Semiparametric Efficiency Bounds for Microeconometric Models: A Survey

Semiparametric Efficiency Bounds for Microeconometric Models: A Survey

Thomas A. Severini

Northwestern University
Evanston, IL-60201
USA
severini@northwestern.edu

Gautam Tripathi

University of Luxembourg
Luxembourg
gautam.tripathi@uni.lu

the essence of knowledge

Boston – Delft

Foundations and Trends® in Econometrics

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
USA
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is T. A. Severini and G. Tripathi, Semiparametric Efficiency Bounds for Microeconometric Models: A Survey, Foundations and Trends® in Econometrics, vol 6, nos 3–4, pp 163–397, 2013

ISBN: 978-1-60198-735-8

Foundations and Trends® in Econometrics

Volume 6 Issues 3–4, 2013

Editorial Scope

Foundations and Trends® in Econometrics will publish survey and tutorial articles in the following topics:

- Econometric Models:
  - Identification
  - Model Choice and Specification Analysis
  - Non-linear Regression Models
- Simultaneous Equation Models
- Estimation Frameworks
- Biased Estimation
- Computational Problems
- Microeconometrics
- Treatment Modeling
- Discrete Choice Modeling
- Models for Count Data
- Duration Models
- Limited Dependent Variables
- Panel Data
- Time Series Analysis:
  - Dynamic Specification
  - Inference and Causality
  - Continuous Time Stochastic Models
  - Modeling Non-linear Time Series
  - Unit Roots
  - Cointegration
- Latent Variable Models
- Qualitative Response Models
- Hypothesis Testing
- Econometric Theory:
  - Interactions-based Models
  - Duration Models
- Financial Econometrics
- Measurement Error in Survey Data
- Productivity Measurement and Analysis
- Semiparametric and Nonparametric Estimation
- Bootstrap Methods
- Nonstationary Time Series
- Robust Estimation

Information for Librarians

Foundations and Trends® in Econometrics, 2013, Volume 6, 4 issues. ISSN paper version 1551-3076. ISSN online version 1551-3084. Also available as a combined paper and online subscription.

Foundations and Trends® in
Econometrics
Vol. 6, Nos. 3–4 (2013) 163–397

DOI: 10.1561/0800000019

Semiparametric Efficiency Bounds for Microeconometric Models: A Survey

Thomas A. Severini[1] **and Gautam Tripathi**[2]

[1] *Department of Statistics, Northwestern University, Evanston, IL-60201, USA, severini@northwestern.edu*
[2] *Faculty of Law, Economics and Finance, University of Luxembourg, L-1511 Luxembourg, gautam.tripathi@uni.lu*

Abstract

In this survey, we evaluate estimators by comparing their asymptotic variances. The role of the efficiency bound, in this context, is to give a lower bound to the asymptotic variance of an estimator. An estimator with asymptotic variance equal to the efficiency bound can therefore be said to be asymptotically efficient. These bounds are also useful for understanding how the features of a given model affect the accuracy of parameter estimation.

Contents

1

Introduction

In order to determine whether a finite dimensional parameter in a semiparametric model has been efficiently estimated by a $n^{1/2}$-consistent estimator, where n denotes the sample size, one compares the asymptotic variance of the estimator with a benchmark variance. This benchmark variance, referred to as the efficiency bound, is a lower bound for the asymptotic variance of a large class of $n^{1/2}$-consistent estimators under certain regularity conditions. The aforementioned estimator is therefore said to be asymptotically efficient if its asymptotic variance equals the efficiency bound; otherwise, it is said to be asymptotically inefficient.

Apart from their obvious use in recognizing efficient estimators, another useful feature of calculating the efficiency bounds is that in many cases the calculation process is constructive enough to help construct asymptotically efficient estimators. Semiparametric models may also depend upon infinite dimensional parameters, for example, densities, conditional expectations, or other unknown functional forms, that can only be estimated at rates slower than the $n^{1/2}$-rate.

However, certain features of these unknown functions, for example, their linear functionals, can often be estimated by $n^{1/2}$-consistent estimators. Knowledge of efficiency bounds for estimating linear functionals of unknown functions allow us to measure the relative difficulty in estimating different features of these functions, thus revealing what may be learned from the data about the functions themselves. This is especially useful for the so called "ill-posed" models that have lately attracted much attention in microeconometrics, where the unknown function(s) take on endogenous arguments.

Due to the many and varied uses of efficiency bounds, it is not surprising that there is a vast literature in econometrics and statistics on calculating them. In this survey, we review some of this literature in a unified manner using the approach of Severini and Tripathi (2001). The review presented here is based on several references. For instance, Wong (1992) gives a detailed account of efficiency bounds in parametric models by connecting the seminal contributions made by Fisher (1925), LeCam (1953), Bahadur (1964), and Hájek (1970). Semiparametric efficiency bounds were introduced by Stein (1956), and discussed in papers by, among others, Levit (1974, 1975), Koshevnik and Levit (1976), Begun et al. (1983), Chamberlain (1986, 1987, 1992a,b), Cosslett (1987), van der Vaart (1989, 1991), Newey (1990c), Ai and Chen (2012), and the references therein. Book-length treatments of these topics can be found in Ibragimov and Has'minskii (1981), Pfanzagl and Wefelmeyer (1982), van der Vaart (1988, 1998), Groeneboom and Wellner (1992), and Bickel et al. (1993). Additional references will be given as the survey progresses.

