DBJ Discussion Paper Series, No.1209

Nonparametric Multiplicative Bias Correction for Kernel-Type Density Estimation Using Positive Data

Masayuki Hirukawa
(Setsunan University)
And
Mari Sakudo
(Development Bank of Japan)

December　2012

Nonparametric Multiplicative Bias Correction for Kernel-Type Density Estimation Using Positive Data

Masayuki Hirukawa*
Setsunan University

Mari Sakudo†
Development Bank of Japan

March 2012

Abstract

This paper improves convergence properties of asymmetric kernel density estimators for nonnegative economic and financial variables via two classes of multiplicative bias correction ("MBC") techniques. It is demonstrated that under sufficient smoothness of the true density, each MBC technique reduces the order of magnitude in bias, whereas the order of magnitude in variance remains unchanged. Accordingly, the mean integrated squared error of each MBC estimator achieves a faster convergence rate of $O\left(n^{-8/9}\right)$ when best implemented, where n is the sample size. Furthermore, the estimator always generates a nonnegative density estimate by construction. Plug-in smoothing parameter choice methods are also proposed to implement the MBC estimators using the Gamma and Modified Gamma kernels. Finite sample performance of the estimators are examined via Monte Carlo simulations, and the estimators are applied to estimating income distributions.

Keywords: Asymmetric kernel; boundary effect; higher-order bias kernel; income distribution.
JEL Classification Codes: C13; C14.
MSC 2000 Codes: 62G7; 62G20.

*Faculty of Economics, Setsunan University, 17-8 Ikeda Nakamachi, Neyagawa, Osaka, 572-8508, Japan; phone: (+81)72-839-8095; fax: (+81)72-839-8138; e-mail: hirukawa@econ.setsunan.ac.jp; web: http://www.setsunan.ac.jp/~hirukawa/.

†Research Institute of Capital Formation, Development Bank of Japan, 9-3, Otemachi 1-chome, Chiyoda-ku, Tokyo 100-0004, Japan; phone (+81)3-3244-1241; fax (+81)3-3270-7084; e-mail: mari-sakudo@gmail.com.

1 Introduction

The aims of this paper are to improve convergence properties of kernel density estimators for nonnegative economic and financial variables via a couple of bias correction techniques, and to apply the proposed estimators for estimating income distributions. Although we focus primarily on income distributions, our proposal is expected to fit well with other distributions such as the loss distribution (= the distribution of a payment to the insured) in actuarial science and financial risk management, the distribution of important financial variables such as short-term interest rates,[1] and even the baseline hazard in financial duration analysis.

To illustrate income distribution estimation, let $\{X_i\}_{i=1}^n$ be a random sample drawn from a univariate distribution with density f. Income distributions are empirically characterized by two stylized facts, namely, (i) a natural boundary at the origin (i.e. f has support on $[0, \infty)$) and (ii) a mode near the boundary and a long tail with sparse data. To capture these stylized facts, researchers often fit log-normal and Pareto distributions for the region near the boundary and the right tail, respectively. However, the imposition of a misspecified model leads to inconsistent estimates and misleading inference, as well as to disputable evaluations of inequality measures.

This motivates us to adopt nonparametric kernel methods for their flexibility in curve fitting. Nevertheless, standard kernel smoothing with symmetric kernels must be modified to accommodate the stylized facts. First, when the support of the density has a boundary, boundary correction methods (e.g. Müller, 1991; Jones, 1993; Jones and Foster, 1996; Zhang, Karunamuni and Jones, 1999; Hall and Park, 2002) should

[1] CIR (Cox, Ingersoll and Ross, 1985) and Ahn and Gao's (1999) inverse Feller processes are often used as data generating processes of short-term interest rates. It is known that time-invariant distributions of these processes are gamma and inverse gamma distributions, respectively. These distributions possess the same stylized facts as described shortly.

be employed. Second, global smoothing with a single bandwidth may not work well. If a short bandwidth is used to capture the mode near the origin, the density estimate over the tail region tends to be wiggly. On the other hand, if a long bandwidth is chosen to preserve the shape of the tail part, the mode near the origin is considerably smoothed away. Variable bandwidth methods (e.g. Abramson, 1982; Terrell and Scott, 1992) have been proposed as a remedy for this issue.

Recently, asymmetric kernel functions have emerged as a viable alternative to boundary correction methods.[2] For an asymmetric kernel function $K_{j(x,b)}(\cdot)$ indexed by j that depends on a design point $x > 0$ and a smoothing parameter $b > 0$, the density estimator can be expressed as

$$\hat{f}_{j,b}(x) = \frac{1}{n} \sum_{i=1}^{n} K_{j(x,b)}(X_i). \tag{1}$$

Throughout, $K_{j(x,b)}(\cdot)$ refers to the Gamma ("G"; Chen, 2000), Modified Gamma ("MG"; Chen, 2000), Inverse Gaussian ("IG"; Scaillet, 2004), Reciprocal Inverse Gaussian ("RIG"; Scaillet, 2004), Log-Normal ("LN"; Jin and Kawczak, 2003),[3] and Birnbaum-Saunders ("BS"; Jin and Kawczak, 2003) kernels. Functional forms of these kernels are presented in Table 1.

TABLE 1 AND FIGURE 1 ABOUT HERE

Asymmetric kernels can be viewed as a combination of a boundary correction device and a 'variable bandwidth' method. Because these kernels have support

[2]Strictly speaking, asymmetric kernel functions should be referred to as kernel-type weighting functions. In a slightly different context, Gouriéroux and Monfort (2006) and Jones and Henderson (2007) argue that unlike the cases of symmetric kernels, the roles of the data point X and the design point x in asymmetric kernels are not exchangeable, which leads to lack of normalization in density estimation using these kernels. Nevertheless, we follow the adopted convention in the literature for these kernel-type functions.

[3]Our definition of the Log-Normal kernel slightly differs from the original one in Jin and Kawczak (2003). This definition ensures that the leading variance term of the density estimator (1) becomes $n^{-1}b^{-1/2}f(x)/(2\sqrt{\pi}x)$ for a design point $x > 0$ so that $x/b \to \infty$ as $b \to 0$.

on $[0,\infty)$, they eliminate boundary effects by construction. Besides, the kernels have a number of important advantages in the analysis of economic and financial data. First, shapes of the asymmetric kernels vary according to the position at which smoothing is made; in other words, the amount of smoothing changes in a locally adaptive manner. Figure 1 plots shapes of six asymmetric kernels for four different design points (x = 0.5, 1.0, 2.0, and 4.0), where the smoothing parameter value is fixed at $b = 0.4$. The adaptive smoothing property in asymmetric kernels bears some similarities to the variable bandwidth methods. However, a single smoothing parameter suffices for local adaptive smoothing, which makes asymmetric kernels much more appealing in empirical work. Second, asymmetric kernels achieve the optimal rate of convergence (in mean integrated squared error ("MISE") sense) within the class of nonnegative kernel estimators. Third, unlike the case with symmetric kernels, the variances of asymmetric kernel estimators tend to decrease as the design point moves away from the boundary. Note that the variance reduction for large x is gained at the expense of increasing bias.

Provided that f is twice continuously differentiable, for a design point $x > 0$ so that $x/b \to \infty$ as $b \to 0$, the leading bias and variance terms of the asymmetric kernel density estimator (1) can be approximated by $Bias\left\{\hat{f}_{j,b}(x)\right\} \sim a_{1,j}(x,f)\,b$ and $Var\left\{\hat{f}_{j,b}(x)\right\} \sim n^{-1}b^{-1/2}v_j(x)\,f(x)$, where $a_{1,j}(x,f)$ is a kernel-specific function that depends on x and derivatives of f (see Table 2 below for explicit forms), and

$$v_j(x) = \frac{1}{2\sqrt{\pi}x^{r_j}} \text{ with } r_j = \begin{cases} 1/2 & \text{for } j = G, MG, RIG \\ 1 & \text{for } j = LN, BS \\ 3/2 & \text{for } j = IG \end{cases}. \tag{2}$$

Observe that $v_j(x)$ is proportional to x^{-r_j}, which is the source of shrinking variance with the position of x. This property is equivalent to the strategy of using longer bandwidths over the tail region where the data are sparse, and it is particularly advantageous for estimating income distributions.

Although asymmetric kernels are relatively new in the literature, several papers report favorable evidence from applying them to empirical models in economics and finance. A non-exhaustive list includes: (i) estimation of recovery rate distributions on defaulted bonds (Renault and Scaillet, 2004), (ii) income distribution estimation (Bouezmarni and Scaillet, 2005; Hagmann and Scaillet, 2007); (iii) actuarial loss distribution estimation (Hagmann and Scaillet, 2007; Gustafsson *et al.*, 2009); (iv) hazard estimation (Bouezmarni and Rombouts, 2008); (v) regression discontinuity design (Fé, 2010); (vi) realized integrated volatility estimation (Kristensen, 2010); and (vii) estimation of diffusion models (Gospodinov and Hirukawa, 2012).

This paper demonstrates that the convergence rate of the asymmetric kernel density estimator (1) can be improved via two classes of well-known bias correction techniques. To be more specific, each technique leads the bias convergence to be accelerated to $O(b^2)$ under sufficient differentiability of f, while the order of magnitude in variance is maintained, i.e. the variance is still $O\left\{\left(nb^{1/2}\right)^{-1}\right\}$ if $x/b \to \infty$, and $O\left\{\left(nb^{r_j+1/2}\right)^{-1}\right\}$ if $x/b \to \kappa$ for some $\kappa > 0$ and the kernel-specific exponent r_j given in (2). In the cases of symmetric second-order kernels, this kind of rate improvements can be typically achieved by employing higher-order kernels.[4] To the best of our knowledge, equivalent techniques are yet to be proposed for asymmetric kernels; if any, they would generate negative density estimates over some parts of the support, which cause difficulty in interpretation in practice. Instead, the rate improvement is attained by applying two classes of multiplicative bias correction ("MBC") techniques that are proposed by Terrell and Scott (1980) and Jones, Linton and Nielsen

[4] Jones and Foster (1993) provide an excellent review on the methods of generating higher-order kernels from a given symmetric second-order kernel.

(1995) originally for symmetric kernel density estimation.[5] For each of two MBC estimators, the MISE takes the form of $O\left(b^4 + n^{-1}b^{-1/2}\right)$. Therefore, when best implemented, each estimator can achieve the convergence rate of $O\left(n^{-8/9}\right)$ in MISE, which is faster than $O\left(n^{-4/5}\right)$, the MISE-optimal convergence rate within the class of nonnegative kernel estimators. Besides, two MBC estimators still maintain the same attractive properties as the bias-uncorrected estimator (1) has. They are free of boundary bias and always generate nonnegative density estimates everywhere by construction. Their variances still tend to decrease as the design point moves away from the boundary. Moreover, to implement the two classes of MBC estimators employing the Gamma and Modified Gamma kernels, this paper proposes plug-in methods of choosing the smoothing parameter b with a gamma density used as a reference.

In the closely related literature, Hagmann and Scaillet (2007) apply a semi-parametric MBC technique in the spirit of Hjort and Jones (1996) (called local multiplicative bias correction ("LMBC")) to asymmetric kernel density estimation. Gustafsson *et al.* (2009) propose another semi-parametric MBC technique called local transformation bias correction ("LTBC"), which basically follows the idea of Rudemo (1991). In both LMBC and LTBC, asymmetric kernels are employed at the bias correction step after initial parametric density estimation. A key difference is that the bias correction is made for the original data in LMBC and for the data transformed on the unit interval in LTBC. However, unlike the MBC estimators proposed in this paper, neither LMBC nor LTBC improves the bias in order of magnitude. Furthermore, Chen (1999) proposes another asymmetric kernel (the Beta kernel), which yields a boundary-bias-free density estimator on the unit interval. Two MBC

[5] The forms of bias reductions achieved through MBC techniques are analogous between symmetric and asymmetric kernel cases. However, none of asymmetric kernels can be expressed in the form of $K(\cdot/b)/b$, and thus it is worth emphasizing that mathematics and proof strategies for yielding the results are totally different.

techniques considered in this paper are shown to be applied for this kernel; see a companion paper (Hirukawa, 2010) for details.

The remainder of this paper is organized as follows. Section 2 introduces two classes of MBC density estimators and develops their asymptotic properties. Section 3 proposes plug-in methods of choosing the smoothing parameter b for Gamma kernel-based MBC estimators, and conducts Monte Carlo simulations to check finite sample properties of the estimators. Section 4 applies the MBC estimators for estimating income distributions from the U.S. and Brazilian data sets. Section 5 summarizes the main results of the paper. All proofs are given in the Appendix.

This paper adopts the following notational conventions. $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} \exp(-y)\, dy$ $(\alpha > 0)$ denotes the gamma function. The expression '$X \overset{d}{=} Y$' reads "A random variable X obeys the distribution Y." Lastly, the expression '$X_n \sim Y_n$' is used whenever $X_n/Y_n \to 1$ as $n \to \infty$.

2 MBC Estimators Using Asymmetric Kernels

2.1 Definitions of Two MBC Estimators

Two MBC techniques considered in this paper are subclasses of higher-order bias kernel density estimation methods (Jones and Signorini, 1997), and these are originally designed for symmetric kernels. We now extend the techniques to density estimation using asymmetric kernels with support on $[0, \infty)$.

In the spirit of Terrell and Scott (1980, abbreviated as "TS" hereafter), the first class of MBC techniques constructs a multiplicative combination of two density estimators employing the same kernel but different smoothing parameters. Let $\hat{f}_{j,b/c}(x)$ be the density estimator using asymmetric kernel j and smoothing parameter b/c, where $c \in (0, 1)$ is some predetermined constant that does not depend on the design point x. Then, the TS-MBC asymmetric kernel density estimator can be defined as

$$\tilde{f}_{TS,j}(x) = \left\{\hat{f}_{j,b}(x)\right\}^{\frac{1}{1-c}} \left\{\hat{f}_{j,b/c}(x)\right\}^{-\frac{c}{1-c}}. \quad (3)$$

TS originally proposed this method as an additive bias correction to the logarithms of densities, and later it is generalized and reinterpreted as an MBC technique by Koshkin (1988) and Jones and Foster (1993), respectively.

The second class of MBC techniques due to Jones, Linton and Nielsen (1995, abbreviated as "JLN" hereafter) utilizes a single smoothing parameter b. In light of the identity $f(x) = \hat{f}_{j,b}(x)\left\{f(x)/\hat{f}_{j,b}(x)\right\}$, the JLN-MBC asymmetric kernel density estimator can be defined as

$$\tilde{f}_{JLN,j}(x) = \hat{f}_{j,b}(x)\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{K_{j(x,b)}(X_i)}{\hat{f}_{j,b}(X_i)}\right\}. \quad (4)$$

Recognize that the term inside the bracket is a natural nonparametric estimator of the bias-correction term $f(x)/\hat{f}_{j,b}(x)$. This MBC technique is also applied in nonparametric regression (Linton and Nielsen, 1994), hazard estimation (Nielsen, 1998; Nielsen and Tanggaard, 2001), and spectral matrix estimation (Xiao and Linton, 2002; Hirukawa, 2006).[6] Also observe that both $\tilde{f}_{TS,j}(x)$ and $\tilde{f}_{JLN,j}(x)$ always generate non-negative density estimates everywhere by construction.

