Journal of Statistics: Advances in Theory and Applications
Volume 15, Number 1, 2016, Pages 35-49
Available at http://scientificadvances.co.in
DOI: http://dx.doi.org/10.18642/jsata_7100121633

LATENTLY REGRESSING POVERTY ON INCOME AND INEQUALITY

GORDON BECHTEL

Warrington College of Business
University of Florida
Gainesville, Florida 32611-7155
USA
e-mail: bechtel@ufl.edu

Abstract

Poverty is reduced by mean income about twice as much as it is increased by income standard deviation. This result holds when poverty varies over congressional districts in a spatial regression and over years in a temporal regression. These population-weighted regressions are equivalent to latent regressions over millions of Floridians.

1. Introduction

Skewed and high-variance income distributions are long standing topics treated in the classic economics text by Nobel Laureate Samuelson and his co-author Nordhaus [1]. They assumed (p. 47) that a fundamental function of government is to correct market economies through "techniques such as income redistribution to reflect society's concerns for the poor or hapless". Unemployment was also characterized by these authors (pp. 207-209) beyond its loss of household income and purchasing power:

2010 Mathematics Subject Classification: 62A01, 62J05, 62P20, 91B64, 91B82.
Keywords and phrases: American community survey, economic indicators, experiential distributions, latent and population-weighted regressions, spatial and temporal poverty.
Received February 15, 2016

However large the economic costs of unemployment, a recounting of dollars does not adequately convey the human, social, and psychological toll that persistent periods of involuntary unemployment bring. … recent studies indicate that unemployment leads to a deterioration of both physical and psychological health – higher levels of heart disease, alcoholism, and suicide. … other studies indicate that involuntary joblessness is a highly traumatic event for many people.

Income inequality, a perennial problem for capitalist economies, has now become a global issue. Aly [2] views this inequality as the major causal factor in the 2011 Arab Spring *and* the 2011 Occupy Wall Street demonstrations in the United States. In 2012, the latter movement morphed into the 99 Percent Spring, protesting the fact that 1% of Americans earn 21% of all U.S. income [3]. Cutting across this 99 percent, Zogby [4] described *The Demographics of Hunger in America*:

Residents of large cities (15%) reported going hungry, as did 10% of people living in union households, 14% of the Creative Class, 17% of NASCAR fans, 13% of the Investor Class, 15% of those living with children under 17, 29% of current NRA members, 14% of weekly Wal-Mart Shoppers, 14% of Tea Party supporters, 20% of Occupy Wall Street supporters, and 16% of those who identified as Lesbian Gay Bisexual Transgender. Hunger is not concentrated in one section of town nor is anyone automatically immune.

Public and media apathy to this situation is revealed by recent polling and election campaigns in the United States [5, p. 250]:

Less than fifty percent of Americans favour income redistribution… surveys show that "the gap between the rich and poor" is rated near the bottom of lists of national problems presented to American respondents … This opinion data may account for the fact that a moderator from the Public Broadcasting Service put no question about poverty to either presidential candidate in their debate on the economy on October 3, 2012. *Just before this debate,* on September 12, 2012, the United States Census Bureau released its report stating that 15% of the population, or 46.2 million Americans, fall below the poverty line …(italics added)

The present paper regresses poverty in Florida on the competing effects of income and income inequality. In a departure from classical economic analysis, we define inequality as the standard deviation of a population's income distribution. This parameter is more straightforward than the well-known Gini coefficient, which is unavailable for small areas.

Section 2 equates a population's mean income, income standard deviation, and percent poverty to the first moments of three otherwise unknowable distributions. The macro indicators defining these three moments are provided by the American Community Survey (ACS) of the United States Census Bureau. Sections 3 and 4 demonstrate that poverty is reduced by income level (mean) about twice as much as it is increased by income inequality (standard deviation). This holds when poverty varies over congressional districts in a spatial regression and over years in a temporal regression. These results rest on Bechtel [6], who showed that slopes of population-weighted regressions equal those of latent regressions over millions of individuals in states, countries, regions, or cross-national designations such as the G20.

