Journal of Statistical Science and Application, February 2017, Vol. 5, No. 1-2, 16-29
doi: 10.17265/2328-224X/2017.0102.002

Estimating Poverty in Greece in Small Geographical Areas

Stefanos G. Giakoumatos
Professor of Statistics, Dept of Accounting and Finance, Technological Educational Institute of Peloponnese
E-mail: s.giakoumatos@teipel.gr
Malapani Eleni
Ph.D Candidate, Panteion University

In this paper, the Small Area methods are presented. We use this method to produce estimates of poverty in Greece, at county (NUTS3) level. This is succeeded by combining survey data from EU-SILC 2013 with auxiliary data derived from the 2011 Census. In the application section, we adopt the Fay and Herriot model and we provide estimates for the percentage of Greek population under the poverty line in 2013 and also the mean equivalized income. Using Fay and Herriot model the percentages of poverty and the mean equivalized income changed in several areas. Also, the standard deviations of the estimates were improved in all prefectures.

Keywords: Small area model, EBLUP, Fay and Herriot model, poverty, EU-SILC, Census

Introduction

In the last years many surveys have been conducted concerning the income and the living conditions of households in every country. It is widely acceptable that while those surveys examine a large number of variables, they do not produce reliable results about areas where the sample is small. On the other hand the data from the census in every decade do not cover all the time periods and many times do not examine enough variables concerning the households in order to result in the desirable estimations. Nevertheless, the combination of the above data, that is the surveys on the households and on the census, has been noted to have given reliable estimators of the variables of the households we are interested. In the present study we combine the above data types in order to derive reliable results on the poverty level in Greece on county (NUTS 3) level for the year 2012.

There are many different approaches in defining poverty as it is composite and multidimentional. One of those approaches is the objective poverty. According to this approach, individuals, families and groups in the population can be said to be in poverty when they lack the resources to obtain the types of diet, participate in the activities and have the living conditions and amenities which are customary, or at least widely encouraged or approved, in the society to which they belong" (Townsend, 1979, p 31).

Three ingredients are required in computing a poverty measure. First one has to choose the relevant dimension and indicator of well-being. Second one has to select a poverty line, that is, a threshold below which a given household or individual will be classified as poor. Finally one has to select a poverty measure to be used for reporting for the population as a whole or for a population subgroup only.

When measured, poverty may be absolute or relative. Absolute poverty refers to the set of resources a person must acquire in order to maintain a "minimum standard of living". Relative poverty is concerned with how well off an individual is with respect to others in the same society. In theory, therefore, while an absolute poverty line is a measure that could, adjusting for price fluxes, remain stable over time, a relative poverty line is one that could be expected to shift with the overall standard of living in a given society. In EU and in the most of the developed countries, the relative poverty indices are used. According to Eurostat the poverty line is calculated with its relative concept (poor in relation to others) and it is defined at 60% of the median total equivalised disposable income of the household, using modified OECD equivalised scale.

Small Area Methods and Models

Introduction

In recent years, the demand for small area estimates of some socioeconomic indicators (poverty, unemployment,…) has greatly increased worldwide. An area (domain) is regarded as small if the domain-specific sample is not large enough to support direct estimates of adequate precision. The official statistical office of the states produce estimates on a national or at the best case on a regional level. In order for these characteristics to be estimated, large samples are required on a national level (so that to provide enough sampling units on a local area and to produce reliable estimation) which equals a geometrical rise of the cost, a case that the states do not want or cannot cover. A way to resolve this problem is by using small area estimation (SAE) (Rao 2003).

Small area estimation (SAE) is widely used for producing estimates of population parameters for areas (domains) with small, or even zero, sample size. In those areas, direct estimators that only relies on domain-specific observations may lead to estimates with large sampling variability. When direct estimation is not possible, one has to rely upon indirect estimators. Indirect estimators borrow strength by using values of the variable of interest, y, from related areas and/or time periods and thus increase the effective sample size. These values are brought into the estimation process through a model (either implicit or explicit) that provides a link to related areas and/or time periods through the use of the supplementary information related to y, such as recent census database and current administrative records.

Small Area Models

Explicit linking models based on random area-specific effects that account for between area variation beyond that is explained by auxiliary variables included in the model will be called "Small Area Models". Indirect estimators based on small area models will be called "model-based estimators", (Rao, 2003). We classify small area models into two brad types:

Area level models that relate small area direct estimators to area-specific covariates. Such models are necessary if unit (or element) level data are not available, (Fay and Herriot, 1979).

Unit level models that relate the unit values of a study variable to unit-specific covariates, (Battese, Harter and Fuller, 1988)

Fay and Herriot Model

Fay and Herriot model was introduced by Fay and Herriot (1979) to obtain small area estimators of median income in some places in the United States. This model is widely used area level model in SAE, is the basic tool when only aggregated auxiliary data at the area level are available. The SAE under this model is one of the most popular method used by private and public agencies because of its flexibility in combining different sources of information and explaining different sources of errors

Fay and Herriot model uses mixed (random) effects models for SAE (F-H 1979, Battese 1988). A mixed effects model consists of a fixed effects part and a random effects part with the latter accounting for between area variations beyond that explained by the auxiliary variables included in the fixed part of the model.

We assume that $\theta_i = g(\bar{Y}_i)$ is a known function of $\bar{Y}_i$ and $z_i = \left(z_{1i}, z_{2i,}, \ldots, z_{pi}\right)$ is the known auxiliary vector for the i-th area, i=1,…,m.

The function $g(.)$ is related to area specific auxiliary data z_i, through a linear model $\theta_i = z_i^T \beta + b_i v_i$, i=1,…,m, where the b_i 's are known positive constants and $\beta = (\beta_1, \beta_2, ..., \beta_p)^T$ is the px1 vector of regression coefficients. The v_i 's are area specific random effects assumed to be independent and identically distributed (iid) with $E_m(v_i) = 0$ and $V_m(v_i) = \sigma_v^{\ 2}\ (\geq 0)$.

