Small Area Estimation of employment and unemployment for Local Labour Market Areas in Italy

Silvia Loriga, Alessandro Martini

Istat

2017 Workshop on Labour Force Survey Methodology

Copenaghen, 4 – 5 May 2017

Istat

Topics

- Disparities in Regional labour market In Italy
- Labour Market areas in Italy
- The new SAE estimation methodology
- Consistency between direct/SAE estimates
- Conclusions

Direct estimates in IT-LFS

LFS in Italy (IT-LFS) is designed as a quarterly survey, with rotational pattern 2-2-2.

In the cross-sectional perspective it provides:

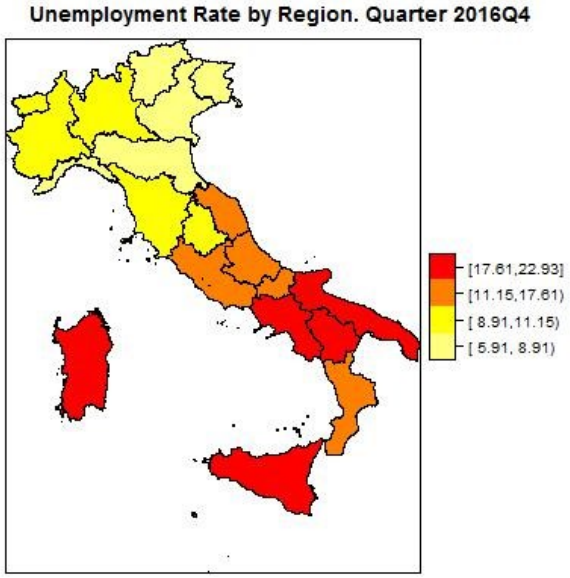
- Monthly estimates, for the whole country;
- Quarterly quarterly figures at NUT2 level, for the 21 Regions;
- Yearly figures as NUTS3 level, for the 110 Provinces, as average of the quarterly estimates

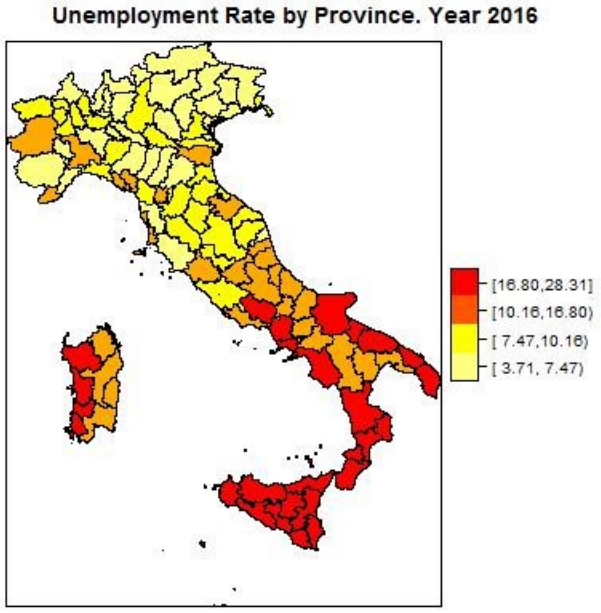
Calibration guarantees the full consistence between Monthly and Quarterly estimates

Territorial Disparities in Italian Labour Market

Italy is one of the European countries with higher dispersion of employment and unemployment rates.

Italy's southern regions suffer from high unemployment rates and low employment rates





