Int. J. Nonlinear Anal. Appl. In Press, 1–8
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/ijnaa.2023.29180.4080

Small area estimation of labor force indicators using the multinomial logit mixed model

Anita Abounoori*, Mohammadreza Faghihi Habibabadi

Department of Statistics, Faculty of Mathematical Science, Shahid Beheshti University, Tehran, Iran

(Communicated by Gholamreza Askari)

Abstract

Small area estimation methods have been considered in various fields, especially medicine, agriculture, economics, social sciences, and political science. These methods have many applications in providing reliable statistics for small-sample or non-sample statistical areas. In estimating the small area, there are two approaches: the basic design and the basic model. In this paper, a model-based approach to labor force indicators is considered using a multinomial mixed Logit model. The practical application of the method proposed in this article is to estimate the total number of employees, unemployed and unemployment rate using household income and expenditure data for Semnan province by cities; Semnan, Shahroud, Damghan and Garmsar concerning the period 2011-2016. Finally, we have found the estimates of unemployment rate for Garmsar (9.14), Semnan (9.84), Damghan (11.29), and Shahroud (12.40) in 2016. The more distance from Tehran (the Iranian Capital), the more is the unemployment rate!

Keywords: Small Area Estimation, Labor Force, Logit Mixed Model, Iran
2020 MSC: 62P05

1 Introduction

If it is not possible to conduct a census of a statistical population for any reason, the statistical characteristics of that community cannot be calculated. In this case, sampling can be used instead of a census and statistical indicators can be estimated. In this regard, different methods can be used to estimate. One of the methods used to estimate is the Small Area Estimation (SAE) method. If the sampling at the level of a large set is done in such a way that the size of the sample in its subsets is not large enough, it is not possible to estimate the statistical indicators at that level with necessary validity applying the usual statistical methods. Therefore, the method of small area estimation is proposed in [4] Small area estimation methods are an example of statistical sampling methods that have been considered by researchers in various fields of science over the past few decades and have been widely used in providing reliable statistics for small-sample or nonsample statistical areas. The survey shows, only since 1990 have some textbooks included chapters to introduce SAE principles. The application of this estimation method is used in cases where the sampling of a large community is done in such a way that the size of the sample corresponding to each area of that community is not large enough. Various small area estimation methods have been proposed to address this shortcoming. In most European countries, the labor force indicator is estimated by sampling. Some theoretical and

*Corresponding author
Email addresses: a.abounoori@mail.sbu.ac.ir (Anita Abounoori), m.faghihi@sbu.ac.ir (Mohammadreza Faghihi Habibabadi)

Received: November 2022 *Accepted:* January 2023

empirical studies concerning SAE can be found in; [13], or in [6, 7, 10, 11, 14]. In the area of small area estimation, data is often available for many small areas and for different time periods.

Based on the data available concerning both cross section and time period (i.e. panel data), [1, 2, 3, 5, 9, 12, 16, 18, 20] provided a simple way to obtain cross-sectional information over time period by introducing models that include random effects and time effects.

2 A brief literature survey

2.1 Small area estimator

Small area does not refer to the size of the area, but to the number of samples in the area. The needs for SAE have led to reduction of traditional direct methods and so the application of indirect estimator (model based) has increased. A direct estimator refers to an estimate that can only be obtained using observed data or based on field-specific data. Direct estimators are typically design-based using sampling weights. Often the sample is not large enough to be used to produce acceptable direct estimators for all areas. Also, due to the cost and time spent on such issues, it may not be possible to increase the sample size as needed. Therefore, in order to obtain sufficiently accurate estimators for small areas, it is necessary to use indirect estimators. Thus, the indirect estimator refers to the estimator which, in addition using the observed data in the sample, also uses the values studied in the relevant areas and times and auxiliary data at the community level. In general, three types of indirect estimators can be introduced in the small area domain:

One is the indirect estimator of the area (cross section), this estimator uses only the values of the variables for other regions and does not use the values of time periods. The other one is indirect time estimator, which is estimator using only the values of the variable for the concerned area (not the other areas) in previous time periods. Finally indirect estimator of time and space (cross-sectional and temporal), this estimator uses the variable values in other regions and previous periods.