Normality of the random effects v_i is also often used, but it is possible to make robust inferences by relaxing the normality assumption. The parameter $\sigma_v^{\ 2}$ is a measure of homogeneity of the areas after counting for the covariates z_i.

In some applications, not all areas are selected in the sample. We assume that we have M areas in the population and only m areas are selected in the sample. We assume the population model $\theta_i = z_i^T \beta + b_i v_i$, i=1,…,M (1). We also assume that the sample areas obey the population model. We want to estimate the population mean of the i-th area. For making inferences about $\bar{Y}_i$ under model (1) we assume that:

- The direct estimators $\hat{\bar{Y}}_i$ are available
- $\hat{\theta}_i = g(\hat{\bar{Y}}_i) = \theta_i + e_i$, i=1,…,m (2) (as in the James-Stein method), where the sampling errors e_i are independent with $E_p(e_i / \theta_\iota) = 0$ and $V_p(e_i / \theta_\iota) = \psi_i$. The sampling variances, ψ_i, are known.

Combining model (1) and (2) we obtain the Fay and Herriot model $\hat{\theta}_i = z_i^T \beta + b_i v_i + e_i$, i=1,…,m. We assume that v_i and e_i are independent.

Application

Research Characteristics

In our research we make an estimation at NUTS 3 area (in Greek Nomos) using the model of Fay and Herriot. The variables of interest are:

- The percentage of Greek people under the poverty line and
- The average disposable income

We have derived the data of our sample from the EU Survey of income and Living Conditions EU-SILC 2013, while the data of the auxiliary variables from the 2011 Census database. The auxiliary variables were the following two:

- The percentage of people per county with a lower educational level (X1)

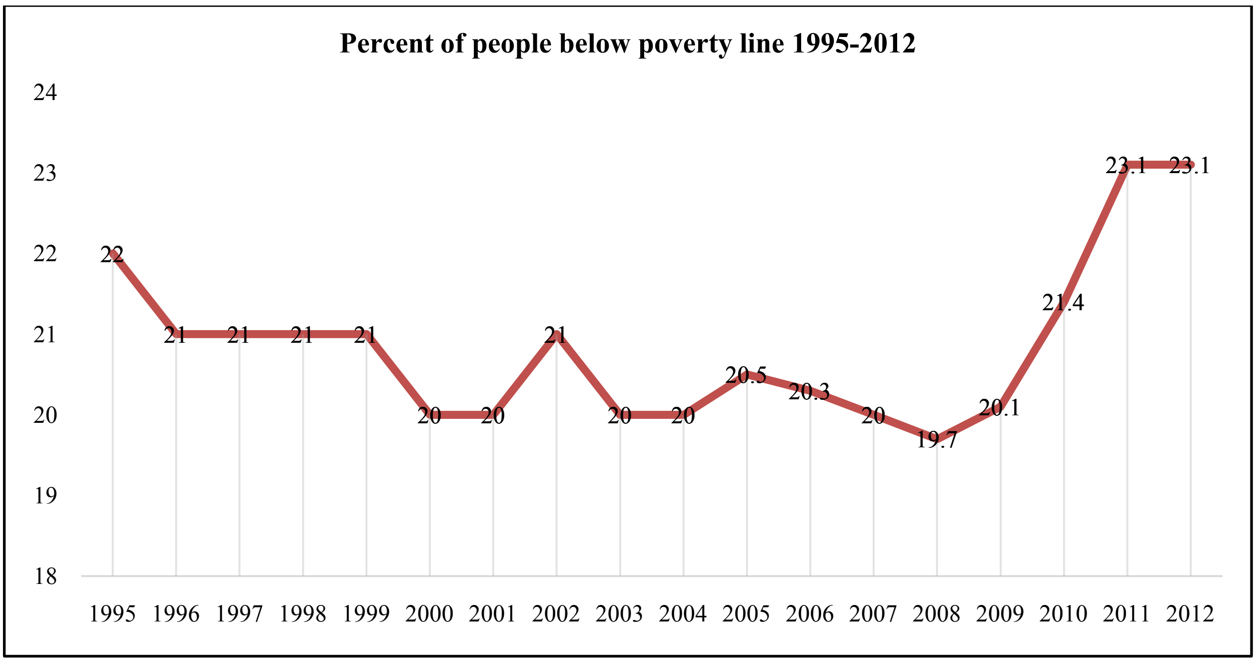
- The percentage of inactive people per NUTS 3 (individuals who are not interested in working) (X2)

We have also used the relative poverty lines. Poverty line is the level of income under which the individual is considered poor. We consider that this is (according to OECD) 60 % of the median total equivalized disposable income of the household. As an equivalent available individual income is considered the total available income of household after it has been divided with the equivalent size of the household. The equivalent size of household is calculated according to the modified scale of OECD. Equivalent size refers to OECD modified scale gives weight 1.0 to the first adult of the household, 0.5 to other persons above the age of 14 and 0.3 to every child under the age of 14 of the household. The income components included in the survey are :

- Income from work
- Income from property
- Social transfers and pensions
- Monetary transfers from other households and
- Imputed income from the use of company car

EU -SILC Greece 2013

The EU Survey of income and Living Conditions EU-SILC is conducted on an annual basis in all the EU since 2003 with the responsibility of the Eurostat. The aim of EU-SILC is to gather reliable and comparable data on the income, the living conditions, on the labour of people and of households in the EU states. Collecting the necessary data has been achieved through questionnaires answered by a representative sample of households in each member state of the EU. The year in reference in the present study is 2012, the final sample was of 7349 households and 18030 people (15318 age 16+). According to the EU-SILC the line of poverty rises in 5023 euro per person annually and the 23.1% of the total population is placed under the poverty line. The chart below demonstrates the evolution of poverty since 1995-2012.



For the production of our results we used the programming language R.