Given our research interests, we confine ourselves to surveying the efficiency bounds literature for microeconometric models. Efficiency bounds can be calculated for time-series models as well, cf., for example, Hansen et al. (1988), but will not be covered by this survey. For the most part, we will restrict ourselves to the case where observed data is collected by random sampling, although we also look at efficiency bound calculations for some models that are estimated using stratified samples. The latter requires additional care because, depending on the nature of the sampling scheme, the observations may be independently but not identically distributed (i.n.i.d.).

The topics covered and the extent of details provided in this survey are highly idiosyncratic. Although we have tried to be relatively broad in our coverage, we have given the most detailed treatment only for those models we have investigated in our research. Indeed, much of the material in this survey is from our own papers, although we have tried to revise earlier treatment and add extra material in the form of additional explanation or examples whenever we could. For instance, Sections 3, 4, 5, 6, 7, 10 are drawn from Severini and Tripathi (2001), Section 8.3 from Tripathi (2000), Sections 14.1 and 14.2 from Tripathi (2011a,b), Section 15.2.1 from Devereux and Tripathi (2009), and Section 16 from Severini and Tripathi (2012a). We focus only on efficiency bound calculations. Construction of efficient estimators, the main reason why these bounds are calculated, is not touched upon in this survey although we do try and provide some selective references to this literature whenever possible.

The following notation is used throughout the survey. Additional notation will be introduced when required. By "vector," we mean a column vector. Given a set A, we use $\mathbb{1}_A$ to denote its indicator function. When thought of as an event, the indicator of A is written as $\mathbb{1}(A)$. The symbols $\bar{A}$ and $\mathrm{cl}(A)$ both denote the closure of A in some norm topology made explicit in the context. The set of real-valued functions on $\mathbb{R}^d$ which are square integrable with respect to the Lebesgue measure on $\mathbb{R}^d$ is denoted by $L_2(\mathbb{R}^d; \mathrm{Leb}^d)$, where Leb^d is the Lebesgue measure on $\mathbb{R}^d$. The Lebesgue measure on $\mathbb{R}$ is simply $\mathrm{Leb} := \mathrm{Leb}^1$. Similarly, $L_2(Z; P_Z)$ is the set of real-valued functions of a random variable (or random vector) Z that are square-integrable with respect to P_Z, the distribution of Z. When there is no ambiguity regarding the probability distribution, $L_2(Z; P_Z)$ is written simply as $L_2(Z; P)$ or even $L_2(Z)$. The support of Z is denoted by $\mathrm{supp}(Z)$. The operator $\mathscr{P}_A$ denotes orthogonal projection onto $A \subset L_2(Z; P)$ using the inner product $\langle a, b \rangle_P := \mathbb{E}_P[ab]$, where $\mathbb{E}_P$ indicates that expectation is with respect to the probability measure P. Similarly, $\mathscr{P}_{A^\perp} := I - \mathscr{P}_A$ denotes orthogonal projection onto $A^\perp$, the orthogonal complement of A, where I is the identity operator. The inner product $\langle \cdot, \cdot \rangle_P$ induces the P-norm $\| \cdot \|_{2,P} := \langle \cdot, \cdot \rangle_P^{1/2}$.

The euclidean norm of a matrix M is $\|M\| := \sqrt{\mathrm{trace}(M'M)}$. If D is an operator, for example, a matrix, then its domain, range, and null space are $\mathcal{D}(D)$, $\mathcal{R}(D)$, and $\mathcal{N}(D)$, respectively. Functional notation, where arguments taken by functions are suppressed, is used extensively whenever there is no danger of confusion.

References

Ahn, S. C. and P. Schmidt (1995), 'Efficient estimation of models for dynamic panel data'. *Journal of Econometrics* **68**, 5–27.

Ai, C. (1997), 'A semiparametric maximum likelihood estimator'. *Econometrica* **65**, 933–963.

Ai, C. and X. Chen (2003), 'Efficient estimation of models with conditional moment restrictions containing unknown functions'. *Econometrica* **71**, 1795–1843.

Ai, C. and X. Chen (2005), 'On efficient sequential estimation of semi-nonparametric moment models'. http://eswc2005.econ.ucl.ac.uk/ESWC/2005/prog/viewpaper.asp?pid=2673.

Ai, C. and X. Chen (2007), 'Estimation of possibly misspecified semiparametric conditional moment restriction models with different conditioning variables'. *Journal of Econometrics* **141**, 5–43.

