Statistica Sinica **10**(2000), 1231-1241

DENOISED LEAST SQUARES ESTIMATORS: AN APPLICATION TO ESTIMATING ADVERTISING EFFECTIVENESS

Zongwu Cai*, Prasad A. Naik† and Chih-Ling Tsai†

** University of North Carolina and † University of California*

Abstract: It is known in marketing science that an advertiser under- or overspends millions of dollars on advertising because the estimation of advertising effectiveness is biased. This bias is induced by measurement noise in advertising variables, such as awareness and television rating points, which are provided by commercial market research firms based on small-sample surveys of consumers. In this paper, we propose a denoised regression approach to deal with the problem of noisy variables. We show that denoised least squares estimators are consistent. Simulation results indicate that the denoised regression approach outperforms the classical regression approach. A marketing example is presented to illustrate the use of denoised least squares estimators.

Key words and phrases: Advertising, linear smoother, noisy data, threshold, wavelets.

1. Introduction

When analyzing real data, ordinary least squares (OLS) estimators are most often used for simple regression models. It is known that these estimators have nice properties (*e.g., consistency*) if the relationship between dependent and independent variables is linear and the classical error assumptions (*i.e.*, normality, independence and homoscedasticity) are satisfied. In practice, however, both dependent and independent variables may be a function of another variable (*e.g.,* time) and both variables may contain measurement noise. In this case, OLS estimators are not consistent (see Carroll, Ruppert and Stefanski (1995), and Cheng and Van Ness (1999, pp. 3, 11)). A natural approach to overcome this difficulty is first to denoise the data and then to fit a simple regression model to the denoised data. We call this a denoised regression model, and the resulting estimators denoised least squares (DLS) estimators.

There are several methods available in the literature which can be used to screen out noise, but we use WaveShrink, developed recently by Donoho and Johnstone (1994, 1995a and 1995b). It has been found particularly useful for high frequency series and for sharp and short aberrant series. These appear often in applied fields such as marketing (Blattberg and Neslin (1990)), medicine and biology (Aldroubi and Unser (1996)), and image processing (Prasad and

Lyengar (1997)). The essential of wavelets theory for statistical applications and data analysis can be found in Hubbard (1996) and Ogden (1997).

In the area of marketing, removing noise from data is important because advertisers or advertising agencies face serious monetary issues. Typically an advertiser spends hundreds of millions of dollars to generate awareness for his/her products and services. The advertiser tracks awareness generated by an advertising campaign, and television rating points bought by his advertising agency, to determine the effectiveness of advertising. Figures 1(a) and 1(b) show scatterplots (denoted by the symbol $\circ$) of Awareness and Television Rating Points for Cadbury's Dairy Milk chocolate brand in the U. K. over a period of 128 weeks. These measures are based on a small-sample survey of consumers' responses and contain measurement noise.

Measurement noise induces positive or negative bias in the estimate of advertising effectiveness. If advertising effectiveness is under-estimated (negative bias), then the advertiser spends less on advertising, losing valuable sales opportunities, and the advertising agency loses potential revenue from commissions on buying media time and space. If advertising effectiveness is over-estimated (positive bias), then an advertiser overspends on advertising, making less profit, and the advertising agency runs the risk of adversely affecting the valuable client-agency relationship. Thus, regardless of under- or over- estimation of advertising effectiveness, advertisers are likely to make expensive sub-optimal marketing decisions if the unreliability in advertising data is not controlled (see Naik and Tsai (2000)).

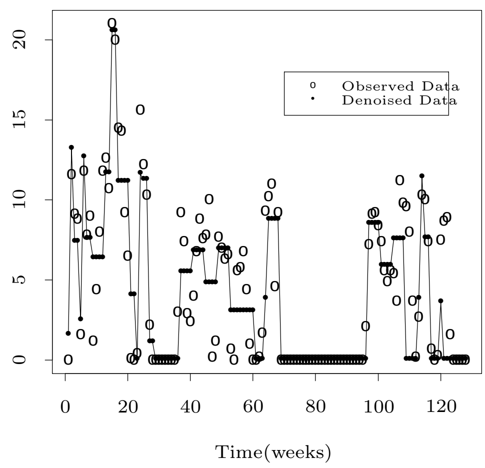

Figure 1(a). The observed and denoised advertising schedules for Cadbury's Dairy Milk chocolate brand over 128 weeks.

