Econometric Theory, **16**, 2000, 729–739. Printed in the United States of America.

LOCAL SEMIPARAMETRIC EFFICIENCY BOUNDS UNDER SHAPE RESTRICTIONS

GAUTAM TRIPATHI
University of Wisconsin

Consider the model $y = \mathbf{x}'\boldsymbol{\beta}_0 + f^*(\mathbf{z}) + \varepsilon$, where $\varepsilon \stackrel{\mathrm{d}}{=} \mathrm{N}(0, \sigma_0^2)$. We calculate the smallest asymptotic variance that $n^{1/2}$ consistent regular ($n^{1/2}$CR) estimators of $\boldsymbol{\beta}_0$ can have when the only information we possess about f^* is that it has a certain shape. We focus on three particular cases: (i) when f^* is homogeneous of degree r, (ii) when f^* is concave, (iii) when f^* is decreasing. Our results show that in the class of all $n^{1/2}$CR estimators of $\boldsymbol{\beta}_0$, homogeneity of f^* may lead to substantial asymptotic efficiency gains in estimating $\boldsymbol{\beta}_0$. In contrast, at least asymptotically, concavity and monotonicity of f^* do not help in estimating $\boldsymbol{\beta}_0$ more efficiently, at least for $n^{1/2}$CR estimators of $\boldsymbol{\beta}_0$.

1. INTRODUCTION

The objective of this paper is to determine the minimum asymptotic variance, hereafter called the efficiency bound, for estimating finite dimensional parameters in an important class of econometric models. These models, referred to as partially linear models (henceforth PLM's), take the form $y = \mathbf{x}'\boldsymbol{\beta}_0 + f^*(\mathbf{z}) + \varepsilon$. Here y is the response variable, $(\mathbf{x}, \mathbf{z})$ the covariates, $\boldsymbol{\beta}_0$ a finite dimensional parameter of interest, and ε an unobserved "error" term. The functional form of f^*, although unknown, is known to possess certain shape properties such as homogeneity, concavity, or monotonicity. These properties are often the restrictions that economic theory imposes upon unknown functional forms. We let $\mathcal{F}$ denote the set of certain shape restricted functions. Semiparametric estimation of $\boldsymbol{\beta}_0$, when $\mathcal{F}$ is just a set of functions satisfying some smoothness properties, has been studied extensively. However, when dealing with economic data, the functions in $\mathcal{F}$ often have to satisfy certain shape restrictions besides smoothness. Although these restrictions have been extensively studied in economic theory, they have received only limited use in econometric practice notwithstanding their tremendous usefulness. As pointed out in Matzkin (1994, p. 2525),

I thank Joel Horowitz and two anonymous referees for comments that greatly improved this paper. I am also indebted to Rosa Matzkin and Tom Severini for their constant support and advice. A previous version of this paper was circulated under the title "Semiparametric Efficiency Bounds under Shape Restrictions." Financial support from the University of Wisconsin Graduate School is gratefully acknowledged. Address correspondence to: Gautam Tripathi, Department of Economics, University of Wisconsin, Madison, WI 53706, USA; e-mail gtripath@ssc.wisc.edu.

© 2000 Cambridge University Press 0266-4666/00 $9.50

https://doi.org/10.1017/S0266466600165053 Published online by Cambridge University Press

imposing shape restrictions may lead to a reduction in the variance of parameter estimators. Omitting shape restrictions, when economic theory demands otherwise, would typically lead to inefficient estimation procedures, thus reducing the power of subsequent statistical analysis.

A very readable account of computing efficiency bounds in semiparametric models can be found in Newey (1990). However, as this paper illustrates, much of the research to date has concentrated upon developing efficiency bounds for distribution free models, i.e., models in which the distribution of the error terms is unknown (Chamberlain, 1986; Cosslett, 1987). Where shape restrictions have been involved, they have been imposed on the error distribution (Newey, 1988), rather than on the unknown function. In fact, to the best of my knowledge, this paper is the first of its kind to develop efficiency bounds for models where the shape restrictions are imposed on the unknown functional form rather than on the distribution of the error term. The bounds that we obtain are local in nature because we have specified the distribution of the error term to be Gaussian. As Rilstone (1993, p. 361) points out, this may not be as severe a restriction as one might think. Furthermore, imposition of normality is a convenient assumption that allows us to focus on the main object of our investigation, namely, shape restrictions and their effect upon efficiency bounds. The results in this paper should be of particular interest to practitioners in the field, because they provide some new insights into incorporating shape restrictions in estimation procedures.

The paper is organized as follows. Section 2 lists the maintained assumptions. Efficiency bounds for $\boldsymbol{\beta}_0$, when f^* is homogeneous of degree r, are described in Section 3. A simple example is also provided to illustrate possible efficiency gains due to homogeneity. Section 4 (resp. Section 5) explores the case when f^* is concave (resp. decreasing). The form of the efficiency bounds obtained in these two cases is used as evidence to conclude that in the class of all $n^{1/2}$ consistent regular (henceforth $n^{1/2}$CR) estimators of $\boldsymbol{\beta}_0$, concavity and monotonicity of f^* do not help in estimating $\boldsymbol{\beta}_0$ more efficiently in large samples, i.e., asymptotically. Section 6 concludes. Section 7 provides the notation for the Appendixes and a useful definition. A quick review of calculating efficiency bounds, along with all computational and technical details for the different cases, is relegated to the Appendixes.

1.1. Notation

Let $\mathcal{S}$ be a convex, compact subset of $\mathbb{R}^q$, $\mathrm{C}^2(\mathcal{S})$ the set of all real valued twice continuously differentiable functions on $\mathcal{S}$, and $\mathrm{L}^2(\mathcal{S})$ the set of all square integrable (w.r.t. the probability distribution on $\mathcal{S}$) functions on $\mathcal{S}$. $\mathcal{F}_h \subset \mathrm{C}^2(\mathcal{S})$ is the set of all $\mathrm{C}^2(\mathcal{S})$ functions that are also homogeneous of degree r, and $\mathcal{F}_c \subset \mathrm{C}^2(\mathcal{S})$ is the set of all C^2 functions that are concave on $\mathcal{S}$. In the one-dimensional case $q = 1$, $\mathcal{S}$ becomes a compact interval in $\mathbb{R}$, $\mathrm{C}^1(\mathcal{S})$ denotes the

https://doi.org/10.1017/S0266466600165053 Published online by Cambridge University Press

set of all continuously differentiable functions on $\mathcal{S}$, and $\mathcal{F}_d = \{f \in C^1(\mathcal{S}) : f'(z) \leq 0 \ \forall z \in \mathcal{S}\}$ is the set of all C^1 decreasing functions on $\mathcal{S}$. All vectors are expressed in boldface and are treated as column vectors. The usual Euclidean norm is denoted by $\|\cdot\|$.