Application Results

Poverty estimates. From the results below (see Figure 1,2,3,4), we observe that using Fay and Herriot model the percentages of poverty changed in several areas, including Imathia, Kefalonia, Lasithi, Samos and Thesprotia. Also, using the model of Fay and Herriot standard deviations of the estimates were improved in all prefectures. Especially in counties with small sample (30-70) as Thesprotia, Samos, Grevena, Lefkada, Chios and Lasithi the difference in standard deviations between direct and Fay and Herriot estimator was great. In these areas the direct estimation method gave large standard deviations and the Fay and Herriot method much smaller. Features mention Thesprotia where the direct estimator gave poverty rate 57.13 % while the Fay and Herriot estimator 29.15 % with standard deviations 17.24 % and 5.92 % respectively . Finally note that in areas such as the prefecture of Attica , Thessaloniki etc. where the sample size is large , the differences between direct and Fay and Herriot estimator are too small to negligible both poverty rates and the corresponding standard deviations.

Income estimates. Respectively with poverty rates, we notice several differences in income between direct and Fay and Herriot estimator (see Figure 5,6,7,8), in regions such as Evritania , Laconia , Lefkada , Lasithi , Chios , Samos and Grevena. Even greater are the differences between the two estimators as reference standard deviations in the above areas . The standard deviations of Fay and Herriot estimator is for all counties smaller than the direct estimator.

Conclusions

In this paper, we adopt the small area methods to produce estimates for the poverty in Greece into a small geographical areas. In detail, we apply the Fay and Herriot model using as independent (auxiliary) variables data from the Census of 2011. For this first approach, we use only two independent variables in order to produce estimates, but the results are very promising. The reduction of the standard deviation of the estimates was large in many areas where the standard deviation for the direct estimates was high (see Figure 3, 4, 7, 8). Therefore, the accuracy of the estimators for the poverty was increased using the small are methods. For further research, the author plan to produce a model that include a larger set of auxiliary information for the census and other administrative sources in order to improve the accuracy and conclude to the final SAE model for the estimation of Greek Poverty Indices.

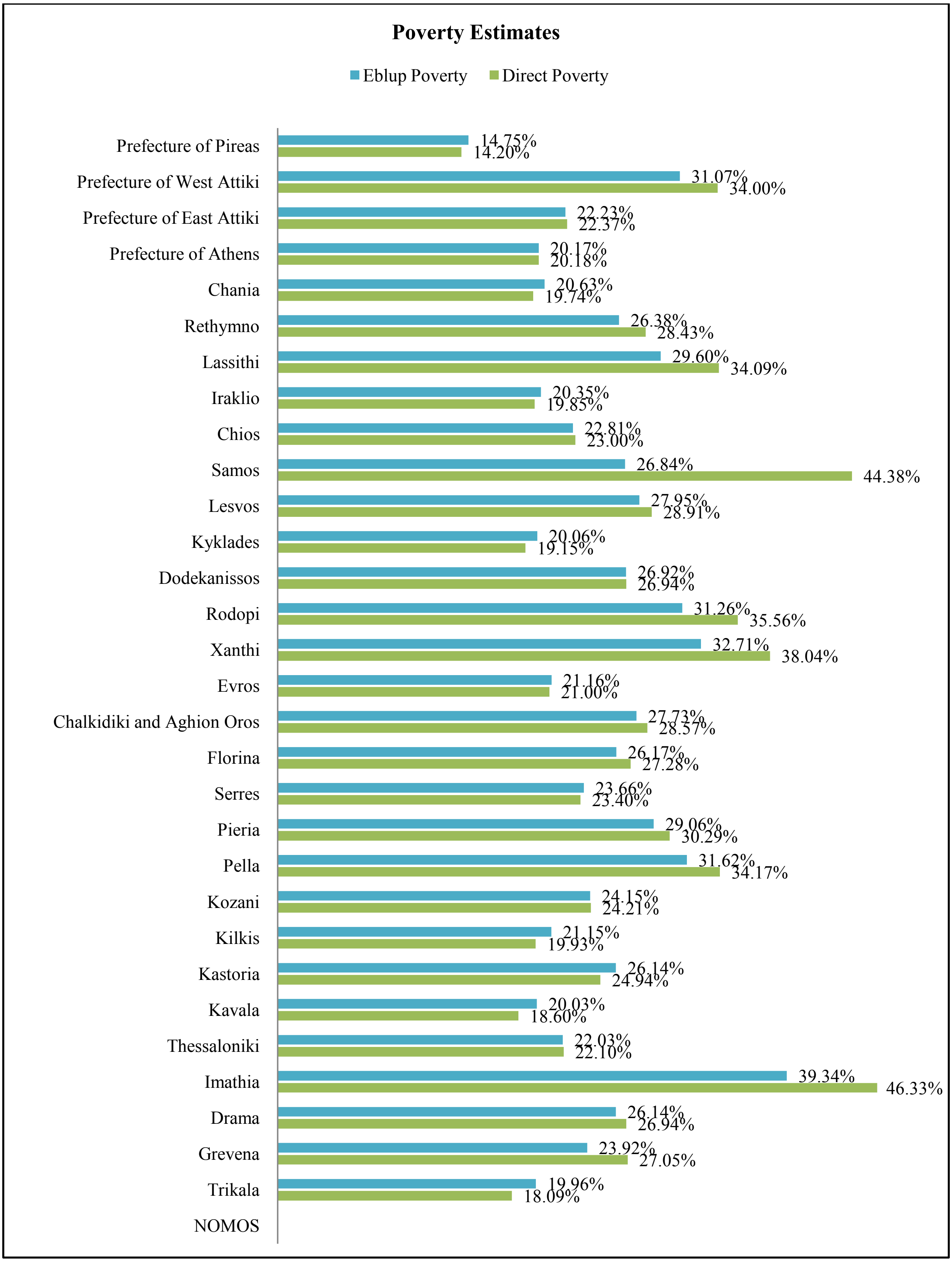


Figure 1. Poverty Estimates (First part), Source: Author's elaboration.