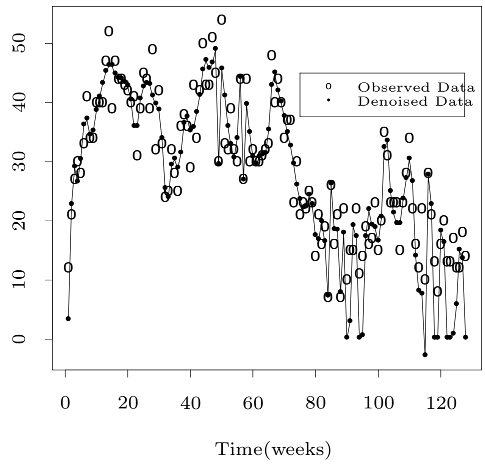

Figure 1(b). The observed and denoised awareness responses over 128 weeks.

The purpose of this paper is to provide a solution to the problem of controlling measurement noise in advertising data such as those shown in Figure 1. A natural approach is first to use WaveShrink to denoise both variables, and then to fit them with a linear regression model.

We propose the DLS estimator and study its consistency as well as apply it to the area of marketing science. Section 2 shows that DLS estimators are consistent. Section 3 presents a Monte Carlo study to compare the performance of the OLS and DLS parameter estimators by using WaveShrink to denoise both variables. The results show that DLS outperforms OLS. We illustrate this methodology by analyzing the aforementioned commercial advertising data. Section 4 gives concluding remarks. Finally, the Appendix presents the technical proof of our result.

2. DLS Estimates

2.1. Model structure and DLS estimates

Suppose that $\{(\xi_i, \eta_i) : 1 \le i \le n\}$ are unobservable "true" variables satisfying a linear relationship

$$\eta_i = \alpha + \beta\, \xi_i, \tag{2.1}$$

where α and β are unknown constants, $\xi_i = \xi(t_i)$, $\eta_i = \eta(t_i)$, and $\{t_i : 1 \le i \le n\}$ are design points such that $0 \le t_1 \le \cdots \le t_n \le 1$. However, what we are able to observe are $\{(x_i, y_i) : 1 \le i \le n\}$, the true variables plus additive measurement errors $\{(\delta_i, \varepsilon_i) : 1 \le i \le n\}$, such that

$$x_i = \xi(t_i) + \delta_i \qquad \text{and} \qquad y_i = \eta(t_i) + \varepsilon_i. \tag{2.2}$$

Here we assume that δ_i and ε_i are mutually independent and identically distributed (i.i.d.) random errors with mean zero and variances of σ_δ^2 and σ_ε^2, respectively. It is worth noting that (2.1) and (2.2) are different from the model structure of errors-in-variables since both ξ and η are functions of the third variable t_i.

For given observed data $x = (x_1, \ldots, x_n)'$ and $y = (y_1, \ldots, y_n)'$, we first apply the denoising operator H_x and H_y to obtain the denoised data:

$$x^* = H_x \circ x \qquad \text{and} \qquad y^* = H_y \circ y, \tag{2.3}$$

where $x^* = (x_1^*, \ldots, x_n^*)'$, $y^* = (y_1^*, \ldots, y_n^*)'$, and H_x and H_y are $n \times n$ smoothing matrices (either linear, *e.g.* smoothing spline, kernel or locally weighted, or nonlinear, *e.g.* wavelet). We then fit the denoised data via a linear regression model

$$y_i^* = \alpha^* + \beta^* x_i^* + \varepsilon_i^*, \tag{2.4}$$

where α^* and β^* are unknown parameters and $\varepsilon^* = (\varepsilon_1^*, \ldots, \varepsilon_n^*)'$ is distributed with mean zero and variance $\sigma_{\varepsilon^*}^2 \Sigma$. Σ may be a function of H_x and H_y and it usually does not have a closed-form expression. Hence we focus on studying the properties of least squares estimators of α^* and β^*.

Clearly (2.4) gives the denoised regression model and the resulting DLS estimators of α^* and β^* are, respectively, $\hat{\alpha}^* = \overline{y}^* - \hat{\beta}^* \, \overline{x}^*$ and $\hat{\beta}^* = S_{x^*y^*}/S_{x^*}^2$, where $\overline{x}^* = \sum_{i=1}^n x_i^*/n$, $\overline{y}^* = \sum_{i=1}^n y_i^*/n$, $S_{x^*y^*} = \sum_{i=1}^n (x_i^* - \overline{x}^*)(y_i^* - \overline{y}^*)/n$, and $S_{x^*}^2 = \sum_{i=1}^n (x_i^* - \overline{x}^*)^2/n$.