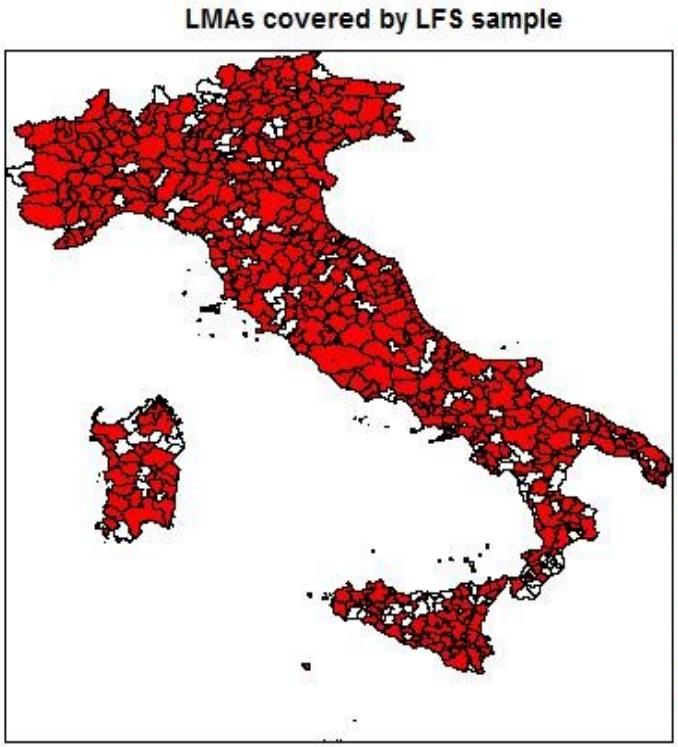
Sample size it is not adequate to ensure the reliability of the direct estimates for all subpopulations of interest

Istat

Labour Market Areas in Italy

- LMAs are sub-regional geographical areas where the bulk of the labour force lives and works;
- are defined on a functional basis, the key criterion is the proportion of commuters who cross the LMA boundary;
- LMAs are developed through an allocation process based on the analysis of commuting patterns;
- 2011 LMAs are based on commuting data stemming from the 15th Population Census;
- they are 611 distinct areas, aggregation of municipalities.

2011 LMAs in Italy



Applying this new definition of LMAS we find that for 2015 150 LMAs, 23.7% of the total, are not covered by LFS sample.

LFS-SAE historical background

Until 1990

- Sub-regional Labour market indicators were obtained increasing sample size.

1996-2002

- Estimates for the 1996-2002 period were defined according to the geography of the LMAs, according to the results of the 1991 Census;
- Design based composite estimator.

2004-2013

- Definition of LMAs according to the results of the 2001 Census;
- Model Based EBLUP estimator was applied to improve the accuracy and the overall quality of the estimates

2015

- In 2015 a working group has been defined in ISTAT in order to review the estimation methodology according to the new definition of LMAs

Estimation Methodology – Model Specification

We have applyed unit level models considering:

- individual survey data since 2004Q1 to 2016Q4;
- auxiliary information refers to population totals for each small area (sex by 14 age groups)
- employment indicator at previous census;
- Dummy variables for 2^{th}-4^{th} interviews to take into account rotational scheme

Estimation Methodology – Model Specification

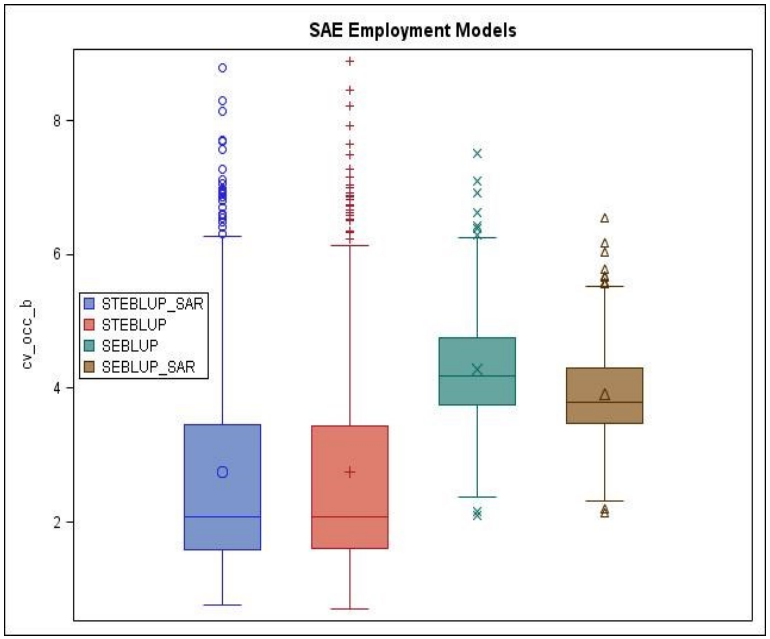
For the back recalculation of the time series 2004-2016 we have compared the results of different SAE methods:

