Family of Generalized Gamma Kernels: A Unified Approach to the Asymptotics on Asymmetric Kernels*

Masayuki Hirukawa† Setsunan University

Mari Sakudo‡ Development Bank of Japan

February 2013
Very Preliminary

Abstract

Unlike symmetric kernels, exploiting the asymptotics on asymmetric kernels has relied on kernel-specific arguments. Toward a unified approach to their asymptotics, this paper proposes a generic form of asymmetric kernels that consists of a set of common conditions. The generic kernel, called a family of Generalized Gamma kernels, is built on the Generalized Gamma density function, and incorporates the Modified Gamma kernel as a special case. As other special cases, two new kernels, namely, the Weibull and Nakagami-m kernels, are also proposed. The density estimator using Generalized Gamma kernels is shown to preserve the appealing properties that the Gamma and Modified Gamma kernels possess. Furthermore, this paper investigates three extensions of the density estimation including multiplicative bias correction.

Keywords: Asymmetric kernel; bias reduction; boundary correction; density estimation; Generalized Gamma kernels; Modified Gamma kernel; Nakagami-m kernel; Weibull kernel.
JEL Classification Codes: C13; C14.
MSC 2000 Codes: 62G7; 62G20.

---

*We would like to thank Bruce Hansen, Yoshihiko Nishiyama, Katsumi Shimotsu, and participants at Summer Workshop on Economic Theory 2012 and seminars at Kyoto University and the University of Tokyo for constructive comments and suggestions. The first author gratefully acknowledges financial support from Japan Society of the Promotion of Science (grant number 23530259). The views expressed herein and those of the authors do not necessarily reflect the views of the Development Bank of Japan.

†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, 9-7, Otemachi 1-chome, Chiyoda-ku, Tokyo 100-8178, Japan; phone (+81)3-3244-1241; fax (+81)3-3270-7084; e-mail: mari-sakudo@gmail.com.

1 Introduction

This paper is concerned with kernel-smoothed probability density estimation for the distributions having support on $[0, \infty)$. Although our main focus is on economics and finance, our proposal may be found applicable to other subjects. Researchers and policy makers are often interested in the distributions of incomes, wages, short-term interest rates, trading volumes of financial assets, and insurance claims. Their distributions are empirically characterized by two stylized facts, namely, (i) a natural boundary at the origin and (ii) the concentration of observations near the boundary and a long tail with sparse data. There are two typical approaches to estimating them. One is to fit parametric models to the original data (e.g. Cowell, Ferreira and Litchfield, 1998), and the other is to apply nonparametric kernel methods to the log-transformed data (e.g. DiNardo, Fortin and Lemieux, 1996). However, the former is subject to the possibility of imposing misspecified models, whereas the latter makes it difficult to imagine the shape of the distribution in the original scale.

This motivates us to adopt kernel density estimation in the original scale. Nevertheless, to accommodate the stylized facts, two distinct modifications are required for standard smoothing techniques with symmetric kernels. For (i), boundary correction methods should be employed. There is indeed huge literature on the methods; see, for instance, Section 3 of Karunamuni and Alberts (2005) for a concise review. For (ii), 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. A remedy for this issue is variable bandwidth methods. Recently, asymmetric kernels with support on $[0, \infty)$ (e.g. Chen, 2000; Jin and Kawczak, 2003; Scaillet,

2004) have emerged as a viable alternative to boundary correction methods. Several papers (e.g. Bouezmarni and Scaillet, 2005; Hagmann and Scaillet, 2007; Bouezmarni and Rombouts, 2008, 2010; Fé, 2012; Gospodinov and Hirukawa, 2012; Malec and Schienle, 2012) provide empirical evidence that the kernels tend to work well for the distributions having the stylized facts because of their property as a combination of a boundary correction device and adaptive smoothing with effect similar to variable bandwidth methods.

A cumbersome aspect of asymmetric kernels, however, is that the approaches to exploiting asymptotic properties of asymmetric kernel estimators have been kernel-specific and thus diversified. In contrast, symmetric kernels are built on a set of common conditions, and asymptotic properties of the kernel estimators can be delivered through manipulating the conditions. Then, this paper attempts to create a 'mold' or a generic form of asymmetric kernels that consists of a set of common conditions, so that asymptoptic properties of estimators using the kernels can be implied directly by the conditions.

Toward such a unified approach, we must also take it into account that while an apparent shape restriction (= symmetry about the origin) is imposed on symmetric kernels, there are no rules on shapes of asymmetric kernels; actually, any shapes are admissible as long as they have support on $[0, \infty)$. Therefore, our pursuit for the generic form starts with choosing a functional form of the kernel, or equivalently, the distribution that generates the kernel. A closer look at the literature reveals that in almost all cases the Gamma and Modified Gamma kernels by Chen (2000) are reported to exhibit superior finite-sample performance. Judging from the good finite-sample performance and mathematical tractability of the Gamma function, it is reasonable to find a candidate of the distribution among 'close-cousins' of Gamma distributions. Then, the generic form is built on the density function of the Generalized Gamma distribution (Stacy, 1962). The generic kernel, called a family of Generalized Gamma kernels, incorporates the Modified Gamma kernel as a special case. As two other special cases, we also propose the Weibull and Nakagami-m kernels, which are generated from the Weibull and Nakagami-m (Nakagami, 1943, 1960) distributions, respectively.

As is usual with symmetric kernels, convergence properties of the probability density estimator using Generalized Gamma kernels can be exploited by manipulating the common conditions. The density estimator is shown to maintain the same attractive properties as other asymmetric kernel density estimators have. First, by construction, it is free of boundary bias and always generates a nonnegative density estimate everywhere. Second, Generalized Gamma kernels vary their shapes according to the design point at which smoothing is made; in other words, the amount of smoothing changes in a locally adaptive manner. Third, when best implemented, the estimator attains Stone's (1980) optimal convergence rate in the mean integrated squared error within the class of nonnegative kernel estimators. Fourth, the variance of the estimator tends to decrease as the design point moves away from the boundary. This property is particularly advantageous to estimating distributions that have a long tail with sparse data, such as income distributions.

Furthermore, this paper investigates three extensions of density estimation using Generalized Gamma kernels. While the first extension, applicability of two multiplicative bias correction techniques studied in Hirukawa and Sakudo (2012), relies exclusively on independent data, the remaining two extensions are based on weakly dependent data. After the Generalized Gamma density estimator is shown to admit the same first-order bias and variance approximations as those for random samples, it is demonstrated that the density estimator is consistent even when the density becomes unbounded at the boundary. It is known that the Gamma and Modified Gamma kernels have all these appealing properties. Our aim is to demonstrate that Generalized Gamma kernels are endowed with the properties in general.

The remainder of this paper is organized as follows. Section 2 provides the definition of a family of Generalized Gamma kernels and introduces three special cases. Convergence properties of the Generalized Gamma density estimator are also provided. Section 3 investigates three extensions of Generalized Gamma density estimation. Section 4 conducts Monte Carlo simulations to examine finite-sample properties of the density estimator. Section 5 summarizes the main results of the paper. All proofs are given in the Appendix.

This paper adopts the following notational conventions. $\Gamma(a) = \int_0^\infty y^{a-1} \exp(-y)\, dy$ $(a > 0)$ is the gamma function; $\mathbf{1}\{\cdot\}$ signifies an indicator function; $\lfloor \cdot \rfloor$ denotes the integer part; and $(s.o.)$ is the smaller-order term. The expression '$X \overset{d}{=} Y$' reads "A random variable X obeys the distribution Y." The expression '$X_n \sim Y_n$' is used whenever $X_n/Y_n \to 1$ as $n \to \infty$. Lastly, following the convention in the literature on asymmetric kernels, in order to describe different asymptotic properties of the estimator across positions of the design point $x > 0$, we denote by "interior x" and "boundary x" a design point x that satisfies $x/b \to \infty$ and $x/b \to \kappa$ for some $0 < \kappa < \infty$ as $n \to \infty$, respectively.

2 A Family of Generalized Gamma Kernels

2.1 Definition

Let Y be distributed by the Generalized Gamma distribution $GG(\alpha, \beta, \gamma)$ due to Stacy (1962). Then, Y has the probability density function ("pdf")

$$p(y; \alpha, \beta, \gamma) = \frac{\gamma y^{\alpha-1} \exp\{-(y/\beta)^{\gamma}\}}{\beta^{\alpha}\Gamma(\alpha/\gamma)} \mathbf{1}\{y \geq 0\}. \tag{1}$$

It is known that the m-th (uncentered) moment of Y is given by

$$E(Y^m) = \beta^m \frac{\Gamma\{(\alpha+m)/\gamma\}}{\Gamma(\alpha/\gamma)}. \tag{2}$$

We now provide the definition of a family of Generalized Gamma kernels that consists of several common conditions. Before proceeding, it could be beneficial to relate the definition to probability density estimation. Below we present a set of regularity conditions for asymmetric kernel density estimation.

Assumption 1. The random sample $\{X_i\}_{i=1}^{n}$ is drawn from a univariate distribution with a probability density function f having support on $[0, \infty)$.

Assumption 2. f is twice continuously differentiable.

Assumption 3. The smoothing parameter $b\,(= b_n > 0)$ satisfies $b + (nb)^{-1} \to 0$ as $n \to \infty$.

To generate a kernel from the pdf (1), we allow (α, β, γ) to be a function of the design point $x > 0$ and the smoothing parameter b, as in Chen (2000), Jin and Kawczak (2003) and Scaillet (2004). To put it another way, whenever we refer to the triplet (α, β, γ), it should be interpreted as a short-handed notation of $(\alpha, \beta, \gamma) = (\alpha_b(x), \beta_b(x), \gamma_b(x))$ unless otherwise noted. Let the crude Generalized Gamma kernel be $K_{GG0}(u;x,b) := p(u;\alpha,\beta,\gamma)$. Given the random sample $\{X_i\}_{i=1}^n$, we have the density estimator $\hat{f}_{GG0}(x) = (1/n)\sum_{i=1}^n K_{GG0}(X_i;x,b)$.

Let $\vartheta_x \stackrel{d}{=} GG(\alpha,\beta,\gamma)$. Under Assumptions 2-3, a second-order Taylor expansion of $E\left\{\hat{f}_{GG0}(x)\right\}$ around $\vartheta_x = x$ yields $E\left\{\hat{f}_{GG0}(x)\right\} = f(x) + E(\vartheta_x - x) f'(x) + (1/2) E(\vartheta_x - x)^2 f''(x) + (s.o.)$. It follows that unless $E(\vartheta_x) = x$ exactly (at least for interior x), the leading bias of $\hat{f}_{GG0}$ would contain the term involving f', which is less desirable. Although there are numerous choices of (α,β,γ) that can achieve $E(\vartheta_x) = x$, we adopt the simplest resolution that we set $\beta = x$ for interior x and employ the pdf of $GG(\alpha, \beta\Gamma(\alpha/\gamma)/\Gamma\{(\alpha+1)/\gamma\}, \gamma)$ (which can be obtained by the change of variable $Z := [\Gamma(\alpha/\gamma)/\Gamma\{(\alpha+1)/\gamma\}]\, Y$ in (1)) as the kernel. In the end, we reach the following definition of a family of Generalized Gamma kernels.

Definition 1. *Let $(\alpha,\beta,\gamma) = (\alpha_b(x), \beta_b(x), \gamma_b(x)) \in \mathbb{R}_+^3$ be a continuous function of the design point x and the smoothing parameter b. For such (α,β,γ), consider the pdf of $GG(\alpha, \beta\Gamma(\alpha/\gamma)/\Gamma\{(\alpha+1)/\gamma\}, \gamma)$, i.e.*

$$K_{GG}(u;x,b) = \frac{\gamma u^{\alpha-1} \exp\left[-\left\{\frac{u}{\beta\Gamma\left(\frac{\alpha}{\gamma}\right)/\Gamma\left(\frac{\alpha+1}{\gamma}\right)}\right\}^{\gamma}\right]}{\left\{\beta\Gamma\left(\frac{\alpha}{\gamma}\right)/\Gamma\left(\frac{\alpha+1}{\gamma}\right)\right\}^{\alpha}\Gamma\left(\frac{\alpha}{\gamma}\right)} \mathbf{1}\{u \geq 0\}. \tag{3}$$

This pdf is said to be a family of Generalized Gamma ("GG") kernels if it satisfies each of the following conditions:

Condition 1. $\beta = \begin{cases} x & \text{for } x \geq C_1 b \\ \varphi_b(x) & \text{for } x \in [0, C_1 b) \end{cases}$*, where $0 < C_1 < \infty$ is some constant, the function $\varphi_b(x)$ satisfies $C_2 b \leq \varphi_b(x) \leq C_3 b$ for some constants $0 < C_2 \leq C_3 < \infty$, and the connection between x and $\varphi_b(x)$ at $x = C_1 b$ is smooth.*

Condition 2. *$\alpha, \gamma \geq 1$, and for $x \in [0, C_1 b)$, α satisfies $1 \leq \alpha \leq C_4$ for some constant $1 \leq C_4 < \infty$. Moreover, connections of α and γ at $x = C_1 b$, if any, are smooth.*

Condition 3. $M_b(x) := \frac{\Gamma\left(\frac{\alpha}{\gamma}\right)\Gamma\left(\frac{\alpha+2}{\gamma}\right)}{\left\{\Gamma\left(\frac{\alpha+1}{\gamma}\right)\right\}^2} = \begin{cases} 1+\psi(x)\,b+o(b) & \text{for } x \geq C_1 b \\ O(1) & \text{for } x \in [0, C_1 b) \end{cases}$, *where the function* $\psi(x)$ *is continuous in* $x \in \mathbb{R}_+$.