2.2. Consistency

Here we study consistency of DLS estimators. Let $S_{\xi\eta} = \sum_{i=1}^n (\xi_i - \overline{\xi})(\eta_i - \overline{\eta})/n$, $S_\xi^2 = \sum_{i=1}^n (\xi_i - \overline{\xi})^2/n$, and $S_\eta^2 = \sum_{i=1}^n (\eta_i - \overline{\eta})^2/n$, and let $||.||_{2,n}^2$ denote the usual squared l_n^2 norm. Then we have $S_{\xi\eta} = \beta \, S_\xi^2$ and $S_\eta^2 = \beta^2 \, S_\xi^2$ since the true model is $\eta_i = \alpha + \beta \, \xi_i$. We need the following assumptions.

Assumption 1. S_ξ^2 converges to σ_ξ^2 for some $0 < \sigma_\xi^2 < \infty$ as $n \to \infty$.

Assumption 2. $||x^* - \xi||_{2,n} = o_p(n^{1/2})$ and $||y^* - \eta||_{2,n} = o_p(n^{1/2})$.

Note that Assumption 2 is satisfied by well-known nonparametric estimators (see Remarks 2 and 4 below for detail).

Theorem 1. *Under Assumptions* 1 *and* 2*,* $\hat{\alpha}^*$ *and* $\hat{\beta}^*$ *are consistent estimators of* α *and* β*, respectively.*

Remark 1. Consider the fixed design regression model with $\xi_i = \xi(t_i)$, where $\{t_i\}$ are equi-spaced design points on interval $[a, b]$. Assumption 1 is satisfied if $0 < \int_a^b (\xi(t) - \int_a^b \xi(t)dt)^2 < \infty$.

Remark 2. In wavelet regression models, Assumption 2 is fulfilled by equation (7) and Corollary 1 in Donoho and Johnstone (1994).

Remark 3. Under some regularity conditions, the optimal mean square error (MSE) of the local polynomial regression estimator of $\xi(t_i^*)$ (or $\eta(t_i^*)$) is of order $O(n^{-4/5})$ for $p = 0$ or 1, and $O(n^{-8/9})$ for $p = 2$ or 3, where p is the order of local polynomial function and $\{t_i^*\}$ are given real numbers in the interval $[a,\, b]$. Hence, Assumption 2 is satisfied. Detailed information about local polynomial regression can be found in Fan and Gijbels (1996, Chapter 3).

Remark 4. Under some regularity conditions, the MSE of the cubic smoothing spline estimator of $\xi(t_i^*)$ (or $\eta(t_i^*)$) is of order $O(n^{-8/9})$. Hence, Assumption 2 is satisfied. Detailed information about the cubic smoothing spline can be found in Eubank (1988, Chapter 5).

2.3. Generalization to multiple regression models

Our approach can be generalized to multiple regression models in which either all or some independent variables are measured with errors. Without loss of generality, we assume that there is only one extra variable, z_i, added into model (2.1). We first study the case in which the z_i, $i = 1, \ldots, n$, are not functions of the third variable. Hence, the true model is $\eta_i = \alpha + \beta\,\xi_i + \gamma\, z_i$ and the fitted model is $y_i^* = \alpha^* + \beta^*\, x_i^* + \gamma^*\, z_i + e_i^*$, where y_i^* and x_i^* are defined in equation (2.3) and $e^* = (e_1^*, \ldots, e_n^*)'$ is distributed with mean zero and variance $\sigma^2_{e^*}\Sigma^*$. Let $\phi^* = (\alpha^*, \beta^*, \gamma^*)'$, $\phi = (\alpha, \beta, \gamma)'$, and $X^* = (\mathbf{1},\ x^*,\ z)$, where $z = (z_1, \ldots, z_n)'$. Then the least squares estimate of ϕ^* is $\hat{\phi}^* = (X^{*\prime}X^*)^{-1}X^{*\prime}\, y^*$. To obtain the asymptotic property of ϕ^*, we need one additional assumption.

Assumption 3. $\sum_{i=1}^n z_i^2/n$ is bounded.

