

Integrazione di dati provenienti da più fonti per il calcolo di indicatori della dimensione del benessere economico per domini non pianificati

Isabella Corazziari¹ Michele D'Alò¹ Maria Giovanna Ranalli²

¹DCME - Direzione centrale per la metododologia e il disegno dei processi statistici MEB - Servizio Disegno dei processi e supporto al sistema dei registri

²Università degli Studi di Perugia

03 Settembre 2020



- 1 Esigenze conoscitive
- 2 Metodologie applicate
- 3 Risultati ottenuti
- 4 Problematiche più rilevanti
- 5 Attività future

Obiettivo

Delineare il processo alla base dell'implementazione di uno strumento statistico utile alla produzione e alla diffusione di stime di indicatori socio-economici ad un livello territoriale molto dettagliato al fine di supportare le politiche urbane che tengano conto delle specifiche caratteristiche locali

- ★ **COERENZA**: Stime diffuse dall'ISTAT.
- ★ **Input**: Roma Capitale (indicatori BES).
- ★ **Sviluppare metodi di stima per piccole aree** per la produzione di indicatori socio-economici dalle principali indagini condotte dall'ISTAT (**FL**, **EUSILC**, **AVQ**).
- ★ **Integrazione micro e macro** dei dati di indagine con dati di fonte amministrativa.
- ★ **Livello territoriale per la sperimentazione**: comuni metropolitani + Province.

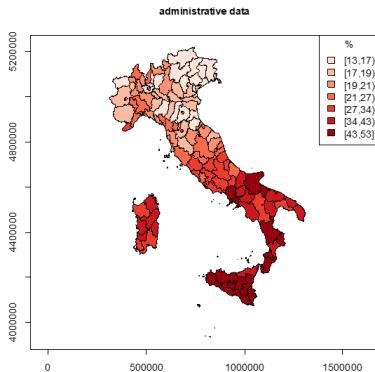
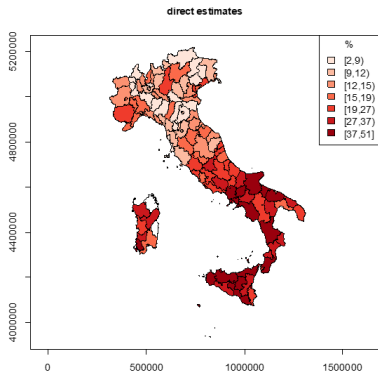
Il progetto di laboratorio si è incentrato su alcuni indicatori EUSILC

- ★ **Rischio di povertà** – media di variabile indicatrice derivata da reddito eq. disponibile.
- ★ **Diseguaglianza del reddito** – QSR, rapporto di due quantili derivata da reddito eq. disponibile.
- ★ **Bassa intensità lavorativa** – incidenza che dipende da condizione occupazionale
- ★ **Severa deprivazione materiale** – incidenza che dipende dalla rilevazione di altre nove variabili.
- ★ **AROPE** – Unione di Rischio di povertà, Bassa intensità lavorativa, Severa deprivazione materiale.
- ★ **Indice di bassa qualità dell'abitazione** – Percentuale di persone che vivono in abitazioni sovraffollate e che presentano almeno uno tra i seguenti problemi: a) problemi strutturali dell'abitazione (soffitti, infissi, ecc.); b) non avere bagno/doccia con acqua corrente; c) problemi di luminosità.