Condition 4. $H_b(x) := \frac{\Gamma\left(\frac{\alpha}{\gamma}\right)\Gamma\left(\frac{2\alpha}{\gamma}\right)}{2^{1/\gamma}\Gamma\left(\frac{\alpha+1}{\gamma}\right)\Gamma\left(\frac{2\alpha-1}{\gamma}\right)} = \begin{cases} 1+o(1) & \text{for interior } x \\ O(1) & \text{for boundary } x \end{cases}$.

Condition 5. $A_{b,\nu}(x) := \left\{\frac{\gamma\Gamma\left(\frac{\alpha+1}{\gamma}\right)}{\beta}\right\}^{\nu-1} \frac{\Gamma\left\{\frac{\nu(\alpha-1)+1}{\gamma}\right\}}{\nu^{\frac{\nu(\alpha-1)+1}{\gamma}}\left\{\Gamma\left(\frac{\alpha}{\gamma}\right)\right\}^{2\nu-1}} \sim \begin{cases} V_I(\nu)\,(xb)^{\frac{1-\nu}{2}} & \text{for interior } x \\ V_B(\nu)\,b^{1-\nu} & \text{for boundary } x \end{cases}$, $\nu \in \mathbb{R}_+$, *where constants* $0 < V_I(\nu), V_B(\nu) < \infty$ *depend only on* ν.

Conditions 1-2 form a legitimate kernel from the GG pdf. It follows from $\varphi_b(x) = O(b)$ uniformly over the $O(b)$ region in Condition 1 that the kernel is well-defined in the vicinity of the origin. While $\alpha \geq 1$ in Condition 2 also ensures boundedness of the kernel near the origin, $\gamma \geq 1$ controls the tail-behavior of the kernel and establishes an exponential rate of the tail decay. The condition also allows for each of α and γ to be a piecewise function of (x, b) like β, where the connection is made at $x = C_1 b$; the common connection point simplifies asymptotic analyses substantially. In addition, smooth connection requirements in Condition 1-2 are inspired by the construction of the Modified Gamma kernel; see p.473 of Chen (2000) for details. Moreover, Condition 3 and Conditions 4-5 are the requirements for valid approximations to the bias and variance of the GG density estimator

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

respectively. Although all the common conditions appear to be high-level ones, it is not hard to find a few special cases that satisfy them. Examples of GG kernels are provided shortly.

2.2 Convergence Properties of Probability Density Estimators Using GG Kernels

2.2.1 Local Property

Before providing examples of GG kernels, readers may wonder whether the kernels are a truly legitimate one, i.e. they can yield a consistent density estimator. To answer this question, we begin this section with presenting the theorem on approximations to the bias and variance of $\hat{f}_{GG}(x)$.

Theorem 1. *Under Assumptions 1-3, the bias of* $\hat{f}_{GG}(x)$ *can be approximated by* $Bias\left\{\hat{f}_{GG}(x)\right\} \sim B_1(x,f)\,b$*, where*

$$B_1(x,f) = \begin{cases} \frac{1}{2}x^2\psi(x)\,f''(x) & \text{for } x \geq C_1 b \\ \xi_b(x)\,f'(x) & \text{for } x \in [0, C_1 b) \end{cases},$$

and $\xi_b(x) = \{\varphi_b(x) - x\}/b = O(1)$. *On the other hand, the variance of* $\hat{f}_{GG}(x)$ *can be approximated by*

$$Var\left\{\hat{f}_{GG}(x)\right\} \sim \begin{cases} \frac{1}{nb^{1/2}}V_I(2)\frac{f(x)}{\sqrt{x}} & \text{for interior } x \\ \frac{1}{nb}V_B(2)\,f(x) & \text{for boundary } x \end{cases}.$$

The theorem states that Conditions 1-5 do lead to familiar properties of asymmetric kernel density estimators. By construction, $\hat{f}_{GG}(x)$ is free of boundary bias and non-negative everywhere. The bias of $\hat{f}_{GG}(x)$ is $O(b)$, and its variance is $O\left\{\left(nb^{1/2}\right)^{-1}\right\}$ for interior x and $O\left\{(nb)^{-1}\right\}$ for boundary x.

2.2.2 Global Property

If $\int_0^\infty \{x^2\psi(x)\,f''(x)\}^2\,dx$ and $\int_0^\infty \{f(x)/\sqrt{x}\}\,dx$ are both finite, then applying the trimming argument in Chen (2000, p.476) approximates the mean integrated squared error ("MISE") of $\hat{f}_{GG}(x)$ as

$$MISE\left\{\hat{f}_{GG}(x)\right\} \sim \frac{b^2}{4}\int_0^\infty \{x^2\psi(x)\,f''(x)\}^2\,dx + \frac{V_I(2)}{nb^{1/2}}\int_0^\infty \frac{f(x)}{\sqrt{x}}dx. \qquad (4)$$

The smoothing parameter value that minimizes the right-hand side of (4) is

$$b_{GG}^{**} = \left[\frac{V_I(2) \int_0^\infty \{f(x)/\sqrt{x}\}\, dx}{\int_0^\infty \{x^2 \psi(x) f''(x)\}^2\, dx} \right]^{2/5} n^{-2/5}.$$

Therefore, when best implemented, the approximation to the MISE becomes

$$MISE^{**}\left\{\hat{f}_{GG}(x)\right\} \sim \frac{5}{4} \left[\int_0^\infty \{x^2 \psi(x) f''(x)\}^2\, dx \right]^{1/5} \left[V_I(2) \int_0^\infty \frac{f(x)}{\sqrt{x}} dx \right]^{4/5} n^{-4/5}.$$

Note that $O\left(n^{-4/5}\right)$ is the optimal convergence rate of the MISE within the class of nonnegative kernel estimators in Stone's (1980) sense. Furthermore, the variance coefficient $V_I(2) f(x)/\sqrt{x}$ decreases as the design point x moves away from the boundary. This property is particularly advantageous to estimating the distributions that have a long tail with sparse data, such as those of the economic and financial variables introduced in Section 1.

2.2.3 A Note on Implementation

Choosing the smoothing parameter b is an important practical issue. Below we consider a very simple, Silverman's (1986) rule-of-thumb type choice rule. Although the rule basically relies on the MISE (4) as the criterion, we make a few modifications. First, we specify $\psi(x) = C_\psi x^{-1}$ for some constant $0 < C_\psi < \infty$. Indeed, all examples of GG Kernels in the next section satisfy this specification; see the proof of Theorem 2 in Appendix for details. Second, the unknown f is replaced by a known reference density. For simplicity, we choose the pdf of $G(\mu, \omega)$, i.e. $g(x) = x^{\mu-1} \exp(-x/\omega) \mathbf{1}\{x \geq 0\} / \{\omega^\mu \Gamma(\mu)\}$, as the reference. Third, the criterion is modified to the asymptotic weighted mean integrated squared error ("AWMISE")

$$AWMISE\left\{\hat{f}_{GG}(x)\right\} := \frac{b^2}{4} \int_0^\infty \{x^2 \psi(x) g''(x)\}^2 w(x)\, dx + \frac{V_I(2)}{nb^{1/2}} \int_0^\infty \frac{g(x)}{\sqrt{x}} w(x)\, dx,$$

where the weighting function $w(x) \geq 0$ must be chosen to ensure finiteness of two integrals. Given the specifications of $\psi(x)$ and $g(x)$, it turns out that $w(x) = x^3$

fulfills this requirement. Then, the AWMISE is simplified to

$$AWMISE\left\{\hat{f}_{GG}(x)\right\} = b^2\left\{\frac{C_\psi^2 C_\mu \Gamma(2\mu)}{4^\mu \Gamma^2(\mu)}\right\} + \frac{1}{nb^{1/2}}\left\{\frac{V_I(2)\,\omega^{5/2}\Gamma(\mu+5/2)}{\Gamma(\mu)}\right\},$$

where

$$\begin{aligned} C_\mu &= \frac{1}{4}(\mu-2)^2(\mu-1)^2 - (\mu-2)(\mu-1)^2(\mu) + \frac{1}{2}(3\mu-4)(\mu-1)(\mu)\left(\mu+\frac{1}{2}\right) \\ &\quad -(\mu-1)(\mu)\left(\mu+\frac{1}{2}\right)(\mu+1) + \frac{1}{4}(\mu)\left(\mu+\frac{1}{2}\right)(\mu+1)\left(\mu+\frac{3}{2}\right). \end{aligned}$$

As a consequence, the AWMISE-optimal smoothing parameter is given by

$$b_{GG}^{\dagger\dagger} = \left\{\frac{4^{\mu-1}V_I(2)\,\omega^{5/2}\Gamma(\mu)\Gamma(\mu+5/2)}{C_\psi^2 C_\mu \Gamma(2\mu)}\right\}^{2/5} n^{-2/5}.$$

In practice, parameters (μ,ω) need to be replaced by their method-of-moments estimates $(\hat{\mu},\hat{\omega})$.

It would be possible to choose b in a more sophisticated manner. For instance, replacing two integrals in (4) with their nonparametric estimators using GG kernels, we could derive an analog to the solve-the-equation plug-in method by Sheather and Jones (1991). Alternatively, we may rely on the cross-validation method (e.g. Bouezmarni and Rombouts, 2010, p.250). Investigating these methods is left for our future research.

2.3 Examples of GG Kernels

This section introduces three special cases of GG kernels. It is worth emphasizing that there could be many other examples belonging to this family. As Stacy (1962, p.1187) states, for instance, functions of a standard normal variate (e.g. its positive even powers, its modulus, and all positive powers of its modulus) will generate GG kernels. The sole reason why the three kernels are listed as examples is that for each of these kernels, approximations to the gamma functions that appear in $M_b(x)$,

$H_b(x)$ and $A_{b,\nu}(x)$ in Definition 1 are readily available; see the proof of Theorem 2 in Appendix for details.

2.3.1 Examples

Modified Gamma Kernel. The Modified Gamma ("MG") kernel in Chen (2000) turns out to be an immediate example of GG kernels. Put

$$(\alpha, \beta) = \begin{cases} \left(\frac{x}{b}, x\right) & \text{for } x \geq 2b \\ \left(\frac{1}{4}\left(\frac{x}{b}\right)^2 + 1, \frac{x^2}{4b} + b\right) & \text{for } x \in [0, 2b) \end{cases}$$

and $\gamma = 1$ in (3). Then, the GG kernel collapses to

$$K_{GG}(u; x, b) = \frac{u^{\alpha-1} \exp\{-u/(\beta/\alpha)\}}{(\beta/\alpha)^{\alpha} \Gamma(\alpha)} \mathbf{1}\{u \geq 0\}, \tag{5}$$

which is the pdf of the Gamma distribution $G(\alpha, \beta/\alpha)$. Observe that $\alpha = \rho_b(x)$ in Chen (2000, p.473) and $\beta/\alpha = b$. It follows that the above kernel finally reduces to the MG kernel

$$K_{MG(\rho_b(x),b)}(u) = \frac{u^{\rho_b(x)-1} \exp(-u/b)}{b^{\rho_b(x)} \Gamma\{\rho_b(x)\}} \mathbf{1}\{u \geq 0\}.$$

Weibull Kernel. To obtain the Weibull ("W") kernel, let

$$(\alpha, \beta) = \begin{cases} \left(\sqrt{\frac{2x}{b}}, x\right) & \text{for } x \geq 2b \\ \left(\frac{1}{2}\left(\frac{x}{b}\right) + 1, \frac{x^2}{4b} + b\right) & \text{for } x \in [0, 2b) \end{cases}$$

and $\gamma = \alpha$ in (3). Then, the GG kernel becomes

$$K_{GG}(u; x, b) = \frac{\alpha u^{\alpha-1} \exp\left[-\left\{\frac{u}{\beta/\Gamma(1+1/\alpha)}\right\}^{\alpha}\right]}{\{\beta/\Gamma(1+1/\alpha)\}^{\alpha}} \mathbf{1}\{u \geq 0\}.$$

Because the right-hand side is the pdf of the Weibull distribution $W\left(\alpha, \beta/\Gamma\left(1+1/\alpha\right)\right)$, the W kernel can be defined as[1]

$$\begin{aligned}
& K_{W(\alpha,\beta/\Gamma(1+1/\alpha))}\left(u\right) \\
= \; & \frac{\alpha}{\beta/\Gamma\left(1+1/\alpha\right)}\left\{\frac{u}{\beta/\Gamma\left(1+1/\alpha\right)}\right\}^{\alpha-1}\exp\left[-\left\{\frac{u}{\beta/\Gamma\left(1+1/\alpha\right)}\right\}^{\alpha}\right]\mathbf{1}\left\{u\geq 0\right\}.
\end{aligned}$$

Nakagami-m Kernel. Use exactly the same (α, β) as for the MG kernel but put $\gamma = 2$ in (3). Then, the GG kernel reduces to

$$K_{GG}\left(u;x,b\right)=\frac{2u^{\alpha-1}\exp\left[-\left\{u/\left(\beta\Gamma\left(\frac{\alpha}{2}\right)/\Gamma\left(\frac{\alpha+1}{2}\right)\right)\right\}^{2}\right]}{\left\{\beta\Gamma\left(\frac{\alpha}{2}\right)/\Gamma\left(\frac{\alpha+1}{2}\right)\right\}^{\alpha}\Gamma\left(\frac{\alpha}{2}\right)}\mathbf{1}\left\{u\geq 0\right\},$$

which is the pdf of the Nakagami-m distribution $NM\left(\alpha/2, (\alpha/2)\left[\beta\Gamma\left(\alpha/2\right)/\Gamma\left\{\left(\alpha+1\right)/2\right\}\right]^{2}\right)$ due to Nakagami (1943, 1960).[2] This distribution is frequently applied in telecommunications engineering as the distribution that can describe signal intensity of short-wave fading. In the end, the Nakagami-m ("NM") kernel is defined as

$$\begin{aligned}
& K_{NM\left(\alpha/2,(\alpha/2)[\beta\Gamma(\alpha/2)/\Gamma\{(\alpha+1)/2\}]^{2}\right)}\left(u\right) \\
= \; & \frac{2\left(\frac{\alpha}{2}\right)^{\alpha/2}}{\left[\left(\frac{\alpha}{2}\right)\left\{\beta\Gamma\left(\frac{\alpha}{2}\right)/\Gamma\left(\frac{\alpha+1}{2}\right)\right\}^{2}\right]^{\alpha/2}\Gamma\left(\frac{\alpha}{2}\right)}u^{2\left(\frac{\alpha}{2}\right)-1}\exp\left[-\frac{\frac{\alpha}{2}}{\left(\frac{\alpha}{2}\right)\left\{\beta\Gamma\left(\frac{\alpha}{2}\right)/\Gamma\left(\frac{\alpha+1}{2}\right)\right\}^{2}}u^{2}\right]\mathbf{1}\left\{u\geq 0\right\}.
\end{aligned}$$

FIGURE 1 ABOUT HERE

[1] This definition differs from the one given in Kuruwita, Kulasekera and Padgett (2010, Table 1), who construct their Weibull kernel from a different motivation. Apparently, their definition

$$K_{W}\left(u;x,b\right)=\frac{1}{bx}\left(\frac{u}{x}\right)^{1/b-1}\exp\left\{-\left(\frac{u}{x}\right)^{1/b}\right\}\mathbf{1}\left\{u\geq 0\right\}$$

tends to be unbounded near the origin.