Under Assumptions 1, 2 and 3, $\hat{\phi}^*$ is a consistent estimate of ϕ. Detailed proof can be obtained from the first author upon request.

Next we consider the case when z_i is a function of the third variable t_i. Specifically, $z_i = \tau_i + \nu_i$, $\tau_i = \tau(t_i)$ and $\{\nu_i\}$ are i.i.d. random variables with mean zero and variance σ^2_ν. In addition, the ν_i are independent of $\{(\delta_i,\, \varepsilon_i)\}$. The true model is $\eta_i = \alpha + \beta\,\xi_i + \gamma\,\tau_i$, and the fitted model is $y_i^* = \alpha^* + \beta^*\, x_i^* + \gamma^*\, z_i^* + \widetilde{e_i}^*$, where z_i^* is the ith component of the $n \times 1$ vector $z^* = H_z z$, H_z is the $n \times n$ smoothing matrix, and $\widetilde{e}^* = (\widetilde{e_1}^*, \ldots, \widetilde{e_n}^*)'$ is distributed with mean zero and variance $\sigma^2_{\widetilde{e}^*}\hat{\Sigma}^*$. Let ϕ and ϕ^* be defined as above and $\widetilde{X}^* = (\mathbf{1},\ x^*,\ z^*)$. Then the least squares estimator of ϕ^* is $\widetilde{\phi}^* = (\widetilde{X}^{*\prime}\widetilde{X}^*)^{-1}\widetilde{X}^{*\prime}\, y^*$. To obtain the asymptotic property, we need two additional assumptions.

Assumption 4. $S_\tau^2 = \sum_{i=1}^n(\tau_i - \overline{\tau})^2/n$ converges to σ_τ^2 for some $0 < \sigma_\tau^2 < \infty$ as $n \to \infty$.

Assumption 5. $||z^* - \tau||_{2,n} = o_p(n^{1/2})$.

Under Assumptions 1, 2, 4 and 5, $\widetilde{\phi}^*$ is a consistent estimate of ϕ. Detailed proof for this case can also be obtained from the first author upon request.

3. Numerical Properties

In this section, we illustrate the proposed method via two examples. In the first example, we use simulated data to evaluate the performance of DLS estimators. We then apply the method to the advertising data mentioned in the introduction.

3.1. A simulated example

Here we present the results of a Monte Carlo study that examines the performance of the OLS and DLS estimators for a simple regression model in two situations: the signal-to-noise ratios for independent and dependent variables are the same or are different. The purpose of using different signal-to-noise ratios (snr) is to assess the effect of snr_x against snr_y on parameter estimation. We generated the true model as follows: $\eta(t_i) = \alpha + \beta\,\xi(t_i)$, with $\alpha = \beta = 1$ and $t_i = i/n$, where $\xi(\cdot)$ is the mean function given in Donoho and Johnstone (1994) (Blocks, Bumps, HeaviSine and Doppler). The signal-to-noise ratios are denoted by $\text{snr}_x = SD(\xi)/\sigma_\delta$ and $\text{snr}_y = SD(\eta)/\sigma_\varepsilon$. Since all the Donoho and Johnstone functions give similar results in our simulation studies, only the Doppler case is presented here.

For given snr_x and snr_y, with $n = 1024$, $\{x_i\}$ and $\{y_i\}$ are generated from equations (2.1) and (2.2), where $\{\delta_i\}$ are i.i.d. from $N(0,\, \sigma_\delta^2)$, $\{\varepsilon_i\}$ are i.i.d. from $N(0,\, \sigma_\varepsilon^2)$ and they are independent. For the first case when $\text{snr}_x = \text{snr}_y$, 1000 realizations are generated and snr_x ranges from 0.1 to 8 in increments of 0.1. For the second case, we draw 1000 realizations by assuming that $\text{snr}_y = 10 - \text{snr}_x$ and snr_x ranges from 0.1 to 8 in increments of 0.1. Both variables x and y are denoised via wavelet computations using the GAUSS-TSM package with the Daubechies Least Asymmetric (8) wavelet. Although we use the default thresholds set by the universal method with shrinkage applied to only the 6 finest levels, other choices of wavelet thresholding can be considered (see Nason (1996) and Hurvich and Tsai (1998)).