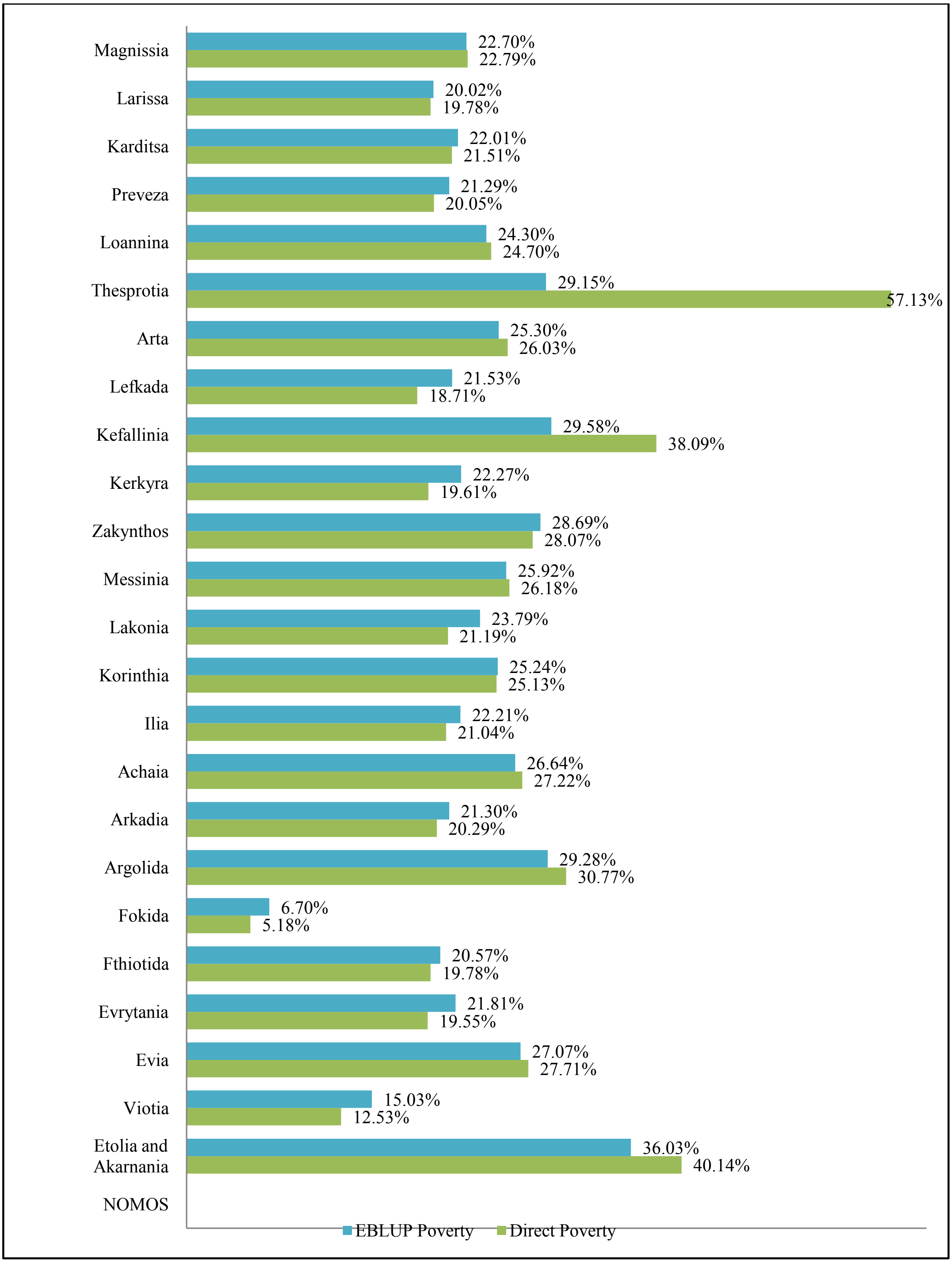


Figure 2. Poverty Estimates (Second part).

