actualización de la metodología utilizada por el Ministerio, destacando:
 - Indicadores de calidad para evaluar las estimaciones directas de Casen que son utilizadas como insumo en las estadísticas SAE
 - Estimación de Intervalos de confianza utilizando la estimación del ECM mediante Bootstrap
 - Estimación de la varianza de estimaciones directas utilizando la FGV
 - Recomendaciones para el diseño muestral de Casen
- El documento con la revisión de resultados de esta revisión se puede obtener desde el sitio web del Observatorio Social ([link](#))



MODELO DE ESTIMACIÓN



- La estimación SAE pondera dos estimaciones:

Estimación sintética: Utiliza registros administrativos y del Censo de Población para estimar un modelo de regresión que explique la pobreza comunal. Con los parámetros $\hat{\beta}$ se obtiene la estimación sintética.

Estimación directa: Corresponde a la estimación que proviene de Casen. Asume que esta estimación tiene un error de muestreo.



MODELO DE ESTIMACIÓN



- El modelo de áreas de Fay-Herriot (1979) plantea una estimación en 2 niveles:
 - **Nivel 1: Modelo de vínculo (sintética)**
$$Y_d = \mathbf{x}_d' \boldsymbol{\beta} + u_d, \text{ con } u_d \sim^{iid} (0, \sigma_u^2)$$
 - **Nivel 2: Modelo de muestreo (directa)**
$$\hat{Y}_d^{dir} = Y_d + e_d, \text{ con } e_d \sim^{ind} (0, \psi_d)$$
- A partir de estas 2 ecuaciones se especifica el modelo de estimación de área de F-H:
 - $$\hat{\mathbf{F}}_d^{FH} = \hat{\mathbf{Y}}_d \cdot \hat{\mathbf{Y}}_d^{dir} + (1 - \hat{\mathbf{Y}}_d) \cdot \hat{\mathbf{Y}}_d^{syn}, \text{ con } d = 1, \dots, D$$

Donde:

$\hat{\mathbf{Y}}_d^{syn} = \mathbf{x}_d' \hat{\boldsymbol{\beta}}$, corresponde a la estimación sintética usando variables auxiliares

$$\hat{\mathbf{Y}}_d = \widehat{\sigma}_u^2 / (\widehat{\sigma}_u^2 + \hat{\psi}_d)$$

- Los parámetros de interés que requieren ser estimados son: $(\hat{\boldsymbol{\beta}}, \widehat{\sigma}_u^2, \hat{\psi}_d)$



MODELO DE ESTIMACIÓN - MDSF



- Las estimaciones de pobreza por ingresos que realiza el Ministerio involucran los siguientes pasos:
 - Transformación arcoseno de las tasas de pobreza directas
 - Generación de base de datos administrativos para la estimación sintética
 - Suavización de factores de expansión (modelo Potter-1993)
 - Calidad de las estimaciones directas
 - Estimación varianza directa usando Función Generalizada de Varianza ($\hat{\psi}_d$)
 - Selección del modelo sintético ($\hat{\beta}, \hat{\sigma}_u^2$)
 - Transformación inversa de estimaciones y cálculo de IC en base al ECM usando Bootstarp
 - Calibración (*benchmark*) con estimaciones regionales
 - Validación del modelo
- Hasta las estimaciones SAE 2017 el Ministerio utilizó una programación propia que realizaba todos los pasos. Desde 2020 se utiliza el paquete EMDI (Estimating and Mapping Disaggregated Indicators) de R, que permite centrar el trabajo en la calidad de los insumos que entran al modelo.



CALIDAD DE LAS ESTIMACIONES



- Para obtener estimaciones de buena calidad, se requiere utilizar insumos adecuados (Eurostat, 2019):
 - **Información administrativa** disponible, actualizada y de calidad, para los niveles de desagregación de interés. Se debe explorar la capacidad explicativa de estas variables con el fenómeno que se busca estimar.
 - **Estimaciones directas** de calidad, con tamaños muestrales apropiados y un diseño muestral, que entregue estimaciones con cierto nivel de precisión.