Figure 2(a) presents estimated values of the two parameters when $\text{snr}_x = \text{snr}_y$. In this case $\hat{\beta}^*_{DLS}$ outperforms $\hat{\beta}_{OLS}$. As snr_x gets larger, both $\hat{\beta}_{OLS}$ and $\hat{\beta}^*_{DLS}$ more closely resemble the true parameter, $\beta = 1$. In order to further explore the impact of snr_x versus snr_y in parameter estimation, we next consider the case $\text{snr}_y = 10 - \text{snr}_x$. Figure 2(b) depicts values for the two parameter estimates as snr_x increases. It shows clearly that $\hat{\beta}^*_{DLS}$ outperforms $\hat{\beta}_{OLS}$. In summary, the signal-to-noise ratio of the independent variable plays a more important role than the signal-to-noise ratio of the dependent variable for estimating β. This finding is not surprising, since large variations in a dependent variable have only a small impact on the contribution of a strong signal of the independent variable to the regression line. In conclusion, both simulation studies show that we should use the denoised least squares approach to estimate unknown parameters.

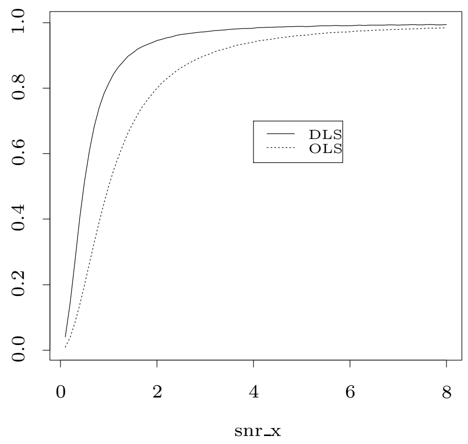

Figure 2(a). Parameter estimates of β, $\hat{\beta}^*_{\mathrm{DLS}}$ (solid line) and $\hat{\beta}_{\mathrm{OLS}}$ (dotted line), when $\mathrm{snr}_x = \mathrm{snr}_y$.

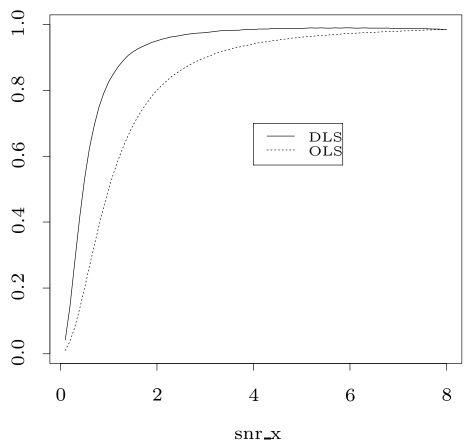

Figure 2(b). Parameter estimates of β, $\hat{\beta}^*_{\mathrm{DLS}}$ (solid line) and $\hat{\beta}_{\mathrm{OLS}}$ (dotted line), when $\mathrm{snr}_y = 10 - \mathrm{snr}_x$, and snr_x ranges from 0.1 to 8.

3.2. A real example

We now continue the marketing example mentioned in the introduction to illustrate our proposed method. These data were collected by Millward Brown International for an advertising tracking study of Cadbury's Dairy Milk chocolate brand, one of the major chocolate brands in England. Detailed descriptions of these data can be found in Brown (1986), Migon and Harrison (1985) and Colman and Brown (1983). In marketing science, it is of interest to study the linear relationship between the percentage of awareness of the advertisement (y) and the Television Rating Points (x) (see Naik, Mantrala and Sawyer (1998)). Here, both x and y are functions of time and they were measured over 128 weeks from August 1977 to January 1980.

In practice, Television Rating Points (TRP) is an estimate of the frequency of exposure to an advertisement and awareness index is the percentage of people who have seen that advertisement based on a small sample survey of consumers. As Winer (1993, p.4) notes, "... the advertising measurement system, while the best available, is fallible...". Hence, we denoise both x and y variables before studying their relationship. This rationale is also supported by the low estimated signal-to-noise ratios, $\hat{snr}_x = 1.78$ and $\hat{snr}_y = 2.60$.