2.2 Asymptotic Properties of MBC Estimators

2.2.1 Asymptotic Results

To develop convergence properties of MBC estimators, we make the following assumptions:

[6] Because the correction term $f(x)/\hat{f}_b(x)$ reminds us of prewhitening in time series analysis, this bias correction technique is referred to as "nonparametric prewhitening" in Xiao and Linton (2002) and Hirukawa (2006).

Assumption 1. f has four continuous and bounded derivatives, and $f(x) > 0$ for a given design point $x > 0$.

Assumption 2. The smoothing parameter $b = b_n$ satisfies $b \to 0$ and $nb^{r_j+5/2} \to \infty$ as $n \to \infty$.

The smoothness condition on f in Assumption 1 is standard for consistency of density estimators using fourth-order kernels, whereas the positivity of $f(x)$ is required for MBC. Assumption 2 implies that the convergence rate of the smoothing parameter b is slower than $O\left\{n^{-1/(r_j+5/2)}\right\}$. We require this condition to control the order of magnitude in remainder terms when approximating the bias of each of TS- and JLN-MBC estimators. It will be shown shortly that the MSE-optimal smoothing parameters for these estimators are $b^* = O\left(n^{-2/9}\right)$ if the design point x satisfies $x/b \to \infty$, and that $b^\dagger = O\left\{n^{-1/(r_j+9/2)}\right\}$ if $x/b \to \kappa$ for some $\kappa > 0$; these convergence rates are indeed within the required range.

The approximation to the bias of each of two MBC estimators is built on a fourth-order Taylor expansion of $E\left\{\hat{f}_{j,b}(x)\right\}$. Under Assumptions 1 and 2, we have

$$E\left\{\hat{f}_{j,b}(x)\right\} = f(x) + a_{1,j}(x,f)\,b + a_{2,j}(x,f)\,b^2 + o\left(b^2\right),$$

where $a_{1,j}(x,f)$ and $a_{2,j}(x,f)$ are kernel-specific functions that depend on the design point x and derivatives of f. Using properties of the random variable corresponding to each kernel, we can specify explicit forms of $a_{1,j}(x,f)$ and $a_{2,j}(x,f)$ as in Table 2.

TABLE 2 ABOUT HERE

This paper refers to the position of x as "interior x" if $x/b \to \infty$, and "boundary x" if $x/b \to \kappa$. We now present two theorems on the approximations to bias and variance terms of two MBC estimators.

Theorem 1. *If Assumptions 1 and 2 hold, then the bias of the TS-MBC estimator using kernel j can be approximated by*

$$Bias\left\{\tilde{f}_{TS,j}(x)\right\} \sim \frac{1}{c(1-c)}p_j(x)\,b^2 := \frac{1}{c(1-c)}\left[\frac{1}{2}\left\{\frac{a_{1,j}^2(x,f)}{f(x)}\right\} - a_{2,j}(x,f)\right]b^2$$

for $a_{1,j}(x,f)$ and $a_{2,j}(x,f)$ in Table 2. For $v_j(x)$ and r_j defined in (2), the variance of the TS-MBC estimator can be approximated by

$$Var\left\{\tilde{f}_{TS,j}(x)\right\} = \begin{cases} n^{-1}b^{-1/2}\lambda(c)\,v_j(x)\,f(x) + o\left\{\left(nb^{1/2}\right)^{-1}\right\} & \text{for interior } x \\ O\left\{\left(nb^{r_j+1/2}\right)^{-1}\right\} & \text{for boundary } x \end{cases},$$

where

$$\lambda(c) = \frac{\left(1+c^{5/2}\right)(1+c)^{1/2} - 2\sqrt{2}c^{3/2}}{(1+c)^{1/2}(1-c)^2}.$$

Theorem 2. *If Assumptions 1 and 2 hold, then the bias of the JLN-MBC estimator using kernel j can be approximated by*

$$Bias\left\{\tilde{f}_{JLN,j}(x)\right\} \sim q_j(x)\,b^2 := -f(x)\,a_{1,j}(x,h)\,b^2,$$

where $a_{1,j}(x,h)$ is obtained by replacing $f = f(x)$ in $a_{1,j}(x,f)$ with

$$h = h(x,f) := \frac{a_{1,j}(x,f)}{f(x)}.$$

For $v_j(x)$ and r_j defined in (2), the variance of the JLN-MBC estimator can be approximated by

$$Var\left\{\tilde{f}_{JLN,j}(x)\right\} = \begin{cases} n^{-1}b^{-1/2}v_j(x)\,f(x) + o\left\{\left(nb^{1/2}\right)^{-1}\right\} & \text{for interior } x \\ O\left\{\left(nb^{r_j+1/2}\right)^{-1}\right\} & \text{for boundary } x \end{cases}.$$

2.2.2 Discussions

Bias and Variance. Because the support of asymmetric kernels matches that of the true density f, both TS- and JLN-MBC estimators are free of boundary bias. More importantly, these estimators reduce the order of magnitude in bias from $O(b)$

to $O(b^2)$, while their variances are $O\left\{\left(nb^{1/2}\right)^{-1}\right\}$ for interior x and $O\left\{\left(nb^{r_j+1/2}\right)^{-1}\right\}$ for boundary x. Observe that orders of variances remain unchanged from those for the corresponding bias-uncorrected estimator (1). We can also compare $p_j(x)$ and $q_j(x)$ in leading bias terms with the corresponding ones for symmetric second-order kernels. As stated in Jones and Signorini (1997, Sections 3.2-3.3), when the symmetric kernels are employed, the term corresponding to $p_j(x)$ for TS-MBC is a linear combination of $f''''(x)$ and $\{f''(x)\}^2/f(x)$, and the term corresponding to $q_j(x)$ for JLN-MBC is proportional to $f(x)\{f''(x)/f(x)\}''$. The reason why $p_j(x)$ and $q_j(x)$ take more complicated forms is that while odd-order moments of symmetric kernels are exactly zero, those of asymmetric kernels around the design point x are often $O(b)$ or $O(b^2)$. As a result, extra density derivatives are included in $p_j(x)$ and $q_j(x)$.

The variance of JLN-MBC estimators is first-order asymptotically equivalent to that of the corresponding bias-uncorrected estimator (1) for interior x. While the LMBC density estimator by Hagmann and Scaillet (2007) also yields the same leading variance term, this technique does not reduce the bias in order of magnitude. In contrast, when the JLN-MBC is applied for the density estimation using a symmetric second-order kernel, the leading variance term tends to be larger (although not inflated in order of magnitude) because the multiplier in the variance term involves the roughness (or squared integral) of the 'twiced' kernel (Stuetzle and Mittal, 1979). Moreover, since the multiplier $\lambda(c)$ in the variance for TS-MBC estimators is increasing in $c \in (0,1)$, ranging from 1 to 27/16, the variance of these estimators tends to be larger than that of the bias-uncorrected estimator (1) for interior x. Lastly (but not least importantly), the asymptotic variances of both TS- and JLN-MBC estimators for interior x are proportional to x^{-r_j}, even after MBC is made. Therefore, MBC estimators maintain the same appealing properties as the bias-uncorrected estimator has in estimating the distributions that have long tails with sparse data.

Mean Squared Error ("MSE"). For interior x, the MSEs of $\tilde{f}_{TS,j}(x)$ and $\tilde{f}_{JLN,j}(x)$ can be approximated by

$$\begin{aligned} MSE\left\{\tilde{f}_{TS,j}(x)\right\} &= \frac{p_j^2(x)}{c^2(1-c)^2}b^4 + n^{-1}b^{-1/2}\lambda(c)\,v_j(x)\,f(x) + o\left(b^4 + n^{-1}b^{-1/2}\right), \\ MSE\left\{\tilde{f}_{JLN,j}(x)\right\} &= q_j^2(x)\,b^4 + n^{-1}b^{-1/2}v_j(x)\,f(x) + o\left(b^4 + n^{-1}b^{-1/2}\right). \end{aligned}$$

The MSE-optimal smoothing parameters are

$$\begin{aligned} b^*_{TS,j} &= \left\{c^2(1-c)^2\lambda(c)\right\}^{2/9}\left\{\frac{v_j(x)\,f(x)}{8p_j^2(x)}\right\}^{2/9} n^{-2/9}, \\ b^*_{JLN,j} &= \left\{\frac{v_j(x)\,f(x)}{8q_j^2(x)}\right\}^{2/9} n^{-2/9}, \end{aligned}$$

which yield the optimal MSEs

$$\begin{aligned} MSE^*\left\{\tilde{f}_{TS,j}(x)\right\} &\sim \frac{9}{8^{8/9}}\gamma(c)\,p_j^{2/9}(x)\left\{v_j(x)\,f(x)\right\}^{8/9} n^{-8/9}, \\ MSE^*\left\{\tilde{f}_{JLN,j}(x)\right\} &\sim \frac{9}{8^{8/9}}q_j^{2/9}(x)\left\{v_j(x)\,f(x)\right\}^{8/9} n^{-8/9}, \end{aligned}$$

where

$$\gamma(c) = \left\{\frac{\left(1+c^{5/2}\right)(1+c)^{1/2} - 2\sqrt{2}c^{3/2}}{c^{1/4}(1+c)^{1/2}(1-c)^{9/4}}\right\}^{8/9}.$$

Observe that the MSE-optimal smoothing parameters are $O\left(n^{-2/9}\right) = O\left(h^{*2}\right)$, where h^* is the MSE-optimal bandwidth for fourth-order kernel estimators, or TS- or JLN-MBC estimators using symmetric second-order kernels. As a result, the optimal MSEs of $\tilde{f}_{TS,j}(x)$ and $\tilde{f}_{JLN,j}(x)$ for interior x become $O\left(n^{-8/9}\right)$, as with MBC estimation using the second-order kernels. The convergence rate is faster than $O\left(n^{-4/5}\right)$, the optimal convergence rate in the MSE of the corresponding bias-uncorrected estimator (1) for interior x. On the other hand, for boundary x, the MSEs of $\tilde{f}_{TS,j}(x)$ and $\tilde{f}_{JLN,j}(x)$ are $O\left\{b^4 + n^{-1}b^{-(r_j+1/2)}\right\}$, which yields the MSE-optimal smoothing parameter $b^{\dagger} = O\left\{n^{-1/(r_j+9/2)}\right\}$ and the optimal MSE of $O\left\{n^{-4/(r_j+9/2)}\right\}$. The optimal convergence rate of MSEs is indeed faster than $O\left\{n^{-2/(r_j+5/2)}\right\}$, that of the bias-uncorrected estimator for boundary x.

Global Property. The undesirable convergence rates over boundary regions do not affect the global properties of the MBC estimators. By the trimming argument in Chen (2000), the MISEs of the MBC estimators are

$$\begin{aligned} MISE\left\{\tilde{f}_{TS,j}(x)\right\} &= \frac{b^4}{c^2(1-c)^2}\int_0^{\infty} p_j^2(x)\,dx \\ &\quad + n^{-1}b^{-1/2}\lambda(c)\int_0^{\infty} v_j(x)\,f(x)\,dx + o\left(b^4 + n^{-1}b^{-1/2}\right), \\ MISE\left\{\tilde{f}_{JLN,j}(x)\right\} &= b^4\int_0^{\infty} q_j^2(x)\,dx + n^{-1}b^{-1/2}\int_0^{\infty} v_j(x)\,f(x)\,dx + o\left(b^4 + n^{-1}b^{-1/2}\right), \end{aligned}$$

provided that $p_j^2(x)$, $q_j^2(x)$, and $v_j(x)$ are integrable.[7] The MISE-optimal smoothing parameters are then given by

$$\begin{aligned} b_{TS,j}^{**} &= \left\{c^2(1-c)^2\lambda(c)\right\}^{2/9}\left\{\frac{\int_0^{\infty} v_j(x)\,f(x)\,dx}{8\int_0^{\infty} p_j^2(x)\,dx}\right\}^{2/9} n^{-2/9}, \\ b_{JLN,j}^{**} &= \left\{\frac{\int_0^{\infty} v_j(x)\,f(x)\,dx}{8\int_0^{1} q_j^2(x)\,dx}\right\}^{2/9} n^{-2/9}. \end{aligned}$$

Therefore, the optimal MISEs become

$$\begin{aligned} MISE^{**}\left\{\tilde{f}_{TS,j}(x)\right\} &\sim \frac{9}{8^{8/9}}\gamma(c)\left\{\int_0^{\infty} p_j^2(x)\,dx\right\}^{2/9}\left\{\int_0^{\infty} v_j(x)\,f(x)\,dx\right\}^{8/9} n^{-8/9}, \\ MISE^{**}\left\{\tilde{f}_{JLN,j}(x)\right\} &\sim \frac{9}{8^{8/9}}\left\{\int_0^{\infty} q_j^2(x)\,dx\right\}^{2/9}\left\{\int_0^{\infty} v_j(x)\,f(x)\,dx\right\}^{8/9} n^{-8/9}. \end{aligned}$$

Furthermore, the multiplier $\gamma(c)$ in the optimal MISE for the TS-MBC estimator is minimized at $c^* \approx 0.2636$; this value is exclusively considered in subsequent analyses.

[7] Throughout this paper, $p_{MG}(x)$ and $q_{MG}(x)$ refer to those for interior x (i.e. $x \geq 2b$), whenever the integrated squared bias is considered.

Normalization. Also observe that neither $\tilde{f}_{TS,j}(x)$ nor $\tilde{f}_{JLN,j}(x)$ integrates to one, although the lack of normalization may be ignored in some applications. In general, MBC leads to lack of normalization, even if symmetric second-order kernels are employed; see Section 2.2 of JLN, for example. It is easy to see that the order of magnitude in the leading bias terms of renormalized TS- and JLN-MBC estimators

$$\begin{aligned}\tilde{f}^{R}_{TS,j}(x) &= \frac{\tilde{f}_{TS,j}(x)}{\int_0^\infty \tilde{f}_{TS,j}(x)\,dx},\\ \tilde{f}^{R}_{JLN,j}(x) &= \frac{\tilde{f}_{JLN,j}(x)}{\int_0^\infty \tilde{f}_{JLN,j}(x)\,dx}\end{aligned}$$

remains unchanged. Because

$$\begin{aligned}E\left\{\int_0^\infty \tilde{f}_{TS,j}(x)\,dx\right\} &= 1+\frac{b^2}{c(1-c)}\int_0^\infty p_j(x)\,dx+o\left(b^2\right),\\ E\left\{\int_0^\infty \tilde{f}_{JLN,j}(x)\,dx\right\} &= 1+b^2\int_0^\infty q_j(x)\,dx+o\left(b^2\right),\end{aligned}$$

provided that $p_j(x)$ and $q_j(x)$ are integrable, the leading biases of $\tilde{f}^{R}_{TS,j}(x)$ and $\tilde{f}^{R}_{JLN,j}(x)$ become

$$\begin{aligned}Bias\left\{\tilde{f}^{R}_{TS,1}(x)\right\} &\sim \frac{1}{c(1-c)}\left\{p_j(x)-\int_0^\infty p_j(x)\,dx\right\}b^2,\\ Bias\left\{\tilde{f}^{R}_{JLN,1}(x)\right\} &\sim \left\{q_j(x)-\int_0^\infty q_j(x)\,dx\right\}b^2.\end{aligned}$$

The leading variances are unaffected.