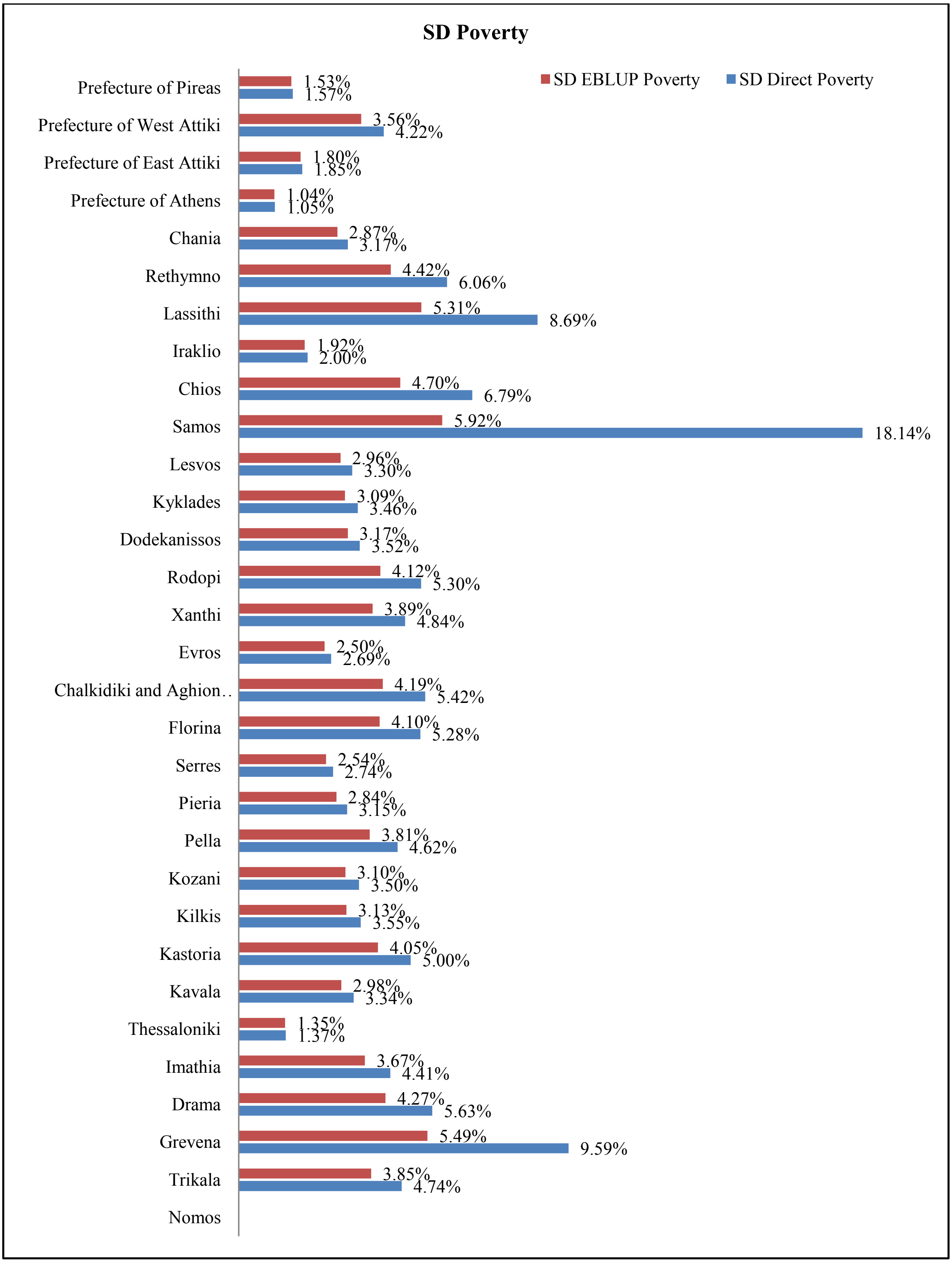


Figure 3. Poverty Standard Deviation (First part). Source: Author's elaboration.

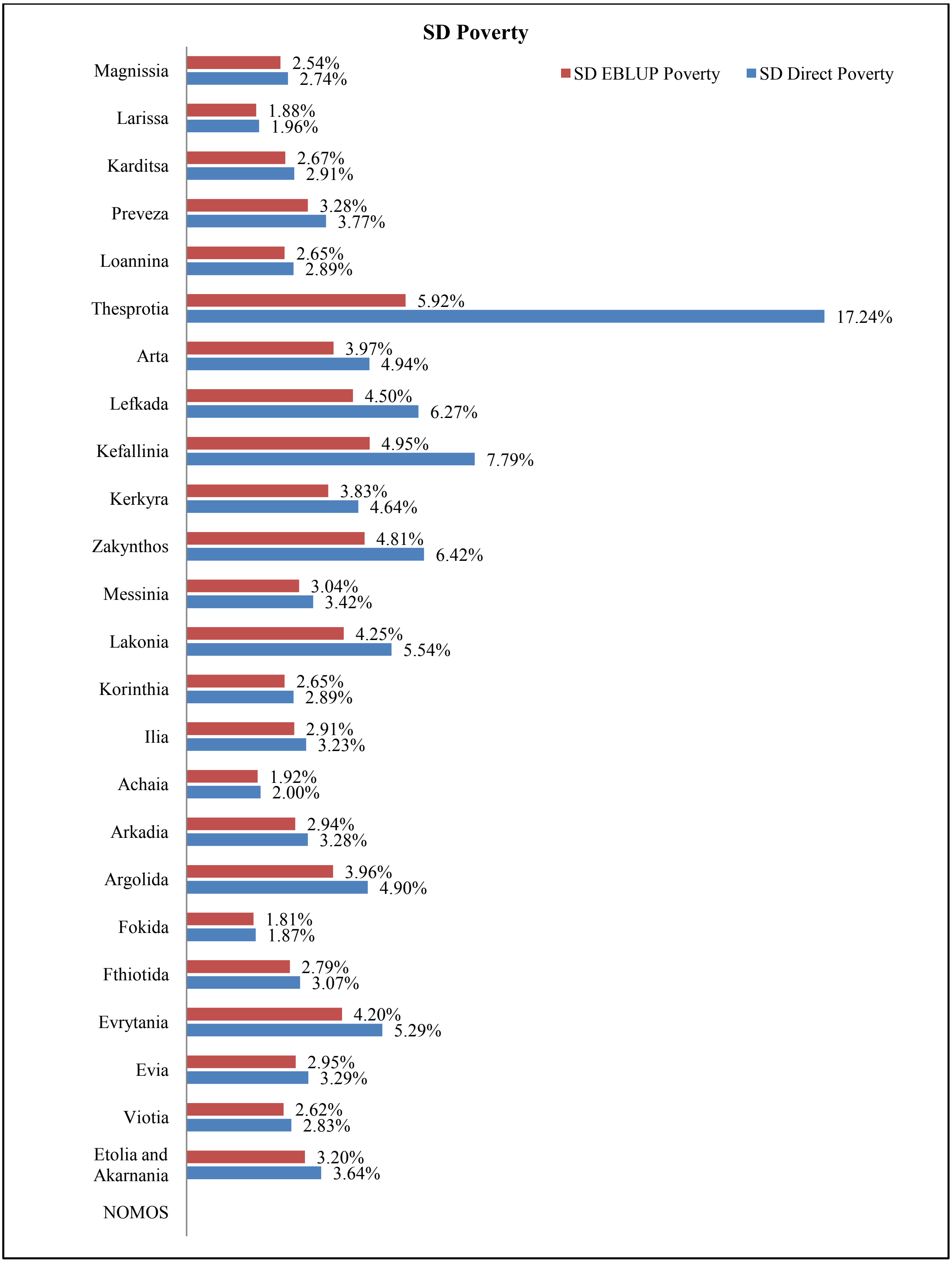


Figure 4. Poverty Standard Deviation (Second part).