2. THE PARTIALLY LINEAR MODEL

Consider the regression $y_i = \mathbf{x}_i'\boldsymbol{\beta}_0 + f^*(\mathbf{z}_i) + \varepsilon_i$, for $i = 1, \ldots, n$. The following assumptions are maintained throughout the paper.

(A.1) The data $(\mathbf{x}_i, y_i, \mathbf{z}_i)$ are independent and identically distributed (i.i.d.) random vectors in $\mathbb{R}^p \times \mathbb{R} \times \mathbb{R}^q$.
(A.2) $\mathbb{E}\|\mathbf{x}\|^2 < \infty$, and $\mathbf{z}$ is drawn from a distribution with support $\mathcal{S}$. The joint probability density function (p.d.f.) of $\mathbf{x}$ and $\mathbf{z}$ exists and is denoted by $g_{\mathbf{x},\mathbf{z}}$.
(A.3) $\varepsilon \overset{d}{=} N(0, \sigma_0^2)$ is unobserved and is independent of $(\mathbf{x}, \mathbf{z})$.
(A.4) $\boldsymbol{\beta}_0 \in \mathbb{R}^p$, and f^* is an element of $\mathcal{F}_h$, or $\mathcal{F}_c$, or $\mathcal{F}_d$.
(A.5) The matrix $\mathbb{E}\{[\mathbf{x} - \mathbb{E}(\mathbf{x}|\mathbf{z})][\mathbf{x}' - \mathbb{E}(\mathbf{x}'|\mathbf{z})]\}$ is nonsingular.

As $\boldsymbol{\beta}_0$ is ancillary to σ_0^2 and $g_{\mathbf{x},\mathbf{z}}$, efficiency bounds for $\boldsymbol{\beta}_0$ will remain unchanged whether σ_0^2 and $g_{\mathbf{x},\mathbf{z}}$ are known or not (Newey, 1990, p. 109). Therefore, to remain focused upon the shape restrictions and to minimize algebraic details, assume w.l.o.g that σ_0^2 and $g_{\mathbf{x},\mathbf{z}}$ are known. Assumption (A.5) ensures identification of $\boldsymbol{\beta}_0$, because $\boldsymbol{\beta}_0$ (without any intercept term) is identified iff $\mathbb{E}\{[\mathbf{x} - \mathbb{E}(\mathbf{x}|\mathbf{z})][\mathbf{x}' - \mathbb{E}(\mathbf{x}'|\mathbf{z})]\}$ is nonsingular.

3. EFFICIENCY BOUNDS WHEN f^* IS HOMOGENEOUS

We now describe the efficiency bounds for estimating $\boldsymbol{\beta}_0$ in the PLM when the only thing we know about f^* is that it is homogeneous[1] of degree r, where r is known; i.e., $f^* \in \mathcal{F}_h$. Keep in mind that merely knowing the efficiency bounds is not enough. To be useful, these bounds must be attainable; i.e., we must be able to construct an estimator of $\boldsymbol{\beta}_0$ with asymptotic variance equal to the efficiency bound. To keep our presentation concise we have excluded the estimation part of the problem from this paper. Following Severini and Wong (1992), in Tripathi (1997) we construct such an estimator and also show that it attains the efficiency bounds.

Because "a nonparametric problem is at least as difficult as any of the parametric problems obtained by assuming we have enough knowledge of the unknown state of nature to restrict it to a finite dimensional set" (Stein, 1956, p. 187), efficiency bounds for $\boldsymbol{\beta}_0$ may be obtained by finding the worst parametric submodel that is contained in the full model. Following the details in Appendixes A and B, efficiency bounds for $\boldsymbol{\beta}_0$, when f^* is homogeneous of degree r, are given by $\sigma_0^2(\mathbb{E}\mathbf{S}\mathbf{S}')^{-1}$ where $\mathbf{S} = \mathbf{x} - \left[z_q^r \mathbb{E}\left(\mathbf{x} z_q^r | (z_1/z_q), \ldots, (z_{q-1}/z_q)\right) / \mathbb{E}\left(z_q^{2r} | (z_1/z_q), \ldots, (z_{q-1}/z_q)\right)\right]$. Note that apart from the projection-based ap-

https://doi.org/10.1017/S0266466600165053 Published online by Cambridge University Press

proach described in Appendix B, efficiency bounds for the homogeneity case can also be obtained by using the result of Chamberlain (1992).[2] As these bounds are obtained by an orthogonal projection onto a proper subspace of $L^2(\mathcal{S})$, not surprisingly,[3] there is an efficiency gain due to homogeneity of f^*. But what about the magnitude of this gain? It is easy to construct examples where large gains in efficiency may be obtained by using the identifying power of homogeneity. However, as the following example demonstrates, large gains in efficiency are possible under homogeneity, even when the parameter of interest is identified.

Example

Let $y = x\beta_0 + f^*(z) + \varepsilon$, where $x, z \stackrel{\mathrm{d}}{=} \mathrm{UIID}(1,2)$, $\varepsilon \stackrel{\mathrm{d}}{=} \mathrm{N}(0,1)$, and f^* is linearly homogeneous. Because f^* is homogeneous of degree one, use $q, r = 1$ in Lemma B.2 (in Appendix B) to see that the efficiency bound for estimating β_0 is 6.1. However, if homogeneity is not imposed upon f^* the bound increases to 12. Thus the loss in efficiency by not imposing homogeneity, when f^* is truly homogeneous, is 96.7%.

4. EFFICIENCY BOUNDS WHEN f^* IS CONCAVE

Following the details in Appendix C, efficiency bounds for estimating $\boldsymbol{\beta}_0$, when $f^* \in \mathcal{F}_c$, are given by $\sigma_0^2(\mathbb{E}[\mathbf{x} - \mathbb{E}(\mathbf{x}|\mathbf{z})][\mathbf{x}' - \mathbb{E}(\mathbf{x}'|\mathbf{z})])^{-1}$. But these are the same bounds that would be obtained if the only thing we knew about f^* was that it was twice continuously differentiable. See, e.g., Newey (1990, p. 109). Hence, in the class of $n^{1/2}$CR estimators of $\boldsymbol{\beta}_0$, computing efficiency bounds for $\boldsymbol{\beta}_0$ when f^* is concave is equivalent to computing efficiency bounds for $\boldsymbol{\beta}_0$ when f^* is just a $C^2(\mathcal{S})$ function. Therefore, at least asymptotically, concavity of f^* does not help us in estimating $\boldsymbol{\beta}_0$ more efficiently. Moreover, this conclusion does not change if "concavity" is replaced by "convexity." Note, however, that as these bounds are asymptotic in nature, this result should not be interpreted as a prescription for disregarding concavity of f^* when estimating $\boldsymbol{\beta}_0$. It may be quite worthwhile to impose concavity upon f^* to obtain better finite sample results for estimating $\boldsymbol{\beta}_0$. Furthermore, even asymptotically, this result leaves open the possibility that we could do better in the presence of concavity if we looked at nonregular estimators of $\boldsymbol{\beta}_0$. As far as I know, these problems are still open and merit further research.