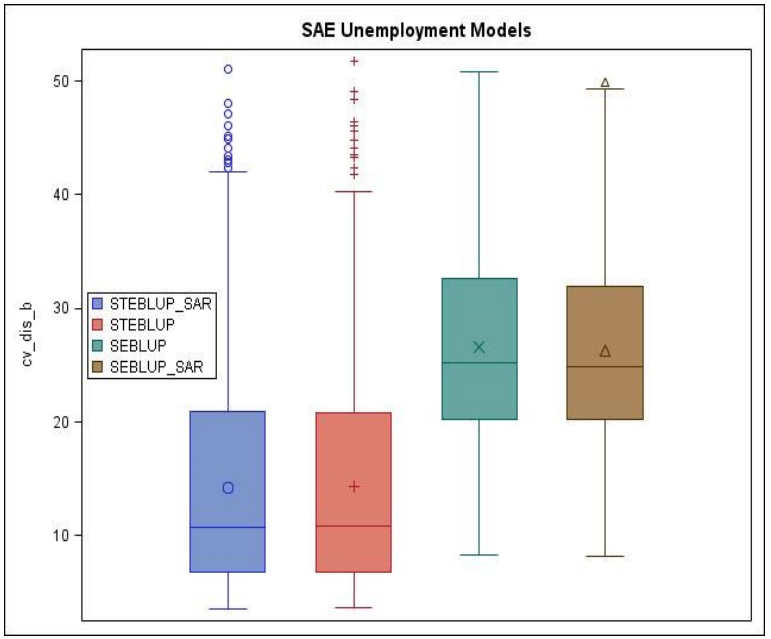
Spatial EBLUP (SEBLUP)
takes into account that the area random effects are spatially correlated, according to two different specifications (Euclidean distance vs neighborhood matrix)

Spatial-time EBLUP (STEBLUP)
a new random effect for the time dimension is defined, as an autoregressive first order AR (1) process.

Estimation Methodology – Model Specification

The box plots in the figure 1 show the Variation Coefficients (CVs) for the different models and specifications of the spatial covariance matrix of the random area effects.





The results show that the STEBLUP estimator has the lowest CVs for both Employement and Unemployement

Reconciliation with Direct estimates

- The results show that the STEBLUP estimator has the lowest CVs for both Employement and Unemployement

- there are not significant differences considering different matrixes for the correlation structure of the area effects

According to these results STEBLUP model with Euclidean matrix for spatial correlation specification was chosed to produce SAE estimates.

Estimation Methodology – Reconciliation with Direct estimates

A benchmarking procedure is needed to take into account contemporaneous constraints according to direct estimates disseminated for planned domains.

MSE estimation is adjusted and includes the benchmarking effect, adding the squared differences between original and benchmarked estimates

$$MSE_{Bench} = MSE_{SAE} + (\check{Y}_{SAE} - \check{Y}_{Bench})^2$$

Benchmarcking procedure tents to inflate the variances of the estimators and can deal to high CVs

Estimation Methodology – Reconciliation with Direct estimates

The first criterion adopted, for provisional 2014 estimation, was :
the benchmarking was imposed for aggregations of LMAs coinciding with a province or a region, or one their aggregation.

The rule, formally elegant, had two flaws:

- created very uneven areas (Trieste, Val d'Aosta, Sicily, Sardinia, part of Puglia, and a big block in the central area of the country) difficult to interpret;

- Over-shrinkage: low SAE estimates variability vs direct estimates

Estimation Methodology – Reconciliation with Direct estimates

The goal was to find a compromise solution to deal with three main issues:

- define benchmarked estimates consistent with LFS direct estimates at at a more detailed NUTS level;
- do not let final CVs increase too much changing base SAE estimates;
- take into account the geographical differences between NUTS classification and LMAs;

.