Section 5 stresses educational factors perpetuating the correlation between income level and inequality and outlines recent economic proposals for relieving inequality. Section 6 discusses the advantages of identifying first moments of latent distributions with established economic indicators and, in the future, with meta-indicators discovered from Big Data.

2. Economic Indicators and Latent Distributions

The American Community Survey (ACS) reports total income (during the last 12 months) for respondent i in congressional district or year c in Florida. In our spatial analysis, this income is summed over all respondents in each household and all households in each district. ACS's five year computations

G_c of mean household income,

I_c of income standard deviation, and

Y_c of percent below the poverty level,

for district c are interpreted as means of experiential distributions ρ, τ, and η over all individuals in this district. In our temporal analysis G_c, I_c, and Y_c are ACS's statewide computations for a single year c.

2.1. Income

Mean household income G_c in congressional district (or year) c is interpreted here as the mean of a latent distribution over this district's (or year's) *entire* population. This experiential distribution carries the felt effects of household income over everyone in district (or year) c. We label this personal welfare experienced by individual i in c as ρ_{ci} and equate the mean over individuals in c to indicator $G_c = \rho_c$.

2.2. Income inequality

Departing from classical economic analysis, we replace the Gini coefficient with simple income variation as our indicator of inequality. This index is more straightforward and available for measuring inequality in small areas such as congressional districts or states. Income standard deviation I_c also has the advantage of being calibrated in the same monetary units as income G_c. We interpret the standard deviation I_c of household income in congressional district (or year) c as the mean of a second latent variable, also distributed over the *entire* population in c. Thus, inequality, as personally experienced by individual i in c, is τ_{ci} whose mean over i in c is $I_c = \tau_c$.

2.3. Poverty

ACS's poverty indicator is the percent of all people in a congressional district or year whose household income in the last 12 months is below the poverty level. We interpret this poverty percent Y_c as the mean of a third population distribution in district (or year) c. This latent variable η spreads a macro poverty indicator into degrees of deprivation (or abundance) experienced at the micro level. We set the indicator Y_c equal to the mean $\eta_{c.}$ of this latent population of individual values η_{ci} in c.

2.4. Partitioning the population variance of $(\rho_{ci}\tau_{ci}\eta_{ci})$

We assume that the within-district (year) sums-of-products (SP) matrix of the vector $(\rho_{ci}\tau_{ci}\eta_{ci})$ is a multiple ω of its observable between SP matrix (cf. Bechtel [6], Section 2; Rao [7], Sections 8c and 8d). Under this assumption, we have C unobservable tri-variate populations, with mean vectors equated to Florida indicators and within SP matrix known up to multiplication by ω. The total SP matrix of these C populations is $(1 + \omega)$ times its between SP matrix. The elements of this between SP matrix are displayed in the 3×3 array:

$$\begin{array}{ccc} \sum_c N_c(G_c - \mu_G)^2 & \sum_c N_c(G_c - \mu_G)(I_c - \mu_I) & \sum_c N_c(G_c - \mu_G)(Y_c - \mu_Y) \\ \sum_c N_c(G_c - \mu_G)(I_c - \mu_I) & \sum_c N_c(I_c - \mu_I)^2 & \sum_c N_c(I_c - \mu_I)(Y_c - \mu_Y) \\ \sum_c N_c(G_c - \mu_G)(Y_c - \mu_Y) & \sum_c N_c(I_c - \mu_I)(Y_c - \mu_Y) & \sum_c N_c(Y_c - \mu_Y)^2, \end{array} \tag{1}$$

where $\mu_G = \sum_c N_c G_c / \sum_c N_c$, $\mu_I = \sum_c N_c I_c / \sum_c N_c$, and $\mu_Y = \sum_c N_c Y_c / \sum_c N_c$ for $c = 1, \ldots, C$ are the means of the latent population $\{\rho_{ci}\tau_{ci}\eta_{ci}\}$. The observable array (1) is used in the following spatial and temporal analyses of poverty.