In marketing, because TRP follows the on-off pattern, known as "advertising pulsing" (see Mahajan and Muller (1986)), it resembles Haar filter. Hence, we use Haar wavelet to denoise the TRP data. Also, awareness grows as TRP increases and declines exponentially when TRP is zero. This pattern of growth and decline in awareness data matches the shape of the Daubechies filter. Therefore, we use Daubechies wavelet to denoise the awareness data. Figure 1(a) displays TRP and the corresponding denoised data (x^*, denoted by the symbol $*$). By the same token, Figure 1(b) gives awareness index for the advertisement and the corresponding denoised data (y^*, indicated by the symbol $*$).

After fitting the observed data $\{(x_i, y_i)\}$ and the denoised data $\{(x_i^*, y_i^*)\}$ by the simple regression model, the resulting slope parameter estimates are $\hat{\beta}_{OLS} = 0.927$ and $\hat{\beta}^*_{DLS} = 1.383$, respectively. Hence, the estimated advertising effectiveness (the slope parameter) shows that the OLS estimate is 33% smaller than the DLS estimate. This large downward bias of OLS is consistent with our simulation findings. Such attenuation of the advertising effectiveness leads advertisers to believe that their advertising campaign is not quite effective, which leads to wrong marketing decisions (*e.g.,* spend less on advertising, change the advertisement copy, or fire the advertising agency). Denoising data and then fitting them by a regression model can improve the estimation of advertising effectiveness and lead to proper marketing decisions.

4. Conclusion

We have proposed a denoised least squares approach to estimate unknown regression parameters in linear regression models. Both simulation results and empirical example show that DLS estimates outperform OLS estimates. In these Monte Carlo studies, we only focus on the equispaced fixed design and then apply WaveShrink method to filter out the noise from Donoho and Johnstone's four well-known mean functions. To further investigate the performance of DLS and OLS, we conducted simulation studies by applying the local polynomial approach with the *Epanechnikov* kernel function to smooth out noise contained in the mean function of the quantile of the standard normal. The results show that DLS outperforms OLS, as found in Section 3.

Acknowledgement

We are grateful to the Editor and to two referees whose insightful comments have improved the paper. Cai's work was supported, in part, by the National Science Foundation grant DMS 0072400 and funds provided by the University of North Carolina at Charlotte.

Appendix. The Proof of Theorem 1

Let $\xi = (\xi_1, \ldots, \xi_n)'$, $\eta = (\eta_1, \ldots, \eta_n)'$, $\overline{\xi} = \sum_{i=1}^n \xi_i/n$, and $\overline{\eta} = \sum_{i=1}^n \eta_i/n$. It is easily seen that

$$n\, S_{x^*}^2 = \sum_{i=1}^n (x_i^* - \xi_i)^2 - n(\overline{\xi} - \overline{x}^*)^2 + 2\sum_{i=1}^n (x_i^* - \xi_i)(\xi_i - \overline{\xi}) + \sum_{i=1}^n (\xi_i - \overline{\xi})^2, \quad \text{(A.1)}$$

$$n\, S_{x^*y^*} = \sum_{i=1}^n (x_i^* - \xi_i)(y_i^* - \eta_i) + \sum_{i=1}^n (x_i^* - \xi_i)(\eta_i - \overline{\eta}) + \sum_{i=1}^n (\xi_i - \overline{\xi})(y_i^* - \eta_i) + n(\overline{x}^* - \overline{\xi})(\overline{\eta} - \overline{y}^*) + \sum_{i=1}^n (\xi_i - \overline{\xi})(\eta_i - \overline{\eta}). \quad \text{(A.2)}$$

By applying the Cauchy-Schwartz inequality to (A.1) and (A.2), we have

$$\left| S_{x^*}^2 - S_\xi^2 \right| \le 2||x^* - \xi||_{2,n}^2/n + 2||x^* - \xi||_{2,n}\, S_\xi/\sqrt{n}, \quad \text{(A.3)}$$

$$\left| S_{x^*y^*} - S_{\xi\eta} \right| \le 2\, ||x^* - \xi||_{2,n}||y^* - \eta||_{2,n}/n + ||x^* - \xi||_{2,n} S_\eta/\sqrt{n} + ||y^* - \eta||_{2,n} S_\xi/\sqrt{n}. \quad \text{(A.4)}$$

Under Assumption 1, (A.3) and (A.4) become

$$|S_{x^*}^2 - S_\xi^2| = O\left(\frac{||x^* - \xi||_{2,n}}{n^{1/2}}\right) \text{ and } |S_{x^*y^*} - S_{\xi\eta}| = O\left(\frac{||x^* - \xi||_{2,n}}{n^{1/2}} + \frac{||y^* - \eta||_{2,n}}{n^{1/2}}\right).$$