Indicatori Archimede rischio di povertà - confronto

★ vari indicatori su <http://amisuradicomune.istat.it/aMisuraDiComune>

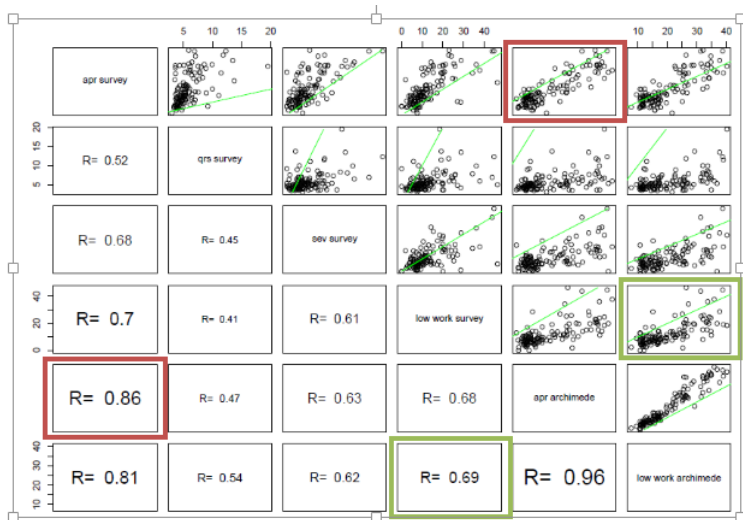
Povertà relativa



★ **sovradimensionamento del fenomeno**

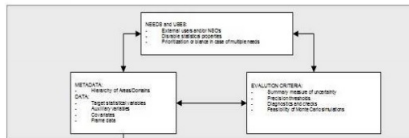
Indicatori Archimede - confronto

★ correlazione dati amministrativi e d'indagine

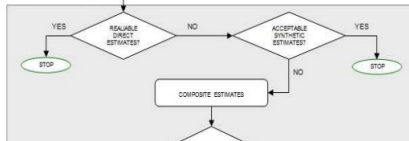


★ PROXY da utilizzare in modelli SAE

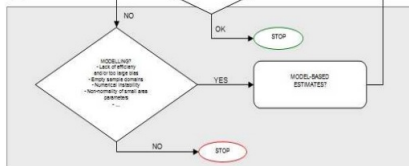
(I) Clarification

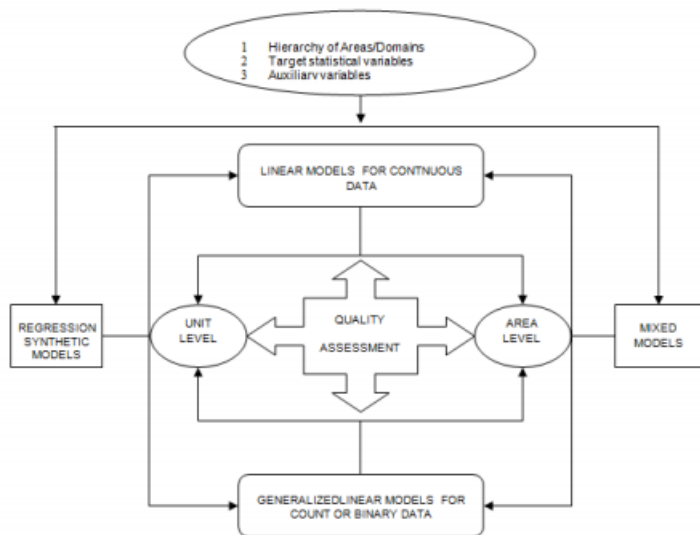


(II) Basic smoothing

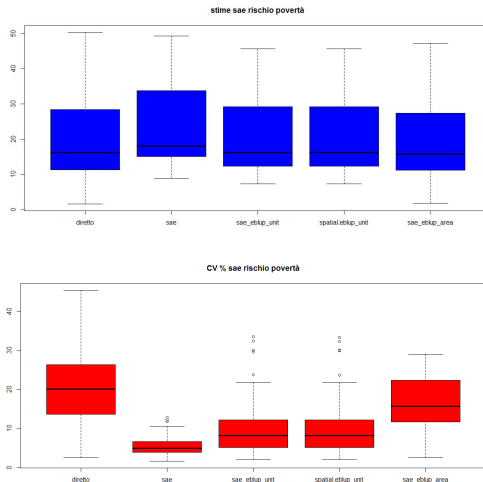


(III) Enhancement





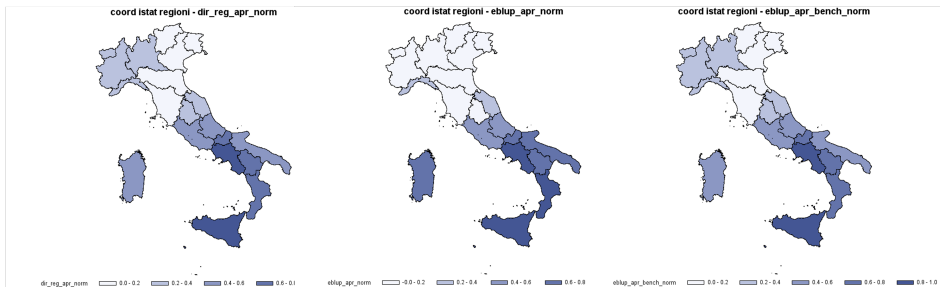
Risultati preliminari - Rischio di povertà anno 2016



