UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE CIÊNCIAS ECONÔMICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

HUDSON DA SILVA TORRENT

# ESTIMAÇÃO NÃO-PARAMÉTRICA E SEMI-PARAMÉTRICA DE FRONTEIRAS DE PRODUÇÃO

Porto Alegre
2010

**HUDSON DA SILVA TORRENT**

# ESTIMAÇÃO NÃO-PARAMÉTRICA E SEMI-PARAMÉTRICA DE FRONTEIRAS DE PRODUÇÃO

Tese submetida ao Programa de Pós-Graduação em Economia da Faculdade de de Ciências Econômicas da UFRGS, como quesito parcial para obtenção do título de Doutor em Economia, com ênfase em Economia Aplicada.

Orientador: Prof. Dr. Flávio A. Ziegelmann

**Porto Alegre**
**2010**

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
Responsável: Biblioteca Gládis W. do Amaral, Faculdade de Ciências Econômicas da UFRGS

T692e Torrent, Hudson da Silva
Estimação não-paramétrica e semi-paramétrica de fronteiras de produção / Hudson da Silva Torrent. – Porto Alegre, 2010.
120 f. : il.

Orientador: Flávio A. Ziegelmann.

Ênfase em Economia Aplicada.

Tese (Doutorado em Economia) - Universidade Federal do Rio Grande do Sul, Faculdade de Ciências Econômicas, Programa de Pós-Graduação em Economia, Porto Alegre, 2010.

1. Séries temporais : Modelo não-paramétrico : Modelo semi-paramétrico I. Ziegelmann, Flávio Augusto. II. Universidade Federal do Rio Grande do Sul. Faculdade de Ciências Econômicas. Programa de Pós-Graduação em Economia. III. Título.

CDU 519.234

**HUDSON DA SILVA TORRENT**

# ESTIMAÇÃO NÃO-PARAMÉTRICA E SEMI-PARAMÉTRICA DE FRONTEIRAS DE PRODUÇÃO

Tese submetida ao Programa de Pós-Graduação em Economia da Faculdade de de Ciências Econômicas da UFRGS, como quesito parcial para obtenção do título de Doutor em Economia, com ênfase em Economia Aplicada.

**Aprovada em 07 de Maio de 2010**

**Banca Examinadora**

Prof. Dr. Marcelo Cunha Medeiros
PUCRJ

Prof. Dr. Ronaldo Dias
UNICAMP

Prof. Dr. Marcelo Savino Portugal
PPGE/UFRGS

Prof. Dr. Flávio A. Ziegelmann - Orientador
PPGE/UFRGS

**Porto Alegre**
**2010**

A Dimas, Beth, Felipe e Alessandra.

# Agradecimentos

Agradeço aos meus pais, Dimas e Beth, e ao meu irmão, Felipe, pelo amor e carinho. E pelo apoio e vibração sempre demonstrados, apesar da saudade incessante que habita nossas vidas.

A Alessandra pelo amor, compreensão e apoio. E por ter estado sempre ao meu lado alegrando e iluminando a minha vida.

A Paulo, Sueli e Adriano por serem para mim uma família.

Ao meu orientador, professor Flávio, por ter ampliado meu horizonte acadêmico, por ter me incentivado sempre, pela convivência sempre agradável e pelo exemplo de pessoa e profissional que é para mim.

Ao professor Carlos Martins pela oportunidade, por ter ampliado meu horizonte acadêmico e pela acolhida durante o sanduíche.

Aos demais professores que fizeram parte desta caminhada e que me influenciaram de diversas formas.

Aos colegas e amigos que me ajudaram em vários momentos. Em especial ao amigo João Frois.

À CAPES e ao CNPq pelo suporte financeiro, sem o qual nada disso seria possível.

