SORT 33 (1) January-June 2009, 49-70

Estimating unemployment in very small areas

Ugarte[1*], M.D., Goicoa[1], T., Militino[1], A.F., and Sagaseta-López, M.[2]

Abstract

In the last few years, European countries have shown a deep interest in applying small area techniques to produce reliable estimates at county level. However, the specificity of every European country and the heterogeneity of the available auxiliary information, make the use of a common methodology a very difficult task. In this study, the performance of several design-based, model-assisted, and model-based estimators using different auxiliary information for estimating unemployment at small area level is analyzed. The results are illustrated with data from Navarre, an autonomous region located at the north of Spain and divided into seven small areas. After discussing pros and cons of the different alternatives, a composite estimator is chosen, because of its good trade-off between bias and variance. Several methods for estimating the prediction error of the proposed estimator are also provided.

MSC: 62D05, 62J12, 62F40

Keywords: Finite population, Prediction theory, Labour Force Survey

1 Introduction

The Spanish Labour Force Survey (SLFS), called "Encuesta de Población Activa" in Spanish or, in short "EPA", is a quarterly survey of households living at private addresses in Spain. Its purpose is to provide information on the Spanish labour market that can then be used to develop, manage, evaluate, and report on labour market policies. It is conducted by the Spanish Statistical Institute (INE). Yet there are multiple aims achieved with this survey, of which the estimation of unemployment is one of the most relevant. The survey follows a stratified two-stage cluster design and, for each province, a separate

[1] Departamento de Estadística e Investigación Operativa, Universidad Pública de Navarra. Campus de Arrosadía, 31006 Pamplona

[2] Instituto de Estadística de Navarra. Avda. Carlos III 36, 2º Dcha, 31003 Pamplona

[*] E-mail: lola@unavarra.es
Received: November 2007
Accepted: July 2008

sample is drawn. In the first stage sampling, 3,588 primary sampling units (PSUs) called “Secciones Censales” in Spanish, are selected with probabilities proportional to size according to the number of households. In the second stage sampling, the secondary sampling units (SSUs) are households and a simple random sampling is applied to draw 18 SSUs from each PSU selected. This sampling design generates self-weighted samples at stratum level and then, every household has the same probability of being drawn. In every Spanish province there is a fixed number of PSUs that have to be selected. Navarre is a small autonomous region (with a single province) located in northern Spain. It has an area of 10,000 km^2 and only 600,000 inhabitants, irregularly distributed in seven small areas. In Navarre, 91 PSUs are selected in the first stage, and, for each PSU, 18 households are drawn obtaining a total of 1,638 households. In this study, interest is focussed on evaluating the performance of design-based, model-assisted, and model-based estimators for estimating unemployment in Navarre using different auxiliary information. We are dealing with a scenario where the number of small areas is very limited (seven small areas), the incidence of the study variable is scarce, and the sample size cannot be modified because it is already determined. To evaluate the alternative estimators, a Monte Carlo study has been conducted drawing 500 samples from the 2001 Spanish Census following the same sampling design of the SLFS. Similar simulation studies over the scenario of the Labour Force Census of companies affiliated to the Social Security system in Catalonia have been recently published by Costa *et al.* (2003). These authors conclude that a composite estimator with estimated weights based on the assumption of heterogeneity of biases and variances across small areas is superior to its competitors: direct and indirect estimators. The sample size effect on these composite estimators has also been studied by Costa *et al.* (2004). Morales *et al.* (2007) have analyzed the performance of model-assisted and model-based estimators on the Spanish Labour Force data from the Canary Islands. In their paper, model-based estimators are more competitive than model-assisted estimators, but they do not consider composite estimators. The impact of supplementary samples sizes on the Spanish Labour Force has also been analyzed by Costa *et al.* (2006), who evaluate the performance of some composite estimators when the sample size of the small areas is boosted.

The rest of this paper is organized as follows. Section 2 describes several design-based, model-assisted, and model-based estimators that might be used for estimating unemployment in very small areas. Section 3 describes a Monte Carlo study, on the scenario of the employment survey in Navarre, and gives the accuracy measures for the estimators presented in Section 2. Section 4 provides three different procedures to calculate the mean squared error (MSE) of the estimator chosen as optimal in Section 3. The paper ends up with some conclusions.

Figure 1: *Small Areas of Navarre*

2 Alternative estimators of unemployment

The variable of interest is the total number of unemployed in the seven small areas (called "comarcas", in Spanish) of Navarre (see Figure 1). Some of these small areas, mainly those located at the north of the province, are scarcely populated. In this section alternative estimators of unemployment are briefly described. They will be compared in the next section through a Monte Carlo study.

2.1 Design-based estimators

In the design-based theory, the variable of interest is a fixed quantity and the probability distribution is induced by the sampling design. It is a distribution-free method mainly focused on obtaining estimates for domains with large samples. A direct estimator only uses observations coming from the domain of interest, whereas an indirect estimator takes information outside of the domain. In the design-based theory, unbiasedness and design-consistency are desirable properties pursued by the majority of estimators. An estimator $\hat{Y}$ of Y is design-unbiased if $E[\hat{Y}] = Y$ and it is design-consistent if it is unbiased and its variance tends to zero as the sample size increases (Rao, 2003). The use of auxiliary information is a common tool for improving the precision of design-based estimators. Here, it consists of age-sex groups (E), with six categories which are a combination of age groups $(16-24, 25-54, > 55)$ and sex; Stratum (S), that represents municipality sizes and takes nine possible values in Spain, although in Navarre only six of those nine possible strata are available: (1) capital of the province, (5) between 20000 and 49999 inhabitants, (6) between 10000 and 19999 inhabitants, (7) between 5000 and 9999 inhabitants, (8) between 2000 and 4999 inhabitants and (9) with less than 2000 inhabitants; educational level (N) has two categories: (1) for illiterate, primary, and secondary school, and (2) for technical workers and professionals; employment status (P) in the Navarre Employment Register (SNE) with two categories: (1) occupied or inactive, and (2) unemployed; claimant of employment (C) taking the value 1 if he/she is registered in the SNE and 0, otherwise.

In this paper, the following design-based estimators are considered:

(a) Two direct estimators: the so-called direct, and the post-stratified estimator
(b) Five indirect estimators: one synthetic, and four composite estimators

Direct estimators use only data in the domain of interest. Although they are design-unbiased, the variability is usually big enough to be considered appropriate in small-area estimation.

The **direct estimator** of the total unemployment Y in the dth small area takes the form

$$\hat{Y}_d^{direct} = \hat{\bar{Y}}_d^{direct} N_d = \frac{\sum_{j=1}^{n_d} w_j y_j}{\sum_{j=1}^{n_d} w_j} N_d, \quad d = 1, \ldots, D,$$

where in area d, N_d is the population aged 16 or more, n_d is the number of people aged 16 or more in the sample, and the sampling weight w_j is the inverse of the inclusion probability for person j. The sampling weights are given by $w_j = N_h/n_h$ for $j = 1, \ldots, n_d$. This means that every person belonging to the same stratum h $(h = 1, \ldots, H)$ has the same weight. Detailed expressions on how to obtain these weights are given by Morales *et al.* (2007). The variable y_j takes the value 1 if person j is unemployed and 0 otherwise. The total number of small areas is denoted by D.