Ai, C. and X. Chen (2012), 'The semiparametric efficiency bound for models of sequential moment restrictions containing unknown functions'. *Journal of Econometrics* **170**, 442–457.

Akhiezer, N. and I. Glazman (1993), *Theory of Linear Operators in Hilbert Space,* vol. 1. Dover Publications.

Amemiya, T. (1984), 'Tobit models: A survey'. *Journal of Econometrics* **4**, 3–61.

Andrews, G. E., R. Askey, and R. Roy (1999), *Special Functions*. Cambridge University Press.

Angrist, J. D. and A. B. Kreuger (1992), 'The effect of age at school entry on educational attainment: An application of instrumental variables with moments from two samples'. *Journal of the American Statistical Association* **87**, 328–336.

Arellano, M. and S. Bond (1991), 'Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations'. *Review of Economic Studies* **58**, 277–297.

Arellano, M. and O. Bover (1995), 'Another look at the instrumental variable estimation of error-components models'. *Journal of Econometrics* **68**, 29–51.

Arellano, M. and C. Meghir (1992), 'Female labor supply and on-the-job search: An empirical model estimated using complementary data sets'. *Review of Economic Studies* **59**, 537–557.

Bahadur, R. R. (1964), 'On Fisher's bound for asymptotic variances'. *Annals of Mathematical Statistics* **35**, 1545–1552.

Begun, J., W. Hall, W. Huang, and J. Wellner (1983), 'Information and asymptotic efficiency in parametric-nonparametric models'. *Annals of Statistics* **11**, 432–452.

Bickel, P. J. (1982), 'On adaptive estimation'. *Annals of Statistics* **10**, 647–671.

Bickel, P. J., C. A. J. Klassen, Y. Ritov, and J. A. Wellner (1993), *Efficient and Adaptive Estimation for Semiparametric Models*. Johns Hopkins Press.

Bickel, P. J. and Y. Ritov (1991), 'Large sample theory of estimation in biased sampling regression models'. *Annals of Statistics* **19**, 797–816.

Blundell, R., X. Chen, and D. Kristensen (2007), 'Semi-nonparametric IV estimation of shape-invariant Engel curves'. *Econometrica* **75**, 1613–1669.

Blundell, R. W. and R. J. Smith (1993), 'Simultaneous microeconometric models with censored or qualitative dependent variables'. In: *Econometrics,* vol. 11 of *Handbook of Statistics*. North-Holland, Amsterdam, pp. 117–143.

Brown, B. W. and W. K. Newey (1998), 'Efficient semiparametric estimation of expectations'. *Econometrica* **66**, 453–464.

Brown, B. W. and W. K. Newey (2002), 'GMM, efficient bootstrapping, and improved inference'. *Journal of Business and Economic Statistics* **20**, 507–517.

Butler, J. S. (2000), 'Efficiency results of MLE and GMM estimation with sampling weights'. *Journal of Econometrics* **96**, 25–37.

Carrasco, M., J.-P. Florens, and E. Renault (2007), 'Linear inverse problems in structural econometrics estimation based on spectral decompisition and regularization'. In: J. J. Heckman and E. E. Leamer (eds.): *Handbook of Econometrics,* vol. 6B. Elsevier Science B.V, pp. 5633–5751.

Chamberlain, G. (1986), 'Asymptotic efficiency in semiparametric models with censoring'. *Journal of Econometrics* **32**, 189–218.

Chamberlain, G. (1987), 'Asymptotic efficiency in estimation with conditional moment restrictions'. *Journal of Econometrics* **34**, 305–334.

Chamberlain, G. (1992a), 'Comment: Sequential moment restrictions in panel data'. *Journal of Business and Economic Statistics* **10**, 20–26.

Chamberlain, G. (1992b), 'Efficiency bounds for semiparametric regression'. *Econometrica* **60**, 567–596.

Chen, T. and T. Parker (2013), 'Semi-parametric efficiency for single index regression models under conditional mean and quantile restrictions'. *Manuscript*.

Chen, X., H. Hong, and E. Tamer (2005), 'Measurement error models with auxiliary data'. *Review of Economic Studies* **72**, 343–366.

Chen, X., H. Hong, and A. Tarozzi (2006), 'Semiparametric efficiency in GMM models with auxiliary data'. *Annals of Statistics* **36**, 343–366.

Chen, X. and D. Pouzo (2009), 'Efficient estimation of semiparametric conditional moment models with possibly nonsmooth residuals'. *Journal of Econometrics* **152**, 46–60.

Chen, X. and D. Pouzo (2012), 'Estimation of nonparametric conditional moment models with possibly nonsmooth generalized residuals'. *Econometrica* **80**, 277–321.

Cochran, W. G. (1963), *Sampling Techniques*. 2nd edition. John Wiley & Sons.

Coppejans, M. (2007), 'On efficient estimation of the ordered response model'. *Journal of Econometrics* **137**, 577–614.

Cosslett, S. R. (1981a), 'Efficient estimation of discrete choice models'. In: C. F. Manski and D. McFadden (eds.): *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge, MA: MIT Press, pp. 51–111.