# Resumo

Existe uma grande e crescente literatura sobre especificação e estimação de fronteiras de produção e, portanto, de eficiência de unidades produtivas. Nesta tese, o foco está sobre modelos de fronteiras determinísticas, os quais são baseados na hipótese de que os dados observados pertencem ao conjunto tecnológico. Dentre os modelos estatísticos e estimadores para fronteiras determinísticas existentes, uma abordagem promissora é a adotada por Martins-Filho e Yao (2007). Esses autores propõem um procedimento de estimação composto por três estágios. Esse estimador é de fácil implementação, visto que envolve procedimentos não-paramétricos bem conhecidos. Além disso, o estimador possui características desejáveis vis-à-vis estimadores para fronteiras determinísticas tradicionais como DEA e FDH. Nesta tese, três artigos, que melhoram o modelo proposto por Martins-Filho e Yao (2007), são propostos. No primeiro artigo, o procedimento de estimação desses autores é melhorado a partir de uma variação do estimador exponencial local, proposto por Ziegelmann (2002). Demonstra-se que estimador proposto é consistente e assintoticamente normal. Além disso, devido ao estimador exponencial local, estimativas potencialmente negativas para a função de variância condicional, que poderiam prejudicar a aplicabilidade do estimador proposto por Martins-Filho e Yao, são evitadas. No segundo artigo, é proposto um método original para estimação de fronteiras de produção em apenas dois estágios. É mostrado que se pode eliminar o segundo estágio proposto por Martins-Filho e Yao, assim como, eliminar o segundo estágio proposto no primeiro artigo desta tese. Em ambos os casos, a estimação do mesmo modelo de fronteira de produção requer três estágios, sendo versões diferentes para o segundo estágio. As propriedades assintóticas do estimador proposto são analisadas, mostrando-se consistência e normalidade assintótica sob hipóteses razoáveis. No terceiro artigo, é proposta uma variação semi-paramétrica do modelo estudado no segundo artigo. Reescreve-se aquele modelo de modo que se possa estimar a fronteira de produção e a eficiência de unidades produtivas no contexto de múltiplos insumos, sem incorrer no *curse of dimensionality.* A abordagem adotada coloca o modelo na estrutura de modelos aditivos, a partir de hipóteses sobre como os insumos se combinam no processo produtivo. Em particular, considera-se aqui os casos de insumos aditivos e insumos multiplicativos, os quais são amplamente considerados em teoria econômica e aplicações. Estudos de Monte Carlo são apresentados em todos os artigos, afim de elucidar as propriedades dos estimadores propostos em amostras finitas. Além disso, estudos com dados reais são apresentados em todos os artigos, nos quais são estimados *rankings* de eficiência para uma amostra de departamentos policiais dos EUA, a partir de dados sobre criminalidade daquele país.

**Palavras-chave:** Modelos não-paramétricos de fronteira. Regressão exponencial local. Regressão linear local. Modelos aditivos. Regressão semi-paramétrica. *Classical Backfitting. Smooth Backfitting.*

# Abstract

There exists a large and growing literature on the specification and estimation of production frontiers and therefore efficiency of production units. In this thesis we focus on deterministic production frontier models, which are based on the assumption that all observed data lie in the technological set. Among the existing statistical models and estimators for deterministic frontiers, a promising approach is that of Martins-Filho and Yao (2007). They propose an estimation procedure that consists of three stages. Their estimator is fairly easy to implement as it involves standard nonparametric procedures. In addition, it has a number of desirable characteristics vis-à-vis traditional deterministic frontier estimators as DEA and FDH. In this thesis we propose three papers that improve the model proposed in Martins-Filho and Yao (2007). In the first paper we improve their estimation procedure by adopting a variant of the local exponential smoothing proposed in Ziegelmann (2002). Our estimator is shown to be consistent and asymptotically normal. In addition, due to local exponential smoothing, potential negativity of conditional variance functions that may hinder the use of Martins-Filho and Yao's estimator is avoided. In the second paper we propose a novel method for estimating production frontiers in only two stages. There we show that we can eliminate the second stage of Martins-Filho and Yao as well as of our first paper, where estimation of the same frontier model requires three stages under different versions for the second stage. We study asymptotic properties showing consistency and $\sqrt{nh_n}$ asymptotic normality of our proposed estimator under standard assumptions. In the third paper we propose a semiparametric variation of the frontier model studied in the second paper. We rewrite that model allowing for estimating the production frontier and efficiency of production units in a multiple input context without suffering the *curse of dimensionality*. Our approach places that model within the framework of additive models based on assumptions regarding the way inputs combine in production. In particular, we consider the cases of additive and multiplicative inputs, which are widely considered in economic theory and applications. Monte Carlo studies are performed in all papers to shed light on the finite sample properties of the proposed estimators. Furthermore a real data study is carried out in all papers, from which we rank efficiency within a sample of USA Law Enforcement agencies using USA crime data.

**Keywords:** Nonparametric frontier models. Local exponential regression. Local linear regression. Additive models. Semiparametric regression. Classical Backfitting. Smooth Backfitting.

# Sumário

# 1 Introdução

Estimação de fronteiras de produção e, consequentemente, de eficiência (ou ineficiência) de processos produtivos tem sido objeto de pesquisa de uma vasta e crescente literatura (Simar and Wilson (2007)). O problema em questão pode ser explicitado da seguinte maneira: seja $x \in \mathbb{R}_+^p$ um conjunto de insumos usados para produzir um único produto $y \in \mathbb{R}_+$[1] sujeito a uma tecnologia ou conjunto de produção definido como $\Psi = \{(y, x) \in \mathbb{R}_+ \times \mathbb{R}_+^p \mid x \text{ produz } y\}$. Uma fronteira de produção associada a $\Psi$ é definda como $\rho(x) = \sup\{y \in \mathbb{R}_+ \mid (y, x) \in \Psi\}$ para todo $x \in \mathbb{R}_+^p$. Portanto, dada uma amostra aleatória, $\chi_n = \{(Y_i, X_i)\}_{i=1}^n$, correspondente a produto e insumos de $n$ firmas que operam sob a mesma tecnologia $\Psi$, deseja-se estimar a fronteira de produção $\rho(x)$ para todo $x \in \mathbb{R}_+^p$. Além disso, para um dado par $(Y_i, X_i) \in \Psi$, a eficiência da firma $i$, $R_i$, é medida por $0 \leq R_i = \frac{Y_i}{\rho(X_i)} \leq 1$. Logo, uma vez que a fronteira esteja estimada, a eficiência também estará.