Furthermore, Assumption 2 implies that $|S^2_{x^*} - S^2_\xi| = o_p(1)$ and $|S_{x^*y^*} - S_{\xi\eta}| = o_p(1)$. Therefore $|S^2_{x^*} - \sigma^2_\xi| = o_p(1)$ and $|S_{x^*y^*} - \beta\ \sigma^2_\xi| = o_p(1)$. Hence, by Slutsky's theorem, we have $\hat{\beta}^* \to \beta$ in probability and $\hat{\alpha}^* \to \alpha$ in probability. This completes the proof.

References

Aldroubi, A. and Unser, M. (1996). *Wavelets in Medicine and Biology.* CRC Press, Florida.

Blattberg, R. C. and Neslin, S. A. (1990). *Sales Promotion–Concepts, Methods, and Strategies.* Prentice Hall, New Jersey.

Brown, G. (1986). Modelling advertising awareness. *J. Roy. Statist. Soc. Ser. D* **35**, 289-299.

Carroll, R. J., Ruppert, D. and Stefanski, L. A. (1995). *Measurement Error in Nonlinear Models.* Chapman and Hall, London.

Cheng, C. L. and Van Ness, J. W. (1999). *Statistical Regression with Measurement Error.* Arnold, London.

Chung, K. L. (1974). *A Course in Probability Theory.* 2nd edition. Academic Press, New York.

Colman, S. and Brown, G. (1983). Advertising tracking studies and sales effects. *J. Market Research Soc.* **25**, 165-183.

Donoho, D. L. and Johnstone, I. M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika* **81**, 425-455.

Donoho, D. L. and Johnstone, I. M. (1995a). Adapting to unknown smoothness via wavelet shrinkage. *J. Amer. Statist. Assoc.* **90**, 1200-1224.

Donoho, D. L. and Johnstone, I. M. (1995b). Wavelet shrinkage: Asymptopia? *J. Roy. Statist. Soc. Ser. B* **57**, 301-369.

Eubank, R. L. (1988). *Spline Smoothing and Nonparametric Regression.* Marcel Dekker, New York.

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modeling and Its Applications.* Chapman and Hall, London.

Hubbard, B. B. (1996). *The World According to Wavelets.* A. K. Peters, M. A.

Hurvich, C. M. and Tsai, C. L. (1998). A crossvalidatory AIC for hard wavelet thresholding in spatially adaptive function estimation. *Biometrika* **85**, 701-710.

Mahajan, V. and Muller, E. (1986). Advertising pulsing policies for generating awareness for new products (with discussions). *J. Marketing Sci.* **5**, 89-111.

Migon, H. S. and Harrison, P. J. (1985). An application of non-linear Bayesian forecasting to television advertising. In *Bayesian Statistics 2* (Edited by J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith), 681-696. Valencia University Press, Amsterdam, North-Holland.

Naik, P. A. and Tsai, C. L. (2000). Controlling measurement errors in models of advertising competition. *J. Marketing Research* **37**, 113-124.

Naik, P. A., Mantrala, M. K. and Sawyer, A. G. (1998). Planning media schedules in the presence of dynamic advertising quality. *J. Marketing Sci.* **17**, 214-235.

Nason, G. P. (1996). Wavelet shrinkage using cross-validation. *J. Roy. Statist. Soc. Ser. B* **58**, 463-479.

Ogden, R. T. (1997). *Essential Wavelets for Statistical Applications and Data Analysis.* Birkhäuser, Boston.

Prasad, L. and Lyengar, S. S. (1997). *Wavelet Analysis with Application to Image Processing.* CRC Press, Florida.

Winer, R. S. (1993). Using single-source scanner data as a natural experiment for evaluating advertising effects. *J. Marketing Sci.* **2**, 15-31.

Department of Mathematics, University of North Carolina, Charlotte, NC 28223, U.S.A.
E-mail: zcai@uncc.edu

Graduate School of Management, University of California, Davis, CA 95616-8609, U.S.A.
E-mail: panaik@ucdavis.edu

Graduate School of Management, University of California, Davis, CA 95616-8609, U.S.A.
E-mail: cltsai@ucdavis.edu

(Received March 1999; accepted March 2000)