CALIDAD DE LAS ESTIMACIONES - RRAA



- Para el modelo sintético de las estimaciones SAE que realiza el Ministerio de Desarrollo Social y Familia, se cuenta con información de registros administrativos que levanta el sector público a nivel de comuna. Además, se dispone de la información del Censo de Población.
- La información de los registros administrativos proviene de distintas fuentes:
 - Directamente de los Ministerios, tales como Ministerio de Salud, de Vivienda, de Educación y del Trabajo
 - Sistemas de información tales como el Registro Social de Hogares (RSH) y del Sistema Nacional de Información Municipal (SINIM)
 - Otras instituciones públicas, tales como el INE, la Superintendencia de Salud, Superintendencia de Pensiones y la Agencia de Calidad de la Educación
- Una vez definido el conjunto de variables disponibles, se analiza el poder explicativo de estas variables ajustando un modelo de regresión para obtener la estimación sintética.
- Un tema importante a destacar es que en cada ejercicio SAE, se ajusta el mejor modelo explicativo, por lo que las variables seleccionadas cambian en cada ejercicio, no permitiendo una comparación en el tiempo. Esto se ha hecho así para priorizar una buena estimación cada año, dado el uso en la asignación de recursos.



Anexo: Covariables Estimación Sintética



	2009	2011	2013	2015	2017	2020
R cuadrado SAE ²⁴	0,80	0,82	0,82	0,84	0,82	0,86
N comunal	322	232	279	234	242	256
Remuneración promedio de los afiliados al AFC y AFP por cobertura (Mintrab)	x				x	
Proporción de población en el 40% más vulnerable según RSH						x
Cobertura del RSH						x
Porcentaje de pobreza histórica (Casen)	x				x	
Porcentaje población rural (Censo 2002)	x					
Tasa asistencia escolar (Sinim)	x					
Cobertura de afiliados a AFC y AFP con remuneración inferior al salario mínimo					x	
Cobertura de afiliados a AFC y AFP respecto del total de ocupados a nivel comunal					x	
Tasa de ocupación formal dependiente						x
Promedio de años de escolaridad (Censo 2017)					x	
Puntaje Simce historia 8vo básico						x
Porcentaje de población perteneciente a pueblo indígena (Censo)		x	x	x	x	x
Porcentaje afiliados Seguro de Cesantía con remuneración imponible menor al salario mínimo (AFC/AFP)		x	x	x		
Porcentaje de población afiliada a Fonasa A o B (Fonasa)		x	x	x		
Porcentaje de población afiliada a Isapre (Isapre)		x	x	x	x	
Diagnóstico nutricional de niños (obesos y sobrepeso)						x
Tasa de hacinamiento medio						x
Tasa de analfabetismo (Censo 2002)	x	x	x	x		

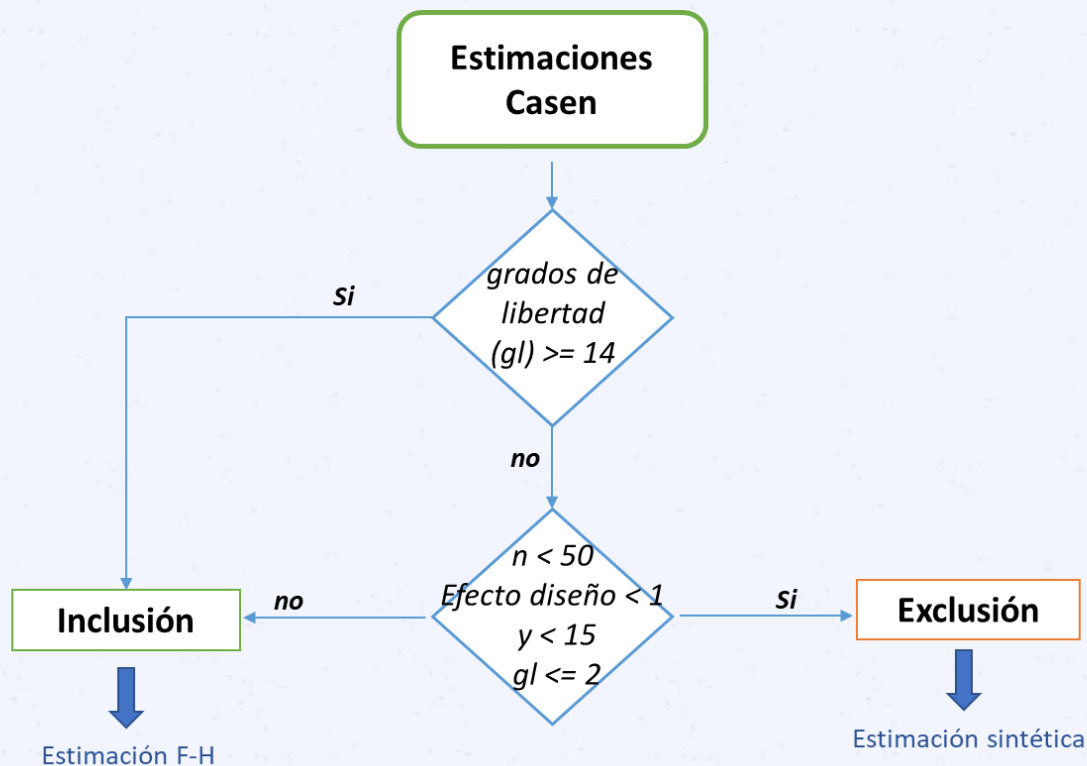
Fuente: Ministerio Desarrollo Social en base a Casen Pandemia 2020.