5. EFFICIENCY BOUNDS WHEN f^* IS DECREASING

Next we look at the case when f^* is decreasing. As a result of the difficulty of working with monotonicity restrictions on f^* in a multivariate setting, we confine ourselves to the one-dimensional case, i.e., $q = 1$. So let $f^* \in \mathcal{F}_d$. Using

https://doi.org/10.1017/S0266466600165053 Published online by Cambridge University Press

the details provided in Appendix D, efficiency bounds for estimating $\boldsymbol{\beta}_0$, when $f^* \in \mathcal{F}_d$, are given by $\sigma_0^2(\mathbb{E}[\mathbf{x} - \mathbb{E}(\mathbf{x}|z)][\mathbf{x}' - \mathbb{E}(\mathbf{x}'|z)])^{-1}$. Therefore, as in Section 4, we conclude that in the class of all $n^{1/2}$CR estimators of $\boldsymbol{\beta}_0$ the knowledge that f^* is decreasing does not help us estimate $\boldsymbol{\beta}_0$ more efficiently, at least asymptotically. This result remains unchanged if the word *decreasing* is replaced by *increasing*. As before, it does not mean that one should disregard monotonicity of f^* while estimating $\boldsymbol{\beta}_0$. The possibility that monotonicity of f^* could help in estimating $\boldsymbol{\beta}_0$ more efficiently, even asymptotically, in some class of nonregular estimators remains to be investigated.

6. CONCLUSION

We have shown that in the class of $n^{1/2}$ consistent regular ($n^{1/2}$CR) estimators of $\boldsymbol{\beta}_0$, concavity and monotonicity of f^*, unlike homogeneity, do not help in estimating $\boldsymbol{\beta}_0$ more efficiently in large samples. Of course, these results are asymptotic in nature. One would expect concavity, and monotonicity of f^* to help in estimating $\boldsymbol{\beta}_0$ more efficiently in finite samples. To gain some intuition with regard to these asymptotic results one can visualize homogeneity as imposing equality restrictions on f^*, whereas concavity and monotonicity impose inequality restrictions on the second and first derivatives, respectively, of f^*. In the case of homogeneity the equality restrictions upon f^* bind on the tangent space (for details, see Appendix B), whereas in the case of concavity and monotonicity these inequality restrictions do not bind (see Appendixes C and D). Because lower bounds for the asymptotic variance of $n^{1/2}$CR estimators of $\boldsymbol{\beta}_0$ are obtained by an orthogonal projection onto the tangent space (see Appendix A), homogeneity of f^* reduces these bounds, whereas concavity and monotonicity of f^* do not. It is useful to note that regular estimators of $\boldsymbol{\beta}_0$ exclude those estimators that impose concavity, or monotonicity, upon f^*. The intuition behind this[4] is that such estimators would require inequality restrictions to be imposed upon estimators of the unknown functional form in the model and this would destroy their regularity. This is best illustrated in parametric models with inequality restrictions on the nuisance parameters, where an estimator of the parameter of interest is obtained by truncating the estimator of the nuisance parameter into its parameter space.[5] Although this procedure destroys regularity by introducing a discontinuity in the asymptotic distribution of the estimator sequence, it leads to improved estimators; i.e., under inequality restrictions one can construct, at least in parametric models, truncated, but nonregular, estimator sequences that improve the performance of any regular estimator of the parameter of interest. However, as van der Vaart (1989, p. 1492) points out, this seems to be an unresolved problem for semiparametric models. In light of these remarks, imposing regularity on estimators of $\boldsymbol{\beta}_0$, when f^* is concave or monotone, may be overly restrictive. On the other hand, it is not clear how the efficiency results of van der Vaart (1989) could be extended in the absence of a

https://doi.org/10.1017/S0266466600165053 Published online by Cambridge University Press

regularity requirement on estimator sequences. Therefore, we leave for future research the issue of whether, and how, concavity and monotonicity of f^* would help in asymptotic estimation of $\boldsymbol{\beta}_0$ within a class of estimators distinct from the usual "$n^{1/2}$ consistent regular" estimators.

7. NOTATION FOR APPENDIXES AND A USEFUL DEFINITION

Let $\nabla^2 f$ be the Hessian and $\|f\|_{\mathrm{C}^2} = \sup_{\mathbf{z}\in\mathcal{S}}|f(\mathbf{z})| + \sup_{\mathbf{z}\in\mathcal{S},1\le j\le q}|(\partial f(\mathbf{z})/\partial z_j)| + \sup_{\mathbf{z}\in\mathcal{S},1\le i,j\le q}|(\partial^2 f(\mathbf{z})/\partial z_i \partial z_j)|$ be the C^2 norm of f. When $q = 1$, the C^1 norm of f is given by $\|f\|_{\mathrm{C}^1} = \sup_{z\in\mathcal{S}}|f(z)| + \sup_{z\in\mathcal{S}}|f'(z)|$. The expressions $\mathcal{F}_h$ and $\mathcal{F}_c$ are closed (w.r.t. $\|\cdot\|_{\mathrm{C}^2}$) convex cones in $\mathrm{C}^2(\mathcal{S})$, whereas $\mathcal{F}_d$ is a closed (w.r.t. $\|\cdot\|_{\mathrm{C}^1}$) convex cone in $\mathrm{C}^1(\mathcal{S})$. If E is a subset of a linear space X, $\operatorname{lin} E$ denotes the linear subspace generated by E; i.e., $\operatorname{lin} E$ is the set of all finite linear combinations of elements of E. $\bar{E}$ is the L^2-closure of E.

DEFINITION 1. *Let M be a subset of a Banach space $(X, \|\cdot\|_X)$. A vector $\dot{x} \in X$ is said to be tangent to M at the point $x_0 \in M$ if there exist a $t_0 > 0$ and a mapping $t \mapsto r(t)$ of the interval $(0, t_0)$ into X such that $x_0 + t\dot{x} + r(t) \in M$ for all $t \in (0, t_0)$, and $\|r(t)\|_X/t \to 0$ as $t \downarrow 0$. The set of vectors tangent to M at the point $x_0 \in M$ is denoted by $T(M, x_0)$ and is a closed (in the topology generated by $\|\cdot\|_X$) nonempty cone in X. By a "cone" we mean that if $\dot{x} \in T(M, x_0)$, then $\lambda\dot{x} \in T(M, x_0)$ for all $\lambda > 0$. The expression $T(M, x_0)$ (resp. $\overline{\operatorname{lin} T(M, x_0)}$) is referred to as the tangent cone (resp. tangent space) to M at x_0. For more on this see Krabs* (1979).