[2] The pdf of $NM\left(\mu,\omega\right)$ $\left(\mu\geq 1/2, \omega>0\right)$ is

$$p\left(y;\mu,\omega\right)=\frac{2\mu^{\mu}}{\omega^{\mu}\Gamma\left(\mu\right)}y^{2\mu-1}\exp\left(-\frac{\mu}{\omega}y^{2}\right)\mathbf{1}\left\{y\geq 0\right\}.$$

It is widely recognized that the Nakagami-m distribution was first proposed in Nakagami (1960). In reality, however, the distribution was originally studied as early as in Nakagami (1943).

2.3.2 Asymptotic Results on Density Estimators Using the MG, W and NM Kernels

Figure 1 plots the shapes of the MG, W and NM kernels for four different design points ($x = 0, 1, 2, 4$) at which the smoothing is performed. For reference, the Gamma ("G") kernel (Chen, 2000) is also drawn in each panel.[3] It is worth noting that for all plotted functions, the value of the smoothing parameter is fixed at $b = 0.2$. When smoothing is made at the origin (Panel (a)), the NM kernel collapses to a half-normal pdf, whereas all others reduce to an exponential pdf. As the design point moves away from the boundary (Panels (b)-(d)), shapes of each kernel vary and get flatter; in other words, each kernel changes the amount of smoothing in an adaptive manner. At the same time, the shapes get closer to a symmetric one; in particular, shapes of the NM kernel become almost symmetric for a large x, as its functional form suggests.

Convergence properties of density estimators using the MG, W and NM kernels are presented in Theorem 2 below. Obviously, the functional form of (α, β, γ) for each of the three kernels satisfies Conditions 1-2. Hence, to demonstrate Theorem 2, it suffices to check that Conditions 3-5 hold for each kernel.

Theorem 2. *Let $\hat{f}_j(x)$ be the probability density estimator using the kernel $j \in \{MG, W, NM\}$. Then, under Assumptions 1-3, the bias and variance of $\hat{f}_j(x)$ can be approximated by $Bias\left\{\hat{f}_j(x)\right\} \sim B_{1j}(x, f)\, b$ and*

$$Var\left\{\hat{f}_j(x)\right\} \sim \begin{cases} \frac{1}{nb^{1/2}} V_{I,j}(2) \frac{f(x)}{\sqrt{x}} & \text{for interior } x \\ \frac{1}{nb} V_{B,j}(2) f(x) & \text{for boundary } x \end{cases},$$

where $V_{B,j}(2)$ can be found in Appendix, and explicit forms of $B_{1j}(x, f)$ and $V_{I,j}(2)$ are as follows:

[3] This kernel does not belong to GG kernels, because it can be obtained by setting $(\alpha, \beta) = (x/b + 1, x + b)$ in (5).

j	$B_{1j}(x,f)$ $x \geq 2b$	$x \in [0,2b)$	$V_{I,j}(2)$
MG	$\frac{x}{2} f''(x)$	$\xi_b(x) f'(x)$	$\frac{1}{2\sqrt{\pi}}$
W	$\frac{\pi^2}{24} x f''(x)$	$\xi_b(x) f'(x)$	$\frac{1}{2\sqrt{2}}$
NM	$\frac{x}{4} f''(x)$	$\xi_b(x) f'(x)$	$\frac{1}{\sqrt{2\pi}}$

Moreover, $\xi_b(x) = \{(1/2)(x/b) - 1\}^2 = O(1)$ *in this case.*

It follows from Theorem 2 that the mean squared error ("MSE") of each density estimator for interior x can be approximated by

$$\begin{aligned}
MSE\left\{\hat{f}_{MG}(x)\right\} &\sim \frac{x^2}{4}\{f''(x)\}^2 b^2 + \frac{1}{nb^{1/2}}\frac{f(x)}{2\sqrt{\pi}\sqrt{x}}, \\
MSE\left\{\hat{f}_{W}(x)\right\} &\sim \left(\frac{\pi^2}{24}\right)^2 x^2\{f''(x)\}^2 b^2 + \frac{1}{nb^{1/2}}\frac{f(x)}{2\sqrt{2}\sqrt{x}}, \\
MSE\left\{\hat{f}_{NM}(x)\right\} &\sim \frac{x^2}{16}\{f''(x)\}^2 b^2 + \frac{1}{nb^{1/2}}\frac{f(x)}{\sqrt{2\pi}\sqrt{x}}.
\end{aligned}$$

The smoothing parameter values that minimize these approximations are

$$\begin{aligned}
b^*_{MG} &= \left[\frac{f(x)}{2\sqrt{\pi}\{f''(x)\}^2}\right]^{2/5} x^{-1}n^{-2/5}, \\
b^*_{W} &= \left[2^{-7/2}\left(\frac{24}{\pi^2}\right)^2\frac{f(x)}{\{f''(x)\}^2}\right]^{2/5} x^{-1}n^{-2/5}, \\
b^*_{NM} &= \left[\frac{2^{3/2}f(x)}{\sqrt{\pi}\{f''(x)\}^2}\right]^{2/5} x^{-1}n^{-2/5}.
\end{aligned}$$

Then, the optimal-MSEs of $\hat{f}_{MG}(x)$ and $\hat{f}_{NM}(x)$ have the relation

$$\begin{aligned}
MSE^*\left\{\hat{f}_{MG}(x)\right\} &\sim MSE^*\left\{\hat{f}_{NM}(x)\right\} \\
&\sim \frac{5}{4^{7/5}}\left(\frac{1}{\pi}\right)^{2/5}\{f(x)\}^{4/5}\{f''(x)\}^{2/5} n^{-4/5}, \qquad (6)
\end{aligned}$$

which is also the optimal-MSE of the density estimator using the Gaussian kernel. In other words, when best implemented, density estimators using these kernels become first-order asymptotically equivalent, and both kernels on $[0,\infty)$ are in a sense equivalent to the Gaussian kernel on $(-\infty,\infty)$. Furthermore, two MSE-optimal smoothing parameters satisfy $b^*_{NM} = 2b^*_{MG}$. Indeed, when the smoothing parameter value for the NM kernel is set equal to twice the value for the MG kernel, the shapes of these two kernels resemble each other (although they are not presented to save space). In contrast, when best implemented, the MSE of $\hat{f}_W(x)$ can be approximated by

$$MSE^* \left\{\hat{f}_W(x)\right\} \sim \frac{5}{4^{7/5}}\left(\frac{\pi^2}{24}\right)^{2/5} \{f(x)\}^{4/5}\{f''(x)\}^{2/5} n^{-4/5}. \tag{7}$$

Comparing the factors of (6) and (7) reveals that $\left(5/4^{7/5}\right)(1/\pi)^{2/5} \approx 0.454178\ldots$ and $\left(5/4^{7/5}\right)(\pi^2/24)^{2/5} \approx 0.503178\ldots$ Therefore, we can see that $\hat{f}_W(x)$ is slightly inefficient than $\hat{f}_{MG}(x)$ and $\hat{f}_{NM}(x)$ under the best-case scenario.

3 Extensions of GG Density Estimation

This section investigates three extensions of GG density estimation. After studying applicability of multiplicative bias correction techniques for independent observations, we consider two more extensions for weakly dependent observations, namely, validity of first-order approximations to the bias and variance of the GG density estimator stated in Theorem 1, and weak consistency of the density estimator when the true density is unbounded at the origin. The G and MG kernels are known to possess all these appealing properties. We can see that the properties essentially inhere in GG kernels.

3.1 Nonparametric Multiplicative Bias Correction for GG Density Estimation

We start with examining whether two classes of nonparametric multiplicative bias correction ("MBC") techniques studied in Hirukawa (2010) and Hirukawa and Sakudo (2012) for asymmetric kernel density estimators can be applicable to GG density estimators in general. The first class of MBC estimators, originally proposed by

Terrell and Scott (1980), is defined as

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

where, $\hat{f}_{GG,b}(x)$ and $\hat{f}_{GG,b/c}(x)$ signify GG density estimators using smoothing parameters b and b/c, and $c \in (0,1)$ is some predetermined constant that does not depend on the design point x. The second class of MBC estimators due to Jones, Linton and Nielsen (1995) is implied by the identity $f(x) = \hat{f}_{GG,b}(x)\left\{f(x)/\hat{f}_{GG,b}(x)\right\}$ and defined as

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

Both $\tilde{f}_{TS,GG}(x)$ and $\tilde{f}_{JLN,GG}(x)$ always generate nonnegative density estimates everywhere by construction.

To develop convergence properties of MBC estimators, we modify Assumptions 2-3 as follows. Discussions on these assumptions can be found in Hirukawa (2010) and Hirukawa and Sakudo (2012).

Assumption 2a. f has four continuous and bounded derivatives, and $f(x) > 0$ for a given design point $x > 0$.

Assumption 3a. The smoothing parameter b satisfies $b + (nb^3)^{-1} \to 0$ as $n \to \infty$.

The next theorem refers to acceleration in bias convergence via the bias correction methods. The proof is similar to the ones for Theorems 1-2 of Hirukawa and Sakudo (2012), and this it is omitted.

Theorem 3. *If Assumptions 1, 2a and 3a hold, and* $E\left\{\hat{f}_{GG}(x)\right\}$ *admits the expansion* $E\left\{\hat{f}_{GG}(x)\right\} = f(x) + B_1(x,f)\,b + B_2(x,f)\,b^2 + o(b^2)$, *where* $B_1(x,f)$ *(which can be found in Theorem 1) and* $B_2(x,f)$ *are kernel-specific functions depending on*

x and derivatives of f, then the bias and variance of $\tilde{f}_{TS,GG}(x)$ can be approximated by

$$\begin{aligned} Bias\left\{\tilde{f}_{TS,GG}(x)\right\} &\sim \frac{1}{c(1-c)}p(x)\,b^2 := \frac{1}{c(1-c)}\left[\frac{1}{2}\left\{\frac{B_1^2(x,f)}{f(x)}\right\} - B_2(x,f)\right]b^2, \\ Var\left\{\tilde{f}_{TS,GG}(x)\right\} &= \begin{cases} \left(nb^{1/2}\right)^{-1}\lambda(c)\,V_I(2)\,f(x)/\sqrt{x} + o\left\{\left(nb^{1/2}\right)^{-1}\right\} & \text{for interior } x \\ O\left\{(nb)^{-1}\right\} & \text{for boundary } x \end{cases}, \end{aligned}$$

where

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

Moreover, the bias and variance of $\tilde{f}_{JLN,GG}(x)$ can be approximated by

$$\begin{aligned} Bias\left\{\tilde{f}_{TS,GG}(x)\right\} &\sim q(x)\,b^2 := -f(x)\,B_1(x,h)\,b^2, \\ Var\left\{\tilde{f}_{JLN,GG}(x)\right\} &= \begin{cases} \left(nb^{1/2}\right)^{-1}V_I(2)\,f(x)/\sqrt{x} + o\left\{\left(nb^{1/2}\right)^{-1}\right\} & \text{for interior } x \\ O\left\{(nb)^{-1}\right\} & \text{for boundary } x \end{cases}, \end{aligned}$$

where $B_1(x,h)$ is obtained by replacing f in $B_1(x,f)$ with $h = h(x,f) := B_1(x,f)/f(x)$.

As the theorem suggests, whether the two MBC techniques may accelerate the bias convergence from $O(b)$ to $O(b^2)$ depends crucially on whether the second-order term in $Bias\left\{\hat{f}_{GG}(x)\right\}$ is $O(b^2)$. It is worth noting that Conditions 1-5 provide no guidance on the order of magnitude in the second-order bias term. For instance, as indicated in the proof of Theorem 2, the second-order term in $Bias\left\{\hat{f}_W(x)\right\}$ is $O\left(b^{3/2}\right)$. Because both MBC techniques only attain the bias convergence up to $O\left(b^{3/2}\right)$, the theorem excludes such inferior cases.