3. Spatial Poverty

ACS respondents' incomes during the last 12 months are summed over all respondents in each household and all households in each congressional district. ACS's five-year computations of G_c, I_c, and Y_c for district c are interpreted here as means of experiential distributions ρ, τ, and η over all individuals in this district.

3.1. Latent and population-weighted regressions

We now measure the impacts of income and income inequality on poverty. Referring to Section 2, the regression of latent variable η on latent variables ρ and τ is

$$\eta_{ci} = \kappa + \delta\rho_{ci} + \theta\tau_{ci} + \varepsilon_{ci} \quad (i = 1, \ldots, N_c;\ c = 1, \ldots, 27), \tag{2}$$

where ε_{ci} is specification error. The slopes in (2) are OLS identified by

$$(\delta\ \theta)^T = \left(\sum_c \sum_i T_{ci} T_{ci}{}^T\right)^{-1} \sum_c \sum_i T_{ci}(\eta_{ci} - \mu_Y), \tag{3}$$

where $T_{ci} = (\rho_{ci} - \mu_G\ \ \tau_{ci} - \mu_I)^T$. The elements of the matrix $\sum_c \sum_i T_{ci} T_{ci}{}^T$ and vector $\sum_c \sum_i T_{ci}(\eta_{ci} - \mu_Y)$ are $(1 + \omega)$ times the 2×2 and 2×1 arrays

$$\begin{array}{lll} \sum_c N_c(G_c - \mu_G)^2 & \sum_c N_c(G_c - \mu_G)(I_c - \mu_I) & \sum_c N_c(G_c - \mu_G)(Y_c - \mu_Y), \\ \sum_c N_c(G_c - \mu_G)(I_c - \mu_I) & \sum_c N_c(I_c - \mu_I)^2 & \sum_c N_c(I_c - \mu_I)(Y_c - \mu_Y), \end{array} \tag{4}$$

which are gleaned from rows 1 and 2 of array (1). In our spatial analysis, N_c is ACS's five-year estimate of the population size of congressional district c. These 27 population estimates range from 695 to 730 thousand people. Therefore, the summations in (4) run over 19 million Floridians.

The OLS slopes in (3) are immediately given by

$$(\delta\ \theta)^T = \left(\sum_c N_c T_c T_c{}^T\right)^{-1} \sum_c N_c T_c (Y_c - \mu_Y), \tag{5}$$

where $T_c = (G_c - \mu_G\ \ I_c - \mu_I)^T$, and summations run only over districts $c = 1 \ldots 27$. The population-weighted slopes in (5) are computed solely from ACS's five-year computations of N_c, G_c, I_c, and Y_c.

3.2. Spatial results

The vector $(\delta\ \theta)^T$ in (5) is given from the Stata [8] command

regress poverty income inequality [fweight = pop*c*],

where pop*c* is the population of congressional district $c = 1 \ldots 27$. The resulting population-weighted slopes equal those in (2) and (3), which are the slopes of a latent tri-variate regression over the entire Florida population [6]. The negligible sampling variation in the 5-year computations N_c, G_c, I_c, and Y_c, and hence in $(\delta\ \theta)^T$, obviates the need for any standard errors or significance tests.

Because income standard deviation I_c is calibrated in the same monetary units as income G_c, we can directly compare the effects of income and inequality on poverty. The left side of Table 1 shows that spatial poverty is reduced by income level more than twice as much as it is increased by income inequality. However, due to a spatial correlation (over districts) of .93 between income mean and standard deviation, this poverty reduction is offset by poverty increments in districts with higher inequality.

Table 1. Impacts of income and inequality on poverty

Poverty	Spatial slopes	Temporal slopes
Income	– .78	– .62
Inequality	.34	.39

Note: These slopes were computed from formula (5).