NOTES

1. A function $f: \mathcal{S} \to \mathbb{R}$ is said to be homogeneous of degree $r \in \mathbb{R}$ on $\mathcal{S}$, iff $f(\lambda\mathbf{z}) = \lambda^r f(\mathbf{z})$ for all $\lambda > 0$ such that $\lambda\mathbf{z} \in \mathcal{S}$.

2. We thank a referee for pointing this out.

3. Because efficiency bounds for $\boldsymbol{\beta}_0$, when homogeneity is not imposed on f^*, are obtained by an orthogonal projection onto the space $\mathrm{L}^2(\mathcal{S})$ itself.

4. We thank one of the referees for this.

5. See, e.g., Example 2.5.1 in Tripathi (1997), which imposes monotonicity in a simple linear regression model.

6. Because we can always find a "one-dimensional subproblem" that has the same asymptotic variance as any "multidimensional subproblem" (Stein, 1956, p. 188), it suffices to look at "one-dimensional subproblems."

7. To see this let $\{g_n\}$ be a sequence in $\mathcal{G}_h$ such that $g_n \xrightarrow{\mathrm{L}^2} g$. Because $\mathrm{L}^2(\mathcal{S})$ is complete, $g \in \mathrm{L}^2(\mathcal{S})$, and it only remains to show that g is homogeneous of degree r. Because by Chebychev's inequality $g_n \xrightarrow{p} g$, there exists a subsequence $\{g_{n,j}\}$ such that $g_{n,j} \xrightarrow{a.s.} g$. But as each $g_{n,j}$ is homogeneous, g is also homogeneous of degree r.

8. In the proof of Lemma C.1 we show that a slightly stronger result can be obtained for the one-dimensional case; i.e., when $q = 1$, we show $\mathcal{W} = T(\mathcal{F}_c, f^*)$.

https://doi.org/10.1017/S0266466600165053 Published online by Cambridge University Press

REFERENCES

Chamberlain, G. (1986) Asymptotic efficiency in semiparametric models with censoring. *Journal of Econometrics* 32, 189–218.
Chamberlain, G. (1992) Efficiency bounds for semiparametric regression. *Econometrica* 60, 567–596.
Cosslett, S. (1987) Efficiency bounds for distribution free estimators of binary choice and censored regression models. *Econometrica* 55, 559–585.
Krabs, W. (1979) *Optimization and Approximation*. New York: Wiley.
Matzkin, R. (1994) Restrictions of economic theory in nonparametric methods. In R. Engle & D. McFadden (eds.), *Handbook of Econometrics*, vol. IV, pp. 2524–2558. Amsterdam: Elsevier Science B.V.
Newey, W.K. (1988) Efficient estimation of Tobit models under symmetry. In W.A. Barnett, J. Powell, & G. Tauchen (eds.), *Nonparametric and Semiparametric Methods in Econometrics and Statistics. Proceedings of the Fifth International Symposium in Economic Theory and Econometrics*, pp. 291–336. Cambridge: Cambridge University Press.
Newey, W.K. (1990) Semiparametric efficiency bounds. *Journal of Applied Econometrics* 5, 99–135.
Rilstone, P. (1993) Calculating the (local) semiparametric efficiency bounds for the generated regressors problem. *Journal of Econometrics* 56, 357–370.
Severini, T.A. & W.H. Wong (1992) Profile likelihood and conditionally parametric models. *Annals of Statistics* 20 (4), 1768–1802.
Stein, C. (1956) Efficient nonparametric testing and estimation. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, pp. 187–195. Berkeley: University of California Press.
Tripathi, G. (1997) *Semiparametric Efficiency Bounds and Testing in Models with Shape Restrictions*. Ph.D. Thesis, Department of Economics, Northwestern University.
van der Vaart, A. (1989) On the asymptotic information bound. *Annals of Statistics* 17 (4), 1487–1500.

APPENDIX A: CALCULATING EFFICIENCY BOUNDS—A QUICK REVIEW

The basic idea behind computing efficiency bounds in semiparametric models is to reduce a "hard" problem (e.g., one containing infinite dimensional nuisance parameters), into a sequence of easier to solve "one-dimensional subproblems."[6] Using these one-dimensional subproblems, and the results of van der Vaart (1989, Theorem 2.5, p. 1491), we can show that when $f^* \in \mathcal{F}$ (where $\mathcal{F}$ is now a generic symbol for $\mathcal{F}_h$, $\mathcal{F}_c$, or $\mathcal{F}_d$) the lower bounds for the asymptotic variance of $n^{1/2}$CR estimators of $\boldsymbol{\beta}_0$ are determined by projecting the scores w.r.t. $\boldsymbol{\beta}_0$ onto the tangent space $\overline{\text{lin}\, T(\mathcal{F}, f^*)}$. So let $\boldsymbol{\delta}^* = \text{proj}(\mathbf{x}|\overline{\text{lin}\, T(\mathcal{F}, f^*)})$ denote the vector of orthogonal projections of the elements of the random vector $\mathbf{x}$ onto the Hilbert space $\overline{\text{lin}\, T(\mathcal{F}, f^*)}$. For the PLM with Gaussian errors some straightforward calculations (see, e.g., Tripathi, 1997) show that the aforementioned lower bounds are given by $\sigma_0^2(\mathbb{E}\mathbf{S}\mathbf{S}')^{-1}$, where $\mathbf{S} = \mathbf{x} - \boldsymbol{\delta}^*$ and the nonsingularity of $\mathbb{E}(\mathbf{S}\mathbf{S}')$ follows from Assumption (A.5). Therefore, to calculate these bounds, all we need to do is to determine $\boldsymbol{\delta}^*$.

https://doi.org/10.1017/S0266466600165053 Published online by Cambridge University Press

APPENDIX B: BOUNDS UNDER HOMOGENEITY—TECHNICAL DETAILS

For the homogeneity case we make an additional assumption about $\mathcal{S}$, the support of $\mathbf{z}$. Besides being used in the proof of Lemma B.1, the following assumption also ensures that the components of $\boldsymbol{\delta}^*$ obtained in Lemma B.2 are elements of $\mathrm{L}^2(\mathcal{S})$.