Cosslett, S. R. (1981b), 'Maximum likelihood estimation for choice-based samples'. *Econometrica* **49**, 1289–1316.

Cosslett, S. R. (1987), 'Efficiency bounds for distribution free estimators of binary choice and censored regression models'. *Econometrica* **55**, 559–585.

Cosslett, S. R. (1991), 'Efficient estimation from endogenously stratified samples with prior information on marginal probabilities'. *Manuscript*.

Cosslett, S. R. (1993), 'Estimation from endogenously stratified samples'. In: G. Maddala, C. Rao, and H. Vinod (eds.): *Handbook of Statistics,* vol. 11. The Netherlands: Elsevier Science, pp. 1–43.

Cosslett, S. R. (2004), 'Efficient semiparametric estimation of censored and truncated regressions via a smoothed self-consistency equation'. *Econometrica* **72**, 1277–1293.

Cosslett, S. R. (2013), 'Efficient semiparametric estimation for endogenously stratified regression via smoothed likelihood'. *Journal of Econometrics* **177**, 116–129.

Cuzick, J. (1992), 'Efficient estimates in semiparametric additive regression models with unknown error distribution'. *Annals of Statistics* **20**, 1129–1136.

Dalalyan, A. S., G. K. Golubev, and A. B. Tsybakov (2006), 'Penalized maximum likelihood and semiparametric second-order efficiency'. *Annals of Statistics* **34**, 169–201.

Darolles, S., Y. Fan, J.-P. Florens, and E. Renault (2011), 'Nonparametric instrumental regression'. *Econometrica* **79**, 1541–1565.

Darolles, S., J.-P. Florens, and E. Renault (2006), 'Nonparametric instrumental regression'. http://idei.fr/doc/by/florens/renaultdarolles.pdf.

DeMets, D. and M. Halperin (1977), 'Estimation of a simple regression coefficient in samples arising from a subsampling procedure'. *Biometrics* **33**, 47–56.

Devereux, P. and G. Tripathi (2009), 'Optimally combining censored and uncensored datasets'. *Journal of Econometrics* **151**, 17–32.

Edgeworth, F. Y. (1908), 'On the probable errors of frequency-constants'. *Journal of the Royal Statistical Society* **71**, 381–397, 499–512, 651–678.

Edgeworth, F. Y. (1909), 'Addendum on "Probable errors of frequency-constants"'. *Journal of the Royal Statistical Society* **72**, 81–90.

Efromovich, S. (2004), 'Distribution estimation for biased data'. *Journal of Statistical Planning and Inference* **124**, 1–43.

El-Barmi, H. and M. Rothmann (1998), 'Nonparametric estimation in selection biased models in the presence of estimating equations'. *Nonparametric Statistics* **9**, 381–399.

Engl, H. W., M. Hanke, and A. Neubauer (2000), *Regularization of Inverse Problems*. Kluwer.

Fisher, R. A. (1925), 'Theory of statistical estimation'. *Proceedings of the Cambridge Philosophical Society* **22**, 700–725.

Florens, J.-P., J. Johannes, and S. van Bellegem (2012), 'Instrumental regression in partially linear models'. *Econometrics Journal* **15**, 304–324.

Ghosh, J. K. (1994), *Higher Order Asymptotics*. Hayward, CA: Institute of Mathematical Statistics.

Goto, F. (1996), 'Achieving semiparametric efficiency bounds in left-censored duration models'. *Econometrica* **64**, 439–442.

Gourieroux, C. and A. Monfort (1995), *Statistics and Econometric Models,* vol. 2. Great Britain: Cambridge University Press.

Graham, B. S. (2011), 'Efficiency bounds for missing data models with semiparametric restrictions'. *Econometrica* **79**, 437–452.

Groeneboom, P. and J. A. Wellner (1992), *Information Bounds and Nonparametric Maximum Likelihood Estimation*. Birkhäuser, Basel.

Hahn, J. (1994), 'The efficiency bound of the mixed proportional hazard model'. *Review of Economic Studies* **61**, 607–629.

Hahn, J. (1997), 'Efficient estimation of panel data models with sequential moment restrictions'. *Journal of Econometrics* **79**, 1–21.

Hahn, J. (1998), 'On the role of the propensity score in efficient semiparametric estimation of average treatment effects'. *Econometrica* **66**, 315–331.

Hahn, J. (2001), 'The information bound of a dynamic panel logit model with fixed effects'. *Econometric Theory* **17**, 913–932.

Hájek, J. (1970), 'A characterization of limiting distributions of regular estimates'. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **14**, 323–330.

Hajivassiliou, V. A. and P. A. Ruud (1994), 'Classical estimation methods for LDV models using simulation'. In: R. F. Engle and D. L. McFadden (eds.): *Handbook of Econometrics,* vol. IV. The Netherlands: Elsevier Science B.V., pp. 2383–2441.