CALIDAD DE LAS ESTIMACIONES - Directa



- En el caso de las estimaciones directas de Casen, se desarrollaron criterios de calidad que permiten identificar las comunas para las cuales la estimación de Casen es un buen insumo para incorporar a las estimaciones de F-H.
- A las comunas que no pasan los criterios de inclusión o que no tienen muestra Casen se les asigna la estimación sintética.



Número de comunas según los criterios de inclusión

Casen	Incluidas	Excluidas	Total comunas
2011	232	92	324
2013	279	45	324
2015	234	90	324
2017	242	82	324
2020	256	68	324

Fuente: Ministerio Desarrollo Social y Familia

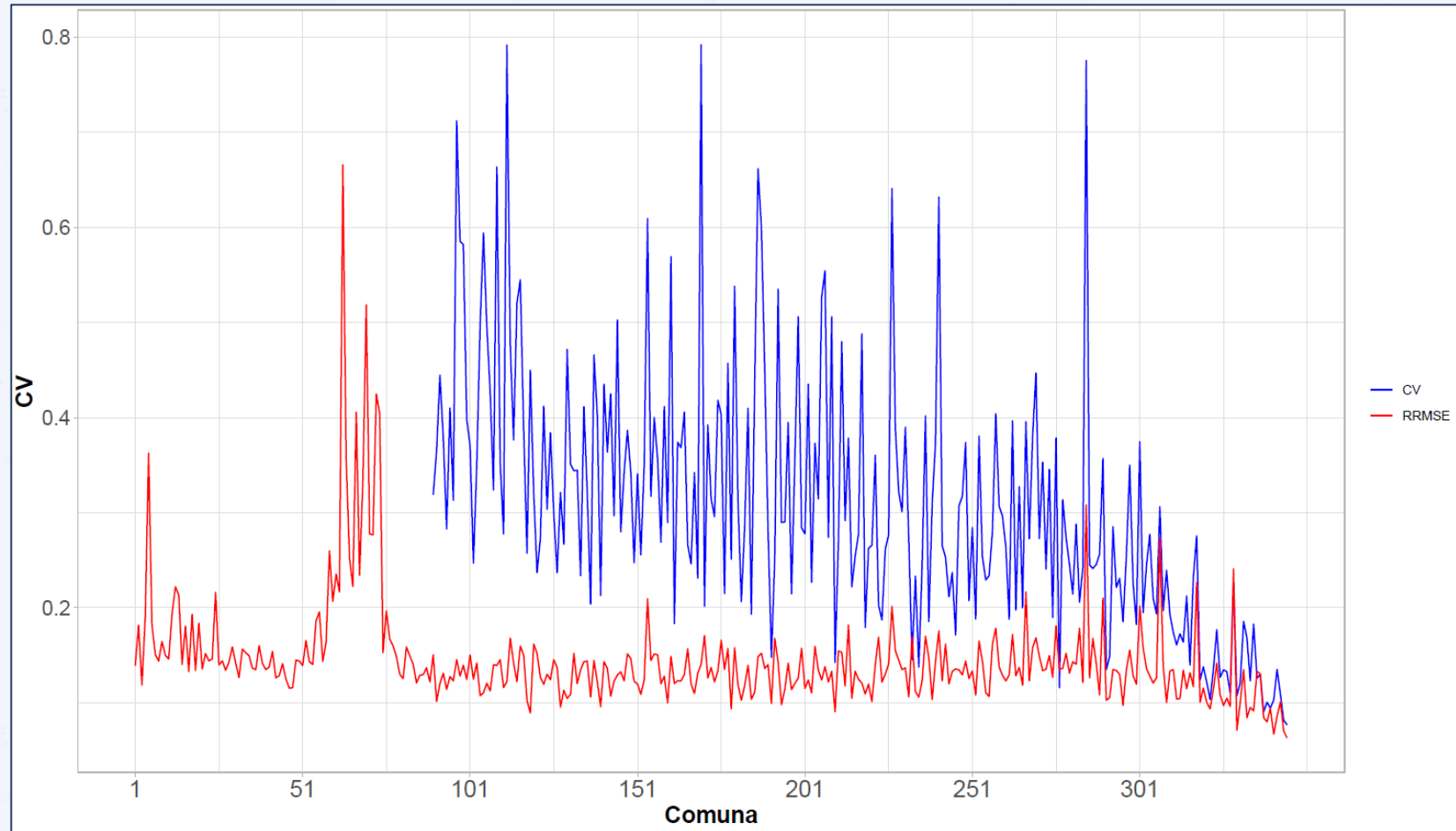


RESULTADOS

<http://observatorio.ministeriodesarrollosocial.gob.cl/pobreza-comunal>

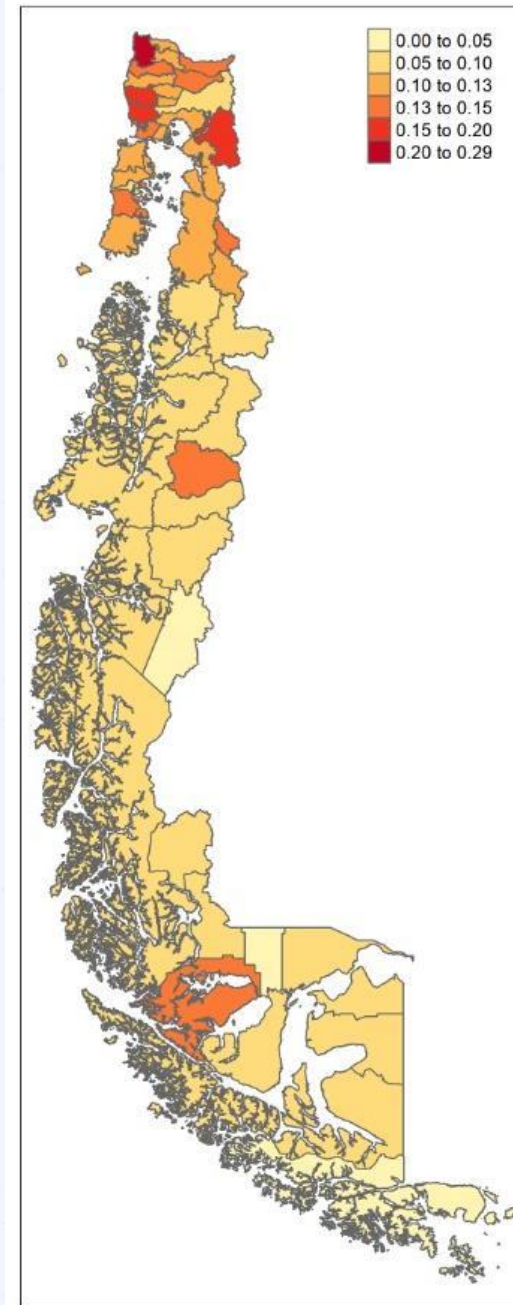
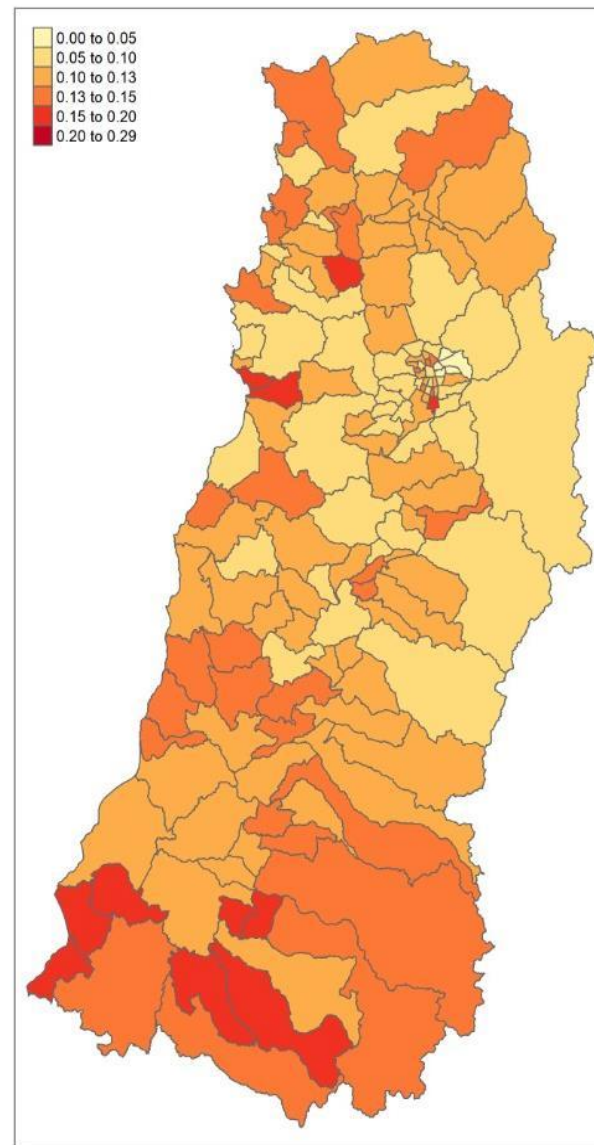
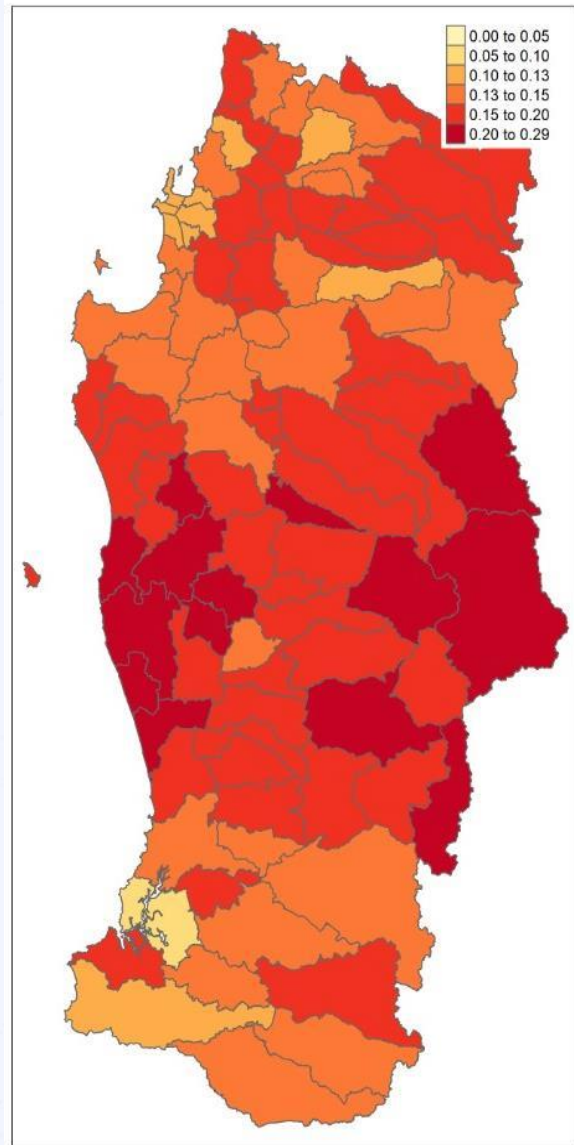
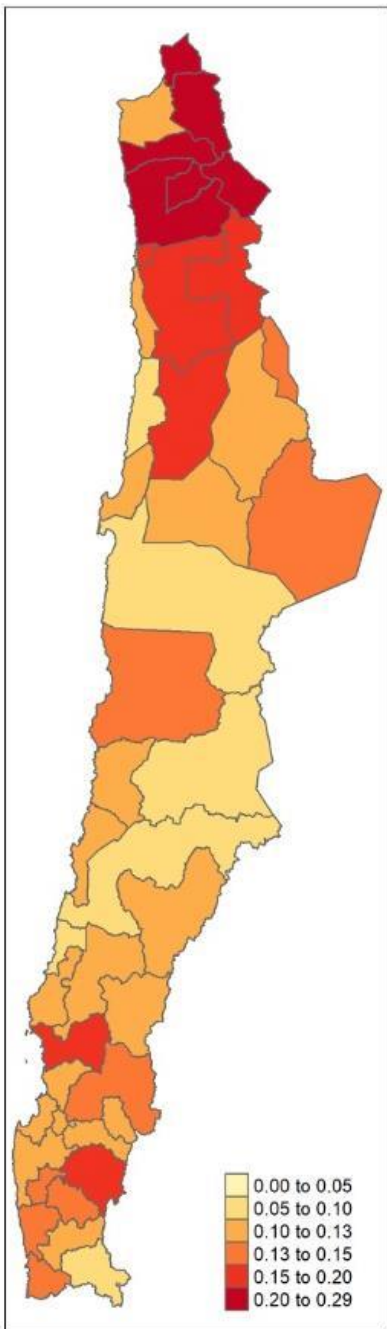