Assumption B.1. $\mathcal{S}$ is such that for any $\mathbf{z} \in \mathcal{S}$, z_q (the last component of $\mathbf{z}$) is bounded away from zero; i.e., $\exists m > 0$ such that $|z_q| \geq m$.

As outlined in Appendix A, efficiency bounds for the homogeneity case are easily obtained once we know $\boldsymbol{\delta}^*$. To do so, we first find $\overline{\operatorname{lin} T(\mathcal{F}_h, f^*)}$ and then calculate $\boldsymbol{\delta}^* = \operatorname{proj}(\mathbf{x}|\overline{\operatorname{lin} T(\mathcal{F}, f^*)})$. Now determination of $T(\mathcal{F}_h, f^*)$ is easy: Because $\mathcal{F}_h$ is a linear subspace of $\mathrm{C}^2(\mathcal{S})$, it is straightforward to see that $T(\mathcal{F}_h, f^*) = \mathcal{F}_h$. Next comes $\overline{\operatorname{lin} T(\mathcal{F}_h, f^*)}$.

LEMMA B.1. *Let $\mathcal{G}_h$ denote the set of all $\mathrm{L}^2(\mathcal{S})$ functions that are also homogeneous of degree r. Then $\overline{\operatorname{lin} T(\mathcal{F}_h, f^*)} = \overline{\operatorname{lin} \mathcal{F}_h} = \mathcal{G}_h$.*

Proof. As $\mathcal{F}_h$ is a linear space, $\operatorname{lin} \mathcal{F}_h = \mathcal{F}_h$. Hence we show that $\bar{\mathcal{F}}_h = \mathcal{G}_h$. Clearly $\bar{\mathcal{F}}_h \subseteq \mathcal{G}_h$, because $\mathcal{F}_h \subseteq \mathcal{G}_h$ and $\mathcal{G}_h$ is closed in the $\mathrm{L}^2(\mathcal{S})$ norm.[7] To show $\mathcal{G}_h \subseteq \bar{\mathcal{F}}_h$, pick any $g \in \mathcal{G}_h$. As g is homogeneous of degree r, write $g(z_1, \ldots, z_q) = z_q^r g\big((z_1/z_q), \ldots, (z_{q-1}/z_q), 1\big)$. Because z_q lies in a compact set, and is bounded away from zero by assumption, we have $\bar{m} \leq |z_q|^r \leq \bar{M}$ for some $(\bar{m}, \bar{M}) > 0$. Therefore, $g(\mathbf{w}, 1) \in \mathrm{L}^2(\mathcal{S})$, where $\mathbf{w} = (w_1, \ldots, w_{q-1})$ and $w_i = z_i/z_q$. For any $\epsilon > 0$, use Lemma A.2 of Chamberlain (1986, p. 215) to find a smooth real valued function f, which has compact support in $\mathbb{R}^{q-1}$, such that $\mathbb{E}[g(\mathbf{w}, 1) - f(\mathbf{w})]^2 < \epsilon$. Hence, using the homogeneity of g and that $|z_q|^r \leq \bar{M}$, we get $\mathbb{E}\big[g(\mathbf{z}) - z_q^r f\big((z_1/z_q), \ldots, (z_{q-1}/z_q)\big)\big]^2 < \bar{M}^2\epsilon$. But as $z_q^r f\big((z_1/z_q), \ldots, (z_{q-1}/z_q)\big)$ is homogeneous of degree r on $\mathcal{S}$, we have a function in $\mathcal{F}_h$ that is arbitrarily close to g in $\mathrm{L}^2(\mathcal{S})$ norm; i.e., $g \in \bar{\mathcal{F}}_h$. ■

So let $\boldsymbol{\delta}^* = (\delta_1^*(\mathbf{z}), \ldots, \delta_p^*(\mathbf{z}))$, where each $\delta_i(\mathbf{z})$ is given by the next result.

LEMMA B.2.

$$\delta_i^*(\mathbf{z}) = \operatorname{proj}(x_i|\mathcal{G}_h) = \frac{z_q^r \mathbb{E}\left(x_i z_q^r \middle| \frac{z_1}{z_q}, \ldots, \frac{z_{q-1}}{z_q}\right)}{\mathbb{E}\left(z_q^{2r} \middle| \frac{z_1}{z_q}, \ldots, \frac{z_{q-1}}{z_q}\right)}.$$

Proof. Clearly $\delta_i^*(\mathbf{z})$ is homogeneous of degree r. Furthermore, $\delta_i^* \in \mathrm{L}^2(\mathcal{S})$ because $\bar{m} \leq |z_q|^r \leq \bar{M}$ for some $(\bar{m}, \bar{M}) > 0$, and $\mathbb{E}\|\mathbf{x}\|^2 < \infty$. Therefore, $\delta_i^* \in \mathcal{G}_h$, and the result follows by using iterated expectations to verify the orthogonality condition of the classical projection theorem. ■

Once $\boldsymbol{\delta}^*$ is obtained, semiparametric efficiency bounds for estimating $\boldsymbol{\beta}_0$ under homogeneity are given by $\sigma_0^2(\mathbb{E}\mathbf{S}\mathbf{S}')^{-1}$, where $\mathbf{S} = \mathbf{x} - \boldsymbol{\delta}^*$.

https://doi.org/10.1017/S0266466600165053 Published online by Cambridge University Press

APPENDIX C: BOUNDS UNDER CONCAVITY—TECHNICAL DETAILS

As before, we begin by finding $T(\mathcal{F}_c, f^*)$ and $\overline{\text{lin}\, T(\mathcal{F}_c, f^*)}$. Notice that because $f^* \in \mathcal{F}_c$, the Hessian $\nabla^2 f^*$ is negative semidefinite. Therefore, two cases arise: first, when $\nabla^2 f^*$ is negative definite it is easy to see that $f^* \in \text{int}(\mathcal{F}_c)$, which implies that $T(\mathcal{F}_c, f^*) = \mathrm{C}^2(\mathcal{S})$. So the more difficult case is when $\nabla^2 f^*$ is negative semidefinite but not negative definite, i.e., when $f^* \in bdry(\mathcal{F}_c)$. To analyze this case, define the nonempty set $\mathcal{S}_0 \subseteq \mathcal{S}$ on which $\nabla^2 f^*$ is negative semidefinite but not negative definite. Also, let $\mathcal{W} = \{f \in \mathrm{C}^2(\mathcal{S}) : f \text{ is concave on } \mathcal{S}_0\}$. Clearly $\mathcal{F}_c \subseteq \mathcal{W}$, and it is easy to see that $\mathcal{W}$ is a closed (in $\|\cdot\|_{\mathrm{C}^2}$) convex cone. We now use Lemmas C.1 and C.2 to characterize the tangent cone and tangent space, respectively.[8]