4. Temporal Poverty

In our time-series regression, the subscript *c* in formula (5) is read as a year (instead of a congressional district). Thus, N_c, G_c, I_c, and Y_c are now ACS's statewide computations for a single year *c* = 2005 … 2013.

The right side of Table 1 confirms the impacts of income level and inequality on poverty. Temporal poverty is reduced by mean household income almost twice as much as it is increased by income variation. However, we now have a temporal correlation (over years) of .88 between income mean and standard deviation. Thus, poverty reduction with increasing income over time is again offset by poverty increases from rising inequality.

5. Alleviating Inequality and Poverty

5.1. Educational solutions

Bechtel [5, p. 244] emphasized the role of the educated in opposing income redistribution:

It is well established that education is strong causal factor in income level [9]. Moreover, Tóth and Keller [10] find an individual's redistributive preference to be negatively related to her (his) personal material status. These two relationships imply that the more educated tend to disagree with income redistribution. This hypothesis is supported by Tóth and Keller in their pan-European regression of individual-level data from the 2009 Euro barometer. We reconfirm it here in our pan-European regression of individual-level data from the 2010 European Social Survey. This negative relationship between education and redistributive preference looms as an impediment to alleviating income inequality because government policy is heavily influenced by the more educated.

In 2010, The College Board [9] reported a strong relationship between education and income. Subsequently, Reardon [11] demonstrated that the low school-readiness of kindergarteners from poor families spawns low reading and mathematics scores of eighth graders from low-income households. The continuing burden of a low-income background into adulthood was then reported by Autor [12], who found that the American college/high school earnings gap (in constant 2012 dollars) increased from $30,298 in 1979 to $58,249 in 2012. The rank of U. S. workers' real weekly earnings (relative to 1963) was: high school dropout < high school graduate < some college < Bachelor's degree <

beyond Bachelor's degree. Educational opportunities offered to the rich and withheld from the poor perpetuate the correlations in Sections 4 and 5 between income means and standard deviations. These correlations impede poverty reduction with rising income, and public policy should be aimed at decoupling income mean and variance. If a nation's income distribution shifted by its mean only, a constant variance would not offset the reduction in national poverty rate. In this ideal situation, the slope of poverty on income standard deviation would be zero, maximizing the beneficial effect of increased income level.

5.2. Economic solutions

A Call for Change in the European Union has recently been issued by the Progressive Economy Initiative [13]:

The EU should declare a new egalitarian ideal. All member states and the EU as a whole will reap political, economic, and social benefits from fighting inequality.

… Taxation, if properly used, can make a decisive contribution in reducing inequality. EU policy should favour progressive taxation of incomes, stiff taxation of inheritance with a strong philanthropic incentive, and taxation of real property and rents. The current reliance on VAT is excessive, regressive, and should be reduced. A range of additional measures are necessary: implementing the financial transactions tax to curb speculation and raise funds for investment, reinforcing transparency obligations, eliminating tax-evading transfer pricing, closing loopholes in national tax systems, putting an end to tax havens.

… The EU should set goals for the reduction of inequality within countries and convergence of income levels across EU member states.

Similar to *A Call for Change*, Galbraith and Varoufakis [14] have presented a *Modest Proposal* for reducing inequality and poverty. The Piketty Group in France advocates more radical tax reform, requiring political and financial structures that require changes in the EU treaty. Piketty and Saez [15] plot rate of return to capital versus growth rate of world output from antiquity to 21st century:

The average rate of return to capital (after tax and capital losses) fell below the growth rate in the 20th century. It may again surpass it in the 21st century, as it did throughout human history except in the 20th century … (p. 841).

These authors (pp. 838-839) also show that in the 20th century western income and wealth inequality fell from their previously high levels. Their remedy to rising inequality in the 21st century

– a universal tax on wealth that takes little from the bottom of the scale and a lot from the top – has drawn fire from left and right, in part because it would be so difficult to impose … – requiring all the world's nations to agree in concert to boost taxes on their wealthiest citizens [16, p. 827].