The **post-stratified estimator** of the total unemployment Y in the dth small area is given by

$$\hat{Y}_d^{post} = \sum_{g=1}^{G} \hat{\bar{Y}}_{dg} N_{dg} = \sum_{g=1}^{G} \frac{\sum_{j=1}^{n_{dg}} w_j y_j}{\sum_{j=1}^{n_{dg}} w_j} N_{dg}, \quad d = 1, \ldots, D, \tag{1}$$

where G is defined by the categories of the different auxiliary variables. For instance, G has six categories when the variable age-sex group (E) is considered. The number of sampled people in the dth region belonging to the gth group is n_{dg} while N_{dg} is the corresponding population value. The post-stratified estimator is a direct estimator that only uses information from the domain of interest, yet it may also be considered as a model-assisted estimator as it can be derived from a linear model where the explanatory variable is a group indicator variable.

The synthetic estimator (González, 1973) is an indirect estimator used in small areas under the assumption that the small areas have the same characteristics as the large area, with regard to the variable of interest. When this does not happen, synthetic estimators are usually biased. The **synthetic estimator** used here takes the form

$$\hat{Y}_d^{synt} = \sum_{g=1}^{G} \hat{\bar{Y}}_g N_{dg} = \sum_{g=1}^{G} \frac{\sum_{j=1}^{n_g} w_j y_j}{\sum_{j=1}^{n_g} w_j} N_{dg}, \quad d = 1, \ldots, D, \tag{2}$$

where n_g is the number of sampled people in the whole province belonging to the gth group.

A natural way to balance the potential bias of a synthetic estimator and the instability of a direct estimator is to take a weighted average of the two estimators, what is called a composite estimator. The name of composite estimator has a more general meaning, corresponding to any kind of linear combination of estimators. In our case, **sample size dependent composite estimators** (Drew *et al.*, 1982) are considered. They are defined as a linear combination of a post-stratified and a synthetic estimator. Namely

$$\hat{Y}_d^{comp} = \lambda_d \hat{Y}_d^{post} + (1 - \lambda_d) \hat{Y}_d^{synt}$$

where

$$\lambda_d = \begin{cases} 1 & \text{if } \hat{N}_d \geq \alpha N_d \\ \dfrac{\hat{N}_d}{\alpha N_d} & \text{otherwise.} \end{cases}$$

$0 \leq \lambda_d \leq 1$, $\hat{N}_d = \sum_{j=1}^{d} w_j$ is a direct expansion estimator of N_d that increases with the domain sample size, and α is chosen to control the contribution of the synthetic estimator, and can take the following values: $\alpha = 2/3, 1, 1.5, 2$. Note that there are four possible composite estimators, one for each value of α. Hereafter in the paper, they will be denoted by composite 1 ($\alpha = 2/3$), 2 ($\alpha = 1$), 3 ($\alpha = 1.5$), and 4 ($\alpha = 2$), respectively. When the sample size increases λ_d is close to 1, and $\hat{Y}_d^{comp}$ is similar to the post-stratified estimator $\hat{Y}_d^{post}$, otherwise more weight is given to the synthetic estimator.

2.2 Model-assisted estimators

These estimators take account of the auxiliary information through the use of regression models as a means to obtain design-consistent estimators (Särndal *et al.*, 1992). They are more efficient than design-based estimators as auxiliary information is explicitly used at the estimation stage. Therefore, an important reduction of bias is attained. The most well-known model-assisted estimators are the generalized regression estimators (GREG), and, here, GREG estimators assisted in three different models are considered:

a linear model (see (3)), a logit model (see (4)), and a logit mixed model (see (5)) . Let us consider the following linear model

$$y_{jd} = \mathbf{x}_{jd}^{\top}\boldsymbol{\beta} + \epsilon_{jd}, \quad j = 1, \ldots, n_d, \quad d = 1, \ldots, D, \tag{3}$$

where for every small area d, y_{jd} takes the value 1 if the jth person is unemployed, and 0 otherwise, $\mathbf{x}_{jd} = (x_{jd,1}, x_{jd,2}, \ldots, x_{jd,p})^{\top}$ is the vector of the p auxiliary variables, and $\epsilon_{jd} \sim N(0, \sigma^2/w_{jd})$.

The GREG estimator of the total number of unemployed in the dth area, assisted in model (3), is given by

$$\hat{Y}_d^{LinearGREG} = N_d \left(\hat{\bar{Y}}_d^{direct} + (\bar{\mathbf{X}}_d - \hat{\bar{\mathbf{X}}}_d^{direct}) \right)^{\top} \hat{\boldsymbol{\beta}},$$

where $\bar{\mathbf{X}}_d = (\bar{X}_{d1}, \bar{X}_{d2}, \ldots, \bar{X}_{dp})^{\top}$ is the vector of the p auxiliary population means. The parameter vector $\boldsymbol{\beta}$ is estimated by generalized least squares with all the province observations.

Assuming that $y_{jd} \sim Bernoulli(P_{jd})$, where P_{jd} is the probability that the jth person in the dth area is unemployed, it seems more appropriate to be assisted in a logit model where

$$\text{logit}(P_{jd}) = \log\left(\frac{P_{jd}}{1 - P_{jd}}\right) = \mathbf{x}_{jd}^{\top}\boldsymbol{\beta}. \tag{4}$$

The GREG estimator of the total number of unemployed in the dth area, assisted in model (4), is now given by

$$\hat{Y}_d^{LogitGREG} = \sum_{j=1}^{N_d} \frac{e^{\mathbf{x}_{jd}^{\top}\hat{\boldsymbol{\beta}}}}{1 + e^{\mathbf{x}_{jd}^{\top}\hat{\boldsymbol{\beta}}}} + \frac{N_d}{\hat{N}_d} \sum_{j=1}^{n_d} w_{jd} \left(y_{jd} - \frac{e^{\mathbf{x}_{jd}^{\top}\hat{\boldsymbol{\beta}}}}{1 + e^{\mathbf{x}_{jd}^{\top}\hat{\boldsymbol{\beta}}}} \right).$$

Usually, $\boldsymbol{\beta}$ is estimated by maximum likelihood (ML) using Fisher or Newton-Raphson algorithms (see for example, McCullagh and Nelder, 1989).

Estimators may be also assisted in mixed models where area random effects are considered. Assuming that $y_{jd}|u_d \sim Bernoulli(P_{jd})$ where u_d is the area random effect with $u_d \sim N(0, \sigma_u^2)$, a logit mixed model takes the form

$$\text{logit}(P_{jd}) = \log\left(\frac{P_{jd}}{1 - P_{jd}}\right) = \mathbf{x}_{jd}^{\top}\boldsymbol{\beta} + u_d. \tag{5}$$

The model fitting was carried out maximizing an approximation of the likelihood integrated over the random effects using adaptive Gaussian quadrature.