LEMMA C.1. *Let $f^* \in bdry(\mathcal{F}_c)$. Then $\mathcal{W} \subseteq T(\mathcal{F}_c, f^*) \subseteq \mathrm{C}^2(\mathcal{S})$.*

LEMMA C.2. *Let $f^* \in \mathcal{F}_c$. Then $\overline{\text{lin}\, T(\mathcal{F}_c, f^*)} = \mathrm{L}^2(\mathcal{S})$.*

It is now immediate that $\boldsymbol{\delta}^* = \text{proj}(\mathbf{x} | \overline{\text{lin}\, T(\mathcal{F}_c, f^*)}) = \mathbb{E}(\mathbf{x}|\mathbf{z})$. Therefore, when f^* is concave, $\mathbf{S} = \mathbf{x} - \mathbb{E}(\mathbf{x}|\mathbf{z})$, and efficiency bounds for estimating $\boldsymbol{\beta}_0$ are given by $\sigma_0^2(\mathbb{E}\mathbf{S}\mathbf{S}')^{-1}$.

Additional notation. Let $\mathcal{U} = \{\boldsymbol{\alpha} \in \mathbb{R}^q : \|\boldsymbol{\alpha}\| = 1\}$ be the unit sphere in $\mathbb{R}^q$. Note that $f \in \mathrm{C}^2(\mathcal{S})$ is concave iff $\boldsymbol{\alpha}'\nabla^2 f(\mathbf{z})\boldsymbol{\alpha} \leq 0$ for all $(\boldsymbol{\alpha}, \mathbf{z}) \in \mathcal{U} \times \mathcal{S}$. $\mathcal{S}_0$, the subset of $\mathcal{S}$ on which $\nabla^2 f^*$ is negative semidefinite but not negative definite can be written as $\mathcal{S}_0 = \{\mathbf{z} \in \mathcal{S} : \boldsymbol{\alpha}'\nabla^2 f^*(\mathbf{z})\boldsymbol{\alpha} = 0 \text{ for some } \boldsymbol{\alpha} \in \mathcal{U}\}$. Using the compactness of $\mathcal{U}$ and continuity of $(\boldsymbol{\alpha}, \mathbf{z}) \mapsto \boldsymbol{\alpha}'\nabla^2 f^*(\mathbf{z})\boldsymbol{\alpha}$, it is easy to see that $\mathcal{S}_0$ is a compact subset of $\mathcal{S}$. Furthermore, let $\mathcal{W}_s = \{f \in \mathrm{C}^2(\mathcal{S}) : \nabla^2 f \text{ is negative definite on } \mathcal{S}_0\}$, $B(\mathbf{z}, \xi)$ an open ball of radius ξ around $\mathbf{z}$, $V_\xi = \cup_{\mathbf{z} \in \mathcal{S}_0} B(\mathbf{z}, \xi)$ an open set containing $\mathcal{S}_0$, and $V_\xi^c = \mathbb{R}^q \backslash V_\xi$.

Proof of Lemma C.1. To show $\mathcal{W} \subseteq T(\mathcal{F}_c, f^*)$ it suffices to prove that $\mathcal{W}_s \subseteq T(\mathcal{F}_c, f^*)$. This follows because $\mathcal{W}_s$ is dense (in C^2 norm) in $\mathcal{W}$ (see Lemma C.4, which follows) and $T(\mathcal{F}_c, f^*)$ is closed in the C^2 norm (Krabs, 1979, p. 154). So pick any $g \in \mathcal{W}_s$. From Lemma C.3 (which follows) we can find a $\lambda_0 > 0$ such that $\nabla^2 g$ is negative definite on $V_{\lambda_0} \cap \mathcal{S}$. Now write $\mathcal{S} = (\mathcal{S} \cap V_{\lambda_0}) \cup (\mathcal{S} \cap V_{\lambda_0}^c)$ and for any $\mathbf{z} \in \mathcal{S}$ define $\zeta_t(\mathbf{z}) = f^*(\mathbf{z}) + tg(\mathbf{z})$. If we can show that $\zeta_t \in \mathcal{F}_c$ then we are done, because this implies that $g \in T(\mathcal{F}_c, f^*)$. So, for any $(\boldsymbol{\alpha}, \mathbf{z}) \in \mathcal{U} \times \mathcal{S}$, write $\boldsymbol{\alpha}'\nabla^2\zeta_t(\mathbf{z})\boldsymbol{\alpha} = \boldsymbol{\alpha}'\nabla^2 f^*(\mathbf{z})\boldsymbol{\alpha} + t\boldsymbol{\alpha}'\nabla^2 g(\mathbf{z})\boldsymbol{\alpha}$. Because $(\mathcal{S} \cap V_{\lambda_0})$ and $(\mathcal{S} \cap V_{\lambda_0}^c)$ are disjoint, first suppose that $\mathbf{z} \in \mathcal{S} \cap V_{\lambda_0}$. In this case $\boldsymbol{\alpha}'\nabla^2 f^*(\mathbf{z})\boldsymbol{\alpha} \leq 0$ because f^* is concave on $\mathcal{S}$, whereas $\boldsymbol{\alpha}'\nabla^2 g(\mathbf{z})\boldsymbol{\alpha} < 0$ because $\nabla^2 g$ is negative definite on $\mathcal{S} \cap V_{\lambda_0}$. Therefore, $\boldsymbol{\alpha}'\nabla^2\zeta_t(\mathbf{z})\boldsymbol{\alpha} < 0$ (i.e., $\zeta_t \in \mathcal{F}_c$) for any $t > 0$. Next suppose that $\mathbf{z} \in \mathcal{S} \cap V_{\lambda_0}^c$, where $\mathcal{S} \cap V_{\lambda_0}^c$ is a compact set. Now $\boldsymbol{\alpha}'\nabla^2 f^*(\mathbf{z})\boldsymbol{\alpha} < 0$ because $\nabla^2 f^*$ is negative definite on $\mathcal{S} \backslash \mathcal{S}_0$, whereas the sign of $\boldsymbol{\alpha}'\nabla^2 g(\mathbf{z})\boldsymbol{\alpha}$ remains indeterminate outside $\mathcal{S}_0$. However, because $\boldsymbol{\alpha}'\nabla^2 f^*(\mathbf{z})\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}'\nabla^2 g(\mathbf{z})\boldsymbol{\alpha}$ are continuous over the compact set $\mathcal{U} \times (\mathcal{S} \cap V_{\lambda_0}^c)$, we can choose a t small enough such that $\boldsymbol{\alpha}'\nabla^2\zeta_t(\mathbf{z})\boldsymbol{\alpha} = \boldsymbol{\alpha}'\nabla^2 f^*(\mathbf{z})\boldsymbol{\alpha} + t\boldsymbol{\alpha}'\nabla^2 g(\mathbf{z})\boldsymbol{\alpha} < 0$; i.e., $\zeta_t \in \mathcal{F}_c$. Combining the two cases we get that $\zeta_t \in \mathcal{F}_c$ for small enough t.