2.3 Further Bias Reduction via Iteration

In principle, further bias reduction is possible after the regularity conditions are properly strengthened. Constructing a multiplicative combination of $(s+1)$ different density estimators, we can extend the TS-MBC estimator to

$$\tilde{f}^{(s)}_{TS,j}(x)=\prod_{r=0}^{s}\left\{\hat{f}_{j,b/c_r}(x)\right\}^{\alpha_{s,r}},$$

where $c_0 \equiv 1$, $c_1, \ldots, c_s \in (0,1)$ are mutually different constants, and the exponent is

$$\alpha_{s,r} = \frac{(-1)^s c_r^s}{\prod_{p=0, p \neq r}^{s} (c_p - c_r)}.$$

Similarly, the sth iterated JLN-MBC estimator can be defined as

$$\tilde{f}_{JLN,1}^{(s)}(x) = \tilde{f}_{JLN,j}^{(s-1)}(x) \left\{ \frac{1}{n} \sum_{i=1}^{n} \frac{K_{j(x,b)}(X_i)}{\tilde{f}_{JLN,j}^{(s-1)}(X_i)} \right\},$$

where $\tilde{f}_{JLN,j}^{(0)}(x) \equiv \hat{f}_{j,b}(x)$. Provided that the true density f is $2(s+1)$ times continuously differentiable, for each of these estimators, it can be shown that the order of magnitude in bias is $O(b^{s+1})$ while the order of magnitude in variance remains unchanged from that of the corresponding bias-uncorrected estimator $\hat{f}_{j,b}(x)$. In particular, it can be shown that $Var\left\{\tilde{f}_{JLN,j}^{(s)}(x)\right\}$ is first-order asymptotically equivalent to $Var\left\{\tilde{f}_{JLN,j}^{(0)}(x)\right\} = Var\left\{\hat{f}_{j,b}(x)\right\}$ for interior x. As a result, their optimal MSEs are $O\left\{n^{-(4s+4)/(4s+5)}\right\}$ and $O\left\{n^{-(2s+2)/(2s+3)}\right\}$ for interior and boundary x, respectively. Accordingly, as the number of iterations increases, the global convergence rate of iterated MBC estimators can be arbitrarily close to the parametric one when best implemented. However, it is doubtful whether there is much gain in practice from these estimators, and thus we do not pursue this issue any further.

3 Finite Sample Performance

3.1 Monte Carlo Setup

We evaluate the finite sample performance of two classes of MBC estimators via Monte Carlo simulations. For estimators involving asymmetric kernels, we concentrate on the Gamma and Modified Gamma kernels due to their popularity in the literature. This simulation study compares of the following five classes of estimators: (i) bias-uncorrected estimators (1) [G, MG]; (ii) LMBC estimator with a gamma start and a log-linear correction factor in Hagmann and Scaillet (2007) [LMBC-G]; (iii) LTBC estimator with the generalized Champernowne start[8] and a log-linear correction factor in Gustafsson *et al.* (2009) [LTBC-C]; (iv) TS-MBC estimators (3) [TS-MBC-G, TS-MBC-MG]; and (vi) JLN-MBC estimators (4) [JLN-MBC-G, JLN-MBC-MG]. The value of the constant c in each TS-MBC estimator is set equal to the MISE-optimal $c^* = 0.2636$. Ten true distributions are considered, as listed in Table 3. All these distributions are popularly chosen as models for the income distribution, the loss distribution and the baseline hazard. For each distribution, $1,000$ data sets of sample size $n = 100$, 200 or 500 are simulated. All density estimates are evaluated on an equally spaced grid of 500 points over the interval $[0,5]$. Following Gustafsson *et al.* (2009), for each estimator $\bar{f}$, we compute three performance measures, namely, the integrated absolute deviation ("IAD"), the root integrated squared error ("RISE"), and the root weighted integrated squared error ("RWISE"), where

$$\begin{aligned}
IAD\left\{\bar{f}(x)\right\} &= \int_0^\infty \left|\bar{f}(x) - f(x)\right| dx, \\
RISE\left\{\bar{f}(x)\right\} &= \sqrt{\int_0^\infty \left\{\bar{f}(x) - f(x)\right\}^2 dx}, \\
RWISE\left\{\bar{f}(x)\right\} &= \sqrt{\int_0^\infty \left\{\bar{f}(x) - f(x)\right\}^2 x^2 dx}.
\end{aligned}$$

IAD and RISE measure the error between the true and estimated densities with equal weights across the support, whereas RWISE focuses on the fit of the tail part. In our reports, the integrals are approximated over the 500 points.

TABLE 3 ABOUT HERE

[8]The density of the generalized Champernowne distribution (Buch-Larsen *et al.*, 2005) is

$$f(x;\alpha,M,c) = \frac{\alpha(x+c)^{\alpha-1}\left\{(M+c)^\alpha - c^\alpha\right\}}{\left\{(x+c)^\alpha + (M+c)^\alpha - 2c^\alpha\right\}^2},\ x \geq 0$$

with parameters $\alpha > 0$, $M > 0$ and $c \geq 0$.

3.2 Choices of Smoothing Parameters

Choosing the smoothing parameter b is an important practical issue. To expedite computations, as in Hirukawa (2010), we develop plug-in methods for TS- and JLN-MBC estimators that use a gamma density as a reference. The plug-in smoothing parameters for $\tilde{f}_{TS,MG}(x)$ and $\tilde{f}_{JLN,G}(x)$ (called "gamma-referenced smoothing parameters" hereafter) are defined as the minimizers of asymptotic weighted mean integrated squared errors ("AWMISEs")

$$\begin{aligned}
\hat{b}_{GR-TS} &= \arg\min_{b} AWMISE\left\{\tilde{f}_{TS,MG}(x)\right\} \\
&= \arg\min_{b} \frac{b^4}{c^2(1-c)^2}\int_0^\infty \tilde{p}_{MG}^2(x)\, w_{TS}(x)\, dx + \frac{n^{-1}b^{-1/2}\lambda(c)}{2\sqrt{\pi}}\int_0^\infty \frac{g(x)}{\sqrt{x}} w_{TS}(x)\, dx, \\
\hat{b}_{GR-JLN} &= \arg\min_{b} AWMISE\left\{\tilde{f}_{JLN,G}(x)\right\} \\
&= \arg\min_{b} b^4\int_0^\infty \tilde{q}_G^2(x)\, w_{JLN}(x)\, dx + \frac{n^{-1}b^{-1/2}}{2\sqrt{\pi}}\int_0^\infty \frac{g(x)}{\sqrt{x}} w_{JLN}(x)\, dx,
\end{aligned}$$

where $g(x) = x^{\alpha-1}\exp(-x/\beta)/\{\beta^\alpha\Gamma(\alpha)\}$ is the density function for the gamma distribution with parameters (α,β), and $\tilde{p}_{MG}(x)$ and $\tilde{q}_G(x)$ can be obtained by replacing $f(x)$ in $p_{MG}(x)$ and $q_G(x)$ with $g(x)$. The weighting functions are chosen as $w_{TS}(x) = x^5$ and $w_{JLN}(x) = x$ to ensure finiteness of integrals. The parameters (α,β) are replaced by their estimates $\left(\hat{\alpha},\hat{\beta}\right)$ via method of moments or maximum likelihood ("ML").[9] Analytical expressions of $\hat{b}_{GR-TS}$ and $\hat{b}_{GR-JLN}$, as well as $\hat{b}_{GR}$ (= the gamma-referenced smoothing parameter for $\hat{f}_{MG}(x)$), are given in the Appendix.

We do not pursue the gamma-referenced smoothing parameter for $\tilde{f}_{TS,G}(x)$; since extra terms are involved in $p_G(x)$, the minimizer of its AWMISE takes a much more complicated form than $\hat{b}_{GR-TS}$. On the other hand, although it is possible to derive

[9] ML estimates are used exclusively in this paper.

the gamma-referenced smoothing parameter for $\tilde{f}_{JLN,MG}(x)$ in a similar way,[10] our preliminary simulation results indicate that the formula frequently generates large values and thus we do not advocate its use. From the viewpoint of practical relevance, $\hat{b}_{GR-TS}$ and $\hat{b}_{GR-JLN}$ are simply employed for $\tilde{f}_{TS,G}(x)$ and $\tilde{f}_{JLN,MG}(x)$ in our simulations, respectively. Similarly, $\hat{b}_{GR}$ is chosen as the smoothing parameter for $\hat{f}_G(x)$.

Besides, a very simple formula is frequently applied in the literature (e.g. Gustafsson *et al.*, 2009). From this viewpoint, a "rule-of-thumb" smoothing parameter is also considered for each estimator. More precisely, we additionally employ: (i) $\hat{b}_{ROT1} = \hat{\sigma}_x n^{-2/5}$ for G, MG and LMBC-G[11]; (ii) $\hat{b}_{ROT2} = \hat{\sigma}_u n^{-2/5}$ for LTBC-C; and (iii) $\hat{b}_{ROT3} = \hat{\sigma}_x n^{-2/9}$ for TS-MBC-G, TS-MBC-MG, JLN-MBC-G, and JLN-MBC-MG, where $\hat{\sigma}_x$ and $\hat{\sigma}_u$ are sample standard deviations of the original data and the transformed data on the unit interval, respectively.

TABLE 4 AND FIGURE 2 ABOUT HERE

3.3 Simulation Results

Table 4 reports simulation averages and standard deviations of performance measures. The gamma-referenced method in general works better than the rule-of-thumb method for a given estimator, and it often substantially reduce the values of performance measures.[12] Therefore, the table presents the results from the gamma-referenced method for all estimators other than two semiparametric estimators (i.e. LMBC-G

[10] Choosing x^5 as the weighting function, we can derive the smoothing parameter as

$$\arg\min_b b^4 \int_0^\infty \tilde{q}_{MG}^2(x)\, x^5 dx + \frac{n^{-1}b^{-1/2}}{2\sqrt{\pi}} \int_0^\infty \frac{g(x)}{\sqrt{x}} x^5 dx = \left\{ \frac{4^\alpha \beta^{9/2} \Gamma(\alpha+9/2)\,\Gamma(\alpha)}{4\sqrt{\pi}\,(\alpha-1)^2 (\alpha-2)^2\, \Gamma(2\alpha)} \right\}^{2/9} n^{-2/9}.$$

[11] Although Hagmann and Scaillet (2007) employ the cross validation method for LMBC-G, we adopt the simple plug-in rule to speed up computations.

[12] Exceptions are Distributions 3 and 4, but gains from the rule-of-thumb method are marginal.

and LTBC-C); in fact, only the rule-of-thumb method is chosen for these estimators.

For each distribution, results are qualitatively similar across three sample sizes. Comparing TS- and JLN-MBC estimators with their corresponding bias-uncorrected estimators, we can see that JLN-MBC estimators often improve all three performance measures, whereas results are mixed for TS-MBC estimators. JLN-MBC outperforms TS-MBC in most cases; only a few exceptions include RISEs for Distributions 5, 7. This finding is consistent with the simulation results reported in Jones and Signorini (1997). In addition, even when the two semiparametric estimators perform better in terms of overall performance measures (i.e. IAD and RISE), JLN-MBC is competitive with them in terms of the performance measure for the tail (i.e. RWISE); see the results for Distributions 5, 6, 7, for instance.

Results do not confirm that for a given estimator, employing the Modified Gamma kernel improves performance measures over the Gamma kernel. A rationale can be found in Figure 1. The Gamma kernel put the maximum weight on the design point x whereas the Modified Gamma not, which explains why the Modified Gamma estimators may not outperform the Gamma estimators globally. In particular, TS-MBC-MG consistently performs inferiorly to TS-MBC-G, except Distribution 5. The poor performance of TS-MBC-MG can be attributed to the following two respects. First, TS-MBC estimation relies on two smoothing parameters b and b/c. Controlling both b and b/c is a cumbersome task. Because $0 < c < 1$, the density estimator using b/c tends to be oversmoothed, which is potentially a source of a large bias in every TS-MBC estimator. On the other hand, if we make b too short in order to have a reasonable length of b/c, additional variability is introduced to the other estimator using b due to undersmoothing. Second, when the Modified Gamma kernel is employed, these two smoothing parameters also play a role of determining the boundary region explicitly (e.g. $[0, 2b)$ for the density estimator using b). Unless b is short enough, there is a relatively small interior region for the density estimator using b/c. This aspect is also thought to worsen the performance measures of TS-MBC-MG. In conclusion, when TS-MBC estimation is applied, it is desirable to use a smoothing parameter choice method that tends to provide a small value consistently. Our experiment indicates that for a given distribution and a given sample size, $\hat{b}_{GR-TS}$ on average yields a smaller value than $\hat{b}_{ROT3}$ does, which explains why the former is preferable in general for TS-MBC estimation over the latter.

It is also worth remarking that performance of the two semiparametric estimators depends on the shape of the distribution. Due to its gamma start, LMBC-G performs well when the distribution looks like the gamma one (e.g. Distributions 1, 2, 7, 8). On the other hand, because the generalized Champernowne distribution is designed to capture the tail of the underlying distribution, the performance of LTBC-C is remarkable for medium- to heavy-tailed distributions (e.g. Distributions 5, 6, 9, 10). However, LMBC-G for Distributions 3 and 4 and LTBC-C for Distributions 1-4 perform inferiorly. To confirm these findings, we now inspect average plots of their density estimates.

Figure 2 presents average plots of density estimates from G, LMBC-G, LTBC-C, TS-MBC-G, and JLN-MBC-G against the true density for selected distributions. The plots are obtained from 1,000 Monte Carlo samples with sample size $n = 100$, and the average for each estimator is taken over 1,000 replications for each grid. For each distribution, the left and right panels correspond to the average plots over the region near the origin and the tail part, respectively. Panels (a) and (d) indicate that LMBC-G and LTBC-C can capture the shapes of Weibull and Pareto distributions nearly perfectly, respectively. In contrast, as suggested by Panel (a), LTBC-C imprecisely captures the shape near the origin of such a distribution as Distribution 1 or 2. Moreover, Distributions 3 and 4 satisfy so-called the shoulder condition $f'(0^+) = 0$. In this case, local concavity of the true density around the origin tends to force the gamma and generalized Champernowne starts to generate a spurious peak near the origin (see Panel (b)), which leads to the poor performance of LMBC-G and LTBC-C. Likewise, local convexity of the true density around the origin in Distribution 6 makes the gamma start for LMBC-G unbounded at zero. Panel (d) indicates that the unboundedness cannot be corrected even after the bias correction step.