The GREG estimator assisted in model (5) of the total number of unemployed in the dth area takes the form

$$\hat{Y}_d^{LogitMixedGREG} = \sum_{i=1}^{N_d} \frac{e^{\mathbf{x}_{jd}^{\mathsf{T}}\hat{\boldsymbol{\beta}}+\hat{u}_d}}{1+e^{\mathbf{x}_{jd}^{\mathsf{T}}\hat{\boldsymbol{\beta}}+\hat{u}_d}} + \frac{N_d}{\hat{N}_d}\sum_{j=1}^{n_d} w_{jd}\left(y_{jd} - \frac{e^{\mathbf{x}_{jd}^{\mathsf{T}}\hat{\boldsymbol{\beta}}+\hat{u}_d}}{1+e^{\mathbf{x}_{jd}^{\mathsf{T}}\hat{\boldsymbol{\beta}}+\hat{u}_d}}\right).$$

The choice between fixed or random effects when using models is not a trivial task neither from a theoretical nor from a practical point of view. In principle, if there are a sensible number of small areas and one expects a different behaviour from them, it makes sense to introduce random effects to avoid model overparameterization. An interesting discussion about the use of fixed or random effects models in a real context is given by Militino *et al.*(2007a).

Table 1: *Mean of the absolute relative bias (MARB), and mean of the relative root mean square error (MRMSE) for the post-stratified, synthetic, and direct estimators evaluated in the eight groups of auxiliary variables.*

		Accuracy Measures of Design-Based Estimators					
		Males		Females		Total	
		MARB	MRMSE	MARB	MRMSE	MARB	MRMSE
Poststratified	E	1.119	47.695	2.022	38.778	1.377	30.964
	EC	6.770	44.306	5.709	36.311	6.152	29.737
	EN	1.521	48.343	3.095	38.445	2.277	30.966
	EP	9.292	44.320	6.178	36.718	7.491	29.855
	ES	2.665	47.862	3.350	39.790	3.039	31.444
	ESC	17.380	45.021	11.367	37.627	13.919	31.656
	ESN	4.562	48.221	5.634	39.322	5.150	31.330
	ESP	17.620	45.703	12.526	38.048	14.688	32.136
Synthetic	E	17.246	22.248	13.585	17.956	13.451	16.811
	EC	12.114	17.867	12.426	16.482	10.746	14.265
	EN	17.869	22.778	13.064	17.679	13.589	16.949
	EP	13.645	18.900	11.526	15.646	10.765	14.171
	ES	6.151	22.312	8.017	18.962	6.022	15.451
	ESC	7.941	22.063	8.471	18.679	6.741	15.283
	ESN	5.419	22.184	8.158	19.150	5.973	15.504
	ESP	7.986	21.791	7.476	18.251	7.534	15.313
Direct		1.043	47.307	1.770	38.789	1.232	30.828

2.3 Model-based estimators

Model-based estimators are essentially based and derived from models. To estimate in a particular area, the models "borrow strength" from other related areas, improving the quality and efficiency of the estimation procedure. In this regard, many classical inferential tools are available in small-area estimation (SAE). Frequently, the goal in SAE is to obtain the best linear unbiased prediction (BLUP) estimators. The BLUP estimators minimize the mean squared error (MSE) among the class of linear unbiased estimators. These estimators usually depend on the covariance matrix of the random effects that can be estimated by several methods as maximum likelihood, restricted maximum likelihood or the method of fitting of constants. When we estimate the variance components and plug these values in the BLUP estimator, the resulting estimator is called empirical BLUP or EBLUP (see for instance Rao, 2003, pp. 95).

Both, model-based and model-assisted estimators use models. However, model-assisted estimators are built to produce design-consistent estimators because they are derived under the design-based theory, while model-based estimators are developed under the prediction theory (see, for instance, Valliant *et al.*, 2000). This means that the sampling scheme is usually ignored in the model-based perspective. There are some recent attempts in the literature to introduce sampling weights in model-based estimators (see, for example, You and Rao, 2002; Militino *et al.*, 2006, Militino *et al.*, 2007b). In summary, model-assisted and model-based procedures produce competitor estimators that are used by statistical offices, and both have mixed reviews, as one may find fans and detractors of both procedures in the literature. An important key-point is the different way of calculating the mean squared prediction error.

The model-based theory, called prediction theory, considers $y_1, \ldots, y_N$ as realizations of the random variables $Y_1, \ldots, Y_N$. Splitting the population of area d in sample (s_d) and non-sample units (r_d), the total of Y in area d, called T_d, can be expressed as

$$T_d = \sum_{j \in s_d} y_{jd} + \sum_{j \in r_d} y_{jd}$$

The task of estimating T_d becomes one of predicting the value of $\sum_{j \in r_d} y_{jd}$ for the non-observed variable $\sum_{j \in r_d} Y_{jd}$, and therefore the estimator is written as the sum of the sample and predicted observations

$$\hat{T}_d = \sum_{j \in s_d} y_{jd} + \sum_{j \in r_d} \hat{Y}_{jd}.$$

If the sampling fractions are negligible, the above predictor can be written as

$$\hat{T}_d = \sum_{j=1}^{N_d} \hat{Y}_{jd}.$$

To obtain estimators under the prediction theory, different models may be used. In this paper, linear models, logit models, and some mixed models have been considered. When the linear model (3) is assumed, and the sampling fractions are negligible, the predictor of the total number of unemployed in area d is given by

$$\hat{T}_d^{linear} = \sum_{j=1}^{N_d} \hat{Y}_{jd} = \mathbf{X}_d \hat{\boldsymbol{\beta}} \tag{6}$$

where $\mathbf{X}_d = (X_{d1}, X_{d2}, \ldots, X_{dp})^{\mathsf{T}}$ is the total population vector of the p covariates in area d. Alternative estimators can be obtained depending on the use of sampling weights to estimate $\boldsymbol{\beta}$, and the inclusion of the areas as random or fixed effects in the model. In this section, fixed effects are also considered because the reduced number of small areas in Navarre produces a lack of significance of the variance component of the random effects in some models. The following alternative estimators based on linear models are considered in the Monte Carlo study

(a) A synthetic estimator, called Linear Synthetic, assuming that in model (3), $\epsilon_{jd} \sim N(0, \sigma^2)$.

(b) A synthetic estimator, called Linear Synthetic W, based on a weighted linear model where $\epsilon_{jd} \sim N(0, \sigma^2/w_{jd})$.

(c) An estimator, called Linear F, based on a linear model with an area fixed effect, and $\epsilon_{jd} \sim N(0, \sigma^2)$.

(d) An estimator, called Linear WF, based on a weighted linear model with an area fixed effect similar to model (c), but also assuming that $\epsilon_{jd} \sim N(0, \sigma^2/w_{jd})$.

When the logit model (4) is assumed, the estimator of the total number of unemployed in the dth area is given by

$$\hat{T}_d^{logit} = \sum_{j=1}^{N_d} \frac{e^{\mathbf{x}_{jd}^{\mathsf{T}} \hat{\boldsymbol{\beta}}}}{1 + e^{\mathbf{x}_{jd}^{\mathsf{T}} \hat{\boldsymbol{\beta}}}}. \tag{7}$$

The following alternative models are also considered:

(e) A synthetic estimator, called Logit Synthetic, based on a logit model that does not incorporate sampling weights in the estimation process.

(f) A synthetic estimator, called Logit Synthetic W, based on a weighted logit model similar to (e) but including weights w_{jd} when estimating $\boldsymbol{\beta}$.