Jones [8] details a methodology for estimating the unemployment rate at a spatial level below that currently available. It suggests that a reasonable estimate of the unemployment rate, based on residents, can be made, using available data, for small areas. He used Cardiff as a case study and showed how, once published data has been manipulated to achieve an age and gender disaggregated population base for electoral wards, estimates of activity rates can be applied to achieve a total resident economically active population. The estimate is then used, in conjunction with the claimant count, to produce a resident-based unemployment rate applicable to small areas within the county. Finally, the results were evaluated against alternative approaches.

Wang and Zhu [19] proposed a new unit level model based on a pairwise penalised regression approach for problems in small area estimation (SAE). Instead of assuming common regression coefficients for all small domains in the traditional model, they applied a new estimator based on a subgroup regression model which allows different regression coefficients in different groups. They used alternating direction method of multipliers (ADMM) algorithm to find subgroups with different regression coefficients. They also considered pairwise spatial weights for spatial areal data. In the simulation study, they compared the performances of the new estimator with the traditional small area estimator. They also applied the new estimator to urban area estimation using data from the National Resources Inventory survey in Iowa.

Salvati, et. al. [17] indicates that data linkage can be used to combine values of the variable of interest from a national survey with values of auxiliary variables obtained from another source, such as a population register, for use in small area estimation. However, linkage errors can induce bias when fitting regression models; moreover, they can create non-representative outliers in the linked data in addition to the presence of potential representative outliers. They adopt a secondary analyst's point of view, assuming that limited information is available on the linkage process, and develop small area estimators based on linear mixed models and M-quantile models to accommodate linked data containing a mix of both types of outliers. They illustrate the properties of these small area estimators, as well as estimators of their mean squared error, by means of model-based and design-based simulation experiments. Moreover, they illustrate the proposed methodology by applying it to linked data from the European Survey on Income and Living Conditions and the Italian integrated archive of economic and demographic micro data in order to obtain estimates of the average equalized income for labor market areas in central Italy.

3 Method of research

In the basic model approach, indirect estimators use models to determine reliable estimators with sufficient level of accuracy for small areas. In other words, the variables and the auxiliary information are used from the related time periods and are entered into the parameter estimation using a link model. Estimators based on small area models are called base model estimators. These models can be divided into two categories: implicit and explicit link models.

3.1 Model specification

Whenever the data in question is one-dimensional (cross-sectional or time series), the Logit model is used, and if it is two-dimensional (cross-sectional and time-series), the multinomial login model is used as follows: If the Y answer variable has J mode (index)

$(j = 1, ..., J)$ and I variable $(i = 1, ..., I)$, vector $(y_{i1}, y_{i2}, ..., y_{i(J-1)}|u)$ given u And m_i you can write the multinomial distribution with the following density function:

$$f(y_{i1}, ..., y_{i(J-1)}|u) = \frac{m_{i!}}{y_{i1}, ..., y_{iJ}} \times P_{i1}^{y_{i1}}, ..., P_{iJ}^{y_{iJ}} \tag{3.1}$$

Assuming $(p_{i1}, p_{i2}, ..., p_{iJ})$ Depending on the auxiliary variables and random effects of the region through the Logit link as follows:

$$\log\left(\frac{P_{ij}}{P_{iJ}}\right) = x_i\beta_j + u \tag{3.2}$$

In which $j = 1, ..., J$, $i = 1, ..., I$, and $u \sim N(0, \varphi)$. Each p_{ij} Indicates the probability of each statistical unit with a feature i in the category j. Now we can explain multinomial mixed Logit model:

Independent can be written as the distribution of several sentences with the following density function:

$$f(y_{di1}, ..., y_{di(J-1)}|u_d) = \frac{m_{di!}}{y_{di1}, ..., y_{diJ}} \times P_{di1}^{y_{di1}}, ..., P_{diJ}^{y_{diJ}} \tag{3.3}$$

Where assumptions are assumed $(p_{di1}, p_{di2}, ..., p_{diJ})$ depending on the auxiliary variables and random effects of the region through the Logit connection as follows:

$$\log\left(\frac{P_{dij}}{P_{diJ}}\right) = x_{di}\beta_j + u_d \tag{3.4}$$

where in $j = 1, ..., J$, $i = 1, ..., I$, $d = 1, ..., D$, and $u_d^{iid} \sim N(0, \varphi)$ is that any p_{dij} Indicates the probability of each statistical unit with a feature i in the area d in the category j. Includes coefficients of explanatory variables for polynomials with categories $j = 1, ..., J$. This model introduces a structure of natural correlation between different categories and within statistical units of each small area. The estimate of this model provides the probabilities of different categories in the range of [0.1] with a total of one.