We can also see from Figure 2 that if G underestimates (overestimates) the density, JLN-MBC-G corrects the estimate in an upward (downward) direction, as reported in Hirukawa (2006, 2010); such a bias correction mechanism is not obvious for TS-MBC-G. Both of TS-MBC-G and JLN-MBC-G can trace out entire shapes of a variety of distributions at a satisfactory level; the chances that they mistakenly capture the shapes would be low, unlike the two semiparametric estimators. While JLN-MBC-G performs superiorly to TS-MBC-G in terms of the three performance measures, it appears that the latter is better at capturing heights of the peaks of distributions than the former. The figure also demonstrates that both estimators can preserve the shapes of tail parts equally well. Overall, simulation results indicate that MBC estimators are viable alternatives to the two semiparametric estimators.

4 Applications to Income Data

4.1 Case #1: U.S. Data

In this section, the MBC density estimators are applied to a couple of income data sets. Our first application focuses on the U.S. income data. We use the data set studied in Abadie (2003), which includes 9,275 observations from the 1991 Survey of

Income and Program Participation (SIPP). The data set is available under the name "`401ksubs.raw`" on the web site for Wooldridge (2001), and the variable "`inc`" (= annual income, in thousands of dollars) is examined in this section.

Table 5 presents descriptive statistics of the data, including some inequality measures, where the data is converted to dollars. Although sample moments and quantiles suggest a right-skewed distribution, it appears that the tail is not extremely heavy. This is confirmed by the Hill plot (= a plot of tail index estimates by Hill, 1975) of the data up until the 1,000th observation from the largest (unreported), which indicates that estimates of the tail index roughly ranges from 4 to 5, except initial few estimates.

TABLE 5 AND FIGURE 3 ABOUT HERE

The income distribution is estimated by G, LMBC-G, TS-MBC-G, and JLN-MBC-G. The original data is converted to 10^5 dollars, and then the resulting density estimates are back-transformed to the ones denominated in dollars. At the initial step, we fit a gamma density $g(x) = x^{\alpha-1}\exp(-x/\beta)/\{\beta^{\alpha}\Gamma(\alpha)\}$ to the converted data by ML, and obtain parameter estimates $\left(\hat{\alpha}, \hat{\beta}\right) = (3.158, 0.124)$. Then, plug-in smoothing parameters $\hat{b}_{GR}$, $\hat{b}_{ROT1}$, $\hat{b}_{GR-TS}$, and $\hat{b}_{GR-JLN}$ are chosen for G, LMBC-G, TS-MBC-G, and JLN-MBC-G, respectively; their values are 0.0030, 0.0062, 0.0071, and 0.0354. Given these numbers, we can draw Figure 3. For reference, the figure presents shapes around the mode (Panel (b)) and on the tail (Panel (c)), as well as the entire shape (Panel (a)). The figure suggests that while estimates from G and TS-MBC-G are almost indistinguishable, the ones from the remaining two are considerably different, in particular, in terms of the location and height of the mode; see Panel (b) for details. Moreover, a closer look at Panel (c) finds that the wiggle over the tail generated by G is substantially smoothed away by JLN-MBC-G; this kind of smoothing mechanism is not observed in TS-MBC-G.

4.2 Case #2: Brazilian Data

Our second application employs the Brazilian income data in 1990. The large micro data set ($n = 71,523$) is collected via the annual National Household Survey (PNAD) by the Brazilian Statistical Office, and it has been also analyzed in Hagmann and Scaillet (2007). Using similar data sets, Cowell, Ferreira and Litchfield (1998) investigate the dynamics of the Brazilian income distribution in 1980s. The data considered is monthly per capita denominated in 1990 cruzeiros. Analyzing this data set is of particular interest, because Brazil has the eighth largest GDP in the world and faces a strong inequality in the income distribution.

Table 6 presents descriptive statistics of the data. The income distribution is empirically known to have a single sharp peak near the origin and a long right tail. Indeed, sample moments and quantiles reveal an extremely right-skewed distribution. The Hill plot of the data up until the 8,000th observation from the largest suggests that estimates of the tail index roughly range from 1.6 to 3, except initial few estimates. Inequality measures much larger than those from the U.S. data also reflect the long, thick tail.

TABLE 6 AND FIGURE 4 ABOUT HERE

As in the previous section, the income distribution is estimated by G, LMBC-G, TS-MBC-G, and JLN-MBC-G. Again, the original data is converted to 10^5 cruzeiros, and then the resulting density estimates are back-transformed to the ones denominated in cruzeiros. Fitting a gamma density $g(x) = x^{\alpha-1}\exp(-x/\beta)/\{\beta^{\alpha}\Gamma(\alpha)\}$ to the converted data yields ML estimates $\left(\hat{\alpha}, \hat{\beta}\right) = (0.887, 0.588)$. Computing values of smoothing parameters $\hat{b}_{GR}$, $\hat{b}_{ROT1}$, $\hat{b}_{GR-TS}$, and $\hat{b}_{GR-JLN}$ as 0.0050, 0.0104, 0.0246, and 0.0533, we can obtain Figure 4. Although the value of $\hat{\alpha}$ implies that unlike the U.S. case, the data would hint an income distribution unbounded at zero, the estimate from LMBC-G does not confirm this gesture. Because of quite a few observations, it seems that there is a consensus on the shape among estimators, except that of the mode. G generates the highest peak. JLN-MBC-G substantially smooths out the mode, and as a result, the peak is lowest. LMBC-G and TS-MBC-G yield a similar shape around the mode, and the height of the peak is in-between. On the other hand, a careful investigation of Panel (c) reveals that G and LMBC-G generate wiggles over the tail, whereas two MBC estimators smooth them out thoroughly.

5 Conclusion

This paper has demonstrated that two well-known MBC techniques designed originally for symmetric kernels can be applied to density estimation using asymmetric kernels that have support on $[0, \infty)$. Under sufficient smoothness of the true density, both bias reduction methods are shown to improve the order of magnitude in bias from $O(b)$ to $O(b^2)$, while the order of magnitude in variance remains unchanged. Two classes of MBC density estimators are by construction nonnegative, and establish a faster convergence rate of $O\left(n^{-8/9}\right)$ in MSE for the interior part when best implemented, as with symmetric second-order kernels. Monte Carlo simulations indicate superior performance of JLN-MBC estimators in particular, compared to corresponding bias-uncorrected estimators. The MBC estimators are applied to estimating income distributions from the U.S. and Brazilian data.

A Appendix

The proof of each theorem requires kernel-specific arguments, which include Taylor expansions and properties of the random variable corresponding to the kernel. To save space, we present only the proofs when the Gamma kernel is employed.

A.1 Proof of Theorem 1

This proof closely follows the one for Theorem 1 of Hirukawa (2010).

Bias. Applying the first equation on p.489 of Hirukawa (2010), we can approximate

$$E\left\{\tilde{f}_{TS,G}(x)\right\} = f(x) + \frac{1}{c(1-c)}\left[\frac{1}{2}\left\{\frac{a_{1,G}^2(x,f)}{f(x)}\right\} - a_{2,G}(x,f)\right]b^2 + o\left(b^2\right) + O\left(n^{-1}b^{-1}\right),$$

where the remainder term $O\left(n^{-1}b^{-1}\right) = o\left(b^2\right)$ by Assumption 2.

Variance. The variance of $\tilde{f}_{TS,G}(x)$ can be expressed as

$$\begin{aligned}
& Var\left\{\tilde{f}_{TS,G}(x)\right\} \\
= & \frac{1}{(1-c)^2}\left[Var\left\{\hat{f}_{G,b}(x)\right\} - 2cCov\left\{\hat{f}_{G,b}(x), \hat{f}_{G,b/c}(x)\right\} + c^2 Var\left\{\hat{f}_{G,b/c}(x)\right\}\right] + O\left(n^{-1}\right).
\end{aligned}$$

For interior x, a similar argument to the one in the proof for Theorem 1 of Hirukawa (2010) yields

$$\begin{aligned}
Var\left\{\hat{f}_{G,b}(x)\right\} &= n^{-1}b^{-1/2}v_G(x)f(x) + o\left(n^{-1}b^{-1/2}\right), \\
Var\left\{\hat{f}_{G,b/c}(x)\right\} &= n^{-1}b^{-1/2}c^{1/2}v_G(x)f(x) + o\left(n^{-1}b^{-1/2}\right), \\
Cov\left\{\hat{f}_{G,b}(x), \hat{f}_{G,b/c}(x)\right\} &= n^{-1}b^{-1/2}\frac{\sqrt{2}c^{1/2}}{(1+c)^{1/2}}v_G(x)f(x) + o\left(n^{-1}b^{-1/2}\right).
\end{aligned}$$

Then, the result immediately follows.

On the other hand, for boundary x, each of $Var\left\{\hat{f}_{G,b}(x)\right\}$, $Var\left\{\hat{f}_{G,b/c}(x)\right\}$ and $Cov\left\{\hat{f}_{G,b}(x), \hat{f}_{G,b/c}(x)\right\}$ is at most $O\left\{(nb)^{-1}\right\}$. Therefore, the stated order of magnitude is also established. ■

A.2 Proof of Theorem 2

This proof closely follows the one for Theorem 2 of Hirukawa (2010). To approximate the variance, the following two lemmata are required:

Lemma 1. *Suppose that Assumptions 1 and 2 hold. Let*

$$\varsigma_j(u) = \frac{2K_{j(x,b)}(u)}{f(u)} - \frac{1}{n}\sum_{i=1}^{n} \frac{K_{j(x,b)}(X_i)\, K_{j(X_i,b)}(u)}{f^2(X_i)}.$$

Then,

$$\varsigma_j(u) \sim \zeta_j(u) = \rho_j \frac{K_{j(x,b)}(u)}{f(u)},$$

where

$$\rho_j = \begin{cases} \begin{cases} 1 & \text{if } u/b \to \infty \\ 2 - \Phi\left(\kappa^{1/2}\right) & \text{if } u/b \to \kappa \end{cases} & \text{for } j = G, MG, RIG \\ 1 & \text{for } j = LN, BS \\ \begin{cases} 1 & \text{if } ub \to 0 \\ 2 - \Phi\left(\lambda^{1/2}\right) & \text{if } ub \to \lambda^{-1} \end{cases} & \text{for } j = IG \end{cases}$$

for some $\lambda > 0$.

Lemma 2. *Suppose that Assumptions 1 and 2 hold. Let $X > 0$ be drawn from the distribution with density $f(x)$ having support on $[0, \infty)$. Then, the transformation $\zeta_j(\cdot)$ defined in Lemma 1 can be approximated by*

$$E\left\{\zeta_j^2(X)\right\} = \begin{cases} b^{-1/2} v_j(x)/f(x) + o\left(b^{-1/2}\right) & \text{for interior } x \\ O\left\{b^{-(r_j+1/2)}\right\} & \text{for boundary } x \end{cases},$$

where $v_j(x)$ and r_j are defined in Theorem 1.

A.2.1 Proof of Lemma 1

It suffices to demonstrate that

$$\frac{1}{n}\sum_{i=1}^{n} \frac{K_{G(x/b+1,b)}(X_i)\, K_{G(X_i/b+1,b)}(u)}{f^2(X_i)} \sim \frac{K_{G(x/b+1,b)}(u)}{f(u)} \times \begin{cases} 1 & \text{if } u/b \to \infty \\ \Phi\left(\kappa^{1/2}\right) & \text{if } u/b \to \kappa \end{cases}.$$

Let $\psi_G(y) := K_{G(x/b+1,b)}(y)/f(y)$. Then, for an arbitrarily large n, we can approximate $(1/n)\sum_{i=1}^{n} K_{G(x/b+1,b)}(X_i) K_{G(X_i/b+1,b)}(u)/f^2(X_i)$ by

$$E\left\{\frac{K_{G(x/b+1,b)}(X_1) K_{G(X_1/b+1,b)}(u)}{f^2(X_1)}\right\} = \int_0^\infty \psi_G(y) \frac{u^{y/b}\exp(-u/b)}{b^{y/b+1}\Gamma(y/b+1)} dy. \tag{A1}$$

Now,

$$\frac{u^{y/b}\exp(-u/b)}{b^{y/b+1}\Gamma(y/b+1)} = \exp\left\{\left(\frac{y}{b}\right)\log u - \frac{u}{b} - \log y - \left(\frac{y}{b}\right)\log b - \log\Gamma\left(\frac{y}{b}\right)\right\},$$

where

$$\log\Gamma\left(\frac{y}{b}\right) = \left(\frac{y}{b}-\frac{1}{2}\right)\log y - \left(\frac{y}{b}-\frac{1}{2}\right)\log b - \frac{y}{b} + \frac{1}{2}\log 2\pi + \frac{b}{12y} + O\left(b^3\right)$$

by Stirling's approximation. Then, the right-hand side of (A1) can be approximated by

$$\frac{b^{-1/2}}{\sqrt{2\pi}}\int_0^\infty \frac{\psi_G(y)}{\sqrt{y}} \exp\left\{\left(\frac{y}{b}\right)(\log u - \log y) - \left(\frac{u-y}{b}\right) - \frac{b}{12y} + O\left(b^3\right)\right\} dy. \tag{A2}$$

By a change of variable $w = (u-y)/b^{1/2}$ and Taylor expansions for the logarithm and exponential functions, (A2) can be further rewritten as

$$\begin{aligned}
& \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u/b^{1/2}} \frac{\psi_G\left(u-b^{1/2}w\right)}{\left(u-b^{1/2}w\right)^{1/2}} \\
& \times \exp\left[\left(\frac{u-b^{1/2}w}{b}\right)\left\{\log u - \log\left(u-b^{1/2}w\right)\right\} - b^{-1/2}w - \frac{b}{12\left(u-b^{1/2}w\right)} + O\left(b^3\right)\right] dw \\
= \; & \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u/b^{1/2}} \frac{\psi_G\left(u-b^{1/2}w\right)}{\left(u-b^{1/2}w\right)^{1/2}} \exp\left(-\frac{w^2}{2u}\right) \\
& \times \left\{1 - \frac{b^{1/2}w^3}{6u^2} - \frac{bw^4}{12u^3} - \frac{b}{12u} + \frac{bw^6}{72u^4} + O\left(b^{3/2}\right)\right\} dw.
\end{aligned} \tag{A3}$$