In contrast, for the MG and NM kernels, each MBC technique improves their bias convergence to $O(b^2)$. Explicit forms of $p(x)$ and $q(x)$ for the NM kernel are

$$p_{NM}(x) = \begin{cases} \frac{x^2}{32}\frac{\{f''(x)\}^2}{f(x)} - \frac{1}{8}\left\{\frac{1}{2}f'(x) + \frac{x}{3}f''(x) + \frac{x^2}{4}f'''(x)\right\} & \text{for } x \geq 2b \\ \frac{1}{2}\left[\frac{\{\xi_b(x)f'(x)\}^2}{f(x)} - \left\{\left(\xi_b(x)+\frac{x}{b}\right)^2 \frac{\Gamma\left(\frac{\xi_b(x)+x/b}{2}\right)\Gamma\left(\frac{\xi_b(x)+x/b}{2}+1\right)}{\left(\Gamma\left(\frac{\xi_b(x)+x/b+1}{2}\right)\right)^2} - 2\left(\frac{x}{b}\right)\xi_b(x) + \left(\frac{x}{b}\right)^2\right\}\right] & \text{for } x \in [0, 2b) \end{cases},$$

$$q_{NM}(x) = \begin{cases} -f(x)\frac{x}{4}\left\{\frac{x}{4}\frac{f''(x)}{f(x)}\right\}'' & \text{for } x \geq 2b \\ -f(x)\xi_b(x)\left\{\frac{\xi_b(x)f'(x)}{f(x)}\right\}'' & \text{for } x \in [0, 2b) \end{cases},$$

where $\xi_b(x)$ can be found in Theorem 2. Those for the MG kernel are available in Hirukawa and Sakudo (2012). It follows that MSEs of $\tilde{f}_{JLN,MG}(x)$ and $\tilde{f}_{JLN,NM}(x)$ for interior x can be approximated by

$$\begin{aligned} MSE\left\{\tilde{f}_{JLN,MG}(x)\right\} &\sim \left[-f(x)\frac{x}{2}\left\{\frac{x}{2}\frac{f''(x)}{f(x)}\right\}''\right]^2 b^4 + \frac{1}{nb^{1/2}}\frac{f(x)}{2\sqrt{\pi}\sqrt{x}}, \\ MSE\left\{\tilde{f}_{JLN,NM}(x)\right\} &\sim \left[-f(x)\frac{x}{4}\left\{\frac{x}{4}\frac{f''(x)}{f(x)}\right\}''\right]^2 b^4 + \frac{1}{nb^{1/2}}\frac{f(x)}{\sqrt{2\pi}\sqrt{x}} \\ &= \frac{1}{16}\left[-f(x)\frac{x}{2}\left\{\frac{x}{2}\frac{f''(x)}{f(x)}\right\}''\right]^2 b^4 + \frac{\sqrt{2}}{nb^{1/2}}\frac{f(x)}{2\sqrt{\pi}\sqrt{x}}. \end{aligned}$$

The smoothing parameter values that minimize these approximations are

$$\begin{aligned} b^*_{JLN,MG} &= \frac{1}{2^{2/3}}\left\{\frac{f(x)}{2\sqrt{\pi}\sqrt{x}}\right\}\left[f(x)\frac{x}{2}\left\{\frac{x}{2}\frac{f''(x)}{f(x)}\right\}''\right]^{-4/9} n^{-2/9}, \\ b^*_{JLN,NM} &= 2^{1/3}\left\{\frac{f(x)}{2\sqrt{\pi}\sqrt{x}}\right\}\left[f(x)\frac{x}{2}\left\{\frac{x}{2}\frac{f''(x)}{f(x)}\right\}''\right]^{-4/9} n^{-2/9}. \end{aligned}$$

Observe that the relation $b^*_{JLN,NM} = 2b^*_{JLN,MG}$ holds once again. Moreover, when best implemented, the optimal-MSEs of two estimators are first-order asymptotically equivalent, i.e.

$$\begin{aligned} MSE^*\left\{\tilde{f}_{JLN,MG}(x)\right\} &\sim MSE^*\left\{\tilde{f}_{JLN,NM}(x)\right\} \\ &\sim \frac{9}{2^{8/3}}\left[f(x)\frac{x}{2}\left\{\frac{x}{2}\frac{f''(x)}{f(x)}\right\}''\right]^{2/9}\left\{\frac{f(x)}{2\sqrt{\pi}\sqrt{x}}\right\}^{8/9} n^{-8/9}. \end{aligned}$$

3.2 Bias and Variance Approximations of GG Density Estimators Using Weakly Dependent Observations

The second extension is concerned with estimating the marginal density from positive time-series data. Examples include estimation of the distribution of important financial variables such as short-term interest rates or trading volumes, and even the baseline hazard in financial duration analysis.

To allow for weakly dependent observations in GG density estimation, we replace Assumptions 1-2 with the regularity conditions below. While similar conditions can be found, for instance, in Bouezmarni and Rombouts (2010), they assume strong mixing processes with an exponentially decaying mixing coefficient and exclusively study the MG kernel.[4] We relax the mixing condition and employ the generic kernel which encompasses the MG kernel.

Assumption 1b. $\{X_i\}$ is a nonnegative, strictly stationary and strong mixing process with the mixing coefficient $\alpha(\ell)$ of size $-(2r-2)/(r-2)$ for some $r>2$.

Assumption 2b. Let $f(\cdot)$ and $f_j(\cdot,\cdot)$ be the marginal and joint densities of X_i and (X_i, X_{i+j}), respectively. Then, f is twice continuously differentiable, and f_j is uniformly bounded.

The next theorem states that results in Theorem 1 are carried through even when positive weakly dependent observations are used.

Theorem 4. *Results in Theorem 1 still hold under Assumptions 1b, 2b and 3.*

[4] Carrasco and Chen (2002) provide the conditions that make GARCH processes stationary and β-mixing with exponential decay, whereas Chen, Hansen and Carrasco (2010) discuss the conditions that can establish β-mixing with exponential decay in scalar diffusion processes. In relation to the latter, Feller's square-root process and its inverse, which have been employed as models of short-term interest rates by Cox, Ingersoll and Ross (1985) and Ahn and Gao (1999), respectively, are examples of exponentially decaying β-mixing processes. Since β-mixing implies strong mixing, their assumption can cover many important applications in economics and finance.

3.3 Weak Consistency of GG Density Estimators for the Density Unbounded at $x = 0$

The third extension is to demonstrate that $\hat{f}_{GG}(0) \xrightarrow{p} \infty$ when the true density $f(x)$ is unbounded at $x = 0$. Often a clustering of observations near the boundary can be found in intraday trading volumes or realized volatilities, for instance. In this case, it is highly likely that the true density has a pole at $x = 0$ (e.g. Malec and Schienle, 2012). Such shapes also appear in many other applications including spectral densities of long memory processes. Provided that the assumption on the smoothing parameter is replaced by Assumption 3c, the theorem below establishes weak consistency of $\hat{f}_{GG}(0)$ to infinity when $f(x)$ has a pole at the origin.

Assumption 3c. The smoothing parameter b satisfies $b + (nb^2)^{-1} \to 0$ as $n \to \infty$.

Theorem 5. *If the true density $f(x)$ is unbounded at $x = 0$ and Assumption 1b and 3c hold, then $\hat{f}_{GG}(0) \xrightarrow{p} \infty$.*

FIGURE 2 ABOUT HERE

Bouezmarni and Scaillet (2005, p.399) find that this property is peculiar to the G and MG kernels and not shared by other asymmetric kernels (e.g. the Inverse Gaussian and Reciprocal Inverse Gaussian kernels by Scaillet, 2004), although they exclusively examine random sampling cases. To illustrate the result in this theorem, we prepare Figure 2, in which the W, NM, MG, and G density estimates based on 500 independent observations drawn from $G(0.75, 1.25)$ are plotted. The figure indicates that estimates using three examples of GG kernels, as well as the G estimate, can indeed capture the shape of the density near the boundary reasonably well.

4 Finite Sample Performance

4.1 Setup

This simulation study compares accuracy of GG density estimators $\hat{f}_j(x)$, $j \in \{MG, W, NM\}$. The G kernel, although it does not belong to GG kernels, is popularly chosen, and thus the G density estimator $\hat{f}_G(x)$ is included as a benchmark in the simulation. 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 Scaillet (2004), as the performance measure for each estimator $\bar{f}$, we compute the root integrated squared error ("RISE") $RISE\left\{\bar{f}(x)\right\} = \sqrt{\int_0^\infty \left\{\bar{f}(x) - f(x)\right\}^2 dx}$. In our report, the integral is approximated over the 500 points. The smoothing parameter b for estimator $\bar{f}$ is chosen as the minimizer of the approximation to $RISE\left\{\bar{f}(x)\right\}$. Lastly, the following six distributions are considered as truths. All these distributions are popularly chosen as models for the income distribution, the actuarial loss distribution and the baseline hazard.

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:* $\left\{2/\left(\sqrt{2\pi}\sigma\right)\right\}\exp\left\{-(x-\mu)^2/(2\sigma^2)\right\}$, $(\mu,\sigma) = (0, 1.5)$.
4. *Log-Normal:* $\left\{1/\left(x\sqrt{2\pi}\sigma\right)\right\}\exp\left\{-(\log x-\mu)^2/(2\sigma^2)\right\}$, $(\mu,\sigma) = (0, 0.75)$.
5. *Burr:* $\alpha\beta x^{\alpha-1}/(1+x^\alpha)^{\beta+1}$, $(\alpha,\beta) = (1.5, 2.5)$.
6. *Generalized Gamma:* $\gamma x^{\alpha-1}\exp\{-(x/\beta)^\gamma\}/\{\beta^\alpha\Gamma(\alpha/\gamma)\}$, $(\alpha,\beta,\gamma) = (5, 2, 2.5)$.

TABLE 1 ABOUT HERE

4.2 Simulation Results

Table 1 presents averages and standard deviations of RISEs and averages of estimated smoothing parameters over $1,000$ Monte Carlo replications. For each distribution, results are qualitatively similar across three sample sizes. Distributions 1-2 are expected to be favorable to the G or MG, and W kernels, respectively. Because the shapes of two densities are similar, the results from three kernels are close for Distribution 1. For Distribution 2, the G and MG estimators slightly outperform the W estimator. This may reflect that the latter becomes slightly less efficient than the former even when best implemented; see Section 2.3.2 for details.

While the NM estimator does not perform well for these distributions, it decisively outperforms other estimators for Distribution 3. The good performance comes from the fact that the distribution is a special case of the Nakagami-m distribution. In contrast, poor performance of all other estimators is attributed to their difficulty in capturing the shape of the pdf near the origin. The pdf is decreasing and satisfies the shoulder condition $f'(0) = 0$. Accordingly, the data suggest concavity of the density in the vicinity of the origin. Then, the G, MG and W kernels tend to misinterpret the local concavity as an indication of a mode over the strictly positive region.

As opposed to previous three distributions, none of the four estimators has clear advantage for Distributions 4-6. While the G estimator exhibits good performance, density estimators using three kernels belonging to the Generalized Gamma family also work reasonably well.

As regards smoothing parameter values, for each combination of the distribution and sample size, we can find a consistent ordering of $\hat{b}_{MG}$, $\hat{b}_G$, $\hat{b}_W$, and $\hat{b}_{NM}$ from the smallest to the largest. It also holds that $\hat{b}_{NM}$ is roughly twice the size of $\hat{b}_{MG}$. This is congruous with what asymptotic results predict; see Section 2.3.2 for details.

5 Conclusion

Toward a unified approach to delivering convergence properties of asymmetric kernel estimators, this paper has proposed a family of GG kernels that consists of a set of common conditions. As special cases, the GG family incorporates not only the MG kernel but also the newly proposed W and NM kernels. Manipulating the conditions demonstrates that the GG density estimator preserves the appealing properties that the G and MG density estimators possess. Furthermore, this paper investigates three extensions of the density estimation including multiplicative bias correction. Monte Carlo simulations indicate good finite-sample properties of GG density estimators.

There are two possible research extensions. First, choice methods of the smoothing parameter b in GG density estimation need to be further investigated. Promising candidates include a more sophisticated plug-in method such as the one by Sheather and Jones (1991) and the cross-validation method. Second, goodness-of-fit tests can be built on GG kernels. Our particular emphasis is on tests for symmetry in unconditional and conditional distributions. Fernandes, Mendes and Scaillet (2011) propose the test statistic for symmetry that is based on the integrated squared discrepancy between two asymmetric kernel density estimates on the left- and right-hand sides from the axis of symmetry, and they report good finite-sample properties of the statistic in terms of size and power. Under the null of symmetry, the pdf on the right-hand side is expected to be decreasing and satisfy the shoulder condition, like the half-normal density. Invoking the Monte Carlo results in Section 4, we anticipate that the NM kernel is likely to exhibit better size properties. Moreover, with the availability of their commands in standard statistical packages, robust regression estimation methods are more popularly applied in empirical economics than before. However, focusing on the fact that conditional symmetry of the error distribution implies consistency of parameter estimators in certain versions of robust estimation, Baldauf and Santos Silva (2012) conduct simulation experiments and warn that violations of conditional symmetry in the error distribution lead to identification failure in robust estimation. In light of this finding, it would be interesting to reexamine adequacy of robust regression estimation techniques via testing for symmetry of conditional distributions of least-squares residuals. These extensions are currently under authors' investigation and will be addressed in separate papers.

A Appendix

In order to approximate the Gamma function, we frequently refer to the following well-known formulae:

Stirling's formula ("SF").

$$\Gamma(z+1) = \sqrt{2\pi} z^{z+1/2} e^{-z} \left\{ 1 + \frac{1}{12z} + \frac{1}{288z^2} + O\left(z^{-3}\right) \right\} \text{ as } z \to \infty.$$

Series expansion of the log Gamma function ("SELG").

$$\log \Gamma(1+z) = -\gamma z + \sum_{k=2}^{\infty} \frac{(-1)^k \zeta(k)}{k} z^k \text{ for } |z| < 1,$$

where (only in this context) $\gamma = \lim_{n\to\infty} \left(\sum_{k=1}^{n} k^{-1} - \log n \right) = 0.5772156649\ldots$ is Euler's constant, and $\zeta(k) = \sum_{n=1}^{\infty} n^{-k}$ $(k > 1)$ is the Riemann zeta function.