Hall, P. and J. L. Horowitz (1996), 'Bootstrap critical values for tests based on generalized method of moments estimators'. *Econometrica* **64**, 891–916.

Hall, P. and J. L. Horowitz (2005), 'Nonparametric methods for inference in the presence of instrumental variables'. *Annals of Statisics* **33**, 2904–2929.

Hansen, L. P. (1982), 'Large sample properties of generalized methods of moments estimators'. *Econometrica* **50**, 1029–1054.

Hansen, L. P., J. C. Heaton, and M. Ogaki (1988), 'Efficiency bounds implied by multiperiod conditional moment restrictions'. *Journal of the American Statistical Association* **83**, 863–871.

Härdle, W. and T. Stoker (1989), 'Investigating smooth multiple regression by the method of average derivatives'. *Journal of the American Statistical Association* **84**, 986–995.

Härdle, W. and A. B. Tsybakov (1993), 'How sensitive are average derivatives?'. *Journal of Econometrics* **58**, 31–48.

Harville, D. A. (1997), *Matrix Algebra from a Statistician's Perspective.* New York: Springer-Verlag.

Hausman, J. A. and D. A. Wise (1981), 'Stratification on endogenous variables and estimation: The Gary income maintenance experiment'. In: C. F. Manski and D. McFadden (eds.): *Structural analysis of discrete data with econometric applications*. Cambridge, MA: MIT Press, pp. 365–391.

Heckman, J. J. (1985), 'Selection bias and self-selection'. In: *The New Palgrave: A Dictionary of Economics*. Stockton, New-York: MacMillan Press, pp. 287–297.

Hellerstein, J. and G. W. Imbens (1999), 'Imposing moment restrictions from auxiliary data by weighting'. *Review of Economics and Statistics* **81**, 1–14.

Hirano, K., G. W. Imbens, G. Ridder, and D. B. Rubin (2001), 'Combining panel data sets with attrition and refreshment samples'. *Econometrica* **69**, 1645–1659.

Hirose, Y. (2007), 'M-estimators in semi-parametric multi-sample models'. *Manuscript*.

Hirose, Y. and A. J. Lee (2008), 'Semi-parametric efficiency bounds for regression models under generalised case-control sampling: The profile likelihood approach'. *Annals of the Institute of Statistical Mathematics* **62**, 1023–1052.

Holt, D., T. M. F. Smith, and P. D. Winter (1980), 'Regression analysis of data from complex surveys'. *Journal of The Royal Statistical Society, Series A* **143**, 474–487.

Honoré, B. and E. Kyriazidou (2000), 'Panel data discrete choice models with lagged dependent variables'. *Econometrica* **68**, 839–874.

Hu, Y. and G. Ridder (2012), 'Estimation of nonlinear models with measurement error using marginal information'. *Journal of Applied Econometrics* **27**, 347–385.

Hua, L. and C. Ping (1993), 'Second order asymptotic efficiency in a partial linear model'. *Statistics and Proability Letters* **18**, 73–84.

Huang, J. (1999), 'Efficient estimation of the partly linear additive Cox model'. *Annals of Statistics* **27**, 1536–1563.

Ibragimov, I. A. and R. Z. Has'minskii (1981), *Statistical Estimation–Asymptotic Theory*. New York: Springer-Verlag.

Ichimura, H. and E. Martinez-Sanchis (2005), 'Identification and estimation of GMM models by a combination of two data sets'. *Manuscript*.

Imbens, G. W. (1992), 'An efficient method of moments estimator for discrete choice models with choice-based sampling'. *Econometrica* **60**, 1187–1214.

Imbens, G. W. and T. Lancaster (1994), 'Combining micro and macro data in microeconomic models'. *Review of Economic Studies* **61**, 655–680.

Imbens, G. W. and T. Lancaster (1996), 'Efficient estimation and stratified sampling'. *Journal of Econometrics* **74**, 289–318.

Jacho-Chávez, D. T. (2009), 'Efficiency bounds for semiparametric estimation of inverse conditional density weighted functions'. *Econometric Theory* **25**, 847–855.

Jewell, N. P. (1985), 'Least squares regression with data arising from stratified samples of the dependent variable'. *Biometrika* **72**, 11–21.

Kagan, A. M., Y. V. Linnik, and C. R. Rao (1973), *Characterization Problems in Mathematical Statistics*. New York: John Wiley & Sons.

Khan, S. and D. Nekipelov (2008), 'A comment on alternative methods for semiparametric efficiency bound calculations'. *Manuscript*.

Khan, S. and E. Tamer (2010), 'Irregular identification, support conditions, and inverse weight estimation'. *Econometrica* **78**, 2021–2042.

Kim, J. and D. Pollard (1990), 'Cube root asymptotics'. *Annals of Statistics* **18**, 191–219.

Kitamura, Y., G. Tripathi, and H. Ahn (2004), 'Empirical likelihood based inference in conditional moment restriction models'. *Econometrica* **72**, 1667–1714.

Klein, R. W. and R. H. Spady (1993), 'An efficient semiparametric estimator for binary response models'. *Econometrica* **61**, 387–421.