(g) An estimator, called Logit F, based on a logit model with an area fixed effect that estimates $\boldsymbol{\beta}$ without sampling weights.

Table 2: *Mean of absolute relative bias (MARB), and mean of the relative root mean square error (MRMSE) for the composite design-based estimators evaluated in the 8 groups of auxiliary variables and the direct estimator for comparison purposes.*

		Accuracy Measures of the Design-Based Estimators					
		Males		Females		Total	
		MARB	MRMSE	MARB	MRMSE	MARB	MRMSE
Composite 1 ($\alpha = 2/3$)	E	1.258	47.030	1.776	38.104	1.099	30.424
	EC	6.221	43.683	5.386	35.673	5.762	29.199
	EN	1.024	47.743	2.751	37.880	1.921	30.477
	EP	8.762	43.634	5.814	36.106	7.107	29.310
	ES	2.624	47.299	3.148	39.141	2.853	30.944
	ESC	16.956	44.488	11.072	37.024	13.606	31.165
	ESN	4.250	47.727	5.323	38.779	4.900	30.894
	ESP	17.242	45.121	12.175	37.461	14.384	31.646
Composite 2 ($\alpha = 1$)	E	1.447	43.943	1.233	35.799	1.047	28.641
	EC	5.174	40.808	4.442	33.337	4.830	27.327
	EN	1.119	44.657	2.170	35.619	1.300	28.721
	EP	7.558	40.686	4.836	33.801	6.156	27.484
	ES	2.676	44.604	2.577	37.063	2.470	29.371
	ESC	15.836	42.124	9.987	34.887	12.583	29.467
	ESN	3.891	45.158	4.517	36.799	4.306	29.384
	ESP	16.355	42.687	11.127	35.349	13.473	29.983
Composite 3 ($\alpha = 1.5$)	E	5.668	33.199	4.089	27.074	4.116	21.754
	EC	3.099	30.150	4.088	24.984	3.277	20.125
	EN	5.470	33.718	3.558	26.895	3.644	21.762
	EP	3.579	29.847	2.925	25.081	2.892	20.035
	ES	3.397	35.194	3.105	29.120	2.858	23.042
	ESC	11.620	33.272	6.644	27.448	8.901	23.004
	ESN	3.717	35.656	3.524	29.021	3.654	23.111
	ESP	12.716	33.759	7.731	27.705	9.990	23.398
Composite 4 ($\alpha = 2$)	E	8.562	27.675	6.429	22.551	6.450	18.567
	EC	4.447	24.432	5.398	20.637	4.636	16.604
	EN	8.570	28.084	5.933	22.301	6.121	18.534
	EP	4.565	24.127	4.812	20.425	4.301	16.425
	ES	4.086	30.124	4.333	24.850	3.586	19.795
	ESC	9.062	28.586	5.686	23.541	6.757	19.704
	ESN	4.101	30.460	4.200	24.800	3.860	19.872
	ESP	10.487	28.967	5.810	23.636	7.954	20.004
Direct		1.043	47.307	1.770	38.789	1.232	30.828

(h) An estimator, called Logit WF, based on a weighted logit model with an area fixed effect but including the sampling weights w_{jd} in the estimation process of $\boldsymbol{\beta}$.

(i) An estimator, called EB Mixed Logit, based on a weighted logit model with an area random effect.

The EB Mixed Logit estimator of the total number of unemployed in the dth area is given by

$$\hat{T}_d^{logitMixed} = \sum_{j=1}^{N_d} \frac{e^{\mathbf{x}_{jd}^{\mathsf{T}}\hat{\boldsymbol{\beta}}+\hat{u}_d}}{1+e^{\mathbf{x}_{jd}^{\mathsf{T}}\hat{\boldsymbol{\beta}}+\hat{u}_d}}.$$

To produce real small-area estimates, official statistical agencies must adjust the small-area estimates to make them coherent with more accurate values at some level

Table 3: *Mean of the absolute relative bias (MARB), and mean of the relative root mean square error (MRMSE) for model-assisted estimators evaluated in the 8 groups of auxiliary variables and the direct estimator for comparison purposes.*

		Accuracy Measures of the Model-Assisted Estimators					
		Males		Females		Total	
		MARB	MRMSE	MARB	MRMSE	MARB	MRMSE
Linear GREG	E	1.011	46.713	1.733	38.126	1.149	30.236
	EC	0.445	43.395	1.724	36.271	0.865	28.998
	EN	0.960	46.686	1.710	38.151	1.129	30.265
	EP	0.870	42.793	1.347	35.975	0.929	28.608
	ES	0.993	46.192	0.983	37.543	0.582	29.893
	ESD	2.010	42.514	1.346	35.480	0.761	28.387
	ESN	1.050	46.162	0.989	37.565	0.564	29.909
	ESP	1.040	42.094	1.150	35.319	1.319	28.209
Logit GREG	E	1.011	46.713	1.733	38.126	1.149	30.236
	EC	0.442	43.616	1.772	36.356	0.948	29.169
	EN	1.006	46.657	1.714	38.138	1.130	30.261
	EP	0.749	42.962	1.345	36.076	0.983	28.819
	ES	0.940	46.531	1.708	37.946	1.034	30.087
	ESD	1.206	43.253	2.007	35.920	1.499	28.859
	ESN	0.963	46.478	1.702	37.963	1.045	30.103
	ESP	3.459	42.909	2.443	35.847	2.604	28.754
Mixed Logit GREG	E	1.005	46.732	1.706	38.158	1.116	30.244
Direct		1.043	47.307	1.770	38.789	1.232	30.828

Table 4: *Mean of the absolute relative bias (MARB), and mean of the relative root mean square error (MRMSE) for the linear model-based estimators evaluated in the eight groups of auxiliary variables. The direct estimator is also included for comparison purposes.*

		Accuracy Measures of the Model-Based Estimators					
		Males		Females		Total	
		MARB	MRMSE	MARB	MRMSE	MARB	MRMSE
Linear Synthetic	E	18.752	23.589	14.035	18.603	14.604	17.884
	EC	12.897	18.428	12.606	16.765	11.382	14.771
	EN	19.429	24.346	13.234	17.931	14.349	17.736
	EP	14.384	19.521	11.862	15.981	11.286	14.709
	ES	6.182	21.700	8.735	18.506	6.218	15.096
	ESC	7.726	19.985	9.338	17.788	7.367	14.315
	ESN	5.539	21.494	8.615	18.489	6.250	15.117
	ESP	7.961	20.056	7.876	17.090	7.479	14.441
Linear Synthetic W	E	17.246	22.248	13.585	17.956	13.451	16.811
	EC	12.202	17.904	12.458	16.441	10.784	14.308
	EN	17.829	22.841	12.893	17.394	13.206	16.702
	EP	13.466	18.755	11.558	15.648	10.704	14.164
	ES	6.155	21.787	8.653	18.554	6.176	15.146
	ESC	7.708	20.096	9.248	17.831	7.303	14.368
	ESN	5.510	21.580	8.529	18.534	6.215	15.169
	ESP	7.937	20.210	7.794	17.159	7.465	14.517
Linear F	E	3.613	43.794	3.815	33.671	2.215	27.742
	EC	6.817	38.015	5.339	30.544	4.062	24.984
	EN	3.736	43.743	3.857	33.716	2.256	27.757
	EP	5.580	38.522	4.792	30.988	3.552	25.310
	ES	3.645	43.192	4.267	33.555	2.346	27.410
	ESD	7.555	37.607	6.019	30.428	4.462	24.806
	ESN	3.757	43.099	4.292	33.602	2.382	27.424
	ESP	6.034	38.043	5.334	30.944	3.827	25.098
Linear WF	E	2.654	43.702	3.298	33.663	1.801	27.683
	EC	6.041	37.972	4.981	30.546	3.652	24.941
	EN	2.730	43.638	3.362	33.700	1.827	27.693
	EP	4.673	38.475	4.312	30.999	3.018	25.286
	ES	3.390	43.279	4.031	33.645	2.169	27.481
	ESC	7.218	37.702	5.797	30.534	4.206	24.874
	ESN	3.500	43.184	4.062	33.687	2.208	27.493
	ESP	5.686	38.153	5.078	31.029	3.536	25.178
Direct		1.043	47.307	1.770	38.789	1.232	30.828