Legendre duplication formula ("LDF").

$$\Gamma(z) \Gamma\left(z + \frac{1}{2}\right) = \frac{\sqrt{\pi}}{2^{2z-1}} \Gamma(2z) \text{ for } z > 0.$$

A.1 Proof of Theorem 1

Bias. Let $\theta_x \stackrel{d}{=} GG\left(\alpha, \beta\Gamma(\alpha/\gamma) / \Gamma\{(\alpha+1)/\gamma\}, \gamma\right)$. Then, a second-order Taylor expansion of $E\left\{\hat{f}_{GG}(x)\right\}$ around $\theta_x = x$ yields $E\left\{\hat{f}_{GG}(x)\right\} = f(x) + E(\theta_x - x) f'(x) +$

$(1/2)\, E(\theta_x - x)^2 f''(x) + (s.o.)$. It follows from (2) that

$$E(\theta_x^m) = \beta^m \frac{\{\Gamma(\alpha/\gamma)\}^{m-1}\Gamma\{(\alpha+m)/\gamma\}}{[\Gamma\{(\alpha+1)/\gamma\}]^m}.$$

In particular, $E(\theta_x) = \beta$ (by construction) and

$$E\left(\theta_x^2\right) = \beta^2 \frac{\Gamma(\alpha/\gamma)\Gamma\{(\alpha+2)/\gamma\}}{[\Gamma\{(\alpha+1)/\gamma\}]^2} = \beta^2 M_b(x).$$

Using Conditions 2-3, we have, for $x \geq C_1 b$, $E(\theta_x) = x$ and $E\left(\theta_x^2\right) = x^2 + x^2\psi(x)\, b + o(b)$, and thus $E(\theta_x - x) = 0$ and $E(\theta_x - x)^2 \sim x^2\psi(x)\, b$. As a consequence, $Bias\left\{\hat{f}_{GG}(x)\right\} \sim (x^2/2)\,\psi(x)\, f''(x)\, b$. On the other hand, for $x \in [0, C_1 b)$, $E(\theta_x) = \varphi_b(x) = O(b)$ and $x = O(b)$ hold, and thus $E(\theta_x - x) = [\{\varphi_b(x) - x\}/b]\, b := \xi_b(x)\, b$, where $\xi_b(x) = O(1)$. Moreover, it follows from $E\left(\theta_x^2\right) = O(b^2)$ that $E(\theta_x - x)^2 = O(b^2)$. Therefore, $Bias\left\{\hat{f}_{GG}(x)\right\} \sim \xi_b(x)\, f'(x)\, b$.

Variance. As usual, we consider the approximation $Var\left\{\hat{f}_{GG}(x)\right\} = (1/n)\, E(K_i^2) + (s.o.)$. A straightforward calculation yields, for $\nu \in \mathbb{R}_+$,

$$\begin{aligned} K_{GG}^{\nu}(u; x, b) &= \left[\left\{\frac{\gamma\Gamma\left(\frac{\alpha+1}{\gamma}\right)}{\beta}\right\}^{\nu-1} \frac{\Gamma\left\{\frac{\nu(\alpha-1)+1}{\gamma}\right\}}{\nu^{\frac{\nu(\alpha-1)+1}{\gamma}}\left\{\Gamma\left(\frac{\alpha}{\gamma}\right)\right\}^{2\nu-1}}\right] \\ &\quad \cdot \frac{\gamma u^{\{\nu(\alpha-1)+1\}-1}\exp\left[-\left\{\frac{u}{\beta\Gamma\left(\frac{\alpha}{\gamma}\right)/\left(\nu^{1/\gamma}\Gamma\left(\frac{\alpha+1}{\gamma}\right)\right)}\right\}^{\gamma}\right]}{\left[\beta\Gamma\left(\frac{\alpha}{\gamma}\right)/\left\{\nu^{1/\gamma}\Gamma\left(\frac{\alpha+1}{\gamma}\right)\right\}\right]^{\nu(\alpha-1)+1}\Gamma\left\{\frac{\nu(\alpha-1)+1}{\gamma}\right\}}\mathbf{1}\{u \geq 0\} \\ &= A_{b,\nu}(x) \cdot \left\{\text{pdf of } GG\left(\nu(\alpha-1)+1, \frac{\beta\Gamma\left(\frac{\alpha}{\gamma}\right)}{\nu^{1/\gamma}\Gamma\left(\frac{\alpha+1}{\gamma}\right)}, \gamma\right)\right\}. \quad \text{(A1)} \end{aligned}$$

Hence, $Var\left\{\hat{f}_{GG}(x)\right\} \sim (1/n)\, A_{b,2}(x)\, E\{f(\eta_x)\}$, where

$$\eta_x \overset{d}{=} GG\left(2\alpha - 1, \frac{\beta\Gamma(\alpha/\gamma)}{2^{1/\gamma}\Gamma\{(\alpha+1)/\gamma\}}, \gamma\right).$$

By the mean-value theorem, $E\{f(\eta_x)\} = f(x) + E(\eta_x - x)\, f'(\bar{\eta}_x)$ for some $\bar{\eta}_x$ joining η_x and x. It follows from (2) that

$$E(\eta_x) = \beta\left[\frac{\Gamma(\alpha/\gamma)\Gamma(2\alpha/\gamma)}{2^{1/\gamma}\Gamma\{(\alpha+1)/\gamma\}\Gamma\{(2\alpha-1)/\gamma\}}\right] = \beta H_b(x).$$

Then, Condition 4, together with Condition 2, implies that $E(\eta_x - x) = o(1)$ so that $E\{f(\eta_x)\} \sim f(x)$ regardless of the position of x. Finally, Condition 5 establishes the approximation to the variance. ■

A.2 Proof of Theorem 2

A.2.1 The MG Kernel

Condition 3. Observe that $M_b(x) = 1 + 1/\alpha$. For $x \geq 2b$, $M_b(x) = 1 + (1/x) b$ so that $\psi(x) = 1/x$. On the other hand, for $x \in [0, 2b)$, $\alpha = (x/b)^2/4 + 1 = O(1)$, and thus $M_b(x) = O(1)$ holds.

Condition 4. Substituting $\alpha = x/b$ into $H_b(x) = 1 - (2\alpha)^{-1}$ yields $H_b(x) = 1 + O(b) = 1 + o(1)$ for interior x. On the other hand, for boundary x, $\alpha = O(1)$ regardless of whether $\alpha = x/b$ (when $x = O(b)$ and $x \geq 2b$) or $\alpha = (x/b)^2/4 + 1$ (when $x = O(b)$ and $x \in [0, 2b)$). It follows that $H_b(x) = O(1)$.

Condition 5. Because $\beta/\alpha = b$ regardless of the position of x, we have

$$A_{b,\nu}(x) = b^{1-\nu} \frac{\Gamma\{\nu(\alpha-1)+1\}}{\nu^{\nu(\alpha-1)+1}\{\Gamma(\alpha)\}^{\nu}}.$$

For interior x, SF yields, as $\alpha = x/b \to \infty$, $\Gamma(\alpha) \sim \sqrt{2\pi}\alpha^{\alpha-1/2}e^{-\alpha}$ and $\Gamma\{\nu(\alpha-1)+1\} \sim \sqrt{2\pi}\{\nu(\alpha-1)\}^{\nu(\alpha-1)+1/2} e^{-\nu(\alpha-1)}$. Then,

$$A_{b,\nu}(x) \sim \frac{b^{1-\nu}(\alpha-1)^{(1-\nu)/2}}{\nu^{1/2}\left(\sqrt{2\pi}\right)^{\nu-1}}.$$

The result immediately follows from defining $V_I(\nu) := \left\{\nu^{1/2}\left(\sqrt{2\pi}\right)^{\nu-1}\right\}^{-1}$ and recognizing that $(\alpha-1)^{(1-\nu)/2} = \alpha^{(1-\nu)/2}(1-1/\alpha)^{(1-\nu)/2} \sim (x/b)^{(1-\nu)/2}$. For boundary x, the result is established by defining

$$V_B(\nu) := \begin{cases} \frac{\Gamma\{\nu(\kappa-1)+1\}}{\nu^{\nu(\kappa-1)+1}\{\Gamma(\kappa)\}^{\nu}} & \text{if } x/b \to \kappa \geq 2 \\ \frac{\Gamma\left(\frac{\nu}{4}\kappa^2+1\right)}{\nu^{\frac{\nu}{4}\kappa^2+1}\left\{\Gamma\left(\frac{\kappa^2}{4}+1\right)\right\}^{\nu}} & \text{if } x/b \to \kappa \in (0,2) \end{cases}.$$

A.2.2 The W Kernel

Condition 3. Observe that $M_b(x) = \Gamma(1+2/\alpha)/\{\Gamma(1+1/\alpha)\}^2$. For $x \geq 2b$, we may pick an arbitrarily small $b > 0$ so that $|2/\alpha| = \left|\sqrt{2b/x}\right| \leq 1$. Then, by SELG and $\zeta(2) = \pi^2/6$, two Gamma functions admit the following approximations:

$$\begin{aligned}
\Gamma\left(1+\frac{2}{\alpha}\right) &= \exp\left\{\log\Gamma\left(1+\frac{2}{\alpha}\right)\right\} = 1 - \frac{2\gamma}{\alpha} + \frac{(\pi^2/3)+2\gamma^2}{\alpha^2} + O\left(\alpha^{-3}\right), \\
\Gamma\left(1+\frac{1}{\alpha}\right) &= \exp\left\{\log\Gamma\left(1+\frac{1}{\alpha}\right)\right\} = 1 - \frac{\gamma}{\alpha} + \frac{(\pi^2/12)+(\gamma^2/2)}{\alpha^2} + O\left(\alpha^{-3}\right).
\end{aligned}$$

Applying a geometric series expansion to the approximation to $\{\Gamma(1+1/\alpha)\}^{-2}$ finally delivers $M_b(x) \sim 1 + (\pi^2/6)/\alpha^2 = 1 + \{\pi^2/(12x)\} b$ so that $\psi(x) = \pi^2/(12x)$. On the other hand, for $x \in [0, 2b)$, $\alpha = (x/b)/2 + 1 = O(1)$, and thus $M_b(x) = O(1)$ is the case.

Condition 4. In this case, $H_b(x) = \left\{2^{1/\alpha}(1-1/\alpha)\Gamma(1+1/\alpha)\Gamma(1-1/\alpha)\right\}^{-1}$. It is easy to see that $H_b(x) = O(1)$ for boundary x. On the other hand, for interior x, $2^{1/\alpha} = \exp\{(1/\alpha)\log 2\} = 1 + O\left(b^{1/2}\right)$, and $(1-1/\alpha)\Gamma(1+1/\alpha)\Gamma(1-1/\alpha) = 1 + O\left(b^{1/2}\right)$ by SELG. Therefore, $H_b(x) = 1 + O\left(b^{1/2}\right) = 1 + o(1)$ holds.

Condition 5. We have

$$A_{b,\nu}(x) = \left(\frac{\alpha}{\beta}\right)^{\nu-1} \frac{\left\{\Gamma\left(1+\frac{1}{\alpha}\right)\right\}^{\nu-1}\Gamma\left(\nu+\frac{1-\nu}{\alpha}\right)}{\nu^{\nu+\frac{1-\nu}{\alpha}}}.$$

For interior x, because $\alpha = \sqrt{2x/b} \to \infty$, we may approximate $\Gamma(1+1/\alpha) \sim 1$, $\Gamma\{\nu+(1-\nu)/\alpha\} \sim \Gamma(\nu)$ and $\nu^{\nu+(1-\nu)/\alpha} \sim \nu^\nu$. In addition, $(\alpha/\beta)^{\nu-1} = 2^{(\nu-1)/2}(xb)^{(1-\nu)/2}$, and thus

$$A_{b,\nu}(x) \sim 2^{\frac{\nu-1}{2}} \frac{\Gamma(\nu)}{\nu^\nu}(xb)^{\frac{1-\nu}{2}} := V_I(\nu)(xb)^{\frac{1-\nu}{2}}.$$

For boundary x, the result is established by defining

$$V_B(\nu) := \begin{cases} \left(\frac{2}{\kappa}\right)^{\frac{\nu-1}{2}} \frac{\left\{\Gamma\left(1+\frac{1}{\sqrt{2\kappa}}\right)\right\}^{\nu-1}\Gamma\left(\nu+\frac{1-\nu}{\sqrt{2\kappa}}\right)}{\nu^{\nu+\frac{1-\nu}{\sqrt{2\kappa}}}} & \text{if } x/b \to \kappa \geq 2 \\ \left\{\frac{2(\kappa+2)}{\kappa^2+4}\right\}^{\nu-1} \frac{\left\{\Gamma\left(1+\frac{2}{\kappa+2}\right)\right\}^{\nu-1}\Gamma\left\{\nu+\frac{2(1-\nu)}{\kappa+2}\right\}}{\nu^{\nu+\frac{2(1-\nu)}{\kappa+2}}} & \text{if } x/b \to \kappa \in (0,2) \end{cases}.$$

A.2.3 The NM Kernel

Condition 3. Observe that $M_b(x) = (\alpha/2)\left[\Gamma(\alpha/2)/\Gamma\{(\alpha+1)/2\}\right]^2$. For $x \geq 2b$, it follows from LDF that $M_b(x) = (\alpha/2)\left[2^{\alpha-1}\{\Gamma(\alpha/2)\}^2/\{\sqrt{\pi}\Gamma(\alpha)\}\right]^2$. Next, SF implies that, as $\alpha = x/b \to \infty$,

$$\Gamma(\alpha) = \sqrt{2\pi}\alpha^{\alpha-1/2}e^{-\alpha}\left\{1+\frac{1}{12\alpha}+O\left(\alpha^{-2}\right)\right\}, \tag{A2}$$

$$\Gamma\left(\frac{\alpha}{2}\right) = \sqrt{2\pi}\left(\frac{\alpha}{2}\right)^{\alpha/2-1/2}e^{-\alpha/2}\left\{1+\frac{1}{6\alpha}+O\left(\alpha^{-2}\right)\right\}. \tag{A3}$$

Then, we find that the approximation of $M_b(x)$ takes a very simple form $M_b(x) = 1+(2\alpha)^{-1}+O(\alpha^{-2}) = 1+(2x)^{-1}b+o(b)$ so that $\psi(x) = (2x)^{-1}$. On the other hand, for $x \in [0,2b)$, $\alpha = (x/b)^2/4+1 = O(1)$, and thus $M_b(x) = O(1)$ holds.