Koenker, R. and G. Bassett (1978), 'Regression quantiles'. *Econometrica* **46**, 33–50.

Koshevnik, Y. A. and B. Y. Levit (1976), 'On a non-parametric analogue of the information matrix'. *Theory of Probability and its Applications* **21**, 738–753.

Krabs, W. (1979), *Optimization and Approximation*. John Wiley and Sons.

Kress, R. (1999), *Linear Integral Equations*. 2nd edition. Springer Verlag.

Kreyszig, E. (1978), *Introductory Functional Analysis with Applications*. John Wiley and Sons.

LeCam, L. (1953), 'On some asymptotic properties of maximum likelihood estimates and related Bayes' estimates'. *University of California Publications in Statistics* **1**, 277–330.

Lee, L.-F. (1995), 'Semiparametric maximum likelihood estimation of polychotomous and sequential choice models'. *Journal of Econometrics* **65**, 381–428.

Lee, S. (2003), 'Efficient semiparametric estimation of a partially linear quantile regression model'. *Econometric Theory* **19**, 1–31.

Levit, B. Y. (1974), 'On optimality of some statistical estimates'. *Proceedings of the Prague Symposium on Asymptotic Statistics* **2**, 215–238.

Levit, B. Y. (1975), 'On the efficiency of a class of nonparametric estimates'. *Theory of probability and its applications* **20**, 723–740.

Lewbel, A. (1998), 'Semiparametric latent variable model estimation with endogenous or mismeasured regressors'. *Econometrica* **66**, 105–121.

Lewbel, A. (2000), 'Semiparametric qualitative response model estimation with unknown heteroscedasticity or instrumental variables'. *Journal of Econometrics* **97**, 145–177.

Lewbel, A. (2012), 'An overview of the special regressor method'. *Manuscript*. http://www2.bc.edu/∼lewbel.

Luenberger, D. G. (1969), *Optimization by Vector Space Methods*. John Wiley and Sons.

Magnac, T. and E. Maurin (2007), 'Identification and information in monotone binary models'. *Journal of Econometrics* **139**, 76–104.

Manski, C. F. (1975), 'Maximum score estimation of the stochastic utility model of choice'. *Journal of Econometrics* **3**, 205–228.

Manski, C. F. (1984), 'Adaptive estimation of non-linear regression models'. *Econometric Reviews* **3**, 145–194.

Manski, C. F. and S. R. Lerman (1977), 'The estimation of choice probabilities from choice based samples'. *Econometrica* **45**, 1977–1988.

Manski, C. F. and D. McFadden (1981), 'Alternative estimators and sample design for discrete choice analysis'. In: C. F. Manski and D. McFadden (eds.): *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge, MA: MIT Press, pp. 2–50.

Matzkin, R. (1994), 'Restrictions of economic theory in nonparametric methods'. In: R. Engle and D. McFadden (eds.): *Handbook of Econometrics,* vol. IV. The Netherlands: Elsevier Science B.V., pp. 2524–2558.

McNeney, B. and J. A. Wellner (2000), 'Application of convolution theorems in semiparametric models with non-i.i.d. data'. *Journal of Statistical Planning and Inference* **91**, 441–480.

Müller, U. U., A. Schick, and W. Wefelmeyer (2004), 'Estimating linear functionals of the error distribution in nonparametric regression'. *Journal of Statistical Planning and Inference* **119**, 75–93.

Nan, B., M. J. Emond, and J. A. Wellner (2004), 'Information bounds for Cox regression models with missing data'. *Annals of Statistics* **32**, 723–753.

Nevo, A. (2003), 'Using weights to adjust for sample selection when auxiliary information is available'. *Journal of Business and Economic Statistics* **21**, 43–52.

Newey, W. K. (1988), 'Efficient estimation of Tobit models under conditional symmetry'. In: W. A. Barnett, J. Powell, and G. Tauchen (eds.): *Nonparametric and Semiparametric Methods in Econometrics and Statistics. Proceedings of the Fifth International Symposium in Economic Theory and Econometrics*. U.S.A.: Cambridge University Press, pp. 291–336.

Newey, W. K. (1990a), 'Adaptive estimation of regression models via moment restrictions'. *Journal of Econometrics* **38**, 301–339.

Newey, W. K. (1990b), 'Efficient instrumental variables estimation of nonlinear models'. *Econometrica* **58**, 809–837.

Newey, W. K. (1990c), 'Semiparametric efficiency bounds'. *Journal of Applied Econometrics* **5**, 99–135.

Newey, W. K. (1993), 'Efficient estimation of models with conditional moment restrictions'. In: G. S. Maddala, C. R. Rao, and H. D. Vinod (eds.): *Handbook of Statistics,* vol. 11. The Netherlands: Elsevier Science B.V., pp. 2111–2245.

Newey, W. K. (1994), 'The asymptotic variance of semiparametric estimators'. *Econometrica* **62**, 1349–1382.