Table 5: *Mean of the absolute relative bias (MARB), and mean of the relative root mean square error (MRMSE) for the logit model-based estimators evaluated in the 8 groups of auxiliary variables. The direct estimator is also included for comparison purposes.*

		Accuracy Measures of the Model-Based Estimators					
		Males		Females		Total	
		MARB	MRMSE	MARB	MRMSE	MARB	MRMSE
Logit Synthetic	E	18.752	23.589	14.035	18.603	14.604	17.884
	EC	13.797	19.060	12.448	16.948	12.382	15.339
	EN	19.181	24.073	13.336	18.038	14.422	17.772
	EP	15.536	20.364	12.043	16.316	12.322	15.382
	ES	6.046	22.259	7.971	18.838	5.973	15.340
	ESC	7.340	21.463	8.351	18.315	6.555	14.909
	ESN	5.526	22.067	8.110	18.861	5.997	15.363
	ESP	7.823	21.481	7.319	17.834	7.297	15.068
Logit Synthetic W	E	17.247	22.248	13.585	17.957	13.451	16.811
	EC	12.885	18.471	12.254	16.535	11.570	14.743
	EN	17.598	22.595	12.966	17.493	13.279	16.736
	EP	14.402	19.518	11.712	15.899	11.437	14.697
	ES	6.056	22.253	7.980	18.839	5.982	15.342
	ESC	7.350	21.460	8.351	18.313	6.555	14.909
	ESN	5.527	22.061	8.129	18.864	6.011	15.369
	ESP	7.833	21.485	7.328	17.839	7.299	15.069
Logit F	E	1.410	46.985	1.755	38.242	1.174	30.481
	EC	1.385	44.263	2.003	36.749	1.120	29.434
	EN	1.411	46.948	1.711	38.282	1.152	30.516
	EP	1.917	44.010	1.664	36.414	1.182	29.161
	ES	0.815	46.880	1.717	38.351	1.036	30.377
	ESC	1.168	44.305	1.789	36.843	1.029	29.413
	ESN	0.815	46.820	1.690	38.389	1.022	30.404
	ESP	1.350	44.034	1.477	36.565	0.981	29.153
Logit WF	E	1.059	46.879	1.842	38.271	1.179	30.410
	EC	1.130	44.315	1.829	36.783	0.922	29.432
	EN	1.042	46.835	1.816	38.310	1.162	30.445
	EP	1.636	43.978	1.567	36.472	1.046	29.150
	ES	0.942	47.038	1.738	38.504	1.077	30.469
	ESC	1.893	44.177	2.131	36.824	1.488	29.364
	ESN	0.954	46.979	1.721	38.538	1.066	30.497
	ESP	4.222	43.623	2.583	36.462	2.644	29.034
EB mixed logit	E	28.909	33.919	17.840	23.217	16.279	21.109
Direct		1.043	47.307	1.770	38.789	1.232	30.828

of aggregation. This adjustment is necessary because when aggregating small-area estimates within the same region (province), the sum of these small-area estimates do not generally coincide with the estimate obtained using an appropriate estimator for the larger region. This adjustment process is called benchmarking and in this study it should be done to the provincial estimate. Ugarte *et al.* (2008) show how to introduce constraints into a linear mixed model to produce final benchmarked estimates in small areas.

3 Monte Carlo simulation study

In this section the performance of the estimators described in Section 2 to estimate unemployment in the small areas of Navarre is evaluated. We have drawn $K = 500$ samples from the 2001 Census following the same sampling design as the SLFS. To assess the estimators bias the mean absolute relative bias ($MARB$) over the D small areas is computed. Namely

$$MARB(\hat{Y}) = \frac{1}{D}\sum_{d=1}^{D}|RB_d(\hat{Y})| \quad \text{where} \quad RB_d(\hat{Y}) = \frac{1}{K}\sum_{k=1}^{K}\frac{\hat{Y}_d(k) - Y_d}{Y_d}100.$$

To evaluate the estimators prediction errors the mean of the square root of the relative mean squared error ($MRMSE$) over the D small areas is considered

$$MRMSE(\hat{Y}) = \frac{1}{D}\sum_{d=1}^{D}RMSE_d(\hat{Y}) \quad \text{where} \quad RMSE_d(\hat{Y}) = \left(\frac{1}{K}\sum_{k=1}^{K}\left(\frac{\hat{Y}_d(k) - Y_d}{Y_d}\right)^2\right)^{\frac{1}{2}}100.$$

The estimator giving a better balance between bias and prediction error will be selected as the best proposal for estimating unemployment in the small areas of Navarre. The expertise of the local statistical office has also been considered and this aspect will be discussed at the end of this section.

Now, let us recall that there are a total of seven small areas in Navarre for which we have evaluated the seven design-based estimators (including the direct), three GREG estimators, and nine model-based estimators. A total of eight combinations of auxiliary variables have been used: (E) age-sex, (EC) age-sex-claimant, (EN) age-sex-educational level, (EP) age-sex-employment status in the SNE, (ES) age-sex-stratum, (ESC), age-sex-stratum-claimant status in the SNE, (ESN) age-sex-stratum-educational level, and (ESP) age-sex-stratum-employment status in the SNE.

Tables 1 and 2 display the MARB and the MRMSE of the design-based estimators for the eight combinations of auxiliary variables. In general, the post-stratified estimator exhibits neither large biases nor low predictions errors as it is expected because it is a direct estimator. However, both post-stratified estimators ESC and ESP present biases around 17% (for males), and around 14% (for females). The synthetic estimators E, EC, EN, and EP present large biases (for both males and females), indicating that the assumption of a similar behaviour of the small areas with respect to the large region does not hold. We also observe that the bias is small when the stratum (S) auxiliary variable is considered. The behaviour of composite 1, composite 2, and composite 3 estimators (see Table 2) is similar to the post-stratified estimator, and they have, in general, small biases, except ESC and ESP that present larger biases. This is not surprising because, as α is not very large, these composite estimators give more weight to the post-stratified component. This also means that the error is high. Composite 4 estimator is slightly more biased than the rest of composite estimators, but its error is reduced up to a half. Then, it might be a good option to achieve a compromise between bias and error.