★ altri stimatori per piccole aree possono essere applicati

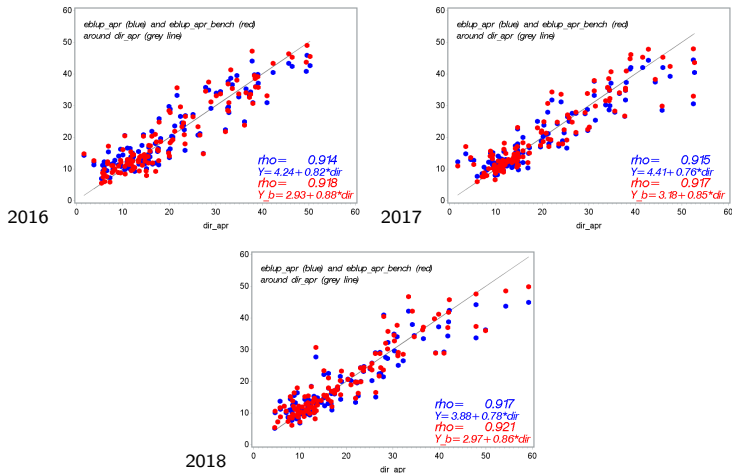
- ★ Le stime per piccole aree sono basate su specifici modelli che permettono tramite la relazione tra la variabile d'interesse e un insieme di variabili ausiliarie di sopperire alla mancanza di informazioni campionarie nelle aree non pianificate.
 - ★ L'aggregazione delle stime SAE per domini non pianificati non coincide con le stime dirette diffuse dall'ISTAT.
- necessità di utilizzare metodi di benchmarking

Rischio di povertà anno 2016 - Coerenza stima eblup unit



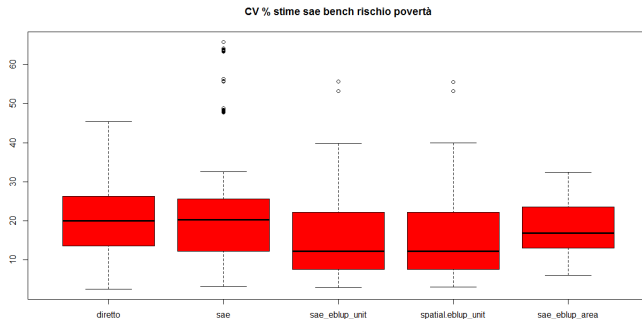
- ★ Le stime sono state normalizzate sottraendo il minimo e dividendo per il range.
- ★ Le mappe mostrano perfetta corrispondenza tra stime regionali ottenute da aggregazione di stime piccole aree dopo benchmark, e stime dirette.
- ★ Prima del benchmark invece si nota una classificazione per difetto di Lombardia, Piemonte, Valle d'Aosta e Umbria; per eccesso per Puglia, Calabria e Sardegna.

Rischio di povertà stime eblup unit



- ★ Le stime 2017 e 2018 sono state calcolate utilizzando l'informazione ausiliaria relativa al 2016
- ★ La relazione tra stime dirette e stime SAE migliora dopo il benchmarking.

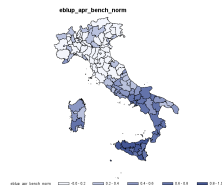
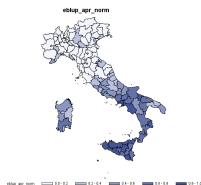
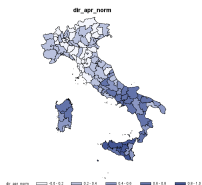
Rischio di povertà stime 2016



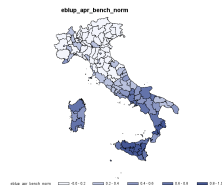
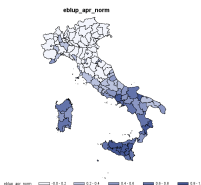
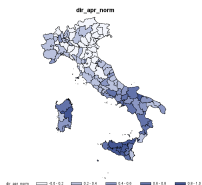
- ★ È necessario aggiungere una componente extra di variabilità alle stime dovuta alla procedura di benchmarking. La conseguenza è che gli errori relativi delle stime SAE aumentano.
- ★ Affinare la procedura di benchmarking
- ★ Migliorare la selezione del modello a livello di unità elementare.