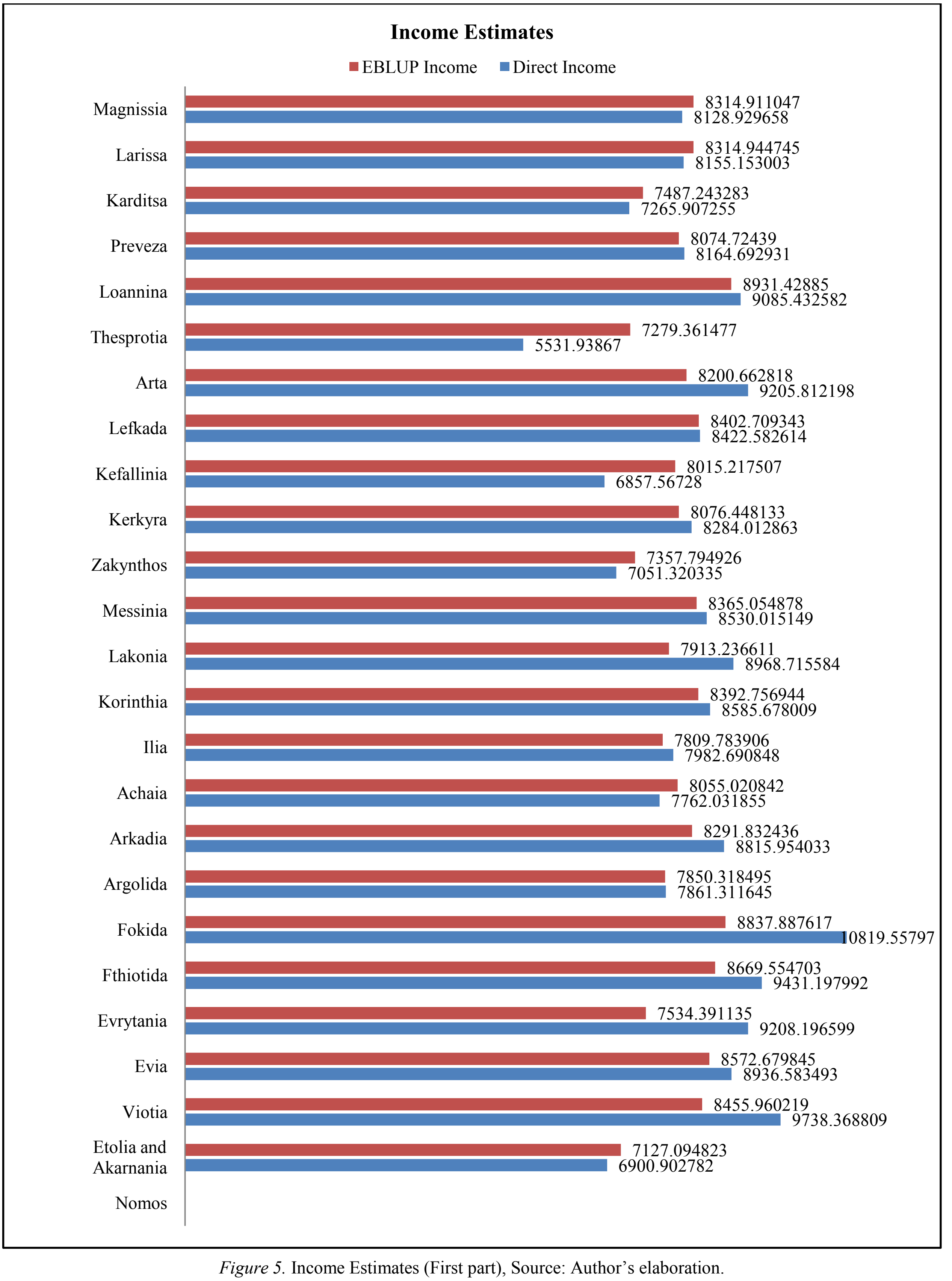


Figure 5. Income Estimates (First part), Source: Author's elaboration.