Table 3 shows the MARB and the MRMSE for the model-assisted estimators (GREG). All of them exhibit practically the same results. They have negligible bias, but the error is not reduced with regard to the direct estimator. Note that the linear GREG estimator is a modified direct estimator, and then it is approximately unbiased. However, it does not increase the effective sample size (see Rao, 2003, chapter 2, p. 20), and then, the error is not decreased. The mixed logit GREG suffers from the same deficiencies as both the linear and the logit GREG.

The results for the model-based estimators are displayed in Tables 4 and 5. The model-based synthetic estimators (linear and logit) are very similar with regard to bias and error. When the auxiliary variable stratum (S) is not considered, the bias is large, but it is notably reduced when stratum is introduced in the models. The error has been decreased, and it is about one half the error of the direct estimator. Those model-based estimators making use of a fixed area effect are practically unbiased, but the error is unacceptably high (as large as the error of the direct estimator).

For an optimal choice between the estimators, we first select the estimator with a smaller MARB – excluding, of course, the classical direct estimator that has an enormous MSE in small areas – and later we select those with a smaller MRMSE. It does not seem reasonable to use just the MRMSE as the single option to choose a sensible estimator, because the statistical office will be reluctant to use an estimator with a large bias. They would agree to accept some bias to reduce variability when estimating in small areas, but not a big amount of bias, because traditionally they are used to work with unbiased estimators. The choice among estimators is then not easy, however, looking at Tables 1, 2, 3 and 4 – the synthetic estimators with auxiliary variables (EC, EN, ES, ESC, ESN), the composite 4 estimators with auxiliary variables (EC and EP), and the linear synthetic estimators with auxiliary variables (ES and ESN) outperform the rest. Figure 2 shows the accuracy measures for these estimators. Unfortunately, the performance of the model-based estimators is not as promising as it should be. In particular, those model-based estimators introducing a fixed area effect exhibit errors as large as the direct estimator. The bias is also too big when introducing a random area effect (as in the EB mixed logit). This error is reduced if the area effect is removed, but in this situation, the bias is large if the auxiliary variable stratum is not included in the model. In this case, the performance of the model-based estimators is pretty similar to the design-based synthetic estimator. A different behaviour is observed for the post-stratified and the composite 1, 2, and 3 estimators. In general, they exhibit low bias (except in two cases with the auxiliary variable stratum), but they present large errors. The best balance between bias and error is achieved with composite 4 estimator. More precisely, the composite 4 EP. The bias is around 5% and the error is one half the error of the direct estimator. Therefore, we consider it a reasonable estimator for estimating unemployment in the small areas of Navarre. In addition, this estimator is very appealing to the local statistical office as it combines two sources of measuring unemployment: namely, data from the SLFS and data from the Navarre Employment Register.

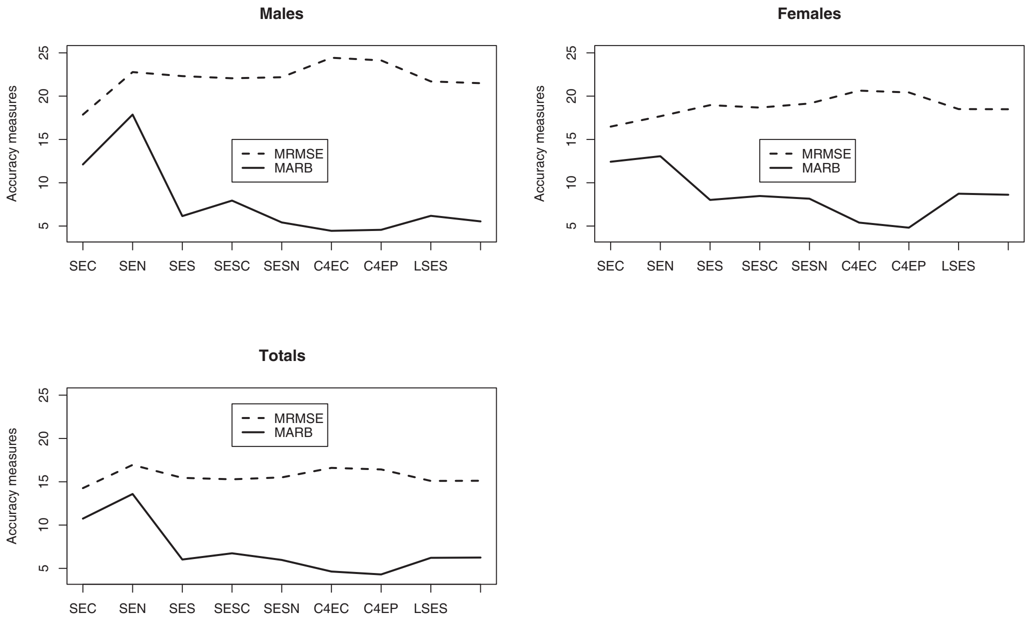

Figure 2: *Mean of the absolute relative bias (MARB) and mean of the relative square error (MRMSE) of the synthetic estimator with auxiliary variables (EC, EN, ES, ESC, ESN), Composite 4 estimator with auxiliary variables (EC and EP) and Logit Synthetic estimator with auxiliary variables (ES, ESN) for males, females, and totals over the seven small areas of Navarre*

4 Estimators of the mean squared error

The main interest of this article is to choose an appropriate estimator to estimate unemployment in very small areas in the province of Navarre, where sometimes theoretical assumptions are not well fulfilled, and practical performance needs to be explored. In the last section composite 4 EP has been chosen as the appropriate estimator, and so this section is devoted to the estimation of its mean squared error. Here, three alternative MSE estimators are derived using three well-known procedures in the literature: the variance linearization method, and two resampling methods, the jackknife and the bootstrap.

4.1 The variance linearization method

The variance linearization method, or delta method, consists of applying a Taylor series expansion to a function of the estimators of the total, and calculating the variance of this function through the variance of its derivatives with regard to these totals (Woodruff, 1971). Let us define the following indicator variables $I_k(h,i,j) = 1$ if person j ($j = 1,\dots,m_{hi}$) of household i, ($i = 1,\dots,n_h$) and stratum h, ($h = 1,\dots,H$) is in group k, and 0 otherwise, $z_{hij} = y_{jd} I_k(h,i,j)$ and $v_{hij} = w_{jd} I_k(h,i,j)$.