Rischio di povertà mappe stime eblup unit 2016, 2017, 2018

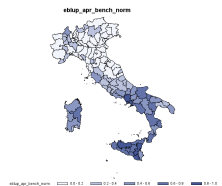
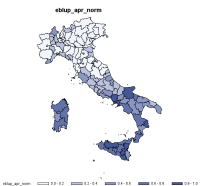
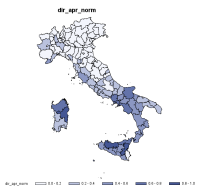
2016



2017



2018

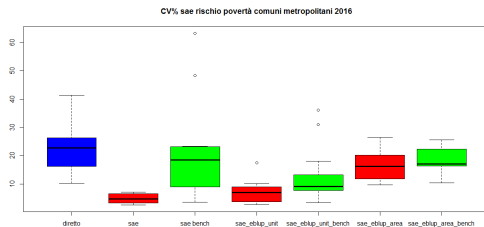
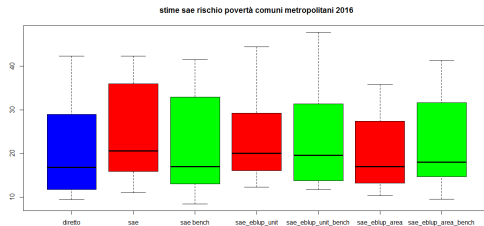


Rischio di povertà stime eblup unit - correlazione ranghi

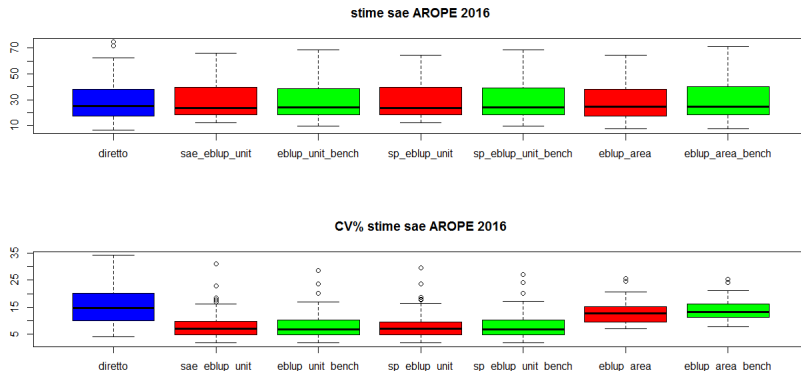
Stimatore sae	Coefficienti di correlazione di Spearman					
	2016		2017		2018	
	dir_apr	dir_apr_norm	dir_apr	dir_apr_norm	dir_apr	dir_apr_norm
eblup_apr	0,86	0,86	0,88	0,88	0,88	0,88
eblup_apr_bench	0,88	0,88	0,90	0,90	0,90	0,90
eblup_apr_norm	0,86	0,86	0,88	0,88	0,88	0,88
eblup_apr_bench_norm	0,88	0,88	0,90	0,90	0,90	0,90

- ★ L'indice di correlazione per ranghi tra stime dirette e stime SAE é molto buona.
- ★ L'indice di correlazione per ranghi di Spearman, calcolato sia sulle stime originarie che su quelle normalizzate