Estimation Methodology – Reconciliation with Direct estimates

Benchmarking has been applied at:

A. Nuts2-Regional Level

B. Nuts3-Province Level

C. Mixed Procedure: combining A&B according to a distance criterion between initial and final CV

CV can decrease for some Provinces

Bechmarking at regional level

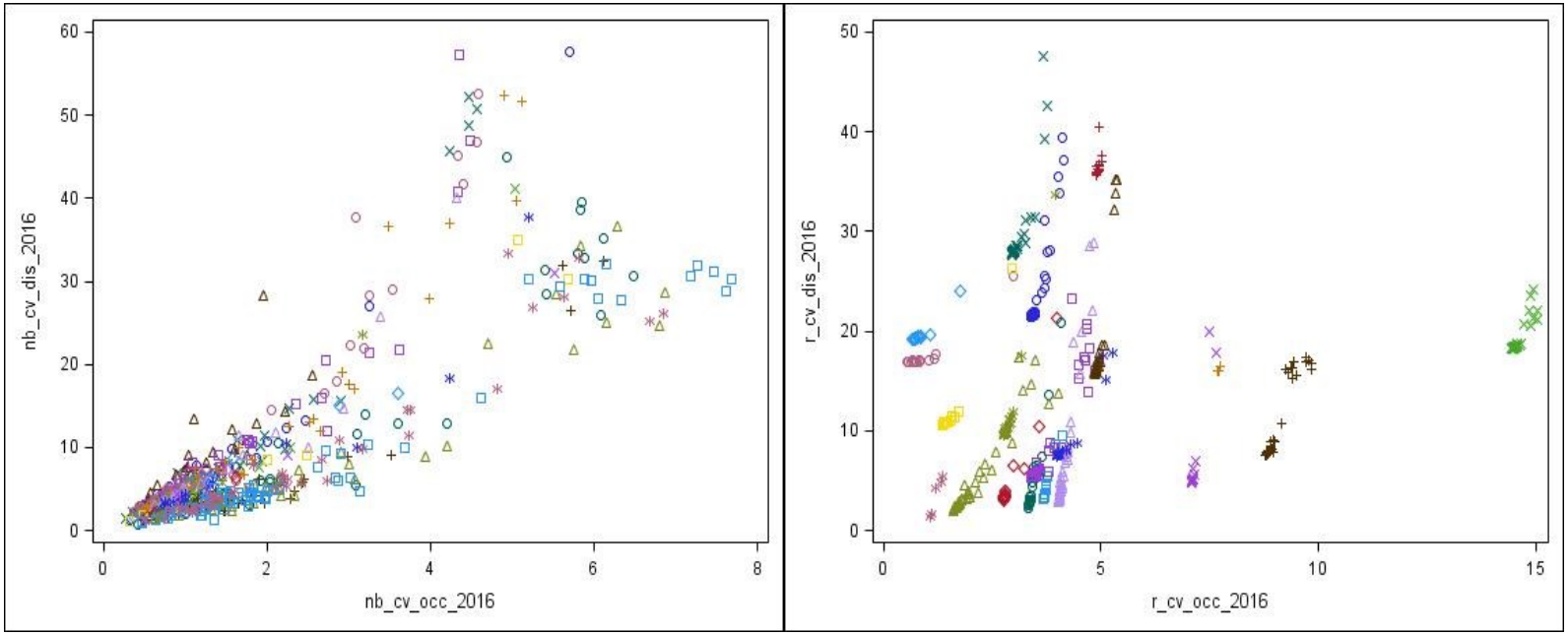
		Regional		Province		Mixed	
		EMP	UNEMP	EMP	UNEMP	EMP	UNEMP
CV	MEAN	5,0	13,7	7,6	19,3	5,4	16
CV	MIN	0,5	1,4	0,3	0,9	0,3	1,0
CV	MAX	15	47,6	44,7	86,4	22,2	86,4

According with these results benchmarking has been applied at regional level, the reconciliation effect on CVs is still relevant, but acceptable.

CVs at Sub-Regional level are too high

Estimation Methodology – Reconciliation with Direct Estimates

Comparison of Unemployment and Employment CVs pre vs post benchmarking grouped by Region.