Post-stratified and synthetic estimators of the mean of Y_d can be written as

$$\widehat{\bar{Y}_d}^k = \left(\sum_{h=1}^{H}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}} v_{hij} z_{hij}\right) / v_{\dots} \text{, where} \qquad v_{\dots} = \sum_{h=1}^{H}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}} v_{hij}, \tag{8}$$

because when k is the group g in the domain d, $\bar{Y}_d^k$ is the post-stratified estimator of the mean (see expression (1)), and when $k = g$, $\bar{Y}_d^k$ is the synthetic estimator of the mean (see expression (2)). For both of them the linearized estimator of the variance is the following

$$\widehat{\text{Var}}_L(\widehat{\bar{Y}_d}^k) = \textstyle\sum_{h=1}^{H} \widehat{\text{Var}}_h(\widehat{\bar{Y}_d}^k), \quad \text{where} \quad \widehat{\text{Var}}_h(\widehat{\bar{Y}_d}^k) = \frac{n_h}{n_h-1}\sum_{i=1}^{n_h}\left(U_{hi.} - \bar{U}_{h..}\right)^2,$$
$$U_{hi.} = \frac{1}{v_{\dots}}\textstyle\sum_{j=1}^{m_{hi}} v_{hij}\left(z_{hij} - \widehat{\bar{Y}_d}^k\right), \quad \text{and} \quad \bar{U}_{h..} = \frac{1}{n_h}\sum_{i=1}^{n_h} U_{hi.}. \tag{9}$$

The estimator of the total of Y in area d, is given by

$$\widehat{Y}_d^k = \sum_{g=1}^{G} \widehat{\bar{Y}_d}^k N_{dg}. \tag{10}$$

Then, the variance linearized estimator is calculated as a weighted sum such that

$$\widehat{\text{Var}}_L(\hat{Y}_d^k) = \sum_{g=1}^{G} \widehat{\text{Var}}_L(\widehat{\bar{Y}_d}^k) N_{dg}^2. \tag{11}$$

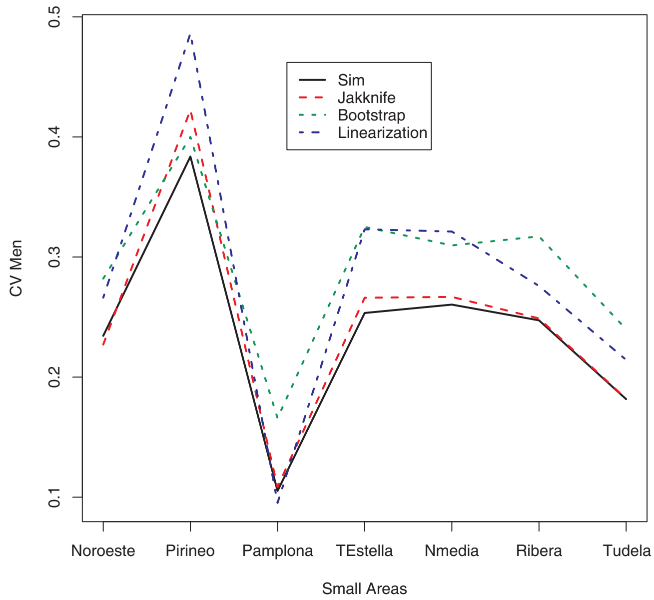

Figure 3: *Coefficient of variation of composite 4 EP (males)*

To calculate the bias of the synthetic estimator given by $\text{Bias}(\hat{Y}_d^{sint}) = -\sum_{j=1}^{N_d} \epsilon_{jd}$, the following estimator is provided (see Ghosh and Särndal, 2001)

$$\widehat{\text{Bias}}(\hat{Y}_d^{sint}) = -N_d \frac{1}{n_d} \sum_{j=1}^{n_d} \hat{\epsilon}_{jd}, \qquad \text{where} \qquad \hat{\epsilon}_{jd} = y_{jd} - \widehat{\bar{Y}}_g. \tag{12}$$

Therefore

$$\widehat{MSE}_L(\hat{Y}_d^{sint}) = \widehat{\text{Var}}_L(\hat{Y}_d^{sint}) + \widehat{Bias}^2(\hat{Y}_d^{sint}),$$

and finally

$$\widehat{MSE}_L(\hat{Y}_d^{comp}) \doteq \lambda_d^2 \widehat{MSE}_L\left(\hat{Y}_d^{post}\right) + (1-\lambda_d)^2 \widehat{MSE}_L\left(\hat{Y}_d^{sint}\right). \tag{13}$$

Note that the $\widehat{MSE}_L\left(\hat{Y}_d^{post}\right)$ is calculated using expression (11) for k defined as a combination of groups of g and d, and the covariance between the post-stratified and the synthetic estimator has been considered negligible.

4.2 The jackknife estimator

The jackknife method was introduced by Quenouille (1949, 1956) as a method to reduce the bias, and later Tukey (1958) proposed its use for estimating variances and confidence intervals. In the jackknife method, we take as many sub-samples as clusters (census sections) are in the sample, because sub-samples are obtained leaving one cluster out every time from the original sample. Let $\hat{Y}^k_{d(hi)}$ be the estimator $\hat{Y}^k_d$ obtained by dropping a cluster i from the hth stratum. Then, original weights w_{jd} must be substituted by $w_{jd(hi)}$, where

$$w_{jd(hi)} = \begin{cases} w_{jd} & \text{if } j \text{ is not in the } h \text{ stratum} \\ 0 & \text{if } j \text{ is in cluster } i \text{ of the } h \text{ stratum} \\ \frac{n_h}{n_h-1} w_{jd} & \text{if unit } j \text{ is in the } h \text{ stratum but not in cluster } i \end{cases} \quad (14)$$

The jackknife estimator of the MSE of $\hat{Y}^k_d$ can be obtained as

$$\widehat{MSE}_{JK}(\hat{Y}^k_d) \doteq \sum_{h=1}^{H} \frac{n_h - 1}{n_h} \sum_{i=1}^{n_h} [\hat{Y}^k_{d(hi)} - \hat{Y}^k_{d(h.)}]^2, \quad (15)$$

where $\hat{Y}^k_{d(h.)} = \frac{1}{n_h} \sum_{i=1}^{n_h} \hat{Y}^k_{d(hi)}$ and superscript k indicates the composite estimator. Note that Expression (15) is approximate as we are ignoring the possible bias of the composite estimator.

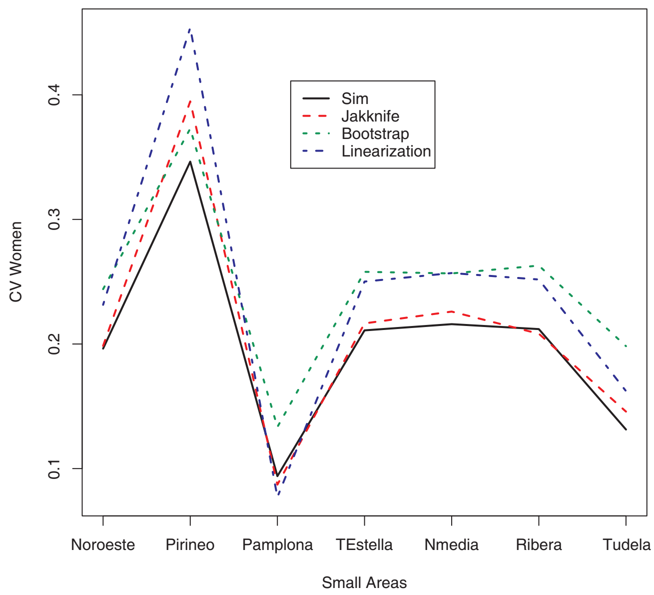

Figure 4: *Coefficient of variation of composite 4 EP (females)*

4.3 The bootstrap estimator

In the bootstrap, the sub-samples are obtained by random sampling with replacement. Similarly to the jackknife, for every sub-sample new weights are defined in each step. The rescaled bootstrap estimator in a stratified random sampling has been provided by Rao and Wu (1988). It assumes the following steps

1. Given the h stratum we have a sample of n_h clusters. From the sample of the h stratum, a sub-sample of $n_h - 1$ clusters is drawn by random sampling with replacement.
2. For every sub-sample r ($r = 1, 2, \ldots, R$) a new weight is defined
$$w_{jd}(r) = w_{jd} \frac{n_h}{n_h - 1} m_i(r), \tag{16}$$
where $m_i(r)$ is the number of times that cluster i is chosen in the sub-sample.
3. Repeat steps (1) and (2) R times.