Rischio di povertà 2016 comuni metropolitani

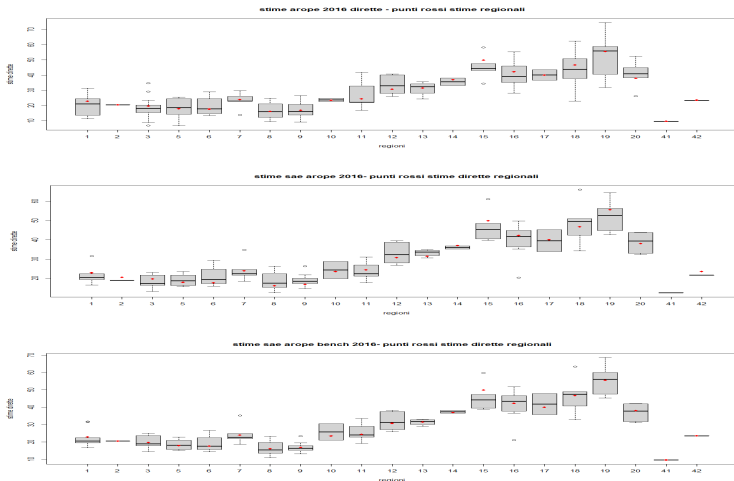


Risultati preliminari - AROPE



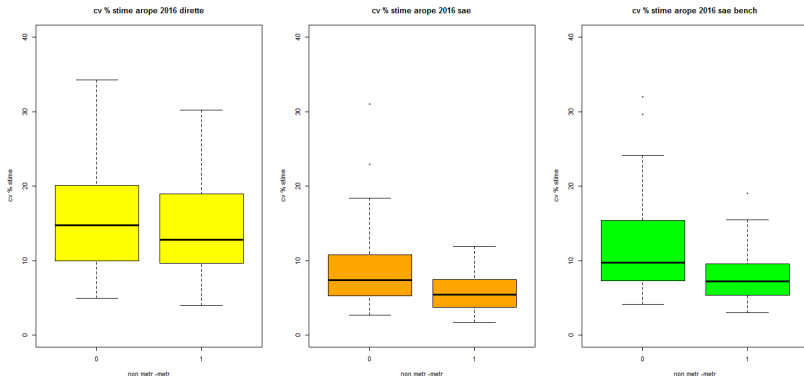
- ★ stime SAE abbastanza simili con cv accettabili
- ★ stime unit in media cv più bassi, valori massimi cv migliori stime aree

AROE distribuzione stime per piccoli domini



- ★ le stime SAE preservano un'accettabile variabilità territoriale anche all'interno delle regioni.

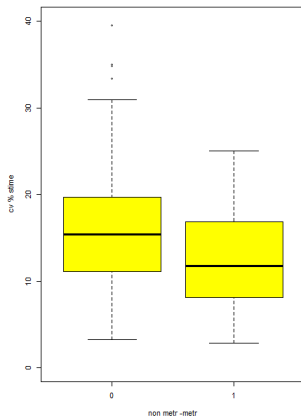
AROPE 2016 - stime province e comuni metropolitani



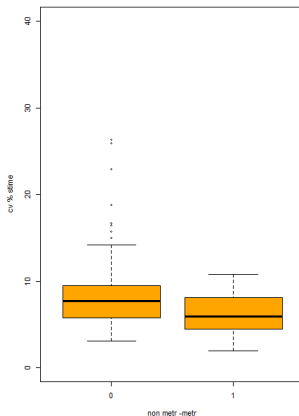
- ★ le stime SAE mostrate sono calcolate con uno stimatore basato su un modello definito a livello di unità
- ★ 1 indica i cv% corrispondenti ai comuni metropolitani; 0 le restanti aree.
- ★ Le stime SAE preservano la variabilità territoriale del fenomeno e permettono dei sostanziali guadagni di efficienza in termini di errore relativo percentuale.