Estimaciones del CV de la estimación directa y SAE



Fuente: Ministerio Desarrollo Social y Familia en base a Casen Pandemia 2020.

RESULTADOS



CIERRE



- Las estimaciones SAE son una herramienta fundamental para el Ministerio de Desarrollo Social y Familia, ya que permite tener insumos para caracterizar la heterogeneidad existente en el país, es relevante en el proceso de descentralización y para la gestión y asignación de recursos de los gobiernos subnacionales.
- La metodología utilizada tiene actualizaciones constantes en la literatura, por lo que es importante una revisión periódica para poder ir incorporando avances y mejores prácticas.
- Una parte importante de la estimación SAE corresponde a los registros administrativos. El Ministerio tiene un proceso continuo de mejora, tanto respecto de nuevos indicadores como de la cobertura de los ya existentes.
- El diseño muestral de Casen también incide en las estimaciones SAE por lo que es importante incorporar las estimaciones SAE como un objetivo de Casen.
- Actualmente el Ministerio realiza estimaciones SAE para la pobreza por ingresos y pobreza multidimensional, sin embargo, a futuro se quiere estudiar la posibilidad de aplicar SAE en otros indicadores a nivel comunal o en grupos de población.





MUCHAS GRACIAS



REFERENCIAS



- Hadam, Sandra, Würz, Nora, & Kreuzmann, Ann-Kristin (2020), Estimating regional unemployment with mobile network data for Functional Urban Areas in Germany. Working paper. Freie Universität.
- Hidiroglou, Michel A. 2019. “Development of a Small Area Estimation System at Statistics Canada.” Survey Methodology 45 (1): 101–26.
- Jiang, J., Lahiri, P. y Wan, S. (2002), A unified jackknife theory for empirical best prediction with Mestimation. The Annals of Statistics. 30 (6), 1782-1810, <https://doi.org/10.1214/aos/1043351257>
- Molina, I. (2019), Desagregación de datos en encuestas de hogares: metodologías de estimación en áreas pequeñas, Series Estudios Estadísticos, No 97, (LC/TS.2018/82/Rev.1), Santiago, Comisión Económica para América Latina y el Caribe, (CEPAL).
- Prasad, N. N. y Rao, J. N. (1990), The estimation of the mean squared error of small-area estimators. Journal of the American Statistical Association, 85, 163–171.
- Rao, J.N.K. y Molina (2015), Small area estimation, Second Ed., Hoboken, NJ: Wiley.