Take another change of variable $v = w/\sqrt{u}$, and let $\phi(\cdot)$ and $\Phi(\cdot)$ denote the standard normal density and distribution functions, respectively. Then, (A3) reduces to

$$\begin{aligned}
& \int_{-\infty}^{\sqrt{u/b}} \psi_G\left(u - b^{1/2}\sqrt{u}v\right)\left(\frac{u}{u-b^{1/2}\sqrt{u}v}\right)^{1/2}\phi(v) \\
& \times \left\{1 - \frac{b^{1/2}v^3}{6\sqrt{u}} - \frac{bv^4}{12u} - \frac{b}{12u} + \frac{bv^6}{72u} + O\left(b^{3/2}\right)\right\} dv \\
\sim \; & \frac{K_{G(x/b+1,b)}(u)}{f(u)} \times \begin{cases} 1 & \text{if } u/b \to \infty \\ \Phi\left(\kappa^{1/2}\right) & \text{if } u/b \to \kappa \end{cases} . \quad \blacksquare
\end{aligned}$$

A.2.2 Proof of Lemma 2

Pick $\delta = b^{1-\epsilon}$ for some $\epsilon \in (0, 1/2)$. Then, by the trimming argument in Chen (2000),

$$E\left\{\zeta_G^2(X)\right\} = \int_0^\delta + \int_\delta^\infty \zeta_G^2(u) f(u)\, du = A_b(x) \int_0^\infty \frac{K_{G(2x/b+1,b/2)}(u)}{f(u)} du + O\left(b^{-\epsilon}\right),$$

where

$$A_b(x) = \frac{b^{-1}\Gamma(2x/b+1)}{2^{2x/b+1}\Gamma^2(x/b+1)} \sim \begin{cases} \frac{b^{-1/2}}{2\sqrt{\pi}x^{1/2}} & \text{if } x/b \to \infty \\ \frac{\Gamma(2\kappa+1)b^{-1}}{2^{2\kappa+1}\Gamma^2(\kappa+1)} & \text{if } x/b \to \kappa \end{cases}.$$

The stated result follows from recognizing that

$$\int_0^\infty \frac{K_{G(2x/b+1,b/2)}(u)}{f(u)} du = E\left\{f^{-1}(\theta_x)\right\} = f^{-1}(x) + o(1),$$

where $\theta_x \overset{d}{=} G(2x/b+1, b/2)$. ■

A.2.3 Proof of Theorem 2

Bias. Write $h(x) = a_{1,G}(x, f)/f(x)$. By the procedures in Section A.2.1 of Hirukawa (2010, pp.490-491) and properties of gamma random variables, we can recognize that

$$E\left\{\tilde{f}_{JLN,G}(x)\right\} = f(x) - f(x)\left\{h'(x) + \frac{x}{2}h''(x)\right\} b^2 + o\left(b^2\right) = f(x) - f(x)\, a_{1,G}(x, h)\, b^2 + o\left(b^2\right).$$

Variance. Following the procedures in Section A.2.2 of Hirukawa (2010, p.492), we can obtain

$$\tilde{f}_{JLN,G}(x) \sim f(x) \frac{1}{n} \sum_{i=1}^n \frac{K_{G(x/b+1,b)}(X_i)}{f(X_i)} \left\{2 - \frac{\hat{f}_{G,b}(X_i)}{f(X_i)}\right\},$$

where

$$\frac{1}{n} \sum_{i=1}^n \frac{K_{G(x/b+1,b)}(X_i)}{f(X_i)} \left\{2 - \frac{\hat{f}_{G,b}(X_i)}{f(X_i)}\right\} = \frac{1}{n} \sum_{i=1}^n \varsigma_G(X_i)$$

for $\varsigma_G(\cdot)$ defined in Lemma 1. Lemma 1 implies that $\varsigma_G(X_i)$ can be approximated by an *iid* random variable $\zeta_G(X_i)$, and thus

$$Var\left\{\tilde{f}_{JLN,G}(x)\right\} \sim f^2(x) \frac{1}{n} Var\left\{\zeta_G(X_1)\right\} = f^2(x) \left[\frac{1}{n} E\left\{\zeta_G^2(X_1)\right\} + O\left(n^{-1}\right)\right].$$

Finally, applying Lemma 2 completes the proof. ■

A.3 Formulae for Gamma-Referenced Smoothing Parameters

The analytical expression of $\hat{b}_{GR-TS}$ is

$$\hat{b}_{GR-TS} = \left\{c^2 (1-c)^2 \lambda(c)\right\}^{2/9} \left\{ \frac{4^{\alpha} \beta^{9/2} \Gamma(\alpha + 9/2) \Gamma(\alpha)}{16\sqrt{\pi} C_{TS}(\alpha) \Gamma(2\alpha)} \right\}^{2/9} n^{-2/9},$$

where

$$\begin{aligned} C_{TS}(\alpha) &= \frac{1}{36} (\alpha - 2)^2 \left(\alpha - \frac{3}{2}\right)^2 (\alpha - 1)^2 \\ &\quad - \frac{1}{6} (\alpha - 2) \left(\alpha - \frac{3}{2}\right) (\alpha - 1)^2 (\alpha) \left(\alpha + \frac{1}{2}\right) \\ &\quad + \frac{1}{9} (\alpha - 2) \left(\alpha - \frac{3}{2}\right) (\alpha - 1) (\alpha) \left(\alpha + \frac{1}{2}\right) (\alpha + 1) \\ &\quad + \frac{1}{4} (\alpha - 1)^2 (\alpha) \left(\alpha + \frac{1}{2}\right) (\alpha + 1) \left(\alpha + \frac{3}{2}\right) \\ &\quad - \frac{1}{3} (\alpha - 1) (\alpha) \left(\alpha + \frac{1}{2}\right) (\alpha + 1) \left(\alpha + \frac{3}{2}\right) (\alpha + 2) \\ &\quad + \frac{1}{9} (\alpha) \left(\alpha + \frac{1}{2}\right) (\alpha + 1) \left(\alpha + \frac{3}{2}\right) (\alpha + 2) \left(\alpha + \frac{5}{2}\right). \end{aligned}$$

On the other hand, $\hat{b}_{GR-JLN}$ takes a much simpler form. It is given by

$$\hat{b}_{GR-JLN} = \left\{ \frac{4^{\alpha} \beta^{5/2} \Gamma(\alpha + 1/2) \Gamma(\alpha)}{4\sqrt{\pi} \Gamma(2\alpha)} \right\}^{2/9} n^{-2/9}.$$

Moreover, the gamma-referenced smoothing parameter for $\hat{f}_{MG}(x)$ is defined as

$$\begin{aligned} \hat{b}_{GR} &= \arg\min_b AWMISE\left\{\hat{f}_{MG}(x)\right\} \\ &= \arg\min_b \frac{b^2}{4} \int_0^\infty x^2 \left\{g''(x)\right\}^2 w(x)\, dx + \frac{n^{-1} b^{-1/2}}{2\sqrt{\pi}} \int_0^\infty \frac{g(x)}{\sqrt{x}} w(x)\, dx, \end{aligned}$$

where the weighting function $w(x)$ is chosen as $w(x) = x^3$ to ensure finiteness of integrals. It follows that $\hat{b}_{GR}$ can be expressed as

$$\hat{b}_{GR} = \left\{ \frac{4^{\alpha} \beta^{5/2} \Gamma(\alpha + 5/2) \Gamma(\alpha)}{8\sqrt{\pi} C(\alpha) \Gamma(2\alpha)} \right\}^{2/5} n^{-2/5},$$

where

$$
\begin{aligned}
C(\alpha) \;=\; & \frac{1}{4}(\alpha-2)^2(\alpha-1)^2-(\alpha-2)(\alpha-1)^2(\alpha)+\frac{1}{2}(\alpha-1)(3\alpha-4)(\alpha)\left(\alpha+\frac{1}{2}\right) \\
& -(\alpha-1)(\alpha)\left(\alpha+\frac{1}{2}\right)(\alpha+1)+\frac{1}{4}(\alpha)\left(\alpha+\frac{1}{2}\right)(\alpha+1)\left(\alpha+\frac{3}{2}\right). \quad \blacksquare
\end{aligned}
$$

References

[1] Abadie, A. (2003): "Semiparametric Instrumental Variable Estimation of Treatment Response Models," *Journal of Econometrics*, 113, 231 - 263.

[2] Abramson, I. S. (1982): "On Bandwidth Variation in Kernel Estimates - A Square Root Law," *Annals of Statistics*, 10, 1217 -1223.

[3] Ahn, D. H., and B. Gao (1999): "A Parametric Nonlinear Model of Term Structure Dynamics," *Review of Financial Studies*, 12, 721 - 762.

[4] Bouezmarni, T., and J. V. K. Rombouts (2008): "Density and Hazard Rate Estimation for Censored and α-Mixing Data Using Gamma Kernels," *Journal of Nonparametric Statistics*, 20, 627 - 643.

[5] Bouezmarni, T., and O. Scaillet (2005): "Consistency of Asymmetric Kernel Density Estimators and Smoothed Histograms with Application to Income Data," *Econometric Theory*, 21, 390 - 412.

[6] Buch-Larsen, T., J. P. Nielsen, M. Guillén, and C. Bollancé (2005): "Kernel Density Estimation for Heavy-Tailed Distributions Using the Champernowne Transformation," *Statistics*, 39, 503 - 518.

[7] Chen, S. X. (1999): "Beta Kernel Estimators for Density Functions," *Computational Statistics & Data Analysis*, 31, 131 - 145.

[8] Chen, S. X. (2000): "Probability Density Function Estimation Using Gamma Kernels," *Annals of the Institute of Statistical Mathematics*, 52, 471 - 480.

[9] Cowell, F. A., F. H. G. Ferreira, and J. A. Litchfield (1998): "Income Distribution in Brazil 1981-1990: Parametric and Non-Parametric Approaches," *Journal of Income Distribution*, 8, 63 - 76.

[10] Cox, J. C., J. E. Ingersoll, Jr., and S. A. Ross (1985): "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53, 385 - 408.

[11] Fé, D. (2010): "An Application of Local Linear Regression with Asymmetric Kernels to Regression Discontinuity Designs," Economics Discussion Paper Series EDP-1016, University of Manchester.

[12] Gospodinov, N., and M. Hirukawa (2012): "Nonparametric Estimation of Scalar Diffusion Models of Interest Rates Using Asymmetric Kernels," *Journal of Empirical Finance*, forthcoming.

[13] Gouriéroux, C., and A. Monfort (2006): "(Non)Consistency of the Beta Kernel Estimator for Recovery Rate Distribution," CREST Discussion Paper n° 2006-31.

[14] Gustafsson, J., M. Hagmann, J. P. Nielsen, and O. Scaillet (2009): "Local Transformation Kernel Density Estimation of Loss Distributions," *Journal of Business & Economic Statistics*, 27, 161 - 175.

[15] Hagmann, M., and O. Scaillet (2007): "Local Multiplicative Bias Correction for Asymmetric Kernel Density Estimators," *Journal of Econometrics*, 141, 213 - 249.

[16] Hall, P., and B. Y. Park (2002): "New Methods for Bias Correction at Endpoints and Boundaries," *Annals of Statistics*, 30, 1460 - 1479.

[17] Hill, B. M. (1975): "A Simple General Approach to Inference About the Tail of a Distribution," *Annals of Statistics*, 3, 1163 - 1174.

[18] Hirukawa, M. (2006): "A Modified Nonparametric Prewhitened Covariance Estimator," *Journal of Time Series Analysis*, 27, 441 - 476.

[19] Hirukawa, M. (2010): "Nonparametric Multiplicative Bias Correction for Kernel-Type Density Estimation on the Unit Interval," *Computational Statistics & Data Analysis*, 54, 473 - 495.

[20] Hjort, N. L., and M. C. Jones (1996): "Locally Parametric Nonparametric Density Estimation," *Annals of Statistics*, 24, 1619 - 1647.

[21] Jin, X., and J. Kawczak (2003): "Birnbaum-Saunders and Lognormal Kernel Estimators for Modelling Durations in High Frequency Financial Data," *Annals of Economics and Finance*, 4, 103 - 124.

[22] Jones, M. C. (1993): "Simple Boundary Correction for Kernel Density Estimation," *Statistics and Computing*, 3, 135 - 146.

[23] Jones, M. C., and P. J. Foster (1993): "Generalized Jackknifing and Higher Order Kernels," *Journal of Nonparametric Statistics*, 3, 81 - 94.

[24] Jones, M. C., and P. J. Foster (1996): "A Simple Nonnegative Boundary Correction Method for Kernel Density Estimation," *Statistica Sinica*, 6, 1005 - 1013.

[25] Jones, M. C., and D. A. Henderson (2007): "Kernel-Type Density Estimation on the Unit Interval," *Biometrika*, 24, 977 - 984.

[26] Jones, M. C., O. Linton, and J. P. Nielsen (1995): "A Simple Bias Reduction Method for Density Estimation," *Biometrika*, 82, 327 - 338.

[27] Jones, M. C., and D. F. Signorini (1997): "A Comparison of Higher-Order Bias Kernel Density Estimators," *Journal of the American Statistical Association*, 92, 1063 - 1073.

[28] Koshkin, G. M. (1988): "Improved Non-Negative Kernel Estimate of a Density," *Theory of Probability and Its Applications*, 33, 759 - 764.

[29] Kristensen, D. (2010): "Nonparametric Filtering of the Realized Spot Volatility: A Kernel-Based Approach," *Econometric Theory*, 26, 60 - 93.

[30] Linton, O., and J. P. Nielsen (1994): "A Multiplicative Bias Reduction Method for Nonparametric Regression," *Statistics & Probability Letters*, 19, 181 - 187.

[31] Müller, H.-G. (1991): “Smooth Optimum Kernel Estimators near Endpoints,” *Biometrika*, 78, 521 - 530.

[32] Nielsen, J. P. (1998): “Multiplicative Bias Correction in Kernel Hazard Estimation,” *Scandinavian Journal of Statistics*, 25, 541 - 553.

[33] Nielsen, J. P., and C. Tanggaard (2001): “Boundary and Bias Correction in Kernel Hazard Estimation,” *Scandinavian Journal of Statistics*, 28, 675 - 698.

[34] Renault, O., and O. Scaillet (2004): “On the Way to Recovery: A Nonparametric Bias Free Estimation of Recovery Rate Densities,” *Journal of Banking & Finance*, 28, 2915 - 2931.

[35] Rudemo, M. (1991): “Comment on “Transformations in Density Estimation” by M. P. Wand, J. S. Marron and D. Ruppert,” *Journal of the American Statistical Association*, 86, 353 - 354.

[36] Scaillet, O. (2004): “Density Estimation Using Inverse and Reciprocal Inverse Gaussian Kernels,” *Journal of Nonparametric Statistics*, 16, 217 - 226.