AROPE 2017- stime province e comuni metropolitani

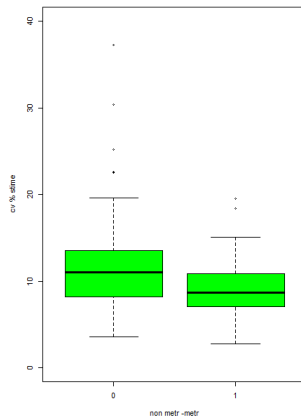
cv % stime arope 2017 dirette



cv % stime arope 2017 sae

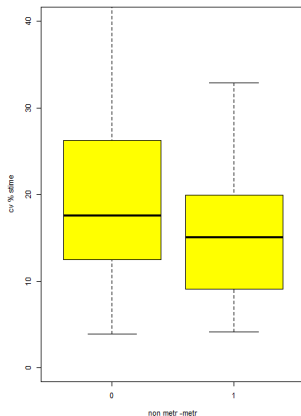


cv % stime arope 2017 sae bench

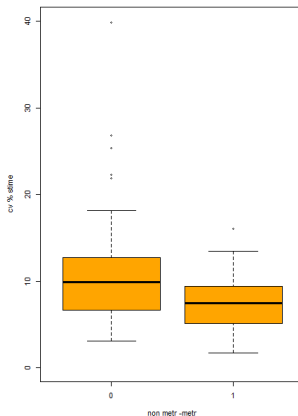


AROPE 2018 - stime province e comuni metropolitani

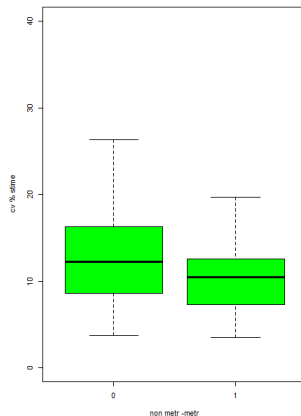
cv % stime arope 2018 dirette



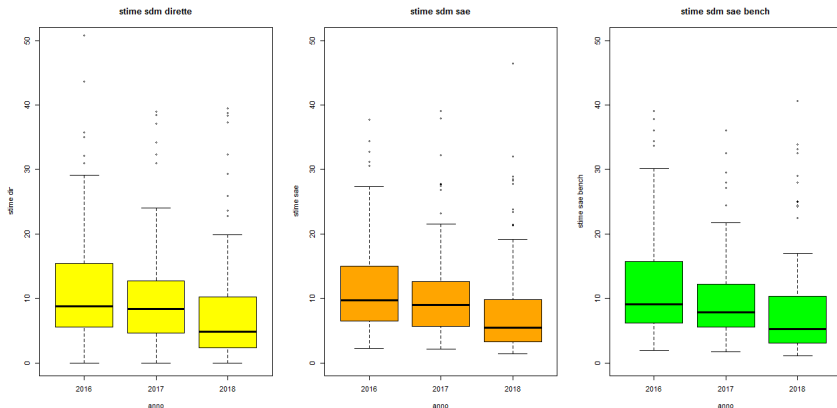
cv % stime arope 2018 sae



cv % stime arope 2018 sae bench

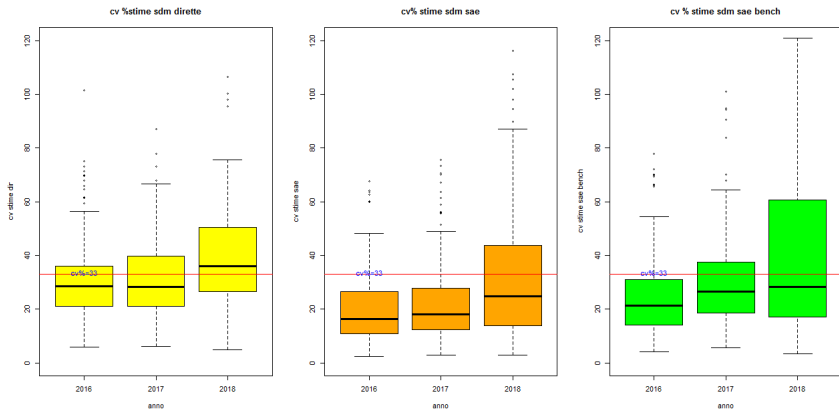


Risultati preliminari - SDM



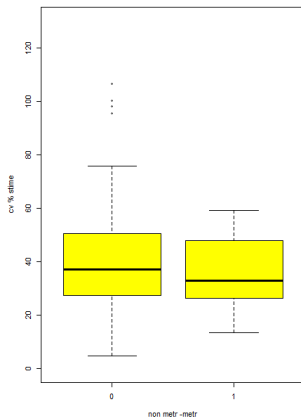
- ★ le stime calcolate con uno stimatore SAE sembrano essere coerenti con delle stime dirette, nonostante il mancato aggiornamento delle informazioni ausiliarie