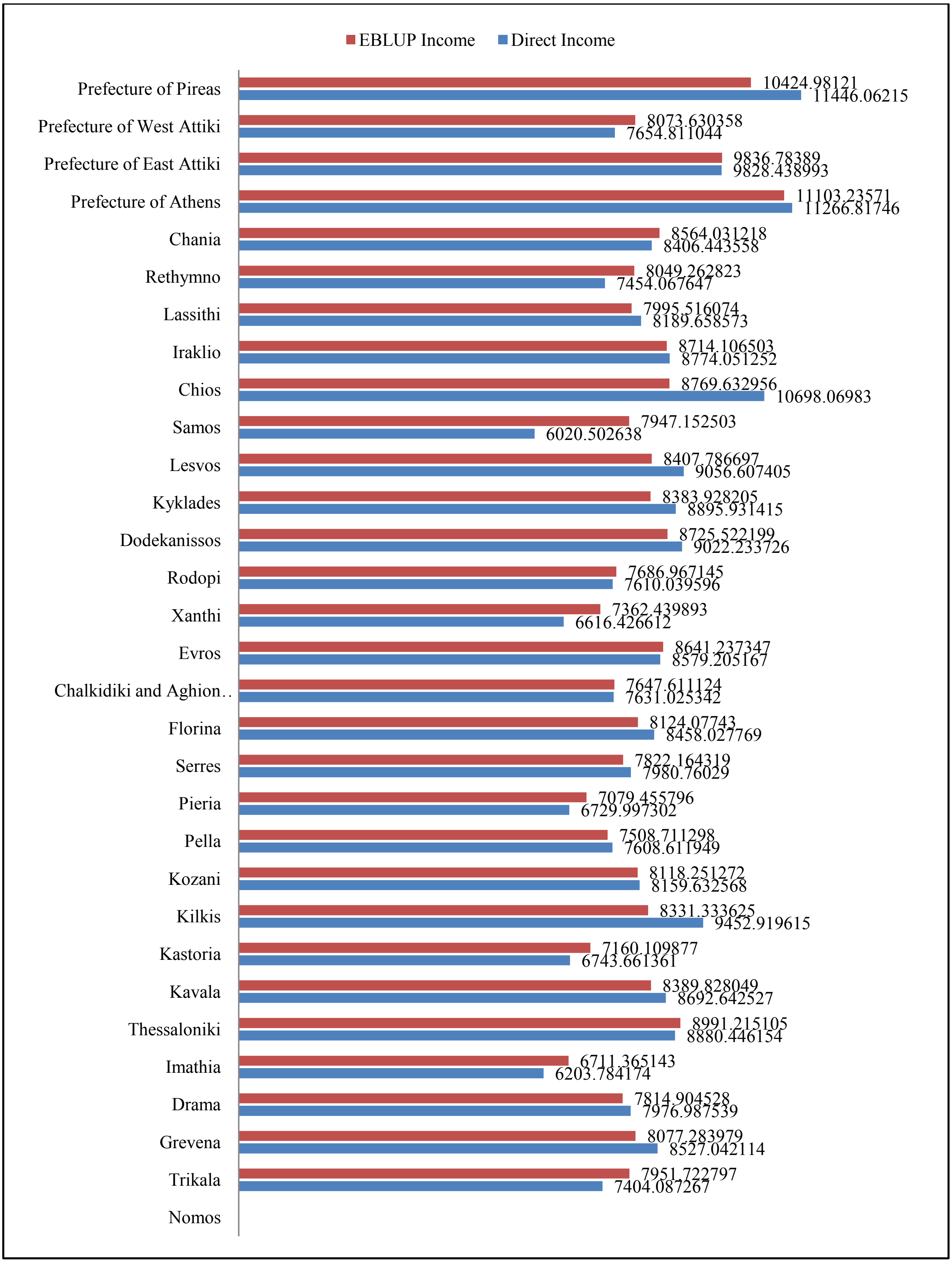


Figure 6. Income Estimates (Second part).

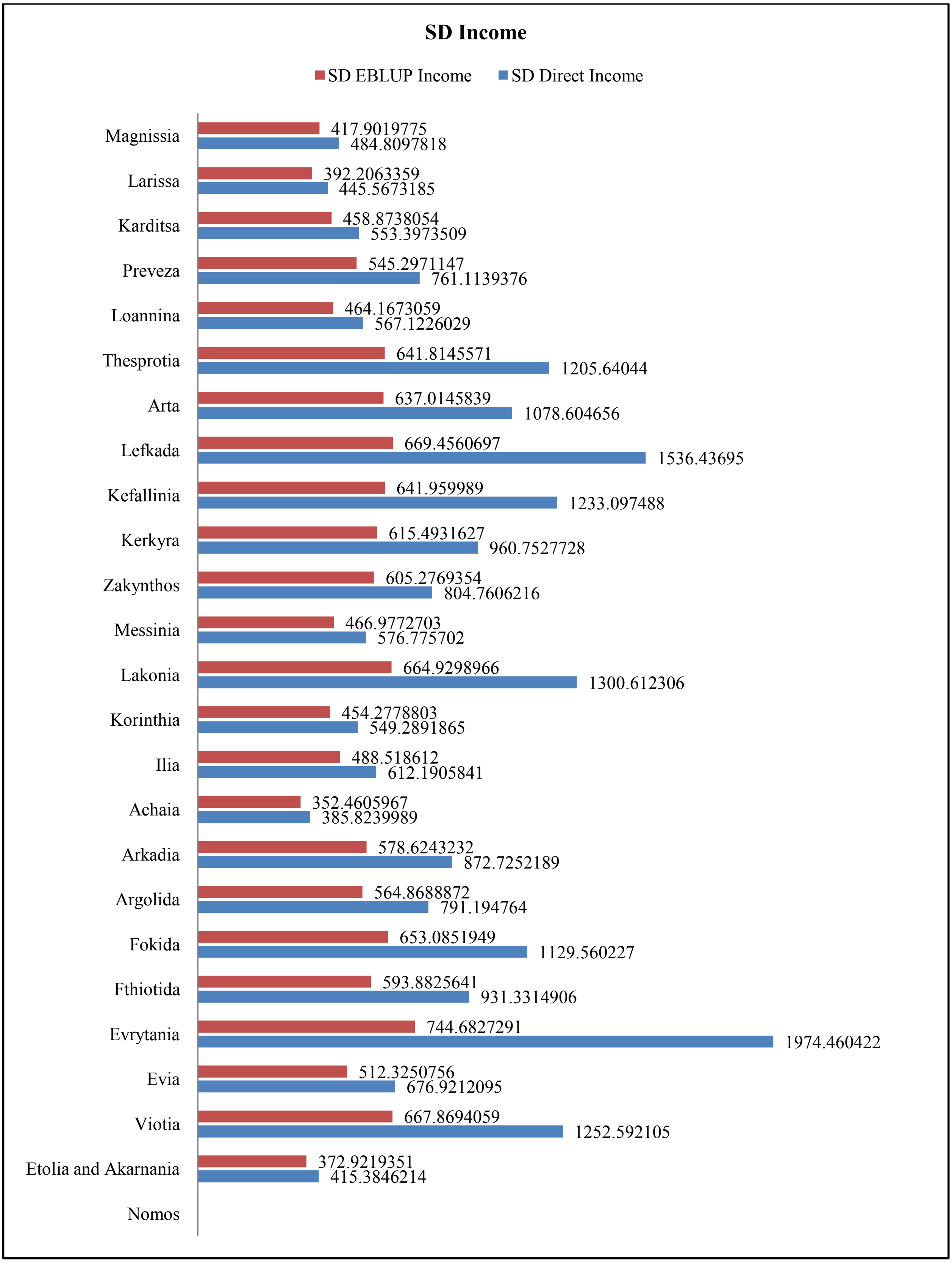


Figure 7. Income Standard Deviation (First part), Source: Author's elaboration.