[37] Stuetzle, W., and Y. Mittal (1979): “Some Comments on the Asymptotic Behavior of Robust Smoothers,” in T. Gasser and M. Rosenblatt (eds.), *Smoothing Techniques for Curve Estimation: Proceedings of a Workshop Held in Heidelberg, April 2 - 4, 1979.* Berlin: Springer-Verlag, 191 - 195.

[38] Terrell, G. R., and D. W. Scott (1980): “On Improving Convergence Rates for Nonnegative Kernel Density Estimators,” *Annals of Statistics*, 8, 1160 - 1163.

[39] Terrell, G. R., and D. W. Scott (1992): “Variable Kernel Density Estimation,” *Annals of Statistics*, 20, 1236 - 1265.

[40] Wooldridge, J. M. (2001): *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

[41] Xiao, Z., and O. Linton (2002): “A Nonparametric Prewhitened Covariance Estimator,” *Journal of Time Series Analysis*, 23, 215 - 250.

[42] Zhang, S., R. J. Karunamuni, and M. C. Jones (1999): “An Improved Estimator of the Density Function at the Boundary,” *Journal of the American Statistical Association*, 94, 1231 - 1241.

Table 1: Functional Forms of Asymmetric Kernels

Kernel (j)	Functional Form $(u \geq 0)$
G	$K_{G(x/b+1,b)}(u) = u^{x/b}\exp(-u/b) / \left\{b^{x/b+1}\Gamma(x/b+1)\right\}$
MG	$K_{MG(\rho_b(x),b)}(u) = u^{\rho_b(x)-1}\exp(-u/b) / \left[b^{\rho_b(x)}\Gamma\{\rho_b(x)\}\right]$, where $\rho_b(x) = \begin{cases} x/b & \text{for } x \geq 2b \\ (1/4)(x/b)^2 + 1 & \text{for } x \in [0, 2b) \end{cases}$
IG	$K_{IG(x,1/b)}(u) = \frac{1}{\sqrt{2\pi b u^3}}\exp\left\{-\frac{1}{2bx}\left(\frac{u}{x} - 2 + \frac{x}{u}\right)\right\}$
RIG	$K_{RIG(1/(x-b),1/b)}(u) = \frac{1}{\sqrt{2\pi b u}}\exp\left\{-\frac{x-b}{2b}\left(\frac{u}{x-b} - 2 + \frac{x-b}{u}\right)\right\}$
LN	$K_{LN(\log x,b)}(u) = \frac{1}{u\sqrt{2\pi b}}\exp\left\{-\frac{(\log u - \log x)^2}{2b}\right\}$
BS	$K_{BS(b^{1/2},x)}(u) = \frac{1}{2x\sqrt{2\pi b}}\left\{\left(\frac{x}{u}\right)^{1/2} + \left(\frac{x}{u}\right)^{3/2}\right\}\exp\left\{-\frac{1}{2b}\left(\frac{u}{x} - 2 + \frac{x}{u}\right)\right\}$

Table 2: Explicit Forms of $a_{1,j}(x,f)$ and $a_{2,j}(x,f)$

Kernel (j)	$a_{1,j}(x,f)$	$a_{2,j}(x,f)$
G	$f'(x) + \frac{x}{2}f''(x)$	$f''(x) + \frac{5}{6}xf'''(x) + \frac{x^2}{8}f''''(x)$
MG	$\begin{cases} \frac{x}{2}f''(x) & \text{for } x \geq 2b \\ \xi_b(x) f'(x) & \text{for } x \in [0, 2b) \end{cases}$, where $\xi_b(x) = \rho_b(x) - \frac{x}{b} = O(1)$	$\begin{cases} \frac{x}{3}f'''(x) + \frac{x^2}{8}f''''(x) & \text{for } x \geq 2b \\ \frac{1}{2}\left\{\xi_b^2(x) + \xi_b(x) + \frac{x}{b}\right\}f''(x) & \text{for } x \in [0, 2b) \end{cases}$
IG	$\frac{x^3}{2}f''(x)$	$\frac{x^5}{2}f'''(x) + \frac{x^6}{8}f''''(x)$
RIG	$\frac{x}{2}f''(x)$	$\frac{1}{2}f''(x) + \frac{x}{2}f'''(x) + \frac{x^2}{8}f''''(x)$
LN	$\frac{x}{2}f'(x) + \frac{x^2}{2}f''(x)$	$\frac{x}{8}f'(x) + \frac{7}{8}x^2f''(x) + \frac{3}{4}x^3f'''(x) + \frac{x^4}{8}f''''(x)$
BS	$\frac{x}{2}f'(x) + \frac{x^2}{2}f''(x)$	$\frac{3}{4}x^2f''(x) + \frac{3}{4}x^3f'''(x) + \frac{x^4}{8}f''''(x)$

Table 3: True Distributions Considered in Monte Carlo Simulations

Distribution	Density Function $f(x)$, $x \geq 0$
1. Gamma	$x^{\alpha-1}\exp(-x/\beta) / \{\beta^\alpha\Gamma(\alpha)\}$, $(\alpha,\beta) = (1.5, 1)$.
2. Weibull	$(\alpha/\beta)(x/\beta)^{\alpha-1}\exp\{-(x/\beta)^\alpha\}$, $(\alpha,\beta) = (1.5, 1.5)$.
3. Half-Normal	$\frac{2}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$, $(\mu,\sigma) = (0, 1.5)$.
4. Half-Logisitic	$\left(\frac{2}{s}\right)\exp\left\{-\left(\frac{x-\mu}{s}\right)\right\} / \left[1 + \exp\left\{-\left(\frac{x-\mu}{s}\right)\right\}\right]^2$, $(\mu, s) = (0, 1)$.
5. Log-Normal	$\frac{1}{x\sqrt{2\pi}\sigma}\exp\left\{-\frac{(\log x - \mu)^2}{2\sigma^2}\right\}$, $(\mu,\sigma) = (0, 0.75)$.
6. Pareto	$\rho\lambda^\rho / (x+\lambda)^{\rho+1}$, $(\lambda,\rho) = (1, 2)$.
7. Burr	$\alpha\beta x^{\alpha-1} / (1 + x^\alpha)^{\beta+1}$, $(\alpha,\beta) = (1.5, 2.5)$.
8. Generalized Gamma	$\gamma x^{\alpha-1}\exp\{-(x/\beta)^\gamma\} / \beta^\alpha\Gamma(\alpha/\gamma)$, $(\alpha,\beta,\gamma) = (5, 2, 2.5)$.
9. Log-Normal and Pareto Mixture 1	$p\left[\frac{1}{x\sqrt{2\pi}\sigma}\exp\left\{-\frac{(\log x-\mu)^2}{2\sigma^2}\right\}\right] + (1-p)\left\{\frac{\rho\lambda^\rho}{(x+\lambda)^{\rho+1}}\right\}$, $(p,\mu,\sigma,\lambda,\rho) = (0.7, 0, 0.5, 1, 2)$.
10. Log-Normal and Pareto Mixture 2	$p\left[\frac{1}{x\sqrt{2\pi}\sigma}\exp\left\{-\frac{(\log x-\mu)^2}{2\sigma^2}\right\}\right] + (1-p)\left\{\frac{\rho\lambda^\rho}{(x+\lambda)^{\rho+1}}\right\}$, $(p,\mu,\sigma,\lambda,\rho) = (0.3, 0, 0.5, 1, 2)$.

Table 4: Averages and Standard Deviations of Performance Measures (Distributions 1 and 2)

Estimator	n = 100			n = 200			n = 500		
	IAD	RISE	RWISE	IAD	RISE	RWISE	IAD	RISE	RWISE
				Distribution #1: Gamma					
G	0.02581	0.03816	0.04787	0.02015	0.03058	0.03668	0.01464	0.02307	0.02624
	(0.00861)	(0.01227)	(0.01737)	(0.00652)	(0.00966)	(0.01281)	(0.00419)	(0.00662)	(0.00823)
MG	0.02651	0.03858	0.05253	0.02044	0.03026	0.03974	0.01460	0.02195	0.02827
	(0.00927)	(0.01192)	(0.02090)	(0.00685)	(0.00911)	(0.01516)	(0.00430)	(0.00596)	(0.00960)
LMBC-G	0.02348	0.03422	0.04386	0.01771	0.02600	0.03289	0.01234	0.01837	0.02300
	(0.00846)	(0.01301)	(0.01666)	(0.00620)	(0.00948)	(0.01228)	(0.00388)	(0.00630)	(0.00774)
LTBC-C	0.02877	0.04478	0.05020	0.02202	0.03509	0.03776	0.01564	0.02593	0.02645
	(0.00954)	(0.01417)	(0.01917)	(0.00681)	(0.01017)	(0.01336)	(0.00425)	(0.00654)	(0.00846)
TS-MBC-G	0.02622	0.03940	0.04893	0.01981	0.03085	0.03579	0.01391	0.02302	0.02428
	(0.00904)	(0.01318)	(0.01879)	(0.00675)	(0.01018)	(0.01348)	(0.00428)	(0.00695)	(0.00855)
TS-MBC-MG	0.02708	0.04084	0.05058	0.02039	0.03167	0.03697	0.01426	0.02307	0.02511
	(0.00920)	(0.01271)	(0.01959)	(0.00671)	(0.00927)	(0.01400)	(0.00421)	(0.00587)	(0.00884)
JLN-MBC-G	0.02322	0.03647	0.04108	0.01779	0.02948	0.03040	0.01267	0.02278	0.02071
	(0.00862)	(0.01302)	(0.01646)	(0.00648)	(0.00995)	(0.01213)	(0.00423)	(0.00694)	(0.00772)
JLN-MBC-MG	0.02403	0.03845	0.04127	0.01857	0.03154	0.03039	0.01345	0.02467	0.02070
	(0.00825)	(0.01043)	(0.01663)	(0.00605)	(0.00754)	(0.01215)	(0.00380)	(0.00472)	(0.00760)
				Distribution #2: Weibull					
G	0.02678	0.03985	0.04856	0.02082	0.03137	0.03795	0.01503	0.02288	0.02766
	(0.00931)	(0.01384)	(0.01819)	(0.00646)	(0.00996)	(0.01295)	(0.00451)	(0.00682)	(0.00910)
MG	0.02755	0.03964	0.05342	0.02102	0.03047	0.04102	0.01491	0.02160	0.02944
	(0.00949)	(0.01307)	(0.02065)	(0.00677)	(0.00932)	(0.01511)	(0.00448)	(0.00610)	(0.00997)
LMBC-G	0.02484	0.03697	0.04420	0.01876	0.02805	0.03391	0.01313	0.01969	0.02420
	(0.00919)	(0.01421)	(0.01692)	(0.00610)	(0.00970)	(0.01155)	(0.00411)	(0.00636)	(0.00825)
LTBC-C	0.02955	0.04596	0.05182	0.02264	0.03579	0.03999	0.01605	0.02599	0.02825
	(0.01036)	(0.01569)	(0.02007)	(0.00716)	(0.01145)	(0.01409)	(0.00447)	(0.00690)	(0.00909)
TS-MBC-G	0.02761	0.04143	0.05005	0.02059	0.03129	0.03740	0.01422	0.02188	0.02577
	(0.00986)	(0.01502)	(0.01946)	(0.00666)	(0.01039)	(0.01349)	(0.00456)	(0.00691)	(0.00918)
TS-MBC-MG	0.02834	0.04209	0.05229	0.02098	0.03145	0.03882	0.01441	0.02179	0.02662
	(0.00985)	(0.01440)	(0.02022)	(0.00675)	(0.00983)	(0.01416)	(0.00443)	(0.00620)	(0.00939)
JLN-MBC-G	0.02387	0.03650	0.04137	0.01830	0.02861	0.03161	0.01302	0.02102	0.02239
	(0.00983)	(0.01490)	(0.01738)	(0.00671)	(0.01050)	(0.01199)	(0.00474)	(0.00718)	(0.00877)
JLN-MBC-MG	0.02404	0.03846	0.03970	0.01827	0.03019	0.02968	0.01291	0.02225	0.02061
	(0.00863)	(0.01203)	(0.01654)	(0.00584)	(0.00811)	(0.01126)	(0.00382)	(0.00497)	(0.00793)

(continued)

Table 4: *Continued* (Distributions 3 and 4)

Estimator	n = 100 IAD	RISE	RWISE	n = 200 IAD	RISE	RWISE	n = 500 IAD	RISE	RWISE
				Distribution #3: Half-Normal					
G	0.02480	0.03708	0.04472	0.01946	0.02914	0.03530	0.01387	0.02115	0.02534
	(0.00874)	(0.01305)	(0.01696)	(0.00627)	(0.00964)	(0.01205)	(0.00408)	(0.00623)	(0.00827)
MG	0.02669	0.03971	0.05043	0.02062	0.03085	0.03915	0.01458	0.02225	0.02780
	(0.00937)	(0.01491)	(0.01951)	(0.00682)	(0.01084)	(0.01400)	(0.00441)	(0.00700)	(0.00959)
LMBC-G	0.02603	0.04441	0.04184	0.01983	0.03372	0.03258	0.01379	0.02383	0.02310
	(0.00945)	(0.01690)	(0.01576)	(0.00659)	(0.01177)	(0.01134)	(0.00409)	(0.00730)	(0.00759)
LTBC-C	0.02744	0.04260	0.04651	0.02091	0.03267	0.03541	0.01485	0.02355	0.02554
	(0.01028)	(0.01679)	(0.01786)	(0.00748)	(0.01302)	(0.01240)	(0.00566)	(0.01110)	(0.00888)
TS-MBC-G	0.02511	0.03809	0.04434	0.01891	0.02851	0.03366	0.01276	0.01951	0.02304
	(0.00974)	(0.01547)	(0.01781)	(0.00690)	(0.01110)	(0.01238)	(0.00437)	(0.00687)	(0.00849)
TS-MBC-MG	0.02768	0.04308	0.04782	0.02061	0.03192	0.03610	0.01392	0.02183	0.02473
	(0.01031)	(0.01773)	(0.01895)	(0.00736)	(0.01262)	(0.01341)	(0.00464)	(0.00787)	(0.00919)
JLN-MBC-G	0.02208	0.03267	0.03936	0.01686	0.02489	0.03051	0.01167	0.01729	0.02157
	(0.00942)	(0.01434)	(0.01711)	(0.00693)	(0.01059)	(0.01275)	(0.00430)	(0.00657)	(0.00826)
JLN-MBC-MG	0.02162	0.03251	0.03660	0.01629	0.02465	0.02782	0.01117	0.01722	0.01913
	(0.00937)	(0.01481)	(0.01587)	(0.00670)	(0.01080)	(0.01134)	(0.00418)	(0.00685)	(0.00733)
				Distribution #4: Half-Logistic					
G	0.02397	0.03452	0.04481	0.01874	0.02710	0.03530	0.01339	0.01969	0.02539
	(0.00844)	(0.01216)	(0.01677)	(0.00620)	(0.00911)	(0.01235)	(0.00409)	(0.00597)	(0.00851)
MG	0.02532	0.03644	0.05041	0.01959	0.02834	0.03913	0.01390	0.02049	0.02785
	(0.00916)	(0.01400)	(0.01972)	(0.00675)	(0.01028)	(0.01455)	(0.00441)	(0.00670)	(0.01005)
LMBC-G	0.02457	0.04005	0.04194	0.01865	0.03025	0.03251	0.01291	0.02130	0.02292
	(0.00912)	(0.01596)	(0.01599)	(0.00642)	(0.01114)	(0.01149)	(0.00402)	(0.00687)	(0.00787)
LTBC-C	0.02685	0.04093	0.04633	0.02043	0.03118	0.03571	0.01412	0.02195	0.02493
	(0.01018)	(0.01645)	(0.01793)	(0.00704)	(0.01145)	(0.01277)	(0.00459)	(0.00773)	(0.00884)
TS-MBC-G	0.02387	0.03453	0.04457	0.01792	0.02584	0.03368	0.01209	0.01770	0.02296
	(0.00944)	(0.01429)	(0.01786)	(0.00681)	(0.01040)	(0.01275)	(0.00436)	(0.00650)	(0.00880)
TS-MBC-MG	0.02618	0.03931	0.04731	0.01945	0.02916	0.03556	0.01315	0.01998	0.02422
	(0.01003)	(0.01683)	(0.01887)	(0.00727)	(0.01212)	(0.01366)	(0.00460)	(0.00759)	(0.00942)
JLN-MBC-G	0.02093	0.03019	0.03826	0.01588	0.02292	0.02923	0.01086	0.01587	0.02013
	(0.00912)	(0.01354)	(0.01654)	(0.00663)	(0.01010)	(0.01182)	(0.00419)	(0.00635)	(0.00781)
JLN-MBC-MG	0.02110	0.03027	0.03790	0.01589	0.02295	0.02873	0.01088	0.01605	0.01964
	(0.00917)	(0.01394)	(0.01624)	(0.00663)	(0.01026)	(0.01169)	(0.00423)	(0.00658)	(0.00783)