Condition 4. $H_b(x) = (\alpha-1/2)\Gamma(\alpha/2)\Gamma(\alpha)/\left[2^{1/2}\Gamma\{(\alpha+1)/2\}\Gamma(\alpha+1/2)\right]$ is $O(1)$ for boundary x because of the same reason as in the proof for the MG kernel. For interior x, using LDF and then $\Gamma(2\alpha) \sim \sqrt{2\pi}(2\alpha)^{2\alpha-1/2}e^{-2\alpha}$ as $\alpha = x/b \to \infty$, as well as (A2)-(A3), yields

$$H_b(x) = \frac{\alpha-\frac{1}{2}}{2^{1/2}}\frac{2^{3\alpha-2}}{\pi}\frac{\left\{\Gamma\left(\frac{\alpha}{2}\right)\right\}^2\Gamma(\alpha)}{\Gamma(2\alpha)} \sim 1-\frac{1}{2\alpha} = 1+O(b) = 1+o(1).$$

Condition 5. We have

$$A_{b,\nu}(x) = \beta^{1-\nu}\frac{2^{\nu-1}}{\nu^{\frac{\nu(\alpha-1)}{2}+\frac{1}{2}}}\frac{\left\{\Gamma\left(\frac{\alpha+1}{2}\right)\right\}^{\nu-1}\Gamma\left\{\frac{\nu(\alpha-1)}{2}+\frac{1}{2}\right\}}{\left\{\Gamma\left(\frac{\alpha}{2}\right)\right\}^{2\nu-1}}.$$

For interior x, by LDF, $A_{b,\nu}(x)$ reduces to

$$A_{b,\nu}(x) = \beta^{1-\nu}\frac{2^{\nu-1}}{\nu^{\frac{\nu(\alpha-1)}{2}+\frac{1}{2}}}\frac{\left(\sqrt{\pi}\right)^{\nu}\{\Gamma(\alpha)\}^{\nu-1}\Gamma\{\nu(\alpha-1)\}}{2^{(\nu-1)(\alpha-1)+\nu(\alpha-1)-1}\Gamma\left\{\frac{\nu(\alpha-1)}{2}\right\}\left\{\Gamma\left(\frac{\alpha}{2}\right)\right\}^{3\nu-2}}.$$

Next, by SF, as $\alpha = x/b \to \infty$, $\Gamma\left\{\nu\left(\alpha-1\right)\right\} \sim \sqrt{2\pi}\left\{\nu\left(\alpha-1\right)\right\}^{\nu(\alpha-1)-1/2} e^{-\nu(\alpha-1)}$ and $\Gamma\left\{\nu\left(\alpha-1\right)/2\right\} \sim \sqrt{2\pi}\left\{\nu\left(\alpha-1\right)/2\right\}^{\nu(\alpha-1)-1/2} e^{-\nu(\alpha-1)/2}$. Substituting these approximations, as well as (A2)-(A3), we finally deduce that

$$A_{b,\nu}\left(x\right) \sim \beta^{1-\nu}\alpha^{\frac{\nu-1}{2}} \frac{e^{\nu/2}}{\nu^{1/2}\left(\sqrt{\pi}\right)^{\nu-1}} \left(1-\frac{1}{\alpha}\right)^{\frac{\nu(\alpha-1)}{2}}.$$

Moreover $\beta^{1-\nu}\alpha^{(1-\nu)/2} = (xb)^{(1-\nu)/2}$ and $(1-1/\alpha)^{\nu(\alpha-1)/2} \sim e^{-\nu/2}$, and thus

$$A_{b,\nu}\left(x\right) \sim \frac{1}{\nu^{1/2}\left(\sqrt{\pi}\right)^{\nu-1}} \left(xb\right)^{\frac{1-\nu}{2}} := V_I\left(\nu\right)\left(xb\right)^{\frac{1-\nu}{2}}.$$

For boundary x, the result is established by defining

$$V_B\left(\nu\right) := \begin{cases} \left(\frac{2}{\kappa}\right)^{\nu-1} \frac{\left\{\Gamma\left(\frac{\kappa+1}{2}\right)\right\}^{\nu-1}\Gamma\left\{\frac{\nu(\kappa-1)+1}{2}\right\}}{\nu^{\frac{\nu(\kappa-1)+1}{2}}\left\{\Gamma\left(\frac{\kappa}{2}\right)\right\}^{2\nu-1}} & \text{if } x/b \to \kappa \geq 2 \\ \left(\frac{8}{\kappa^2+4}\right)^{\nu-1} \frac{\left\{\Gamma\left(\frac{\kappa^2}{8}+1\right)\right\}^{\nu-1}\Gamma\left(\frac{\nu}{8}\kappa^2+\frac{1}{2}\right)}{\nu^{\frac{\nu}{8}\kappa^2+\frac{1}{2}}\left\{\Gamma\left(\frac{\kappa^2}{8}+\frac{1}{2}\right)\right\}^{2\nu-1}} & \text{if } x/b \to \kappa \in (0,2) \end{cases}. \quad \blacksquare$$

A.3 Proof of Theorem 4

We concentrate on the case for interior x; the proof for boundary x is similar and thus omitted. We also employ a short-handed notation $K_i := K_{GG}\left(X_i; x, b\right)$ to save space. Then, it suffices to demonstrate that

$$Var\left\{\sqrt{nb^{1/2}}\hat{f}_{GG}\left(x\right)\right\} = Var\left(b^{1/4}K_i\right) + 2\sum_{\ell=1}^{n-1}\left(1-\frac{\ell}{n}\right) Cov\left(b^{1/4}K_i, b^{1/4}K_{i+\ell}\right) \sim V_I\left(2\right)\frac{f\left(x\right)}{\sqrt{x}}.$$

It follows from Theorem 1 that $Var\left(b^{1/4}K_i\right) \sim V_I\left(2\right) f\left(x\right)/\sqrt{x}$. Hence, we only need to show that

$$\sum_{\ell=1}^{n-1}\left(1-\frac{\ell}{n}\right) Cov\left(b^{1/4}K_i, b^{1/4}K_{i+\ell}\right) = o\left(1\right). \tag{A4}$$

Observe that the absolute value of the left-hand side of (A4) is bounded by

$$\sum_{\ell=1}^{\infty}\left|Cov\left(b^{1/4}K_i, b^{1/4}K_{i+\ell}\right)\right| = \left(\sum_{\ell=1}^{d_n} + \sum_{\ell=d_n+1}^{\infty}\right)\left|Cov\left(b^{1/4}K_i, b^{1/4}K_{i+\ell}\right)\right| = V_1 + V_2 \text{ (say)},$$

where the increasing sequence d_n is specified shortly. We evaluate V_2 first. By Davydov's lemma (e.g. Corollary A.2 of Hall and Heyde, 1980) and the stationarity of X_i,

$$\left|Cov\left(b^{1/4}K_i, b^{1/4}K_{i+\ell}\right)\right| \le 8b^{1/2}\left(E\left|K_i - E\left(K_i\right)\right|^r\right)^{2/r}\alpha\left(\ell\right)^{1-2/r}. \tag{A5}$$

By C_r-inequality and $K_i \ge 0$, $E\left|K_i - E\left(K_i\right)\right|^r \le 2^{r-1}\left[E\left(K_i^r\right) + \left\{E\left(K_i\right)\right\}^r\right]$. Because $E\left(K_i\right) = O\left(1\right)$ and $E\left(K_i^r\right) = O\left\{A_{b,r}\left(x\right)\right\} = O\left(b^{(1-r)/2}\right)$ by (A1), we have

$$E\left|K_i - E\left(K_i\right)\right|^r = O\left(b^{\frac{1-r}{2}}\right). \tag{A6}$$

The size of the mixing coefficient also implies that

$$\alpha\left(\ell\right) \le C_5\ell^{-q} \tag{A7}$$

for some constants $C_5 > 0$ and $q > \left(2 - 2/r\right)/\left(1 - 2/r\right)$. Substituting (A6)-(A7) into (A5) yields $\left|Cov\left(b^{1/4}K_i, b^{1/4}K_{i+\ell}\right)\right| \le cb^{1/r-1/2}\ell^{-q(1-2/r)}$. Hence, $V_2 \le cb^{1/r-1/2}\sum_{\ell=d_n+1}^{\infty}\ell^{-q(1-2/r)}$, where $q\left(1-2/r\right) > 1$ holds by construction. Also define $d_n := \lfloor b^{-a}\rfloor$ for some $a \in \left(\left(1/2\right)\left\{q - \left(1-2/r\right)\right\}^{-1}, 1/2\right)$. Then,

$$\sum_{\ell=d_n+1}^{\infty}\ell^{-q(1-2/r)} \le \int_{d_n}^{\infty}x^{-q(1-2/r)}dx = \frac{d_n^{1-q(1-2/r)}}{q\left(1-2/r\right)-1} = O\left\{b^{a(q(1-2/r)-1)}\right\}, \tag{A8}$$

and thus $V_2 \le O\left\{b^{a(q(1-2/r)-1)-(1/2)(1-2/r)}\right\} \to 0$.

We now turn to V_1. The stationarity of X_i and $K_i \ge 0$ imply that $\left|Cov\left(b^{1/4}K_i, b^{1/4}K_{i+\ell}\right)\right| \le b^{1/2}\left[E\left(K_iK_{i+\ell}\right) + \left\{E\left(K_i\right)\right\}^2\right]$, where both $E\left(K_iK_{i+\ell}\right)$ and $E\left(K_i\right)$ are $O\left(1\right)$. Therefore, $V_1 \le O\left(d_nb^{1/2}\right) = O\left(b^{1/2-a}\right) \to 0$, which establishes (A4). ■

A.4 Proof of Theorem 5

The proof requires four lemmata below. In particular, a Bernstein-type inequality for strong mixing processes in Lemma A4, which restates Theorem 2.1 of Liebscher (1996), constitutes the key part of the proof.

Lemma A1. *Let* $(\alpha_0, \beta_0, \gamma_0) := (\alpha_b(0), \beta_b(0), \gamma_b(0))$. *Then, for any* $\delta > 0$,

$$\int_0^\delta K_{GG}(u; 0, b)\, du = \int_0^\delta \frac{\gamma_0 u^{\alpha_0 - 1} \exp\left[-\left\{u / \left(\beta_0 \Gamma(\alpha_0/\gamma_0) / \Gamma((\alpha_0+1)/\gamma_0)\right)\right\}^{\gamma_0}\right]}{\left\{\beta_0 \Gamma(\alpha_0/\gamma_0) / \Gamma((\alpha_0+1)/\gamma_0)\right\}^{\alpha_0} \Gamma(\alpha_0/\gamma_0)} du \to 1,$$

as $b \to 0$.

Lemma A2. *For some* $\bar{x} \in [0, C_1 b)$, $K_{GG}(u; \bar{x}, b) \le C_6 b^{-1}$, *where*

$$C_6 := \left(\frac{C_4^2}{C_2}\right)(C_4+1)(C_4+2)\max\left\{1, (C_4-1)^{C_4-1}\right\}.$$

Lemma A3. *Let* $\bar{K}_i := K_{GG}(X_i; \bar{x}, b) - E\{K_{GG}(X_i; \bar{x}, b)\}$ *for* $\bar{x}$ *defined in Lemma A2. Then,* $E\left(\sum_{i=1}^m \bar{K}_i\right)^2 \le C_7 m b^{-2}$, *where*

$$C_7 := 2C_6^2\left[1 + 32 C_5^{1-2/r}\left\{1 + \frac{1}{q(1-2/r) - 1}\right\}\right].$$

Lemma A4. (Liebscher, 1996, Theorem 2.1) *Let* $\{Z_i\}$ *be a strictly stationary and strong mixing process with the mixing coefficient* $\alpha(\ell)$ *such that* $E(Z_i) = 0$ *and* $|Z_i| \le S(n), i = 1, \ldots, n$. *Then, for any integer* $1 \le m \le n$ *and for any* $\epsilon > 4mS(n)$,

$$\Pr\left(\left|\sum_{i=1}^n Z_i\right| > \epsilon\right) \le 4\exp\left\{-\frac{\epsilon^2}{64(n/m)\sigma^2(m) + (8/3)\epsilon m S(n)}\right\} + 4\frac{n}{m}\alpha(m),$$

where $\sigma^2(m) := E\left(\sum_{i=1}^m Z_i\right)^2$.