https://doi.org/10.1017/S0266466600165053 Published online by Cambridge University Press

To show the reverse direction, pick any $\dot{f} \in T(\mathcal{F}_c, f^*)$. Then $\dot{f} \in \mathrm{C}^2(\mathcal{S})$ because $\mathcal{F}_c \subset \mathrm{C}^2(\mathcal{S})$, and so $T(\mathcal{F}_c, f^*) \subset \mathrm{C}^2(\mathcal{S})$ always holds. However, we can strengthen this result slightly for the one-dimensional case $q = 1$. To see this, use Definition 1 to conclude that if $\dot{f} \in T(\mathcal{F}_c, f^*)$, then $\exists t_0 > 0$, and a map $t \mapsto r_t$ of $(0, t_0)$ into $\mathrm{C}^2(\mathcal{S})$, such that $f_t(\mathbf{z}) = f^*(\mathbf{z}) + t\dot{f}(\mathbf{z}) + r_t(\mathbf{z}) \in \mathcal{F}_c$ for all $t \in (0, t_0)$, and $r_t/t \xrightarrow{\mathrm{C}^2} 0$ as $t \downarrow 0$; i.e., $\lambda_t(\mathbf{z}) = \big(f_t(\mathbf{z}) - f^*(\mathbf{z})\big)/t \xrightarrow{\mathrm{C}^2} \dot{f}(\mathbf{z})$, as $t \downarrow 0$. When $q = 1$, $d^2 f^*(z)/dz^2 = 0$ on $\mathcal{S}_0$, implying that $d^2\lambda_t(z)/dz^2 \leq 0$ for all $z \in \mathcal{S}_0$ and $t \in (0, t_0)$. Therefore, when $q = 1$, λ_t is a convergent sequence in $\mathcal{W}$. But as $\mathcal{W}$ is closed in the C^2 norm, $\lim_{t\to 0} \lambda_t = \dot{f} \in \mathcal{W}$. Thus for the one-dimensional case we can actually show that $T(\mathcal{F}_c, f^*) = \mathcal{W}$. However, when $q > 1$ it is easy to see that $\nabla^2\lambda_t(\mathbf{z})$ may or may not be negative semidefinite on $\mathcal{S}_0$; i.e., for the case $q > 1$, we cannot guarantee that λ_t is a convergent sequence in $\mathcal{W}$, implying that $\dot{f}$ may not lie in $\mathcal{W}$. ■

Proof of Lemma C.2. When $\nabla^2 f^*$ is negative definite, the result follows directly from the fact that $\mathrm{C}^2(\mathcal{S})$ is dense in $\mathrm{L}^2(\mathcal{S})$ (in the $\mathrm{L}^2(\mathcal{S})$ norm). For a proof of this fact modify the argument in Chamberlain (1986, Lemma A.2, p. 215). So let us look at the case when $\nabla^2 f^*$ is negative semidefinite but not negative definite. Because $\mathcal{W} \subseteq T(\mathcal{F}_c, f^*)$ implies that $\operatorname{lin}\mathcal{W} \subseteq \operatorname{lin} T(\mathcal{F}_c, f^*)$, when $\nabla^2 f^*$ is negative semidefinite but not negative definite, it suffices to show that $\mathrm{L}^2(\mathcal{S}) \subseteq \overline{\operatorname{lin}\mathcal{W}}$. So pick any $f \in \mathrm{L}^2(\mathcal{S})$. Then, for any $\epsilon > 0$, there exists a function $g \in \mathrm{C}^2(\mathcal{S})$ such that $\mathbb{E}[\, f(\mathbf{z}) - g(\mathbf{z})]^2 < \epsilon$. If we can now show that $g \in \operatorname{lin}\mathcal{W}$, we are done. For $\nu > 0$, define $h_\nu(\mathbf{z}) = g(\mathbf{z}) - (2\nu)^{-1}\sum_{i=1}^q z_i^2$, and for any $\boldsymbol{\alpha} \in \mathcal{U}$, let us look at $\boldsymbol{\alpha}'\nabla^2 h_\nu(\mathbf{z})\boldsymbol{\alpha} = \boldsymbol{\alpha}'\nabla^2 g(\mathbf{z})\boldsymbol{\alpha} - \nu^{-1}$. Now $M = \sup_{\boldsymbol{\alpha}\in\mathcal{U}, \mathbf{z}\in\mathcal{S}} |\boldsymbol{\alpha}'\nabla^2 g(\mathbf{z})\boldsymbol{\alpha}|$ is finite, because $(\boldsymbol{\alpha}, \mathbf{z}) \mapsto \boldsymbol{\alpha}'\nabla^2 g(\mathbf{z})\boldsymbol{\alpha}$ is continuous and $\mathcal{U} \times \mathcal{S}$ is compact. If $M = 0$, the Hessian of g vanishes, and thus h_ν is strictly concave for all $\nu > 0$. So suppose that $M \neq 0$ and pick $\nu \leq (2M)^{-1}$. Then $\boldsymbol{\alpha}'\nabla^2 h_\nu(\mathbf{z})\boldsymbol{\alpha} \leq \boldsymbol{\alpha}'\nabla^2 g(\mathbf{z})\boldsymbol{\alpha} - 2M \leq M - 2M < 0$. Hence, whether $M = 0$ or not, h_ν is strictly concave on $\mathcal{S}$ for small enough ν. But we know that strictly concave functions on $\mathcal{S}$ are elements of $\mathcal{W}$. Therefore, $g \in \operatorname{lin}\mathcal{W}$ because it is the difference of two strictly concave functions on $\mathcal{S}$. ■

LEMMA C.3. *For any $g \in \mathcal{W}_s$, there exists a $\lambda_0 > 0$ such that $\nabla^2 g$ is negative definite on $V_{\lambda_0} \cap \mathcal{S}$.*