(continued)

Table 4: *Continued* (Distributions 5 and 6)

Estimator	n = 100			n = 200			n = 500		
	IAD	RISE	RWISE	IAD	RISE	RWISE	IAD	RISE	RWISE
				Distribution #5: Log-Normal					
G	0.03241	0.05338	0.04884	0.02537	0.04239	0.03763	0.01832	0.03127	0.02675
	(0.00966)	(0.01530)	(0.01658)	(0.00682)	(0.01121)	(0.01165)	(0.00484)	(0.00805)	(0.00794)
MG	0.03106	0.04690	0.05380	0.02373	0.03623	0.04115	0.01695	0.02625	0.02916
	(0.01100)	(0.01652)	(0.01996)	(0.00781)	(0.01206)	(0.01410)	(0.00537)	(0.00868)	(0.00948)
LMBC-G	0.02534	0.03937	0.04252	0.01947	0.03074	0.03215	0.01401	0.02261	0.02250
	(0.00900)	(0.01403)	(0.01656)	(0.00614)	(0.00957)	(0.01144)	(0.00442)	(0.00719)	(0.00786)
LTBC-C	0.02853	0.04721	0.04542	0.02163	0.03626	0.03420	0.01551	0.02641	0.02410
	(0.01065)	(0.01764)	(0.01837)	(0.00681)	(0.01170)	(0.01196)	(0.00428)	(0.00738)	(0.00749)
TS-MBC-G	0.03126	0.05085	0.04983	0.02392	0.04030	0.03679	0.01715	0.03044	0.02480
	(0.00977)	(0.01557)	(0.01822)	(0.00671)	(0.01104)	(0.01230)	(0.00465)	(0.00776)	(0.00809)
TS-MBC-MG	0.02949	0.04496	0.05132	0.02200	0.03393	0.03793	0.01503	0.02342	0.02559
	(0.01008)	(0.01584)	(0.01903)	(0.00705)	(0.01109)	(0.01283)	(0.00474)	(0.00790)	(0.00835)
JLN-MBC-G	0.03121	0.05656	0.04158	0.02491	0.04682	0.03147	0.01857	0.03661	0.02201
	(0.00866)	(0.01397)	(0.01441)	(0.00619)	(0.01018)	(0.01012)	(0.00434)	(0.00698)	(0.00686)
JLN-MBC-MG	0.03482	0.06361	0.04563	0.02809	0.05344	0.03517	0.02108	0.04207	0.02498
	(0.00899)	(0.01327)	(0.01575)	(0.00653)	(0.00997)	(0.01092)	(0.00462)	(0.00751)	(0.00736)
				Distribution #6: Pareto					
G	0.03696	0.09475	0.03309	0.02954	0.07801	0.02542	0.02173	0.05865	0.01828
	(0.01109)	(0.03207)	(0.01076)	(0.00807)	(0.02363)	(0.00785)	(0.00522)	(0.01539)	(0.00527)
MG	0.04168	0.08360	0.04994	0.03193	0.06493	0.03754	0.02217	0.04523	0.02593
	(0.01352)	(0.02722)	(0.01791)	(0.00892)	(0.01938)	(0.01140)	(0.00479)	(0.01149)	(0.00631)
LMBC-G	0.02515	0.07257	0.02849	0.01993	0.06229	0.02121	0.01529	0.05188	0.01496
	(0.00880)	(0.03237)	(0.01114)	(0.00669)	(0.02597)	(0.00800)	(0.00508)	(0.02065)	(0.00546)
LTBC-C	0.02604	0.06495	0.02601	0.01969	0.04940	0.01987	0.01395	0.03484	0.01420
	(0.00999)	(0.02871)	(0.01059)	(0.00662)	(0.01927)	(0.00745)	(0.00707)	(0.01671)	(0.00679)
TS-MBC-G	0.03173	0.07971	0.03069	0.02543	0.06773	0.02275	0.01921	0.05443	0.01584
	(0.01225)	(0.03562)	(0.01146)	(0.00912)	(0.02749)	(0.00815)	(0.00602)	(0.01829)	(0.00535)
TS-MBC-MG	0.05906	0.11707	0.06751	0.05020	0.10405	0.05380	0.04019	0.08893	0.03954
	(0.02082)	(0.03067)	(0.03707)	(0.01541)	(0.02371)	(0.02462)	(0.00924)	(0.01636)	(0.01288)
JLN-MBC-G	0.03053	0.07616	0.02850	0.02441	0.06376	0.02138	0.01831	0.04988	0.01510
	(0.00915)	(0.02600)	(0.01046)	(0.00680)	(0.01967)	(0.00731)	(0.00446)	(0.01330)	(0.00494)
JLN-MBC-MG	0.04037	0.10432	0.03655	0.03325	0.09053	0.02823	0.02544	0.07398	0.02011
	(0.00977)	(0.02352)	(0.01086)	(0.00713)	(0.01784)	(0.00734)	(0.00442)	(0.01149)	(0.00451)

(continued)

Table 4: *Continued* (Distributions 7 and 8)

Estimator	n = 100			n = 200			n = 500		
	IAD	RISE	RWISE	IAD	RISE	RWISE	IAD	RISE	RWISE
	Distribution #7: Burr								
G	0.02962	0.06226	0.03435	0.02309	0.04983	0.02657	0.01668	0.03685	0.01906
	(0.00909)	(0.02063)	(0.01080)	(0.00656)	(0.01557)	(0.00773)	(0.00463)	(0.01097)	(0.00531)
MG	0.03053	0.06159	0.03867	0.02342	0.04778	0.02949	0.01649	0.03360	0.02077
	(0.01007)	(0.02013)	(0.01288)	(0.00731)	(0.01475)	(0.00932)	(0.00482)	(0.00987)	(0.00601)
LMBC-G	0.02567	0.05274	0.03060	0.01920	0.03949	0.02315	0.01328	0.02719	0.01626
	(0.00903)	(0.02188)	(0.01027)	(0.00609)	(0.01503)	(0.00717)	(0.00416)	(0.00971)	(0.00499)
LTBC-C	0.02984	0.06927	0.03104	0.02246	0.05283	0.02344	0.01573	0.03752	0.01653
	(0.01032)	(0.02484)	(0.01171)	(0.00682)	(0.01663)	(0.00805)	(0.00438)	(0.01039)	(0.00525)
TS-MBC-G	0.02899	0.06260	0.03413	0.02200	0.04951	0.02535	0.01553	0.03695	0.01730
	(0.00951)	(0.02243)	(0.01134)	(0.00652)	(0.01629)	(0.00785)	(0.00465)	(0.01156)	(0.00534)
TS-MBC-MG	0.02967	0.06355	0.03543	0.02235	0.04894	0.02631	0.01560	0.03518	0.01789
	(0.00984)	(0.02114)	(0.01196)	(0.00688)	(0.01466)	(0.00832)	(0.00454)	(0.00930)	(0.00545)
JLN-MBC-G	0.02798	0.06759	0.02809	0.02246	0.05801	0.02149	0.01704	0.04786	0.01536
	(0.00898)	(0.02098)	(0.00965)	(0.00642)	(0.01607)	(0.00670)	(0.00461)	(0.01150)	(0.00475)
JLN-MBC-MG	0.02765	0.06627	0.02792	0.02230	0.05736	0.02115	0.01696	0.04745	0.01488
	(0.00905)	(0.01481)	(0.01076)	(0.00662)	(0.01026)	(0.00771)	(0.00441)	(0.00679)	(0.00502)
	Distribution #8: Generalized Gamma								
G	0.02875	0.03846	0.09656	0.02226	0.03010	0.07596	0.01584	0.02175	0.05447
	(0.01022)	(0.01360)	(0.03499)	(0.00722)	(0.00982)	(0.02597)	(0.00467)	(0.00654)	(0.01697)
MG	0.02857	0.03792	0.09891	0.02217	0.02972	0.07775	0.01575	0.02141	0.05541
	(0.01014)	(0.01329)	(0.03695)	(0.00725)	(0.00966)	(0.02746)	(0.00469)	(0.00647)	(0.01782)
LMBC-G	0.02314	0.03110	0.07623	0.01755	0.02367	0.05814	0.01242	0.01688	0.04148
	(0.00960)	(0.01307)	(0.03207)	(0.00674)	(0.00923)	(0.02373)	(0.00443)	(0.00619)	(0.01567)
LTBC-C	0.03100	0.04358	0.10924	0.02398	0.03344	0.08365	0.01752	0.02418	0.05973
	(0.01038)	(0.01573)	(0.03970)	(0.00726)	(0.01121)	(0.02827)	(0.00526)	(0.00843)	(0.01993)
TS-MBC-G	0.02925	0.04048	0.10045	0.02174	0.03015	0.07524	0.01486	0.02074	0.05137
	(0.00996)	(0.01426)	(0.03649)	(0.00694)	(0.01006)	(0.02629)	(0.00449)	(0.00658)	(0.01695)
TS-MBC-MG	0.02946	0.04065	0.10183	0.02190	0.03027	0.07628	0.01496	0.02078	0.05198
	(0.00995)	(0.01416)	(0.03692)	(0.00698)	(0.01001)	(0.02670)	(0.00453)	(0.00657)	(0.01727)
JLN-MBC-G	0.02655	0.03656	0.09626	0.02084	0.02892	0.07698	0.01516	0.02121	0.05743
	(0.00916)	(0.01174)	(0.03084)	(0.00650)	(0.00850)	(0.02366)	(0.00427)	(0.00572)	(0.01587)
JLN-MBC-MG	0.02451	0.03376	0.08451	0.01906	0.02652	0.06676	0.01346	0.01903	0.04819
	(0.00803)	(0.01056)	(0.02825)	(0.00544)	(0.00727)	(0.02036)	(0.00349)	(0.00492)	(0.01369)

(continued)

Table 4: *Continued* (Distributions 9 and 10)

Estimator	n = 100			n = 200			n = 500		
	IAD	RISE	RWISE	IAD	RISE	RWISE	IAD	RISE	RWISE
	Distribution #9: Log-Normal and Pareto Mixture 1								
G	0.03273	0.05493	0.04978	0.02607	0.04448	0.03868	0.01891	0.03297	0.02751
	(0.01038)	(0.01681)	(0.01766)	(0.00760)	(0.01239)	(0.01276)	(0.00533)	(0.00901)	(0.00855)
MG	0.03537	0.05600	0.05766	0.02749	0.04423	0.04380	0.01978	0.03232	0.03122
	(0.01237)	(0.01790)	(0.02441)	(0.00893)	(0.01346)	(0.01648)	(0.00581)	(0.00913)	(0.01030)
LMBC-G	0.03328	0.06093	0.04671	0.02685	0.04991	0.03695	0.02002	0.03783	0.02727
	(0.01140)	(0.02112)	(0.01650)	(0.00871)	(0.01555)	(0.01286)	(0.00615)	(0.01105)	(0.00896)
LTBC-C	0.03090	0.05722	0.04487	0.02391	0.04620	0.03344	0.01746	0.03567	0.02362
	(0.00972)	(0.01691)	(0.01659)	(0.00705)	(0.01245)	(0.01206)	(0.00444)	(0.00839)	(0.00757)
TS-MBC-G	0.03104	0.05286	0.04583	0.02422	0.04263	0.03410	0.01772	0.03244	0.02368
	(0.01042)	(0.01708)	(0.01773)	(0.00768)	(0.01268)	(0.01269)	(0.00514)	(0.00879)	(0.00815)
TS-MBC-MG	0.03383	0.05579	0.05325	0.02566	0.04319	0.03909	0.01837	0.03131	0.02766
	(0.01372)	(0.01991)	(0.03012)	(0.00944)	(0.01442)	(0.01771)	(0.00593)	(0.00944)	(0.01049)
JLN-MBC-G	0.03527	0.06194	0.04853	0.02959	0.05328	0.03941	0.02295	0.04296	0.02937
	(0.00969)	(0.01597)	(0.01479)	(0.00705)	(0.01155)	(0.01082)	(0.00477)	(0.00800)	(0.00740)
JLN-MBC-MG	0.03411	0.05945	0.04931	0.02798	0.05017	0.03921	0.02178	0.04032	0.02968
	(0.01075)	(0.01732)	(0.01826)	(0.00801)	(0.01301)	(0.01341)	(0.00524)	(0.00851)	(0.00864)
	Distribution #10: Log-Normal and Pareto Mixture 2								
G	0.03035	0.06028	0.03903	0.02370	0.04917	0.02978	0.01731	0.03698	0.02168
	(0.01012)	(0.01911)	(0.01361)	(0.00693)	(0.01391)	(0.01028)	(0.00438)	(0.00973)	(0.00692)
MG	0.03179	0.05814	0.04909	0.02465	0.04576	0.03715	0.01776	0.03323	0.02613
	(0.00951)	(0.01718)	(0.02032)	(0.00634)	(0.01277)	(0.01326)	(0.00418)	(0.00889)	(0.00815)
LMBC-G	0.02630	0.05372	0.03569	0.02016	0.04180	0.02725	0.01504	0.03072	0.02023
	(0.00941)	(0.02555)	(0.01273)	(0.00677)	(0.01884)	(0.00958)	(0.00442)	(0.01170)	(0.00667)
LTBC-C	0.02809	0.05777	0.03439	0.02109	0.04409	0.02572	0.01512	0.03192	0.01841
	(0.01008)	(0.02282)	(0.01344)	(0.00704)	(0.01654)	(0.00955)	(0.00466)	(0.01093)	(0.00654)
TS-MBC-G	0.02691	0.05156	0.03683	0.02074	0.04218	0.02737	0.01574	0.03360	0.01978
	(0.01020)	(0.01898)	(0.01428)	(0.00668)	(0.01328)	(0.01039)	(0.00407)	(0.00896)	(0.00679)
TS-MBC-MG	0.03684	0.07109	0.04938	0.02996	0.06054	0.03705	0.02337	0.04991	0.02660
	(0.01412)	(0.02206)	(0.03264)	(0.00898)	(0.01555)	(0.01737)	(0.00528)	(0.01103)	(0.00834)
JLN-MBC-G	0.02603	0.05059	0.03504	0.02097	0.04268	0.02719	0.01669	0.03516	0.02042
	(0.00727)	(0.01314)	(0.01240)	(0.00491)	(0.00960)	(0.00923)	(0.00332)	(0.00683)	(0.00633)
JLN-MBC-MG	0.02824	0.06236	0.03374	0.02415	0.05516	0.02671	0.02040	0.04744	0.02101
	(0.00670)	(0.01166)	(0.01234)	(0.00460)	(0.00788)	(0.00877)	(0.00335)	(0.00533)	(0.00579)

Note : Numbers in parentheses are simulation standard deviations.