Risultati preliminari - SDM

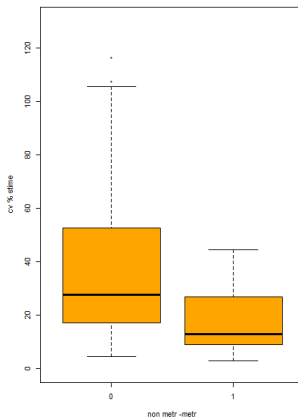


- ★ si ha un miglioramento degli errori relativi
- ★ manca di una proxy da dati amministrativi
- ★ il mancato aggiornamento delle informazioni ausiliarie tende ad avere effetti sul cv, soprattutto sul 2018

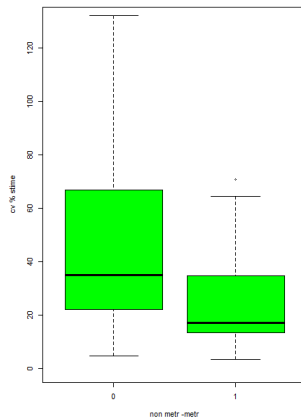
cv % stime sdm 2018 dirette



cv % stime sdm 2018 sae



cv % stime sdm 2018 sae bench



- ★ minore errore relativo delle stime relative ai comuni metropolitani, dovuto probabilmente ad una maggiore copertura campionaria

Risultati preliminari - Altri indicatori

Oltre agli indicatori APR, AROPE e SDM, sono state calcolate le stime SAE per altri indicatori EUSILC

- ★ Risultati simili alla SDM si hanno per [il disagio abitativo](#), indicatore per il quale non è stato possibile costruire alcuna proxy da fonte amministrativa
- ★ Per [la bassa intensità lavorativa](#) la proxy esiste ma è stata costruita in maniera non perfettamente coincidente all'indagine EUSILC. L'indicatore è basato su una specifica sotto-popolazione d'interesse. I risultati sembrano incoraggianti.
- ★ Per [APR e per la disuguaglianza del reddito \(QSR\)](#) è possibile utilizzare metodi che permettono di stimare la distribuzione del reddito nelle piccole aree d'interesse ma sono computazionalmente onerosi.
- ★ L'indicatore [QSR](#) è un indicatore basato sul rapporto tra le code della distribuzione, pertanto anche per i domini pianificati esiste un problema SAE.

Problematiche più rilevanti

- ★ **Complessità della forma funzionale di alcuni indici**, che dipende sia dalla tipologia di indicatori da stimare, sia dagli obiettivi di stima.
 - **Tipologia**, indicatori lineari e indicatori non lineari (come nel caso della povertà)
 - **Obiettivi di stima** stime trasversali, stime longitudinali e stime ordinali.

La complessità degli obiettivi influisce sulla:

- ★ **Disponibilità di dati** - sulla tipologia di dati a cui avere accesso
- ★ Sulla ricerca di soluzioni metodologiche che permettano di fare inferenza simultaneamente su differenti aspetti di interesse.

La disponibilità di dati e la complessità degli obiettivi e dei parametri di interesse influiscono anche sul livello di dettaglio territoriale per il quale è possibile produrre stime sufficientemente attendibili

Problematiche più rilevanti

- ★ **Complessità della forma funzionale di alcuni indici**, che dipende sia dalla tipologia di indicatori da stimare, sia dagli obiettivi di stima.
 - **Tipologia**, indicatori lineari e indicatori non lineari (come nel caso della povertà)
 - **Obiettivi di stima** stime trasversali, stime longitudinali e stime ordinali.

La complessità degli obiettivi influisce sulla:

- ★ **Disponibilità di dati** - sulla tipologia di dati a cui avere accesso
- ★ Sulla ricerca di soluzioni metodologiche che permettano di fare inferenza simultaneamente su differenti aspetti di interesse.