Newey, W. K. (2001), 'Conditional moment restrictions in censored and truncated regression models'. *Econometric Theory* **17**, 863–888.

Newey, W. K. (2004), 'Efficient semiparametric estimation via moment restrictions'. *Econometrica* **72**, 1877–1897.

Newey, W. K. and D. McFadden (1994), 'Large sample estimation and hypothesis testing'. In: R. F. Engle and D. L. McFadden (eds.): *Handbook of Econometrics,* vol. IV. The Netherlands: Elsevier Science B.V., pp. 2111–2245.

Newey, W. K. and J. L. Powell (1990), 'Efficient estimation of linear and type I censored regression models under conditional quantile restrictions'. *Econometric Theory* **6**, 295–317.

Newey, W. K. and J. L. Powell (2003), 'Instrumental variables estimation of nonparametric models'. *Econometrica* **71**, 1557–1569.

Newey, W. K. and T. M. Stoker (1993), 'Efficiency of weighted average derivative estimators and index models'. *Econometrica* **61**, 1199–1223.

Owen, A. (2001), *Empirical Likelihood.* Chapman and Hall/CRC.

Park, B. U., R. C. Sickles, and L. Simar (2003), 'Semiparametric efficient estimation of AR(1) panel data models'. *Journal of Econometrics* **117**, 279–309.

Park, B. U., R. C. Sickles, and L. Simar (2007), 'Semiparametric efficient estimation of dynamic panel data models'. *Journal of Econometrics* **136**, 281–301.

Peng, H. and A. Schick (2005), 'Efficient estimation of linear functionals of a bivariate distribution with equal, but unknown marginals: The least-squares approach'. *Journal of Multivariate Analysis* **95**, 385–409.

Pfanzagl, J. and W. Wefelmeyer (1982), *Contributions to a General Asymptotic Statistical Theory.* New York: Springer-Verlag.

Powell, J. L. (1984), 'Least absolute deviations estimation for the censored regression model'. *Journal of Econometrics* **25**, 303–325.

Powell, J. L. (1986a), 'Censored regression quantiles'. *Journal of Econometrics* **32**, 143–155.

Powell, J. L. (1986b), 'Symmetrically trimmed least squares estimation for tobit models'. *Econometrica* **54**, 1435–1460.

Powell, J. L. (1994), 'Estimation of semiparametric models'. In: R. F. Engle and D. L. McFadden (eds.): *Handbook of Econometrics,* vol. IV. The Netherlands: Elsevier Science B.V., pp. 2443–2521.

Powell, J. L., J. H. Stock, and T. M. Stoker (1989), 'Semiparametric estimation of index coefficients'. *Econometrica* **57**, 1403–1430.

Pratt, J. W. (1976), 'F. Y. Edgeworth and R. A. Fisher on the efficiency of maximum likelihood estimation'. *Annals of Statistics* **4**, 501–514.

Qin, J. (1993), 'Empirical likelihood in biased sample problems'. *Annals of Statistics* **21**, 1182–1196.

Quesenberry, C. P. and N. P. Jewell (1986), 'Regression analysis based on stratified samples'. *Biometrika* **73**, 605–614.

Rao, C. R. (1973), *Linear Statistical Inference and Its Applications*. 2nd edition. New York: John Wiley & Sons.

Ridder, G. and R. Moffitt (2007), 'The econometrics of data combination'. In: J. J. Heckman and E. E. Leamer (eds.): *Handbook of Econometrics,* vol. 6B. Elsevier Science B.V, pp. 5469–5547.

Ridder, G. and T. M. Woutersen (2003), 'The singularity of the information matrix of the mixed proportional hazard model'. *Econometrica* **71**, 1579–1589.

Rilstone, P. (1993), 'Calculating the (local) semiparametric efficiency bounds for the generated regressors problem'. *Journal of Econometrics* **56**, 357–370.

Ritov, Y. (1990), 'Estimation in a linear regression model with censored data'. *Annals of Statistics* **18**, 303–328.

Ritov, Y. and P. J. Bickel (1990), 'Achieving information bounds in non and semiparametric models'. *Annals of Statistics* **18**, 925–938.

Robins, J. M., F. Hsieh, and W. K. Newey (1995), 'Semiparametric efficient estimation of a conditional density with missing or mismeasured covariates'. *Journal of the Royal Statistical Society, Series B* **57**, 409–424.

Robins, J. M. and A. Rotnitzky (1995), 'Semiparametric efficiency in multivariate regression models with missing data'. *Journal of the American Statistical Association* **90**, 122–129.

Robins, J. M., A. Rotnitzky, and L. P. Zhao (1994), 'Estimation of regression coefficients when some regressors are not always observed'. *Journal of the American Statistical Association* **89**, 846–866.

Robinson, P. M. (1987), 'Asymptotically efficient estimation in the presence of heteroskedasticity of unknown form'. *Econometrica* **55**, 875–891.

Robinson, P. M. (1988), 'Root-N-consistent semiparametric regression'. *Econometrica* **56**, 931–954.