Table 5: Descriptive Statistics of the U.S. Income Data

Descriptive statistics:		Quantiles:		Inequality measures:	
Sample size	9,275	5%	12,837	Coefficient of variation	0.61
Mean	39,255	10%	15,300	Gini coefficient	0.32
Standard deviation	24,090	25%	21,660	Theil index	0.16
Skewness	1.60	50%	33,288	Mean log deviation	0.17
Kurtosis	6.68	75%	50,163		
Minimum	10,008	90%	70,878		
Maximum	199,041	95%	88,074		

Table 6: Descriptive Statistics of the Brazilian Income Data

Descriptive statistics:		Quantiles:		Inequality measures:	
Sample size	71,523	5%	4,800	Coefficient of variation	1.74
Mean	52,183	10%	6,056	Gini coefficient	0.59
Standard deviation	90,661	25%	12,166	Theil index	0.68
Skewness	11	50%	26,142	Mean log deviation	0.66
Kurtosis	319	75%	57,000		
Minimum	2	90%	116,130		
Maximum	5,011,000	95%	183,000		

Figure 1: Shapes of Asymmetric Kernels When $b = 0.4$

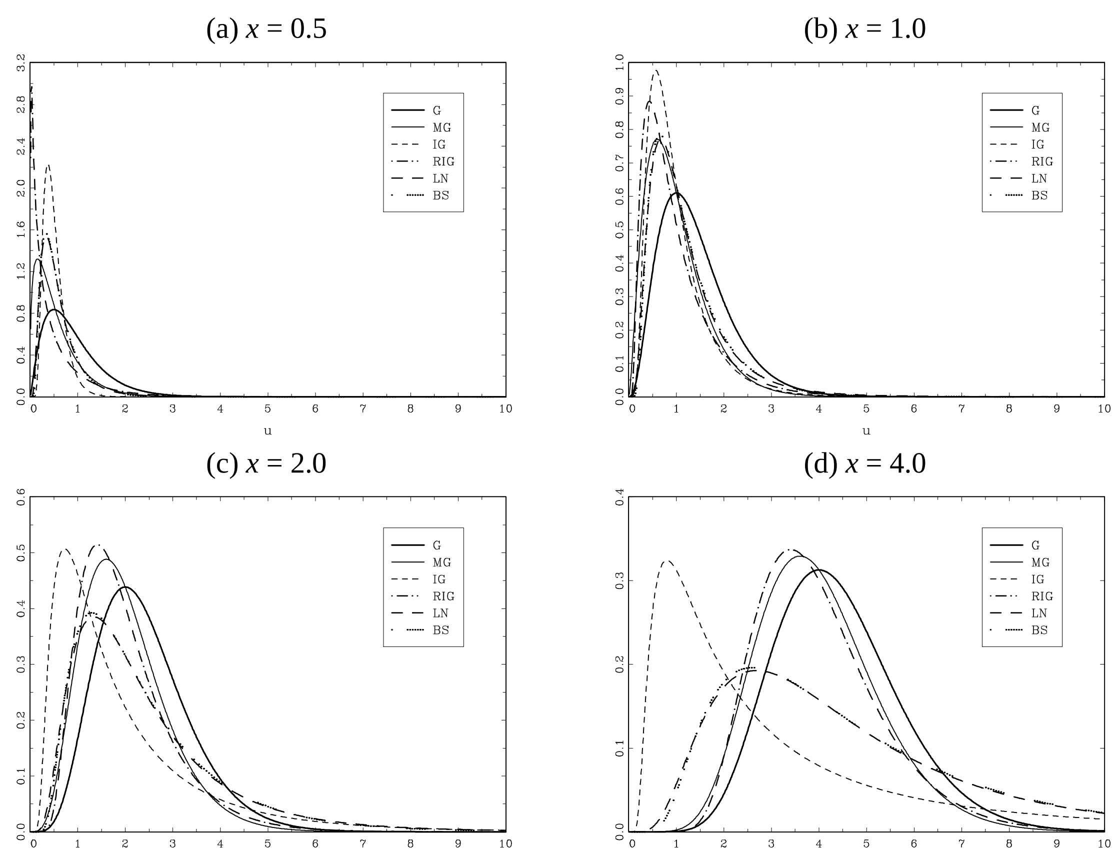
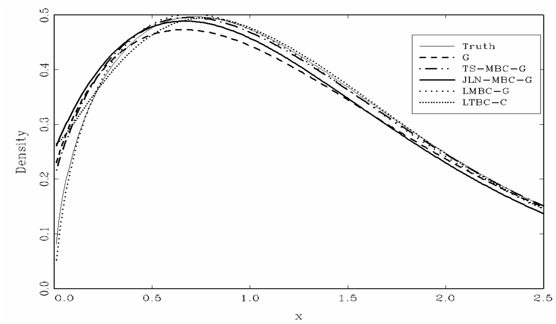


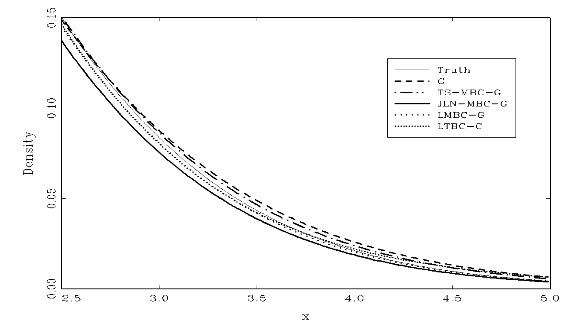
Figure 2: Average Plots of Density Estimates for Selected Distributions ($n = 100$)

(a) Weibull [Distribution 2]:

<< near the origin >>

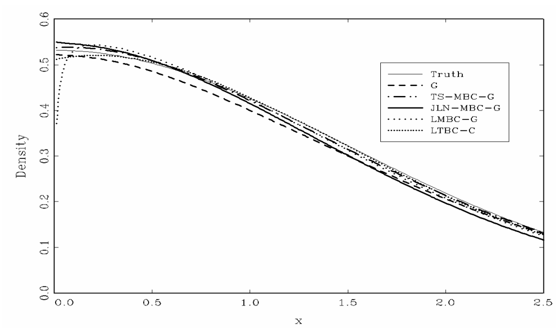


<< on right tail >>

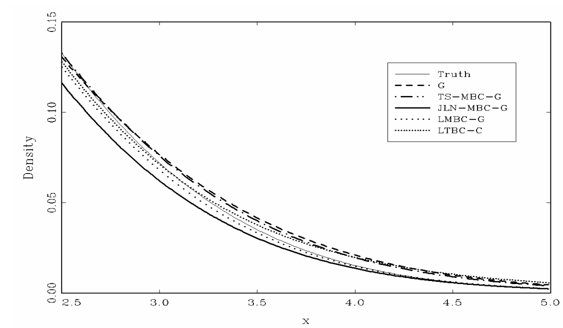


(b) Half-Normal [Distribution 3]:

<< near the origin >>

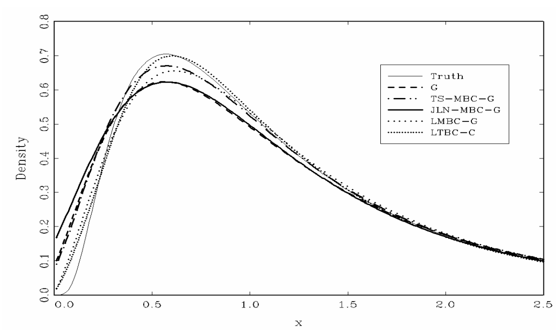


<< on right tail >>

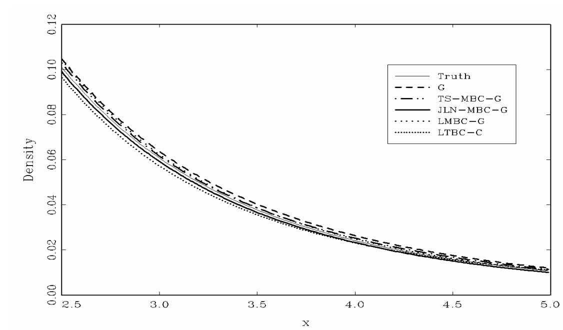


(c) Log-Normal [Distribution 5]:

<< near the origin >>

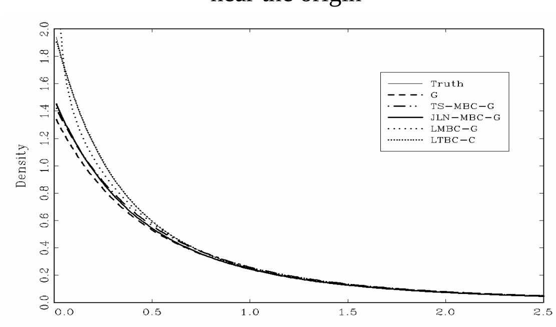


<< on right tail >>



(d) Pareto [Distribution 6]:

<< near the origin >>



<< on right tail >>

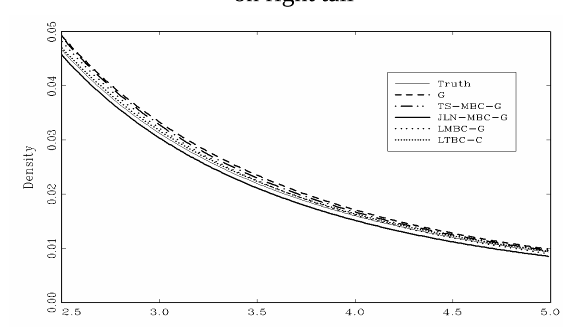
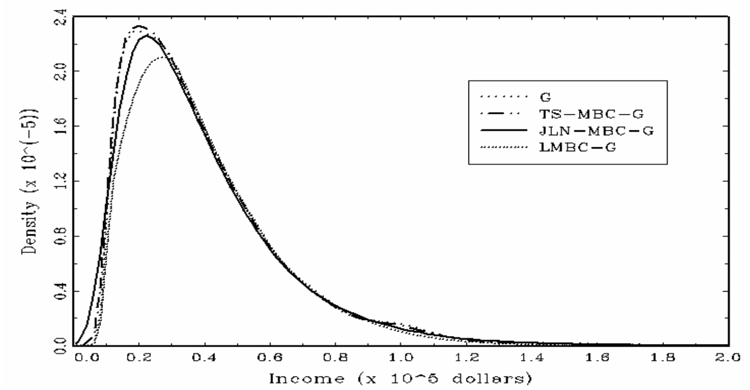
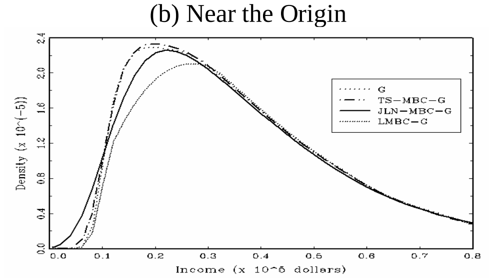


Figure 3: U.S. Income Distribution

(a) Overall



(b) Near the Origin



(c) On Right Tail

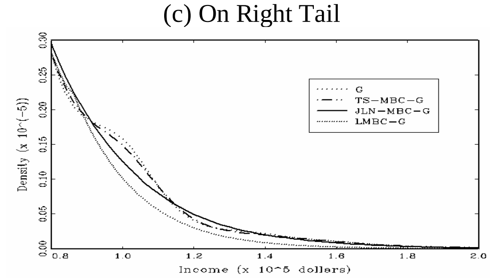
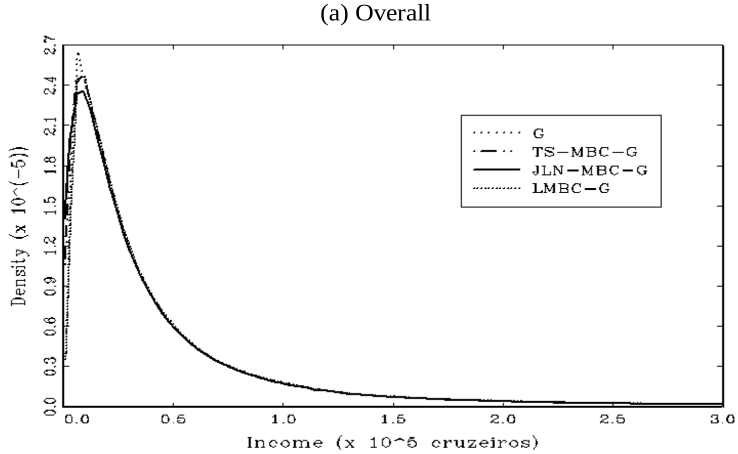
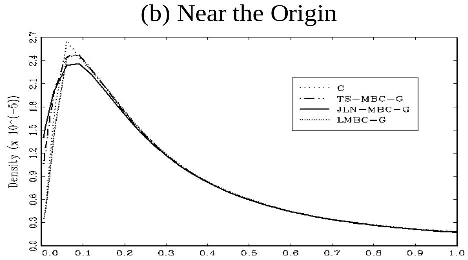


Figure 4: Brazilian Income Distribution

(a) Overall



(b) Near the Origin



(c) On Right Tail