This model is a multinomial mixed Logit model for the following assumptions.

1. $u1$
   $u1 = (u'_{1,1,...}u'_{1,D})', u_{1,d} = (u_{1,d1}, u_{1,d2})'u2$
   $u2 = (u'_{2,1,...}u'_{2,D})', u_{2,d} = (u_{2,d1}, u_{2,d2})', u_{2,dk} = (u_{2,dk1,...}u_{2,dkT})',\ k = 1, 2$
   $u_{2,dt} = (u_{2,d1t}, u_{2,d2t})',\ d = 1, ..., D \quad t = 1, ..., T$
2. $u_1 \in N(0, V_{u1})V_{u1}V_{u1} = diag(\varphi_{11}, \varphi_{12}), 1 \le d \le D3.\ u_{2,dk} \in N(0, V_{2,dk}),\ d = 1, ..., D,\ k = 1, 2,\ V_{u2,dk} = \phi_{2k}\Omega(\varphi_k)$ independent and has a first-ordered variance covariance matrix as follows:

$$\Omega(\varphi_k) = \frac{1}{1-\varphi_k^2}\begin{pmatrix} 1 & \varphi_k & \cdots & \varphi_k^{T-2} & \varphi_k^{T-1} \\ \varphi_k & 1 & \ddots & \ddots & \varphi_k^{T-2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \varphi_k^{T-2} & \ddots & \ddots & 1 & \varphi_k \\ \varphi_k^{T-1} & \varphi_k^{T-2} & \cdots & \varphi_k & 1 \end{pmatrix}$$

Model assumes that the vector is the answer y_{dt} with condition $u_{1,d}$ and $u_{2,dt}$, it is independent of each other and follows the distribution of the following multi-sentences:

$$y_{dt}|u_{1,d}, u_{2,dt} \sim M(n_{dt}, P_{d1t}, P_{d2t}), \quad d = 1, ..., D, \quad t = 1, ..., T \tag{3.5}$$

Parameter η_{dkt} also in the form $\eta_{dkt} = \log\left(\frac{P_{dkt}}{P_{d\Im t}}\right)$ are defined.

Model also assumes that η_{dkt} is equal to: $\eta_{dkt} = x_{dkt}\beta_k + u_{1,dk} + u_{2,dkt}$, $d = 1, ..., D$, $k = 1, 2$, $t = 1, ..., T$ where in $x_{dkt} = (x_{dkt1}, ..., x_{dktrk})'$ and $\beta_k = (\beta_{k1}, ..., \beta_{krk})'$. Finally we have:

$$P_{dkt} = \frac{ext(\eta_{dkt})}{1 + \exp(\eta_{d1t}) + \exp(\eta_{d2t})} \quad d = 1, ..., D \quad k = 1, 2 \quad t = 1, ..., T$$

Indirect estimator (model based estimator):

$$\hat{m}_{dt} = N_{dt}\hat{p}_{dt} \quad d = 1, ..., D \quad t = 1, ..., T$$
$$P_{dkt} = \frac{ext(\eta_{dkt})}{1 + \exp(\eta_{d1t}) + \exp(\eta_{d2t})} \quad d = 1, ..., D \quad t = 1, ..., T$$

3.2 Data collection and description

Since it was not possible to access labor data at the county level in Iran, the cost and household income data were examined. These models now look at household cost and revenue data.