Rudin, W. (1974), *Real and Complex Analysis*. McGraw-Hill.

Samarov, A. (1990), 'On asymptotic efficiency of average derivative estimates'. In: G. Roussas (ed.): *Nonparametric Functional Estimation and Related Topics,* vol. 335 of *NATO ASI Series, Series C*. Kluwer Academic Publishers, pp. 167–172.

Santos, A. (2011), 'Instrumental variable methods for recovering continuous linear functionals'. *Journal of Econometrics* **161**, 129–146.

Sasieni, P. (1992), 'Information bounds for the conditional hazard ratio in a nested family of regression models'. *Journal of the Royal Statistical Society, Series B* **54**, 617–635.

Scott, A. J. and C. J. Wild (1986), 'Fitting logistic models under case-control or choice based sampling'. *Journal of The Royal Statistical Society, Series B* **48**, 170–182.

Sen, P. K. and J. M. Singer (1993), *Large Sample Methods in Statistics: An Introduction with Applications*. London: Chapman & Hall.

Severini, T. A. and G. Tripathi (2001), 'A simplified approach to computing efficiency bounds in semiparametric models'. *Journal of Econometrics* **102**, 23–66.

Severini, T. A. and G. Tripathi (2006), 'Some identification issues in nonparametric linear models with endogenous regressors'. *Econometric Theory* **22**, 258–278.

Severini, T. A. and G. Tripathi (2012a), 'Efficiency bounds for estimating linear functionals of nonparametric regression models with endogenous regressors'. *Journal of Econometrics* **170**, 491–498.

Severini, T. A. and G. Tripathi (2012b), 'Nonparametric estimation of returns to scale'. *Manuscript*.

Severini, T. A. and G. Tripathi (2013), 'Efficiency bounds for binary choice models with "special" regressors'. *Manuscript*.

Small, C. G. and J. Wang (2003), *Numerical Methods for Nonlinear Estimating Equations*. New York: Oxford University Press.

Smith, R. J. (1997), 'Alternative semi-parametric likelihood approaches to generalized method of moments estimation'. *Economic Journal* **107**, 503–519.

Stein, C. (1956), 'Efficient nonparametric testing and estimation'. In: *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability,* vol. 1. Berkeley, CA: University of California Press, pp. 187–195.

Stock, J. H. (1989), 'Nonparametric policy analysis'. *Journal of the American Statistical Association* **84**, 567–575.

Thompson, T. S. (1993), 'Some efficiency bounds for semiparametric discrete choice models'. *Journal of Econometrics* **58**, 257–274.

Tripathi, G. (1999), 'A matrix extension of the Cauchy-Schwarz inequality'. *Economics Letters* **63**, 1–3.

Tripathi, G. (2000), 'Local semiparametric efficiency bounds under shape restrictions'. *Econometric Theory* **16**, 729–739.

Tripathi, G. (2008), 'Inference in conditional moment restriction models when there is selection due to stratification'. *Manuscript*.

Tripathi, G. (2011a), 'Generalized method of moments (GMM) based inference with stratified samples when the aggregate shares are known'. *Journal of Econometrics* **165**, 258–265.

Tripathi, G. (2011b), 'Moment based inference with stratified data'. *Econometric Theory* **27**, 47–73.

van der Vaart, A. (1988), *Statistical Estimation in Large Parameter Spaces*. Amsterdam, The Netherlands: CWI Tract 44, Centrum voor Wiskunde en Informatica.

van der Vaart, A. (1989), 'On the asymptotic information bound'. *Annals of Statistics* **17**, 1487–1500.

van der Vaart, A. (1991), 'On differentiable functionals'. *Annals of Statistics* **19**, 178–204.

van der Vaart, A. (1998), *Asymptotic Statistics*. New York: Cambridge University Press.

van der Vaart, A. and J. A. Wellner (1996), *Weak Convergence and Empirical Processes: With Applications to Statistics*. New York: Springer-Verlag.

Vardi, Y. (1982), 'Nonparametric estimation in the presence of length bias'. *Annals of Statistics* **10**, 616–620.

Wong, W. H. (1986), 'Theory of partial likelihood'. *Annals of Statistics* **14**, 88–123.

Wong, W. H. (1992), 'On asymptotic efficiency in estimation theory'. *Statistica Sinica* **2**, 47–68.

Wong, W. H. and T. A. Severini (1991), 'On maximum likelihood estimation in infinite dimensional parameter spaces'. *Annals of Statistics* **19**, 603–632.

Wooldridge, J. M. (1999), 'Asymptotic properties of weighted M-estimators for variable probability samples'. *Econometrica* **67**, 1385–1406.

Wooldridge, J. M. (2001), 'Asymptotic properties of weighted M-estimators for standard stratified samples'. *Econometric Theory* **17**, 451–470.

Zhang, B. (1999), 'Bootstrapping with auxiliary information'. *The Canadian Journal of Statistics* **27**, 237–249.