A.4.1 Proof of Lemma A1

By the change of variable $v := \left[u / \left\{\beta_0 \Gamma(\alpha_0/\gamma_0) / \Gamma((\alpha_0+1)/\gamma_0)\right\}\right]^{\gamma_0}$, the integral can be rewritten as $\int_0^{C_\delta}\left\{v^{(\alpha_0/\gamma_0)-1}\exp(-v)/\Gamma(\alpha_0/\gamma_0)\right\}dv$, where the integrand is the pdf of $G(\alpha_0/\gamma_0, 1)$, and

$$C_\delta = \left[\frac{\delta}{\beta_0 \Gamma(\alpha_0/\gamma_0) / \Gamma\{(\alpha_0+1)/\gamma_0\}}\right]^{\gamma_0}.$$

Therefore, the proof is boiled down to showing that for any $\delta > 0$, $C_\delta \to \infty$ as $b \to 0$.

Because $\alpha_0 \in [1, C_4]$ and $\gamma_0 \geq 1$, $\alpha_0 = O(\gamma_0)$ or $\alpha_0 = o(\gamma_0)$ must be the case. If $\alpha_0 = O(\gamma_0)$, then $\Gamma(\alpha_0/\gamma_0)$ and $\Gamma\{(\alpha_0+1)/\gamma_0\}$ are both $O(1)$. It follows that $C_\delta = O(b^{-\gamma_0}) \to \infty$. Alternatively, if $\alpha_0 = o(\gamma_0)$, then we may pick an arbitrarily small b so that $|\alpha_0/\gamma_0| \leq 1$ and $|(\alpha_0+1)/\gamma_0| \leq 1$. Using SELG and the property of the Gamma function yields

$$\log \Gamma(z) = -\log(z) - \gamma z + \sum_{k=2}^{\infty} \frac{(-1)^k \zeta(k)}{k} z^k$$

for $z = \alpha_0/\gamma_0, (\alpha_0+1)/\gamma_0$. Then,

$$\begin{aligned}
\Gamma\left(\frac{\alpha_0}{\gamma_0}\right) / \Gamma\left(\frac{\alpha_0+1}{\gamma_0}\right) &= \exp\left\{\log \Gamma\left(\frac{\alpha_0}{\gamma_0}\right) - \log \Gamma\left(\frac{\alpha_0+1}{\gamma_0}\right)\right\} \\
&= \left(1 + \frac{1}{\alpha_0}\right) \exp\left[O\left(\frac{1}{\gamma_0}\right) + O\left\{\left(\frac{\alpha_0}{\gamma_0}\right)^2\right\}\right] = O(1),
\end{aligned}$$

and thus $C_\delta = O(b^{-\gamma_0}) \to \infty$. ■

A.4.2 Proof of Lemma A2

Let $(\bar{\alpha}, \bar{\beta}, \bar{\gamma}) := (\alpha_b(\bar{x}), \beta_b(\bar{x}), \gamma_b(\bar{x}))$. The upper bound can be implied by $K_{GG}(u^*; \bar{x}, b)$, where u^* is the mode. Because the shape of $K_{GG}(u; \bar{x}, b)$ is substantially different between the cases with $\bar{\alpha} = 1$ and $\bar{\alpha} > 1$, we evaluate two cases separately.

When $\bar{\alpha} > 1$, a straightforward calculation yields

$$u^* = \left[\frac{\bar{\beta}\Gamma(\bar{\alpha}/\bar{\gamma})}{\Gamma\{(\bar{\alpha}+1)/\bar{\gamma}\}}\right] \left(\frac{\bar{\alpha}-1}{\bar{\gamma}}\right)^{1/\bar{\gamma}}$$

so that

$$K_{GG}(u^*; \bar{x}, b) = \left(\frac{\bar{\gamma}}{\bar{\beta}}\right) \left[\frac{\Gamma\{(\bar{\alpha}+1)/\bar{\gamma}\}}{\Gamma^2(\bar{\alpha}/\bar{\gamma})}\right] \left(\frac{\bar{\alpha}-1}{\bar{\gamma}}\right)^{(\bar{\alpha}-1)/\bar{\gamma}} \exp\left\{-\left(\frac{\bar{\alpha}-1}{\bar{\gamma}}\right)\right\}. \quad \text{(A9)}$$

Observe that $\Gamma\{(\bar{\alpha}+1)/\bar{\gamma}\}/\Gamma^2(\bar{\alpha}/\bar{\gamma}) = \{\Gamma(2\bar{\alpha}/\bar{\gamma})/\Gamma^2(\bar{\alpha}/\bar{\gamma})\}[\Gamma\{(\bar{\alpha}+1)/\bar{\gamma}\}/\Gamma(2\bar{\alpha}/\bar{\gamma})]$. It follows from Corollary 1 of Cerone (2007) and the property of the Gamma function that for $z > 0$,

$$\frac{\Gamma^2(1+z)}{\Gamma(2+2z)} \geq \frac{1}{(1+z)^{2+z}} \Rightarrow \frac{\Gamma(2z)}{\Gamma^2(z)} \leq \frac{z(1+z)^{2+z}}{2(1+2z)}. \tag{A10}$$

Putting $z = \bar{\alpha}/\bar{\gamma}$ gives

$$\frac{\Gamma(2\bar{\alpha}/\bar{\gamma})}{\Gamma^2(\bar{\alpha}/\bar{\gamma})} \leq \left(\frac{1}{2}\right)\left(\frac{\bar{\alpha}}{\bar{\gamma}+2\bar{\alpha}}\right)\left(1+\frac{\bar{\alpha}}{\bar{\gamma}}\right)^{2+\bar{\alpha}/\bar{\gamma}}. \tag{A11}$$

Moreover, by Theorem 1 of Kečkić and Vasić (1971) and the property of the Gamma function, for $x > y > 0$,

$$\frac{\Gamma(1+x)}{\Gamma(1+y)} \geq \frac{(1+x)^x}{(1+y)^y}\exp(y-x) \Rightarrow \frac{\Gamma(y)}{\Gamma(x)} \leq \frac{x(1+y)^y}{y(1+x)^x}\exp(x-y).$$

Letting $(x, y) = (2\bar{\alpha}/\bar{\gamma}, (\bar{\alpha}+1)/\bar{\gamma})$, we have

$$\frac{\Gamma\{(\bar{\alpha}+1)/\bar{\gamma}\}}{\Gamma(2\bar{\alpha}/\bar{\gamma})} \leq \left(\frac{2\bar{\alpha}}{1+\bar{\alpha}}\right)\frac{\{1+(\bar{\alpha}+1)/\bar{\gamma}\}^{(\bar{\alpha}+1)/\bar{\gamma}}}{(1+2\bar{\alpha}/\bar{\gamma})^{2\bar{\alpha}/\bar{\gamma}}}\exp\left(\frac{\bar{\alpha}-1}{\bar{\gamma}}\right). \tag{A12}$$

Substituting (A11)-(A12) into (A9), rearranging it, and then using $\bar{\alpha} \in (1, C_4]$, $\bar{\beta} \geq C_2 b$ and $\bar{\gamma} \geq 1$, we deduce that

$$\begin{aligned} K_{GG}(u^*; \bar{x}, b) &\leq \left(\frac{1}{\bar{\beta}}\right)\left(\frac{\bar{\gamma}}{\bar{\gamma}+2\bar{\alpha}}\right)\left(\frac{\bar{\alpha}^2}{\bar{\alpha}+1}\right)\left[\frac{(1+\bar{\alpha}/\bar{\gamma})\{1+(\bar{\alpha}+1)/\bar{\gamma}\}}{(1+2\bar{\alpha}/\bar{\gamma})^2}\right]^{\bar{\alpha}/\bar{\gamma}} \\ &\quad \cdot \left(\frac{\bar{\alpha}}{\bar{\gamma}}+1\right)^2\left(\frac{\bar{\alpha}+1}{\bar{\gamma}}+1\right)^{1/\bar{\gamma}}\left(\frac{\bar{\alpha}-1}{\bar{\gamma}}\right)^{(\bar{\alpha}-1)/\bar{\gamma}} \\ &\leq \left(\frac{1}{C_2 b}\right)\cdot 1 \cdot \left(\frac{C_4^2}{C_4+1}\right)\cdot 1 \cdot (C_4+1)^2 \cdot (C_4+2)\cdot \max\left\{1, (C_4-1)^{C_4-1}\right\}. \end{aligned}$$

In sum, as far as $\bar{\alpha} > 1$, $K_{GG}(u; \bar{x}, b) \leq C_6 b^{-1}$, where

$$C_6 := \left(\frac{C_4^2}{C_2}\right)(C_4+1)(C_4+2)\max\left\{1, (C_4-1)^{C_4-1}\right\}.$$

On the other hand, when $\bar{\alpha} = 1$, it follows from (A10) and $u^* = 0$ that

$$\begin{aligned} K_{GG}(u^*; \bar{x}, b) &= \left(\frac{\bar{\gamma}}{\bar{\beta}}\right)\left\{\frac{\Gamma(2/\bar{\gamma})}{\Gamma^2(1/\bar{\gamma})}\right\} \\ &\leq \left(\frac{1}{\bar{\beta}}\right)\left\{\frac{\bar{\gamma}}{2(\bar{\gamma}+2)}\right\}\left(1+\frac{1}{\bar{\gamma}}\right)^{2+1/\bar{\gamma}} \\ &\leq \left(\frac{1}{C_2 b}\right)\cdot\left(\frac{1}{2}\right)\cdot 2^3 = \left(\frac{4}{C_2}\right)b^{-1}. \end{aligned}$$

Note that $C_6 \geq 6/C_2$ holds, which establishes the lemma. ■

A.4.3 Proof of Lemma A3

By the stationarity of Z_i,

$$E\left(\sum_{i=1}^{m}\bar{K}_i\right)^2 \leq mE\left(\bar{K}_i^2\right)+2\sum_{\ell=1}^{m-1}(m-\ell)\,E\left|\bar{K}_i\bar{K}_{i+\ell}\right|, \tag{A13}$$

where, by Lemma A2 and $\int_0^\infty f(u)\,du=1$,

$$E\left(\bar{K}_i^2\right)\leq E\left|K_{GG}(X_i;\bar{x},b)\right|^2+E^2\left|K_{GG}(X_i;\bar{x},b)\right|\leq 2C_6^2b^{-2}. \tag{A14}$$

Following the same manner as in the proof of Theorem 4, we also have $E\left|\bar{K}_i\bar{K}_{i+\ell}\right|\leq 8\left(E\left|\bar{K}_i\right|^r\right)^{2/r}\alpha(\ell)^{1-2/r}$, where, by C_r-inequality, Lemma A2 and $\int_0^\infty f(u)\,du=1$,

$$E\left|\bar{K}_i\right|^r\leq 2^{r-1}\left\{E\left|K_{GG}(X_i;\bar{x},b)\right|^r+E^r\left|K_{GG}(X_i;\bar{x},b)\right|\right\}\leq\left(2C_6b^{-1}\right)^r.$$

Therefore, $E\left|\bar{K}_i\bar{K}_{i+\ell}\right|\leq 32C_5^{1-2/r}C_6^2b^{-2}\ell^{-q(1-2/r)}$ by (A7), and thus

$$\begin{aligned}\sum_{\ell=1}^{m-1}(m-\ell)\,E\left|\bar{K}_i\bar{K}_{i+\ell}\right| &\leq 32C_5^{1-2/r}C_6^2mb^{-2}\sum_{\ell=1}^{\infty}\ell^{-q(1-2/r)}\\ &\leq 32C_5^{1-2/r}C_6^2\left\{1+\frac{1}{q(1-2/r)-1}\right\}mb^{-2},\end{aligned}\tag{A15}$$

where the last inequality follows from (A8). Combining (A13), (A14) and (A15) establishes the result. ■

A.4.4 Proof of Theorem 5

This proof largely follows the one of Proposition 3.3 in Bouezmarni and Van Bellegem (2011). The proof completes if the following statements hold for some $\bar{x}\in[0,C_1b)$:

$$\hat{f}_{GG}(\bar{x}) = E\left\{\hat{f}_{GG}(\bar{x})\right\}+o_p(1). \tag{A16}$$

$$E\left\{\hat{f}_{GG}(\bar{x})\right\} = E\left\{\hat{f}_{GG}(0)\right\}+o(1). \tag{A17}$$

$$E\left\{\hat{f}_{GG}(0)\right\} \to \infty. \tag{A18}$$

Note that (A17) immediately follows from the continuity of $K_{GG}(u;x,b)$ in x.

We demonstrate (A18) first. When $f(x) \to \infty$ as $x \to 0+$, it holds that for any $A > 0$, there is some $\delta > 0$ such that $f(x) > A$ for all $x < \delta$. For the given δ, Lemma A1 implies that

$$E\left\{\hat{f}_{GG}(0)\right\} > \int_0^\delta K_{GG}(u;0,b) f(u)\,du > A\int_0^\delta K_{GG}(u;0,b)\,du \to A,$$

which establishes (A18).