The cost and household income statistics survey project has been implemented in rural areas since 1342 and in urban areas since 1347. Since 1353, in addition to household expenses, income information has also been collected. The overall goal of the Household Cost and Income Statistics Survey is to estimate the average cost and income of an urban household and a rural household across the country and in the provinces. According to the survey of 18809 exemplary households in urban areas and 19337 exemplary households in rural areas of the country, the plan of cost and income of households has been implemented in 2016. The target population of this project includes all normal households living in groups in urban or rural areas. These sample households have been selected from 387 cities in urban areas and 395 cities in rural areas of the whole country. Sampling method is three stages with classification, which is done first, classification of census domains and then selection of domains. In the second stage, urban blocks and rural settlements are selected, and in the third stage, sample households are selected. The number of samples has been optimized according to the purpose of the plan to estimate the average cost and annual income of a household in the province. To achieve estimates that better represent the whole year, samples were distributed between the months of the year for statistics. Since it was possible to access data on household expenses and income data and census data for 4 cities of Semnan province, in this dissertation, 4 cities of Semnan province have been studied. In general, these cities have been studied for a period of 5 years from 2011 to 2016. The variable of the answer is the study of the estimate of the labor force indicators, ie the estimate of the total number of employees, the unemployed and the inactive. The data is now being reviewed. Based on the obtained results, the number of samples taken for three groups of employees, unemployed and inactive people from each city in the cost and household income data in the years 2011 to 2016 for Semnan province in the table. This section analyzes cost and household income data. Data were collected from household income and expenditure schemes for three groups of employees, the unemployed and the inactive. Data are considered for 4 cities of Semnan province. Auxiliary variables are collected from census data from the 1990s and 1995s. These auxiliary variables are the ratio of the education level of men and women aged 10 to 65 years, as well as the marital status of men and women in the same age group. The results of estimating the total number of employees, unemployed, inactive people, as well as the unemployment rate under the multi-component Logit mix model with random area and time effects are given independently (Model 2) in the table. The table also summarizes the results of estimating the total number of employees, the unemployed, inactive people, as well as the unemployment rate under the multicomponent Logit mixed model with random area and time effects (Model 3).

4 Estimation of the Models

Between the two proposed methods, the multivariate Logit mix model with random area and time effects is better than the multi-component Logit model with random area and time independent effects. The sample size in Semnan Province concerning different 5 cities and 6 years (from 2011 to 2016) are indicated in Table 1.

Table 1: Number of sample households concerning Semnan province (Source: Household income and expenditure data for Semnan province by cities)

Number of samples in the design household Budget Surveys	**Year**	**City**	**Number of samples in the design household Budget Surveys**	**Year**	**City**
214	1390	Damghan	395	1390	Shahroud
161	1391	Damghan	382	1391	Shahroud
232	1392	Damghan	363	1392	Shahroud
227	1393	Damghan	401	1393	Shahroud
196	1394	Damghan	375	1394	Shahroud
190	1395	Damghan	334	1395	Shahroud
455	1390	Semnan	148	1390	Garmsar
427	1391	Semnan	129	1391	Garmsar
353	1392	Semnan	196	1392	Garmsar
349	1393	Semnan	202	1393	Garmsar
356	1394	Semnan	219	1394	Garmsar
343	1395	Semnan	201	1395	Garmsar

Using the micro Household Budget Survey data we have estimated the number of unemployed, inactive, and thus employed population using the multinomial mixed logit model. The results are summarized in Table 2.

Table 2: Estimation of the number unemployed (Unemp.), inactive (Inj.), employed (Emp.) at the sample (Sam.) and community (Com.) level based on model 2. (Sources: Estimated using Mixed Logit Model and the data extracted from Household Budget Survey.)

N. Com. Inj. Level	**N. Com. Unemp. Level**	**N. Com. Emp. Level**	**N. Sam. Inj. Level**	**N. Sam. Unemp. Level**	**N. Sam. Emp. Level**	**Year**	**City**
24454	1414	26409	100	6	108	1390	Damghan
31043	1853	20259	94	6	61	1391	Damghan
35698	2132	16218	153	10	69	1392	Damghan
38489	2282	14186	158	10	59	1393	Damghan
39500	2232	14149	139	8	49	1394	Damghan
38314	2077	16432	129	7	54	1395	Damghan
60064	3098	66998	210	11	234	1390	Semnan
82123	4793	47208	262	15	150	1391	Semnan
96805	6113	35294	248	15	90	1392	Semnan
104917	6794	30713	258	16	75	1393	Semnan
106533	6683	33548	259	16	81	1394	Semnan
97129	5342	48810	221	12	110	1395	Semnan
61046	3857	66255	185	11	199	1390	Shahroud
74872	5145	52524	217	14	151	1391	Shahroud
84997	5980	42963	231	16	116	1392	Shahroud
90839	6261	38254	270	18	113	1393	Shahroud
91599	5878	39306	252	16	107	1394	Shahroud
84277	6695	47348	204	16	114	1395	Shahroud
17713	1452	27462	56	5	87	1390	Garmsar
24091	1825	21213	66	5	58	1391	Garmsar
28891	2062	16683	120	8	68	1392	Garmsar
31757	2181	14213	134	9	59	1393	Garmsar
32358	2293	14019	146	10	63	1394	Garmsar
30089	1758	17392	124	7	70	1395	Garmsar