To derive the bootstrap estimator of the MSE of $\hat{Y}_d^k$ we calculate

$$\widehat{MSE}_B(\hat{Y}_d^k) = \frac{1}{R-1} \sum_{r=1}^{R} \left(\hat{Y}_{rd}^{*k} - \hat{Y}_d^k\right)^2, \tag{17}$$

where $\hat{Y}_{rd}^{*k}$ is similar to (10) but using the new weight $w_{jd}(r)$ when estimating the mean (8) and k indicates the composite estimator.

In the Monte Carlo study the composite 4 EP estimator has been computed with 500 samples. For the bootstrap mean squared error estimator, R has taken the values $200, 500, 1000$, and 4000. Small values of R lead to differences in the estimator performance, but values of R equal to 1000 and higher provide similar results. Both Figures 3 and 4 show the coefficients of variation for males and females respectively. The coefficient of variation obtained from the census data is depicted using a continuous line. All of the methods proposed here tend to overestimate the MSE, particularly when the sample size is small. However, the best behaviour corresponds to the jackknife estimator, because the corresponding coefficients of variation are very close to the real ones.

Conclusions

Small area estimation is becoming a challenge in European statistical offices because of the increasing demand of precise estimates at county or regional level. Unfortunately, procedures used in other regions and/or countries seem not to be directly applicable everywhere, because they are based on a large number of small areas, and the availability of the auxiliary information is not the same for every country. In some regions such as Navarre, the task of estimating unemployment in very small areas is not easy, not only because of the reduced number of small areas and the great heterogeneity between them, but also because unemployment has a low incidence in the population. Both reasons can worsen the performance of model-assisted and model-based estimators, which were promising in other scenarios.

In this work a composite 4 EP estimator has shown to be a reasonable alternative for estimating unemployment in Navarre. This estimator comes from a linear combination of a direct estimator (a poststratified estimator) and an indirect estimator (a synthetic estimator). The accuracy measures evaluated through a Monte-Carlo study have shown its good trade-off in terms of bias and MSE. This estimator is easy to calculate and interpret, and the MSE can be derived using jackknife. The composite 4 estimator uses as auxiliary information the age-sex (E) groups and the employment register in Navarre (P). Although it is known that this later register might overestimate unemployment, the combination of the two sources of data to estimate unemployment (namely, the SLFS and the Navarre employment register) is very appealing to the local statistical office.

Acknowledgments

We thank a referee for the helpful comments that have contributed to improving an earlier version of this paper. This research has been supported by the Spanish Ministry of Science and Education (project MTM2005-00511) and the Spanish Ministry of Science and Innovation (MTM2008-03085).

References

Costa, A., Satorra, A. and Ventura, E. (2002). Estimadores compuestos en estadística regional: Aplicación para la tasa de variación de la ocupación en la industria, *Qüestió*, 26, 213-243.

Costa, A., Satorra, A. and Ventura, E. (2003). An empirical evaluation of small area estimators. *SORT (Statistics and Operations Research Transactions)*, 27, 113-135.

Costa, A., Satorra, A. and Ventura, E. (2004). Using composite estimators to improve both domain and total area estimation. *SORT (Statistics and Operations Research Transactions)*, 28, 69-86.

Costa, A., Satorra, A. and Ventura, E. (2006). Improving small area estimation by combining surveys: new perspectives in regional statistics. *SORT (Statistics and Operations Research Transactions)*, 28, 69-86.

Drew, D, Singh, M. P. and Choudhry, G. H. (1982). Evaluation of small area estimation techniques for the Canadian labor Force Survey. *Survey Methodology*, 8, 17-47.

Ghosh, M. and Rao, J. N. K. (1994). Small area estimation: an appraisal. *Statistical Science*, 9, 55-93.

Ghosh, N. and Särndal, C. E. (2001). *Lecture Notes on Estimation for Population Domains and Small Areas*. Statistics Finland, vol. 48.

González, M. E. (1973). Use and evaluation of synthetic estimates. *Proceedings of the Social Statistics Section* 33-36. American Statistical Association. Washington, D.C.

McCullagh P. and Nelder, J. A. (1989). *Generalized Linear Models*. Chapman and Hall.

Militino A. F.,Ugarte, M. D., Goicoa, T. and González-Audícana, M. (2006). Using small area models to estimate the total area occupied by olive trees. *Journal of Agricultural, Biological, and Environmental Statistics*, 11, 450-461.

Militino A. F., Ugarte, M. D. and Goicoa, T. (2007a). A BLUP synthetic versus an EBLUP estimator: an empirical study of a small area estimation problem. *Journal of Applied Statistics*, 34, 153-165.

Militino A. F., Ugarte, M. D. and Goicoa, T. (2007b). Combining sampling and model weights in agricultural small area estimation. *Environmetrics*, 18, 87-99.

Morales, D., Esteban, M. D., Sánchez, A. Santamaría, L., Marhuenda, Y., Pérez, A., Saralegui, J. and Herrador, M. (2007). Estimación en áreas pequeñas con datos de la Encuesta de Población Activa en Canarias. *Estadística Española*, 49, 301-332.

Quenouille, M. (1949). Approximate tests of correlation in time series. *Journal of the Royal Statistical Society Series B*, 11, 18-84.

Quenouille, M. (1956). Notes on bias in estimation. *Biometrika*, 43, 353-360.

Rao, J. N. K. (2003). *Small Area Estimation*. Wiley Series in Survey Methodology.

Rao, J. N. K. and Wu, C. F. J. (1988). Resampling inference with complex survey data. *Journal of the American Statistical Association*, 83, 231-241.

Särndal, C. E., Swensson B. and Wretman, J. H. (1992). *Model Assisted Survey Sampling.* Springer.

Tukey, J. (1958). Bias and confidence in not quite large samples (abstract). *Annals of Mathematical Statistics*, 29, 614.

Ugarte, M. D., Militino, A. F. and Goicoa, T. (2008). Benchmarked estimates in small areas using linear mixed models with restrictions, *Test*, in press. DOI 10.1007/s11749-008-0094-x

Valliant, R., Dorfman, A. H. and Royall, R. M. (2000). *Finite Population Sampling and Inference: a Prediction Approach.* Wiley Series in Survey Methodology.

Woodruff, R. S. (1971). A simple method for approximating the variance of a complicated estimate. *Journal of the American Statistical Association*, 66, 411-414

You, Y. and Rao J. N. K. (2002). A pseudo-empirical best linear prediction approach to small-area estimation using survey weights. *Canadian Journal of Statistics*, 30, 431-439.