Nevertheless, recent proposals for inequality reduction have been echoed by the new Latin American Pope's call for

"The legitimate redistribution of economic benefits by the state, as well as indispensable cooperation between the private sector and civil society" as ways to combat poverty [17].

The feasibility of Pope Francis call is supported by Hvistendahl [18, pp. 834-835] who, like Chomsky [19, p. 38], points to Latin America as leading the way toward alleviating global inequality:

The gulf between income classes in Latin America has gradually narrowed, resulting in an impressive decrease in inequality across the region. … Throughout the 2000s, Ginis fell in 13 of 17 Latin American countries for which the World Bank has reliable data. … As with the Asian "tigers" before them, the economy is growing, while inequality is falling. … From Latin America's success … one thing now seems certain: Where inequality does decline, government involvement is key.

6. Future Directions for Latent Regression

First moments of postulated distributions are supplied here by ACS data on the state of Florida. Bechtel [6] demonstrates the next step in data aggregation by anchoring experiential distributions to national economic indicators and indices. The penultimate leap in aggregation is seen in the advent of Big Data. Thus, it is now possible to equate first moments of unobservable human distributions to supra-macro indicators derived from national and international datasets. This new path suggests further "statistical thinking and new foundational frameworks" that help sort out "the many philosophical issues data science presents …" (Davidian [20]). Davidian's call has been echoed by the American National Science Foundation, who "just released a revised version of the solicitation 'Critical Techniques and Technologies for Advancing Foundations and Applications of Big Data Science' …" [21].

Horrigan [22] (pp. 25-27) has given a perspective on Big Data from the United States Bureau of Labor Statistics:

I begin, and probably at my peril, by attempting to define Big Data. I view Big Data as nonsampled data, characterized by the creation of datasets from electronic sources whose primary purpose is something other than statistical inference. … this type of Big Data typically comprises the universe and, by definition, can represent (nearly) the entire population ….

As examples of "nonsampled universe files" Horrigan mentions daily price indices, point-of-sale retail databases, universe data on hospitals, and corporate data.

In the present paper, experiential population distributions exist in hypothetical files containing "the stuff that cannot be measured" [23]. This postulation is a theory of macro data akin to Coombs' earlier theory of psychological data [24-27]. Means of latent distributions, anchored to established macro indicators, are more palpable and plausible than sample survey means hypothesized to measure particular constructs. Micro scales induced by macro indicators are also practical alternatives to questionnaire scores derived from self-reported experience. In most nations perceptual, attitudinal, and opinion scores are not even available. The host of problems associated with survey measurement and process quality are discussed and illustrated in the volume edited by Lyberg et al. [28].

The population-weighted regression slopes in (5) also evade the data analytic issues daunting survey regression. Regression of sampled micro data is beset by the unresolved competition between randomization-based and model-based regression [29-31]. Both types of regression face problems of measurement error [32], sampling error [33, 34], unit non-response [35], missing data [36], and variance estimation [37, 38]. Unit non-response alone threatens the entire survey industry due to the public's unwillingness to answer mail, telephone, face-to-face, or internet questions [39]. These issues with micro-data sampling are eluded by using the first moments of unknown distributions to produce the regression slopes in (2), (3), and (5). In the future, these first moments can be equated to even better economic indicators discovered by the meta-aggregation of Big Data sets [40-42]. Pfeffermann [43, pp. 428-429, 433] observes

> The UN Statistical Commission established in 2014 a global working group, mandated "to provide strategic vision, direction and coordination of a global programme on Big Data for official statistics" … (United Nations 2014). … *For example,* almost every country in the world publishes monthly a consumer price index (CPI). In most countries, the CPI is based on two separate surveys – a family expenditure survey that is used to estimate the weight attached to each commodity and a survey of stores and sale companies where the prices are recorded. … Can the CPI be computed based on big data? Why not get all sales (quantities) and prices directly from the stores and sales companies? … The use of big data does not require a sampling frame, questionnaires, interviews, and all the other ingredients underlying survey samples, which in the long run could result in large cost reductions. Considering that response rates in traditional surveys are constantly declining, the use of big data … seems inevitable … *Moreover,* Scandinavian countries claim to have sufficiently accurate population records and, hence, Scandinavian countries do not generally carry out special population censuses. … this should be the ultimate target of every country – having sufficiently accurate administrative records so that no population cenuses will be needed. (italics added)