Table 3: Estimation of unemployment rate in Semnan province concerning different cities, based on model 2. (Sources: Estimated using Mixed Logit Model and the data extracted from Household Budget Survey.)

unemployment rate	**Year**	**City**	**unemployment rate**	**Year**	**City**
5.2631	1390	Damghan	5.238	1390	Shahroud
8.9552	1391	Damghan	8.4848	1391	Shahroud
12.6582	1392	Damghan	12.1212	1392	Shahroud
14.4927	1393	Damghan	13.7404	1393	Shahroud
14.035	1394	Damghan	13.0081	1394	Shahroud
11.4754	1395	Damghan	12.3076	1395	Shahroud
4.7891	1390	Semnan	5.4347	1390	Garmsar
9.0909	1391	Semnan	7.9365	1391	Garmsar
14.2857	1392	Semnan	10.5263	1392	Garmsar
17.5824	1393	Semnan	13.2352	1393	Garmsar
16.4948	1394	Semnan	13.6986	1394	Garmsar
9.836	1395	Semnan	9.0909	1395	Garmsar

Table 4: Estimation of the number of unemployed (Unemp.) and inactive (Inj.), employees (Emp.) at the sample (Sam.) and community (Com.) level and based on model 2. (Sources: Estimated using Mixed Logit Model and the data extracted from Household Budget Survey.)

N. Inj. Com. Level	**N. Unemp. Com. Level**	**N. Emp. Com. Level**	**N. Inj. Sam. Level**	**N. Unemp. Sam. Level**	**N. Emp. Sam. Level**	**Year**	**City**
24456	1412	26409	102	5	107	1390	Damghan
31041	1854	20260	92	7	62	1391	Damghan
35698	2140	16210	154	11	67	1392	Damghan
38480	2282	14195	157	10	60	1393	Damghan
39500	2200	14181	139	7	50	1394	Damghan
38300	2091	16432	129	6	55	1395	Damghan
60158	3098	66904	211	9	235	1390	Semnan
82123	4801	47200	263	13	151	1391	Semnan
96805	6117	35290	248	14	91	1392	Semnan
104930	6794	30700	259	17	73	1393	Semnan
106533	6691	33540	259	14	83	1394	Semnan
97129	5340	48812	222	10	111	1395	Semnan
61051	3857	66250	185	10	200	1390	Shahroud
74872	5149	52520	219	13	150	1391	Shahroud
84997	5983	42960	233	16	114	1392	Shahroud
90839	6265	38250	272	16	113	1393	Shahroud
91599	5870	39314	253	14	108	1394	Shahroud
84277	6703	47340	204	16	114	1395	Shahroud
17713	1454	27460	57	4	87	1390	Garmsar
24091	1828	21210	65	6	58	1391	Garmsar
28893	2060	16683	120	7	69	1392	Garmsar
31757	2148	14210	135	10	57	1393	Garmsar
32358	2297	14015	146	9	64	1394	Garmsar
30089	1750	17400	123	6	72	1395	Garmsar

Table 5: Estimation of unemployment rate in Semnan province concerning different cities (model 3). (Sources: Estimated using Mixed Logit Model and the data extracted from Household Budget Survey.)

unemployment rate	**Year**	**City**	**unemployment rate**	**Year**	**City**
5.0753	1390	Damghan	5.5015	1390	Shahroud
8.3838	1391	Damghan	8.9285	1391	Shahroud
11.6621	1392	Damghan	12.2244	1392	Shahroud

13.8496	1393	Damghan	14.0739	1393	Shahroud
13.4301	1394	Damghan	12.9913	1394	Shahroud
11.2886	1395	Damghan	12.403	1395	Shahroud
4.4255	1390	Semnan	5.0287	1390	Garmsar
9.2325	1391	Semnan	7.9347	1391	Garmsar
14.7728	1392	Semnan	10.9907	1392	Garmsar
18.1202	1393	Semnan	13.3219	1393	Garmsar
16.6314	1394	Semnan	14.0816	1394	Garmsar
9.8611	1395	Semnan	9.1383	1395	Garmsar

5 Conclusion

Small area estimation is any of several statistical techniques involving the estimation of parameters for small sub-populations, generally used when the sub-population of interest is included in a larger survey. The term ”small area” in this context generally refers to a small geographical area such as a county. We have used the micro household budget survey data concerning different city of Semnan province for the period 2011-2016 using the multinomial mixed logit model to estimate unemployment rate at city level concerning the small area estimation method. The results indicate that the unemployed rates of Semnan, Shahroud, Damghan, Garmsar in 2016 were 9.86, 12.40, 11.29, and 9.14, respectively. It is interesting to notice that the less distance of these cities from Tehran are Garmsar, Semnan, Damghan and Shahroud, respectively. It seems the less distance to Tehran (the Capital City of Iran) the less unemployed rate! Perhaps there is spillover labor force from different cities to Tehran.