To show (A16), consider $\bar{K}_i$ in Lemma A3. Then, $E\left(\bar{K}_i\right) = 0$ and the same logic as applied for (A14) establishes that $\left|\bar{K}_i\right| \le 2C_6 b^{-1}$. Also pick $b = (\log n/n)^{1/2}$ and $m = \lfloor n^a \rfloor$ for some $a \in (1/(1+q), 1/2)$ for concreteness. Then, for a sufficiently large n, $1 \le m \le n$ holds, and for an arbitrarily chosen $\epsilon > 0$, we also have $n\epsilon > 8C_6 m b^{-1}$. Therefore, for the given ϵ, we may apply Lemmata A3-A4 and (A7) to obtain

$$\begin{aligned}
& \Pr\left(\left|\hat{f}_{GG}(\bar{x}) - E\left\{\hat{f}_{GG}(\bar{x})\right\}\right| > \epsilon\right) \\
= & \Pr\left(\left|\sum_{i=1}^n \bar{K}_i\right| > n\epsilon\right) \\
\le & 4\exp\left\{-\frac{(n\epsilon)^2}{64(n/m)(C_7 m b^{-2}) + (8/3)(n\epsilon) m (2C_6 b^{-1})}\right\} + 4\frac{n}{m}\left(C_5 m^{-q}\right) \\
= & 4\exp\left\{-\frac{3\epsilon^2(nb^2)}{16(12C_7 + C_6\epsilon m b)}\right\} + 4C_5 n m^{-(1+q)}
\end{aligned}$$

We have $mb \le 1$ for a sufficiently large n, and thus the right-hand side is bounded by $O\left\{n^{-(3/16)\epsilon^2/(12C_7+C_6\epsilon)}\right\} + O\left\{n^{-(a(1+q)-1)}\right\} \to 0$, which completes the proof. ■

References

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

[2] Baldauf, M., and J. M. C. Santos Silva (2012): "On the Use of Robust Regression in Econometrics," *Economics Letters*, 114, 124 - 127.

[3] 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.

[4] Bouezmarni, T., and J. V. K. Rombouts (2010): "Nonparametric Density Estimation for Positive Time Series," *Computational Statistics & Data Analysis*, 54, 245 - 261.

[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] Bouezmarni, T., and S. Van Bellegem (2011): "Nonparametric Beta Kernel Estimator for Long Memory Time Series," ECORE Discussion Paper 2011/19.

[7] Carrasco, M., and X. Chen (2002): "Mixing and Moment Properties of Various GARCH and Stochastic Volatility Models," *Econometric Theory*, 18, 17 - 39.

[8] Cerone, P. (2007): "Special Functions: Approximations and Bounds," *Applicable Analysis and Discrete Mathematics*, 1, 72 - 91.

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

[10] Chen, X., L. P. Hansen, and M. Carrasco (2010): "Nonlinearity and Temporal Dependence," *Journal of Econometrics*, 155, 155 - 169.

[11] 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.

[12] 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.

[13] DiNardo, J., N. M. Fortin, and T. Lemieux (1996): "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica*, 64, 1001 - 1044.

[14] Fé, E. (2012): "Efficient Estimation in Regression Discontinuity Designs via Asymmetric Kernels," MPRA Paper No. 38164.

[15] Fernandes, M., E. F. Mendes, and O. Scaillet (2011): "Testing for Symmetry and Conditional Symmetry Using Asymmetric Kernels," Swiss Finance Institute Occasional Paper Series N° 11-32.

[16] Gospodinov, N., and M. Hirukawa (2012): "Nonparametric Estimation of Scalar Diffusion Models of Interest Rates Using Asymmetric Kernels," *Journal of Empirical Finance*, 19, 595 - 609.

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

[18] Hall, P., and C. C. Heyde (1980): *Martingale Limit Theory and Its Application.* New York: Academic Press.

[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] Hirukawa, M., and M. Sakudo (2012): "Nonparametric Multiplicative Bias Correction for Kernel-Type Density Estimation Using Positive Data," Working Paper, Setsunan University.

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

[23] Karunamuni, R. J., and T. Alberts (2005): "A Generalized Reflection Method of Boundary Correction in Kernel Density Estimation," *Canadian Journal of Statistics*, 33, 497 - 509.

[24] Kečkić, J. D., and P. M. Vasić (1971): "Some Inequalities for the Gamma Function," *Publications de l'Institut Mathématique*, 11, 107 - 114.

[25] Kuruwita, C. N., K. B. Kulasekera, and W. J. Padgett (2010): "Density Estimation Using Asymmetric Kernels and Bayes Bandwidths with Censored Data," *Journal of Statistical Planning and Inference*, 140, 1765 - 1774.

[26] Liebscher, E. (1996): "Strong Convergence of Sums of α-Mixing Random Variables with Applications to Density Estimation," *Stochastic Processes and Their Applications*, 65, 69 - 80.

[27] Malec, P., and M. Schienle (2012): "Nonparametric Kernel Density Estimation Near the Boundary," SFB Discussion Paper 2012-047.

[28] Nakagami, M. (1943): "Some Statistical Characters of Short-Wave Fading" (in Japanese), *Journal of the Institute of Electrical Communication Engineers of Japan*, 27, 145 - 150.

[29] Nakagami, M. (1960): "The m-Distribution - A General Formula of Intensity Distribution of Rapid Fading," in W. C. Hoffman (ed.), *Statistical Methods in Radio Wave Propagation: Proceedings of a Symposium Held at the University of California, Los Angeles, June 18 - 20, 1958.* New York: Pergamon Press, 3 - 36.

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

[31] Sheather, S. J., and M. C. Jones (1991): "A Reliable Data-Based Bandwidth Selection Method for Kernel Density Estimation," *Journal of the Royal Statistical Society, Series B*, 53, 683 - 690.

[32] Silverman, B. W. (1986): *Density Estimation for Statistics and Data Analysis.* London: Chapman & Hall.

[33] Stacy, E. W. (1962): "A Generalization of the Gamma Distribution," *Annals of Mathematical Statistics*, 33, 1187 - 1192.

[34] Stone, C. J. (1980): "Optimal Rates of Convergence for Nonparametric Estimators," *Annals of Statistics*, 8, 1348 - 1360.

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

Figure 1: Shapes of Generalized Gamma Kernels When $b = 0.2$

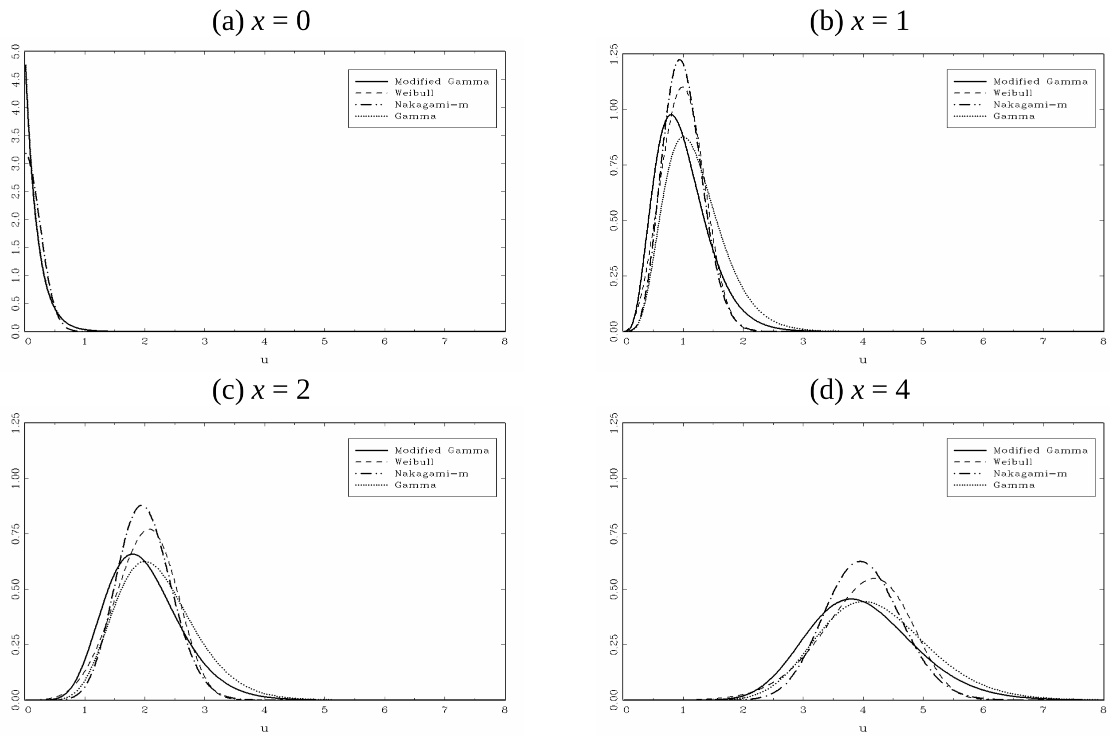

Figure 2: Generalized Gamma Kernel Density Estimates When the True Distribution Is $G\,(0.75, 1.25)$

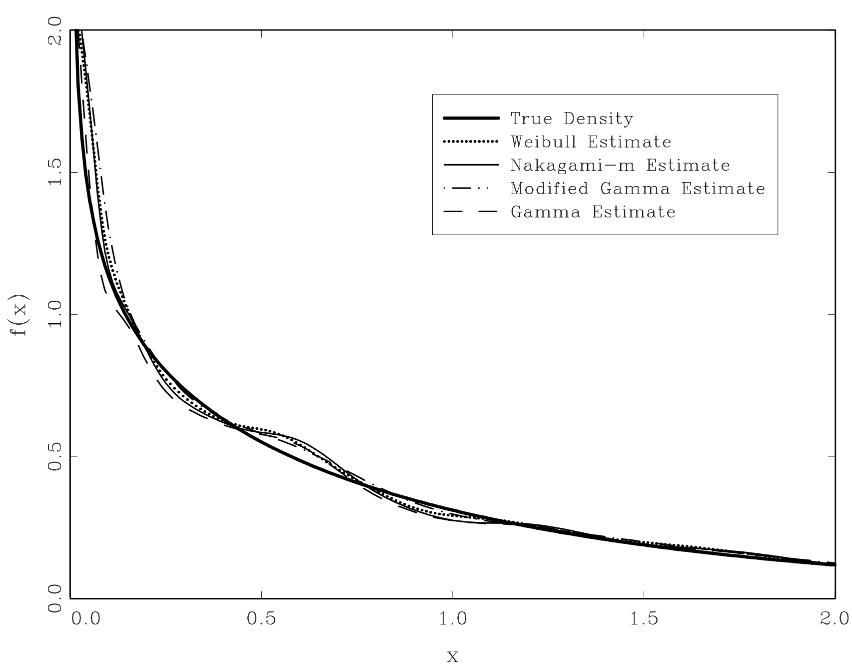

Table 1: Averages of Performance Measures and Smoothing Parameter Values

	$n = 100$		$n = 200$		$n = 500$	
	RISE	$\hat{b}$	RISE	$\hat{b}$	RISE	$\hat{b}$
			1. Gamma			
W	0.0356 (0.0098)	0.2778	0.0294 (0.0081)	0.1701	0.0221 (0.0057)	0.0897
NM	0.0368 (0.0091)	0.3007	0.0306 (0.0076)	0.1861	0.0232 (0.0055)	0.0975
MG	0.0358 (0.0125)	0.1404	0.0290 (0.0098)	0.0962	0.0220 (0.0066)	0.0601
G	0.0362 (0.0112)	0.1712	0.0289 (0.0088)	0.1105	0.0211 (0.0059)	0.0683
			2. Weibull			
W	0.0374 (0.0119)	0.1870	0.0297 (0.0090)	0.1228	0.0214 (0.0058)	0.0809
NM	0.0385 (0.0116)	0.2090	0.0307 (0.0088)	0.1382	0.0222 (0.0058)	0.0911
MG	0.0368 (0.0140)	0.1137	0.0294 (0.0103)	0.0813	0.0218 (0.0069)	0.0526
G	0.0367 (0.0127)	0.1272	0.0286 (0.0092)	0.0915	0.0204 (0.0060)	0.0634
			3. Half-Normal			
W	0.0274 (0.0113)	0.3445	0.0225 (0.0083)	0.2745	0.0172 (0.0056)	0.1936
NM	0.0251 (0.0120)	0.3809	0.0207 (0.0088)	0.3152	0.0158 (0.0059)	0.2354
MG	0.0327 (0.0112)	0.1750	0.0262 (0.0087)	0.1325	0.0193 (0.0057)	0.0911
G	0.0303 (0.0117)	0.2662	0.0245 (0.0087)	0.2039	0.0184 (0.0060)	0.1380
			4. Log-Normal			
W	0.0429 (0.0153)	0.0830	0.0332 (0.0110)	0.0654	0.0242 (0.0080)	0.0471
NM	0.0447 (0.0152)	0.0932	0.0343 (0.0108)	0.0749	0.0245 (0.0078)	0.0562
MG	0.0458 (0.0150)	0.0535	0.0360 (0.0108)	0.0390	0.0263 (0.0075)	0.0261
G	0.0416 (0.0158)	0.0624	0.0324 (0.0114)	0.0480	0.0238 (0.0081)	0.0334
			5. Burr			
W	0.0582 (0.0174)	0.0925	0.0467 (0.0137)	0.0561	0.0334 (0.0092)	0.0342
NM	0.0600 (0.0165)	0.1020	0.0485 (0.0133)	0.0618	0.0347 (0.0091)	0.0376
MG	0.0581 (0.0207)	0.0522	0.0465 (0.0158)	0.0356	0.0343 (0.0106)	0.0225
G	0.0581 (0.0196)	0.0584	0.0453 (0.0147)	0.0395	0.0319 (0.0097)	0.0267
			6. Generalized Gamma			
W	0.0363 (0.0131)	0.0649	0.0290 (0.0094)	0.0473	0.0213 (0.0063)	0.0322
NM	0.0346 (0.0137)	0.0884	0.0276 (0.0099)	0.0658	0.0202 (0.0066)	0.0460
MG	0.0355 (0.0140)	0.0425	0.0282 (0.0100)	0.0317	0.0206 (0.0067)	0.0220
G	0.0346 (0.0138)	0.0452	0.0276 (0.0099)	0.0334	0.0202 (0.0067)	0.0232

Note: "W", "NM", "MG", and "G" denote density estimators using the Weibull, Nakagami-m, Modified Gamma, and Gamma kernels, respectively. "RISE" and "$\hat{b}$" are the root integrated squared error and the smoothing parameter that minimizes the RISE, respectively. Numbers in parentheses are simulation standard deviations.