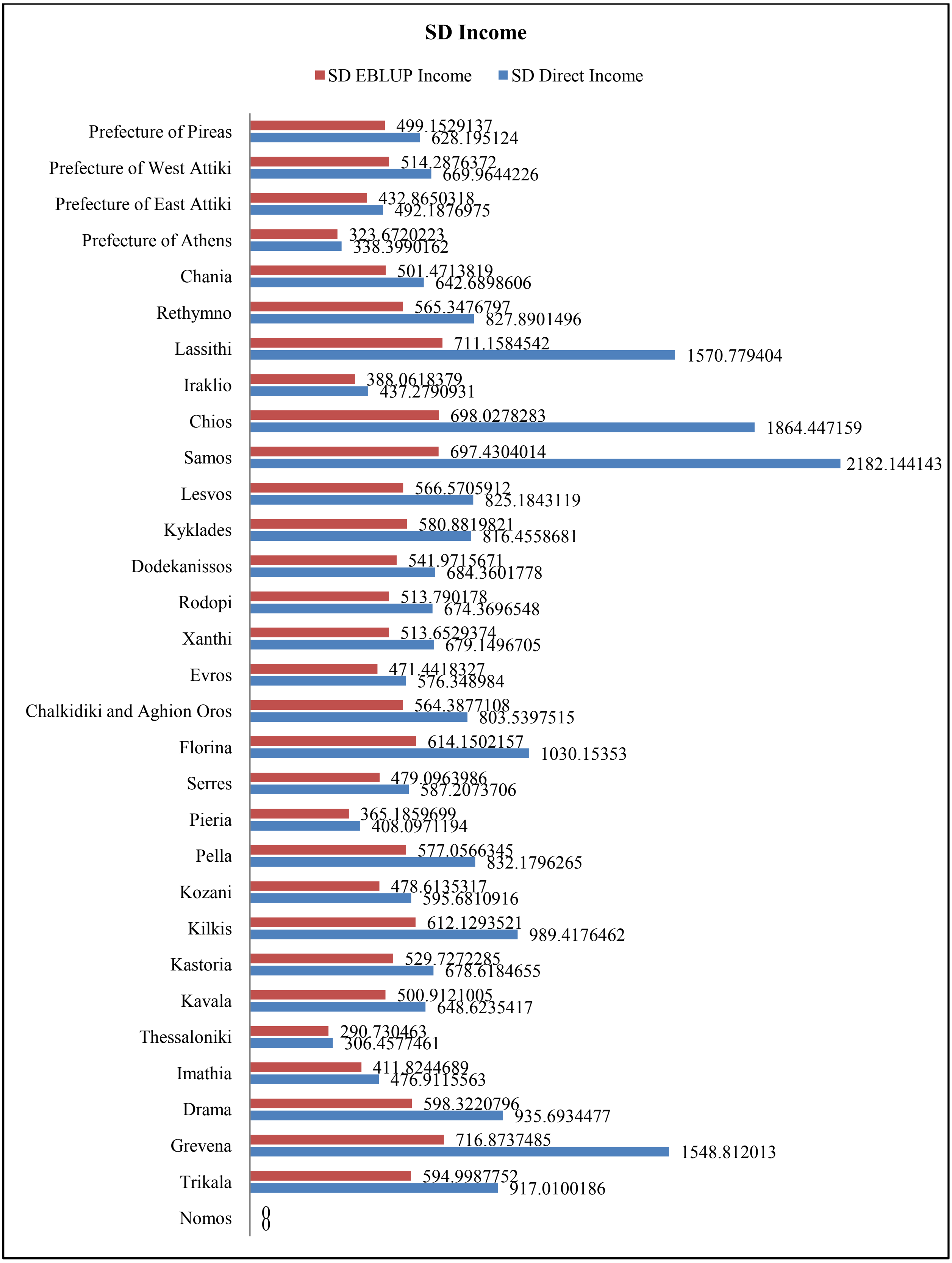


Figure 8. Income Standard Deviation (Second part).

References

Alessandro Tarozzi and Angus Deaton (2009). "Using Census and Survey Data to Estimate Poverty and Inequality for Small Areas". Review of Economics and Statistics, 91(4), 773-792.

Atkinson, A.B. (1991). "Measuring Poverty and Differences in Family Composition", Economica, 59, 1–16.

Battese, G.E., Harter, R.M., and Fuller, W.A. (1988). An error component model for prediction of county crop areas using survey and satellite data. J. Amer. Statist. Assoc. 83, 28-36.

Chandra, H., Sud, UC. and Gupta V.K. (2013). Small Area Estimation under Area Level Model Using R Software.

Cochran, W. G. (1977),Sampling Techniques, Third Edition, New York: JohnWiley & Sons, Inc.

DATTA, G. S. (2009). Model-Based Approach to Small Area Estimation. In: Handbook of Statistics 29B; Sample Surveys: Inference and Analysis. Eds. D. Pfeffermann and C.R. Rao. North Holland. pp. 251-288.

Fay and Herriot, 1979 R.E. Fay, R.A. Herriot Estimates of income for small places: an application of James–Stein procedures to census data, J. Amer. Statist. Assoc., 74 (1979), pp. 269–277.

Ghosh, M. and Rao, J.N.K. (1994). Small Area Estimation: An Appraisal.

Hagenaars, A., K. de Vos and M.A. Zaidi (1994), Poverty Statistics in the Late 1980s: Research Based on Micro-data, Office for Official Publications of the European Communities. Luxembourg.

Isabel Molina, J. N. K. Rao. 2010. Small area estimation of poverty indicators. Canadian Journal of Statistics 38:3, 369-385.

Rao, J. N. K. (2003). Small Area Estimation. Wiley, Hoboken, New Jersey.

Rao, J.N.K. (1999). Some Recent Advances in Model-Based Small Area Estimation. Survey Methodology. 25, 175-186.

Ravallion, Martin (1994), Poverty Comparisons, Fundamentals of Pure and Applied Economics Volume 56, Chur, Switzerland: Harwood Academic Publishers.

Robert Haveman & Barbara Wolfe, 1990. "The Economic Well-Being of the Disabled: 1962-84," Journal of Human Resources, University of Wisconsin Press, vol. 25(1), pages 32-54.

Sen, A. (1979). Equality of What? The Tanner Lectures on Human Values.

Sen, Amartya K. (1992), Inequality Re-examined, Oxford: Clarendon Press. Stanford University, May 22, 1979.

Townsend, P. (1979) Poverty in the United Kingdom,Harmondsworth, Penguin Books Ltd.