

**MODEL-ASSISTED ESTIMATION FOR COMPLEX SURVEYS  
USING PENALIZED SPLINES**

by

**F.J. Breidt, G. Claeskens and J.D. Opsomer**  
Colorado State University, Université Catholique de Louvain, Iowa State  
University

**November 1, 2003**

**ABSTRACT**

Estimation of finite population totals in the presence of auxiliary information is considered. A class of estimators based on penalized spline regression is proposed. These penalized spline survey regression estimators have many appealing operational features: like classical survey regression estimators, they are weighted linear combinations of study variables, with weights that are calibrated to known control totals. Further, they allow straightforward extensions to multiple auxiliary variables (both continuous and categorical) and to complex designs. Under standard design conditions, the penalized spline estimators are design consistent and asymptotically normal, and they admit consistent variance estimation using familiar design-based methods. Data-driven penalty selection using restricted maximum likelihood is considered in the context of unequal probability sampling designs. Simulation experiments show that the estimators are more efficient than parametric regression estimators when the parametric model is incorrectly specified, while being approximately as efficient when the parametric specification is correct. An example using Forest Health Monitoring survey data from the US Forest Service demonstrates applicability of the methodology in the context of a two-phase survey with multiple auxiliary variables.