Proof. Pick any $\boldsymbol{\alpha} \in \mathcal{U}$ and look at the function $(\boldsymbol{\alpha}, \mathbf{z}) \mapsto \boldsymbol{\alpha}'\nabla^2 g(\mathbf{z})\boldsymbol{\alpha}$ on $\mathcal{U} \times \mathcal{S}$. Because $\mathcal{U} \times \mathcal{S}$ is compact, continuity of $(\boldsymbol{\alpha}, \mathbf{z}) \mapsto \boldsymbol{\alpha}'\nabla^2 g(\mathbf{z})\boldsymbol{\alpha}$ implies that it is uniformly continuous on $\mathcal{U} \times \mathcal{S}$ and hence on $\mathcal{U} \times (V_{\bar{\xi}} \cap \mathcal{S})$. Let $-\widetilde{M} = \sup_{\boldsymbol{\alpha}\in\mathcal{U}, \mathbf{z}\in\mathcal{S}_0} \boldsymbol{\alpha}'\nabla^2 g(\mathbf{z})\boldsymbol{\alpha}$, where $\widetilde{M} > 0$. As $\nabla^2 g$ is negative definite on $\mathcal{S}_0$ by assumption, and $\boldsymbol{\alpha}'\nabla^2 g(\mathbf{z})\boldsymbol{\alpha}$ is continuous over the compact set $\mathcal{U} \times \mathcal{S}_0$, we know that $-\widetilde{M}$ exists and is strictly less than zero. Now, for any $\mathbf{z}, \mathbf{z}_1 \in \mathcal{S}$ and $\boldsymbol{\alpha} \in \mathcal{U}$, uniform continuity of $(\boldsymbol{\alpha}, \mathbf{z}) \mapsto \boldsymbol{\alpha}'\nabla^2 g(\mathbf{z})\boldsymbol{\alpha}$ implies that there exists a $\lambda_0 > 0$ (independent of the $\mathbf{z}$'s and $\boldsymbol{\alpha}$) such that $|\boldsymbol{\alpha}'\nabla^2 g(\mathbf{z})\boldsymbol{\alpha} - \boldsymbol{\alpha}'\nabla^2 g(\mathbf{z}_1)\boldsymbol{\alpha}| < \widetilde{M}/2$, whenever $|\mathbf{z} - \mathbf{z}_1| < \lambda_0$. So let $\mathbf{z}_1 \in V_{\lambda_0} \cap \mathcal{S}$. The definition of V_{λ_0} implies that $\mathbf{z}_1 \in B(\mathbf{z}_2, \lambda_0) \cap \mathcal{S}$ for some $\mathbf{z}_2 \in \mathcal{S}_0$, and thus by uniform continuity we have $|\boldsymbol{\alpha}'\nabla^2 g(\mathbf{z}_2)\boldsymbol{\alpha} - \boldsymbol{\alpha}'\nabla^2 g(\mathbf{z}_1)\boldsymbol{\alpha}| < \widetilde{M}/2$; i.e., $\boldsymbol{\alpha}'\nabla^2 g(\mathbf{z}_1)\boldsymbol{\alpha} < \boldsymbol{\alpha}'\nabla^2 g(\mathbf{z}_2)\boldsymbol{\alpha} + (\widetilde{M}/2) \leq \sup_{\boldsymbol{\alpha}\in\mathcal{U}, \mathbf{z}\in\mathcal{S}_0} \boldsymbol{\alpha}'\nabla^2 g(\mathbf{z})\boldsymbol{\alpha} + (\widetilde{M}/2) = -(\widetilde{M}/2) < 0$. Because $\mathbf{z}_1 \in V_{\lambda_0} \cap \mathcal{S}$ and $\boldsymbol{\alpha} \in \mathcal{U}$ were arbitrary, this means that $\nabla^2 g$ is negative definite on $V_{\lambda_0} \cap \mathcal{S}$. ■

LEMMA C.4. *$\mathcal{W}_s$ is dense (in C^2 norm) in $\mathcal{W}$.*

https://doi.org/10.1017/S0266466600165053 Published online by Cambridge University Press

Proof. Pick any $h \in \mathcal{W}$, $(\boldsymbol{\alpha}, \mathbf{z}) \in \mathcal{U} \times \mathcal{S}_0$, and for some $M > 0$ define $g_M(\mathbf{z}) = h(\mathbf{z}) - (2M)^{-1} \sum_{i=1}^{q} z_i^2$. Then $\boldsymbol{\alpha}' \nabla^2 g_M(\mathbf{z}) \boldsymbol{\alpha} = \boldsymbol{\alpha}' \nabla^2 h(\mathbf{z}) \boldsymbol{\alpha} - M^{-1} < 0$, because h is concave on $\mathcal{S}_0$. Furthermore, $\|g_M - h\|_{\mathrm{C}^2} = (2M)^{-1} \|\sum_{i=1}^{q} z_i^2\|_{\mathrm{C}^2} \to 0$ as $M \uparrow \infty$; i.e., $g_M \in \mathcal{W}_s$ and for large M is arbitrarily close to h. ■

APPENDIX D: BOUNDS UNDER MONOTONICITY—TECHNICAL DETAILS

As in Appendix C, we have two cases to consider. First, if the first derivative $f^{*\prime}(z) < 0$ then $f^* \in \mathrm{int}(\mathcal{F}_d)$, and this implies that $T(\mathcal{F}_d, f^*) = \mathrm{C}^1(\mathcal{S})$. The second case arises when $f^{*\prime}$ vanishes for at least one point in $\mathcal{S}$, i.e., when $f^* \in bdry(\mathcal{F}_d)$. So suppose that there exists a nonempty set $\mathcal{S}_0 \subseteq \mathcal{S}$ on which $f^{*\prime}$ is zero and let $\mathcal{M} = \{f \in \mathrm{C}^1(\mathcal{S}) : f \text{ is decreasing on } \mathcal{S}_0\}$. It is straightforward to see that $\mathcal{F}_d \subseteq \mathcal{M}$ and that $\mathcal{M}$ is a closed (in $\|\cdot\|_{\mathrm{C}^1}$) convex cone. Therefore, as in the proof of Lemma C.1 for the case $q = 1$, we can show that $T(\mathcal{F}_d, f^*) = \mathcal{M}$ when $f^* \in bdry(\mathcal{F}_d)$. Furthermore, following the proof of Lemma C.2, it is also straightforward to see that $\overline{\mathrm{lin}}\, T(\mathcal{F}_d, f^*) = \mathrm{L}^2(\mathcal{S})$. Therefore, when f^* is decreasing, $\boldsymbol{\delta}^* = \mathrm{proj}(\mathbf{x}|\mathrm{L}^2(\mathcal{S})) = \mathbb{E}(\mathbf{x}|z)$, and efficiency bounds for $\boldsymbol{\beta}_0$ are given by $\sigma_0^2(\mathbb{E}\mathbf{S}\mathbf{S}')^{-1}$, where $\mathbf{S} = \mathbf{x} - \boldsymbol{\delta}^*$.

https://doi.org/10.1017/S0266466600165053 Published online by Cambridge University Press