Macro-indicator upgrades with Big Data, facilitated by cooperating governments, the United Nations, the World Bank, the International Monitory fund, and the New Development Bank in Shanghai, will help statistics tackle inequality and poverty in the 21st century.

Acknowledgements

The author thanks Bethany Bechtel for her emphasis on the importance of monitoring single spatial entities over time.

References

[1] P. A. Samuelson and W. D. Nordhaus, Economics, 12th Ed., McGraw-Hill: New York, NY, USA, 1985.

[2] H. Aly, Economics prof provides insights on Arab Spring; 2011.

http://osumarion.osu.edu/news/economics

[3] C. Collins, The 99 percent spring; 2012. http://inequality.org/99-percent-spring.

[4] J. Zogby, The demography of hunger in America, Zogby Analytics; 2014, December 17.

http://www.zogbyanalytics.com/news/

[5] G. G. Bechtel, Public opinion about income inequality, Electronic Journal of Applied Statistical Analysis 6 (2013), 238-259.

[6] G. G. Bechtel, National index theory: Chomskyan globalization and well-being, Computational and Applied Mathematics Journal 2 (2016), 6-11.

[7] C. R. Rao, Linear Statistical Inference and its Applications, Wiley and Sons: New York, NY, USA, 1965.

[8] StataCorp. Stata Statistical Software, Release 12, StataCorp LP: College Station, TX, USA, 2011.

[9] College Board, Education Pays, The College Board: New York, NY, USA, 2010.

http://www.collegeboard.com/trends

[10] I. G. Tóth and T. Keller, Income Distributions, Inequality Perceptions, and Redistributive Claims in European Societies, Proceedings of the Annual Conference of the Hungarian Society for Economics, Budapest, Hungary, 2011, December 18-20.

[11] S. F. Reardon, The widening academic achievement gap between the rich and the poor: New evidence and possible explanations, In Wither Opportunity? Rising Inequality, Schools, and Children's Life Chances; G. J. Duncan and R. J. Murnane, Eds., Russell Sage Foundation: New York, NY, USA, 2012, pp. 91-116.

[12] D. H. Autor, Skills, education, and the rise of earnings inequality among the other 99 percent, Science 344 (2014), 843-851.

[13] Progressive Economy, A call for change: From crisis to a new egalitarian ideal for Europe, The Progressive Economy Initiative; (2014).

http://www.progresiveeconomy.eu

[14] J. K. Galbraith and Y. Varoufakis, Wither Europe? The modest camp vs the federalist austerians, OpenDemocracy; 2014, June 11.

https://www.opendemocracy.net

[15] T. Piketty and E. Saez, Inequality in the long run, Science 344 (2014), 838-843.

[16] E. Marshall, Tax man's gloomy message: The rich will get richer, Science 344 (2014), 826-827.

[17] M. Tracy, When the Pope speaks, liberals and conservatives only hear what they want to hear, New Republic; 2014, May 9.

http://www.newrepublic.com

[18] M. Hvistendahl, While emerging economies boom, equality goes bust, Science 344 (2014), 832-835.

[19] N. Chomsky, Hopes and Prospects, Haymarket Books: Chicago, IL, USA, 2010.

[20] M. Davidian, Aren't we data science? Amstat News (July 2013), 3-5.

[21] M. Vogelius, N. Kannan and X. Huo, NSF Big Data funding opportunity for the statistics community, Amstat News (April 2015), 6.

[22] M. W. Horrigan, Big Data: A perspective from the BLS, Amstat News (January 2013), 25-27.