La disponibilità di dati e la complessità degli obiettivi e dei parametri di interesse influiscono anche sul livello di dettaglio territoriale per il quale è possibile produrre stime sufficientemente attendibili

Problematiche più rilevanti

- ★ La metodologia di stima deve essere sottoposta ad un processo di valutazione statistica.
 - Diagnostica e selezione dei modelli per ciascun indicatore
 - Valutazione comparativa dei vari metodi
 - Benchmarking più rifinito che comprenda più vincoli e benchmarking per indicatori non lineari come il QSR.
- ★ Valutazione di esperti tematici
 - nei casi di stima basati su modelli è importante la fase di analisi fatta da esperti tematici (BES), dagli enti locali, etc.
 - analisi nel tempo
 - selezione di informazioni ausiliare rilevanti per ciascun indicatore

Metodi di stima

- ★ Metodologie di stima più complesse
- ★ Modelli di area multivariati (correlazione fra gli indicatori)
- ★ Modelli che sfruttano la correlazione temporale

Dati

- ★ Maggiore e più tempestiva disponibilità di dati
- ★ Tenere conto del nuovo regolamento europeo sulle indagini sociali
- ★ Valutazione delle stime non soltanto in termini di errore relativo
- ★ Applicazione di metodi di stima SAE ad altri indicatori desunti da altre indagini (AVQ, spese, RFL, salute)

Gruppo di lavoro interdipartimentale

- ★ WP1: Acquisizione e analisi delle fonti di dati, validazione dei risultati e progettazione del loro rilascio con il compito di:
 1. fare una ricognizione delle esigenze di stima SAE interna ed esterna all'ISTAT;
 2. fornire le informazioni rilevanti campionarie e proveniente da altre fonti;
 3. analisi tematica dei risultati ottenuti al fine di verificarne la riproducibilità e la conseguente pubblicazione ufficiale

- ★ WP2: Metodologie di stima per indicatori di fonte Spese delle famiglie ed Eusilc con il compito di:
 1. Proporre eventuali ulteriori sviluppi metodologici per la soluzione di particolari problemi di stima.
 2. analisi statistica dei risultati e dei metodi implementati ed utilizzati
 3. Definire una strategia di implementazione che garantisca la riproducibilità del processo di stima e la coerenza tra le stime.

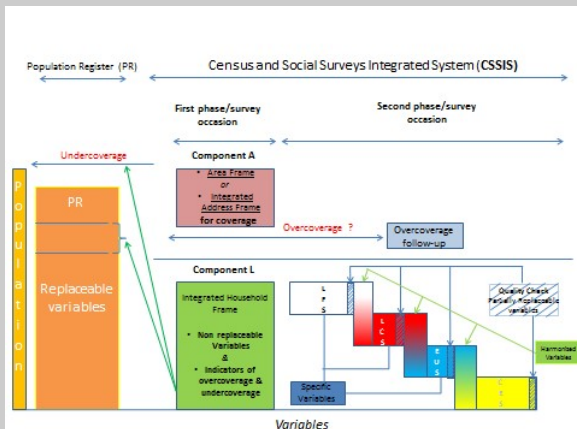
Gruppo di lavoro interdipartimentale

- ★ WP3: Metodologie di stima per indicatori sugli aspetti della vita quotidiana con il compito di:
 1. Proporre metodologie per la produzione di stime dirette ed i relativi errori campionari dei principali parametri delle indagini - modulo d'indagine Censimento permanente;
 2. proporre metodologie specifiche per la stima SAE;
 3. gestire la complessità derivante dalla natura soggettiva di alcune variabili oggetto di studio;
- ★ WP4: Metodologie di stima per indicatori sul mercato del lavoro con il compito di:
 1. di proporre metodologie di stima per piccole aree di tipo amministrativo e funzionale, come smart cities e functional urban area (FUA) e sistemi locali del lavoro (SLL);
 2. Definire una strategia di stima complessiva di implementazione che garantisca la coerenza di stime nel tempo (mensili, trimestrali ed annuali RFLC), stime annuali Censimento permanente e stime SAE nello spazio - coerenza di stime per domini amministrativi e funzionali che spesso si intersecano

Obiettivo di medio periodo

Innestare i processi di stime SAE nel sistema integrato delle indagini sociali con il Censimento permanente.

- ★ L'ISTAT prevede di integrare il sistema delle indagini sociali con il Censimento permanente, attraverso la selezione di moduli campionari dal Master Sample Censuario



Grazie per l'attenzione

