Accepted Manuscript

Small area estimation with multiple covariates measured with errors: A nested error linear regression approach of combining multiple surveys

Gauri S. Datta, Mahmoud Torabi, J.N.K. Rao, Benmei Liu

PII: S0047-259X(17)30766-2
DOI: https://doi.org/10.1016/j.jmva.2018.04.001
Reference: YJMVA 4348

To appear in: *Journal of Multivariate Analysis*

Received date : 11 December 2017

Please cite this article as: G.S. Datta, M. Torabi, J.N.K. Rao, B. Liu, Small area estimation with multiple covariates measured with errors: A nested error linear regression approach of combining multiple surveys, *Journal of Multivariate Analysis* (2018), https://doi.org/10.1016/j.jmva.2018.04.001

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Small area estimation with multiple covariates measured with errors: A nested error linear regression approach of combining multiple surveys

Gauri S. Datta[a,b], Mahmoud Torabi[c], J.N.K. Rao[d], Benmei Liu[e]

[a]*Department of Statistics, University of Georgia, Athens, GA 30602, USA*
[b]*US Census Bureau, Washington, DC 20233, USA*
[c]*Departments of Community Health Sciences and Statistics, University of Manitoba, Winnipeg, Manitoba, Canada R3E 0W3*
[d]*School of Mathematics and Statistics, Carleton University, Ottawa, Ontario, Canada K1S 5B6*
[e]*National Cancer Institute, Washington, DC 20006, USA*

Abstract

Small area estimation has become a very active area of research in statistics. Many models studied in small area estimation focus on one or more variables of interest from a single survey without paying close attention to the nature of the covariates. It is useful to utilize the idea of borrowing strength from covariates to build a model which combines two (or multiple) surveys. In many real applications, there are also covariates measured with errors. In this paper, we study a nested error linear regression model which has multiple unit- or area-level error-free covariates, possibly coming from administrative records, and multiple area-level covariates subject to structural measurement error where the data on the latter covariates are obtained from multiple surveys. In particular, we derive empirical best predictors of small area means and estimators of mean squared prediction error of the predictors of small area means. Performance of the proposed approach is studied through a simulation study and also by a real application.

Keywords: Conditional distribution, Jackknife, Linear mixed model, Mean squared prediction error, Measurement error

1. Introduction

Sample surveys are generally conducted to provide reliable estimates of finite population parameters such as totals, means, counts, quantiles, etc. for the nation, census regions or states. In recent years, there has been increasing demand to get such estimates for smaller sub-populations (small areas), such as counties or age-sex-race demographic groups, due to their growing use in formulating policies and programs, allocating government funds, regional planning, marketing decisions at local level, and other uses. However, sample sizes within small areas are often too small to warrant the use of traditional area-specific direct estimates.

Different methods have been proposed in the literature to produce reliable estimates of characteristics of interest for small areas and to obtain measures of error associated with such estimates. These include, among others, the use of synthetic, composite and/or model-based estimators [5, 10, 11, 13, 15]. Model-based estimators which borrow strength from related areas have been extensively used in small area estimation [15]. In particular, such small area models may be classified into two broad types: (i) area-level models that relate design-based small area direct estimates to area-specific covariates — such models are used if unit-level data are not available; (ii) unit-level models that relate the unit values of a study variable to associated unit-level covariates with known covariates area means (obtained, possibly, from administrative records) and area-specific covariates. A comprehensive account of model-based small area estimation under area-level and unit-level models is given by Rao and Molina [15]. In this paper, we focus on empirical best (EB) predictors of small area means under a unit-level nested error linear regression model with measurement errors in some area-level covariate values.

Email address: Mahmoud.Torabi@umanitoba.ca (Mahmoud Torabi)

Download English Version:

https://daneshyari.com/en/article/7546463

Download Persian Version:

https://daneshyari.com/article/7546463

Daneshyari.com