References

[1] G.S. Datta, P. Lahiri and T. Maiti, *Empirical Bayes estimation of median income of four-person families by state using time series and cross-sectional data*, J. Statist. Plann. Inference **102** (2002), no. 1, 83–97.

[2] G.S. Datta, P. Lahiri, T. Maiti and K.L. Lu, *Hierarchical Bayes estimation of unemployment rates for the states of the US*, J. Amer. Statist. Associat. **94** (1999), no. 448, 1074–1082.

[3] M.D. Esteban, D. Morales, A. Pérez and L. Santamaría, *Small area estimation of poverty proportions under area-level time models*, Comput. Statist. Data Anal. **56** (2012), no. 10, 2840–2855.

[4] R.E. Fay III and R.A. Herriot, *Estimates of income for small places: an application of James-Stein procedures to census data*, J. Amer. Statist. Associat. **74** (1979), no. 366a, 269–277.

[5] M. Ghosh, N. Nangia and D.H. Kim, *Estimation of median income of four person families: a Bayesian time series approach*, J. Amer. Statist. **91** (1996), no. 436, 1423–1431.

[6] M. Ghosh, J.N.K. Rao, *Small area estimation: an appraisal*, Statist. Sci. 9 (1994), no. 1, 55–76.

[7] J. Jiang and P. Lahiri, *Mixed model prediction and small area estimation*, Test: Offic. J. Spanish Soc. Statist. Operat. Res. **15** (2006), no. 1, 96.

[8] C. Jones, *The estimation of rates of unemployment for small spatial units*, Local Econ. **15** (2000), no. 2, 144–158.

[9] Y. Marhuenda, I. Molina and D. Morales, *Small area estimation with spatiotemporal Day–Herriot models*, Comput. Statist. Data Anal. **58** (2013), 308–325.

[10] D. Pfeffermann, *Small area estimation-new developments and directions*, Int. Statist. Rev. **70** (2002), no. 1, 125–143.

[11] D. Pfeffermann, *New important developments in small area estimation*, Statist. Sci. **28** (2013), no. 1, 40–68.

[12] D. Pfeffermann and L. Burck, *Robust small area estimation combining time series and cross-sectional data*, Survey Methodol. **16** (1990), 217–237.

[13] J.N.K. Rao, *Some new developments in small area estimation*, J. e Iranian Statist. Soc. 2 (2003), no. 2, 145–169.

[14] J.N.K. Rao and M. Yu, *Small-area estimation by combining time-series and cross sectional data*, Canad. J. Statist. **22** (1994), no. 4, 511–528.

[15] W. Rudin, *Principles of mathematical analysis*, McGraw-Hill Inc, 1976.

[16] A. Saei and R. Chambers, *Small area estimation under linear and generalized Linear mixed models with time and area effects*, Methodol. Work. Paper no M03/15 Southampton Statistical Sciences Research Institute, University of Southampton, UK, 2003.

[17] N. Salvati, E. Fabrizi, M.G. Ranalli and R.L. Chambers, *Small area estimation with linked data*, J. Royal Statist. Soc. Ser. B **83** (2021), no. 1, 78–107.

[18] M.D. Ugarte, T. Goicoa, A.F. Militino and M. Durbán, *Spline smoothing in small area trend estimation and forecasting*, Comput. Statist. Data Anal. **53** (2009), no. 10, 3616–3629.

[19] X. Wang and Z. Zhu, *Small area estimation with subgroup analysis*, Statist. Theory Related Fields **3** (2019), no. 2, 129–135.

[20] Y. You, J.N.K. Rao and J. Gambino, *Model-based unemployment rate estimation for the Canadian labour force survey: a hierarchical Bayes approach*, Survey Methodol. **29** (2001), no. 1, 25–32.