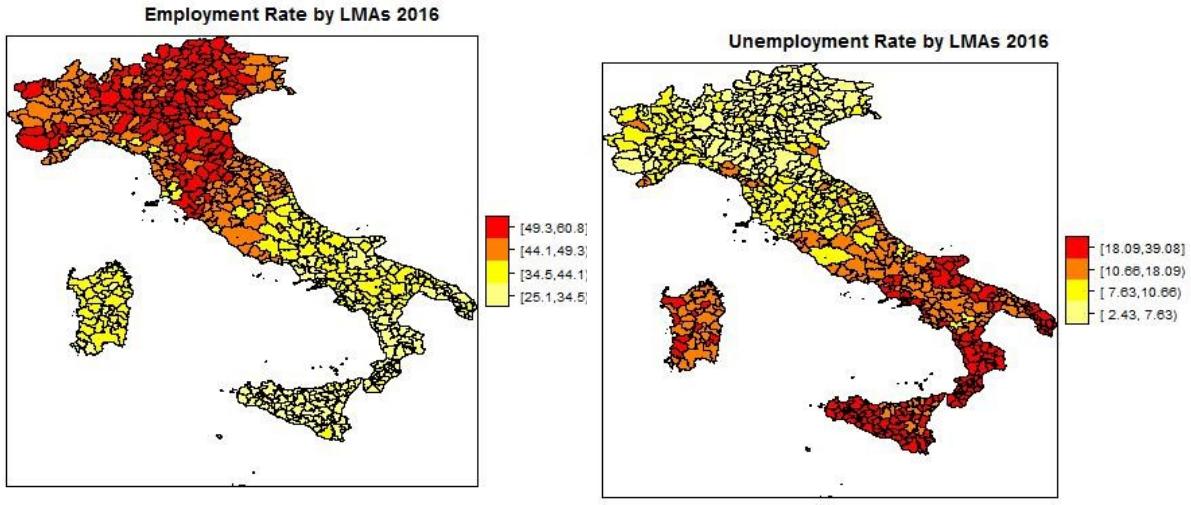


The results show the Benchmarking-level effect, depending on the distance between SAE base estimates and direct ones.

Preliminary results

Back recalculated series refer to population 15+

Employment	Employment rate
Unemployment	Unemployment rate
Labour force	Activity rate



The results show the high disparities in Italian Labour market

Future developments

The increasing availability of administrative data could improve many aspects:

- model specification
- increase the detail of disseminated series
- update the LMA definition;

Conclusions

Findings draw some important and encouraging conclusions in support of the NEW methodology.

- ✓ Estimation methods takes into account both space and time correlation, improving accuracy
- ✓ Accuracy evaluation has been improved
- ✓ Benchmarking can also protect the estimates against potential model misspecification and can be useful for reducing the over-shrinkage of model based small area estimates

On the other hand:

- ✓ Coherence at Province level is not always guaranteed

рахмат
danke
謝謝
ngiyabonga
teşekkür ederim
Баярлалаа
спасибо
faafetai lava
mersi
barka
welalin
tack
spas
kia ora
vinaka
спасибі
blagodaram
dank je
misaotra
matondo
paldies
grazzi
mahalo
tapadh leat
kiitos
dankie
nanni
nandri
dhanyavad
thank you
gracias
хвала
asante
manana
obrigada
tenki
murakoze
bedankt
hvala
maururu
köszönöm
gràcie
bayarlalaa
enkosi
dankon
akun
aciū
djiere dieuf
tau
mochchakkeram
дякую
mamnun
chokrane
dziękuję
chnorakaloutioun
gratias ago
gràcies
sulpáy
taiku
go raibh maith agat
sobodi
děkuji
obrigado
mèsi
didi madloba
sagolun
sukriya
kop khun krap
ありがとう
tanemirt
grazie
arigatô
takk
dakujem
trugarez
najis tuke
kam sah hamnida
rahmat
terima kasih
rahmet
diolch
dhanyavadagalu
shukriya
mercë
мерси
xièxie
감사합니다
ευχαριστώ
merci