[23] Y. Varoufakis, All the good stuff that cannot be measured, 2015a, February 7

https://www.youtube.com

[24] C. H. Coombs, Psychological scaling without a unit of measurement, Psychological Review 57 (1950), 145-158.

[25] C. H. Coombs, A theory of psychological scaling, Engineering Research Institute Bulletin, Number 34, University of Michigan Press: Ann Arbor, MI, USA, 1952.

[26] C. H. Coombs, A Theory of Data, Wiley and Sons: New York, NY, USA, 1964.

[27] G. G. Bechtel, Probabilistic dimensionality: A study of confidence and intention, In Frontiers of Mathematical Psychology: Essays in Honor of Clyde Coombs, D. R. Brown and J. E. K. Smith, Eds., Springer-Verlag: New York, NY, USA, 1991, 80-109.

[28] L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz and D. Trewin, Eds., Survey Measurement and Process Quality; Wiley and Sons: New York, NY, USA, 1997.

[29] A. Chaudhuri and H. Stenger, Survey Sampling: Theory and Methods, 2nd Ed., CRC Press: Boca Raton, FL, USA, 2005.

[30] K. Brewer and T. G. Gregoire, Introduction to survey sampling, In Handbook of Statistics (Sample Surveys: Design, Methods and Applications); D. Pfeffermann and C. R. Rao, Eds., Elsevier: Amsterdam, The Netherlands, 29A (2009), 9-37.

[31] J. D. Opsomer, Introduction to Part 4: Alternative approaches to inference from survey data, In Handbook of Statistics (Sample Surveys: Inference and Analysis); D. Pfeffermann and C. R. Rao, Eds., Elsevier: Amsterdam, The Netherlands, 29B (2009), 3-9.

[32] J. Bound, C. Brown and N. Mathiowetz, Measurement error in survey data, In Handbook of Econometrics; J. J. Heckman and E. Leamer, Eds. Elsevier Science: Amsterdam, The Netherlands, 5 (2001), 3705-3843.

[33] M. E. Thompson, Theory of Sample Surveys, Chapman and Hall: London, UK, 1997.

[34] S. L. Lohr, Sampling: Design and Analysis, 2nd Ed., Brooks/Cole: Boston, MA, USA, 2010.

[35] J. Brick and J. M. Montaquila, Non-response and weighting, In Handbook of Statistics (Sample Surveys: Design, Methods and Applications); D. Pfeffermann and C. R. Rao, Elsevier: Amsterdam, The Netherlands, 29A (2009), 163-185.

[36] R. J. A. Little and D. B. Rubin, Statistical Analysis with Missing Data, Wiley and Sons: New York, NY, USA, 2002.

[37] A. Demnati and J. N. K. Rao, Linearization variance estimators for survey data, Survey Methodology 30 (2004), 17-26.

[38] Z. Prášková and P. K. Sen, Asymptotics in finite population sampling, In Handbook of Statistics (Sample Surveys: Inference and Analysis), D. Pfeffermann and C. R. Rao, Eds. ,Elsevier: Amsterdam, The Netherlands, 29B (2009), 489-522.

[39] K. A. Frankovic, C. Panagopoulos and R. Y. Shapiro, Opinion and election polls, In Handbook of Statistics (Sample Surveys: Design, Methods and Applications); D. Pfeffermann and C. R. Rao, Eds., Elsevier: Amsterdam, The Netherlands, 29A (2009), 567-595.

[40] M. van der Laan, Statistics as a science, not an art: The way to survive in data science, Amstat News (February 2015), 29-30.

[41] L. Li, Statistics training: A big role in Big Data? Amstat News (May 2014), 32-33.

[42] R. Jarmin, J. Lane, A. Marco and I. Foster, Using the classroom to bring Big Data to statistical agencies, Amstat News (November 2014), 12-13.

[43] D. Pfeffermann, Methodological issues and challenges in the production of official statistics, 24th Annual Morris Hansen Lecture, Journal of Survey Statistics and Methodology 3 (2015), 425-467.

■