Na literatura corrente, há duas principais formas de se estimar fronteiras de eficiência. Uma delas é conhecida como Fronteira Determinística, que está fundamentada na hipótese de que os dados observados estão contidos no conjunto de produção, ou seja, $P((Y_i, X_i) \in \Psi) = 1$ para todo $i$, onde $P$ é uma medida de probabilidade. Nesses modelos, qualquer diferença entre o produto realizado $Y_i$ e a fronteira $\rho(X_i)$ é atribuída a ineficiências não observáveis da firma $i$. A segunda forma de se estimar $\rho(\cdot)$ é conhecida como Fronteira Estocástica, pioneiramente desenvolvida por Aigner et al. (1977) e Meeusen e van den Broeck (1977). A abordagem estocástica assume a possibilidade de choques aleatórios no processo de produção e, consequentemente, $P((Y_i, X_i) \notin \Psi) > 0$. Embora a abordagem estocástica pareça ser mais interessante do ponto de vista econométrico, uma limitação existente é que a identificação dos modelos de fronteiras estocásticas exige hipóteses sobremaneira restritivas com respeito à parametrização da distribuição conjunta de $(Y_i, X_i)$ e/ou $\rho(\cdot)$. Essas hipóteses com respeito à parametrização podem levar à má especificação de $\rho(\cdot)$ e invalidar quaisquer propriedades ótimas derivadas dos estimadores propostos (geralmente estimadores de máxima verossimilhança) e, consequentemente, levar a inferências errôneas. Além disso, Baccouche e Kouki (2003) mostraram que as estimativas para os níveis de eficiência e para os rankings de eficiência das firmas são sensíveis à especificação da densidade conjunta de $(Y_i, X_i)$. Consequentemente, diferentes especificações para a densidade podem levar a diferentes conclusões a respeito

[1]Neste trabalho será considerado o caso de um único produto.

da tecnologia e eficiência para uma dada amostra aleatória. Essas deficiências dos modelos de fronteira estocástica têm contribuído para a popularidade das fronteiras determinísticas.

No contexto determinístico, há duas formas bastante populares para se estimar $\rho(\cdot)$. O estimador Free Disposal Hull (FDH), introduzido por Deprins et al. (1984) e o estimador Data Envelopment Analysis (DEA), apresentado em Charnes et al. (1978). Muitos trabalhos aplicam as metodologias FDH e DEA na estimação de fronteiras de produção e estimação de eficiência em uma grande variedade de problemas econômicos. A ideia é estimar o conjunto de produção a partir de uma amostra aleatória observada sem, contudo, assumir alguma estrutura paramétrica sobre a fronteira de produção ou sobre a densidade conjunta de $(Y_i, X_i)$. Nesta seção são apresentadas as versões resumidas desses estimadores seguindo Gijbels et al. (1999) and Park et al. (2000), que obtiveram as distribuições assintóticas para os estimadores DEA e FDH, respectivamente. O estimador FDH é estabelecido sob a hipótese de *free disposability* no conjunto de produção, i.e., se $(x, y) \in \Psi$ então todos os pares $(x', y')$ tais que $x' \geq x$ e $y' \leq y$ pertencem a $\Psi$. O estimador para $\Psi$ é definido como o envelope do conjunto $\chi$, que respeita a hipótese de free disposability:

$$\hat{\Psi}_{FDH} = \{(x, y) \in \mathbb{R}_+^{p+q} | y \leq y_i, x \geq x_i, (x_i, y_i) \in \chi\} \ .$$

Este é o menor conjunto (free disposal) que contém todas as observações. Para o estimador DEA uma hipótese mais forte é assumida: free disposability está associada com a hipótese de que $\Psi$ é convexo. O estimador DEA é então definido como:

$$\begin{aligned} \hat{\Psi}_{DEA} = \{(x, y) \in \mathbb{R}_+^{p+q} | y \leq \sum_{i=1}^{n} \gamma_i y_i; x \geq \sum_{i=1}^{n} \gamma_i x_i \text{ for } (\gamma_1, ..., \gamma_n), \\ \sum_{i=1}^{n} \gamma_i = 1; \gamma_i \geq 0, i = 1, ..., n\}. \end{aligned}$$

Este é o menor conjunto convexo (free disposal) contendo todas as observações. Dadas as hipóteses sobre o conjunto de produção, pode-se estimar a fronteira de produção. O estimador FDH produz uma fronteira de produção monótona não-decrescente, enquanto que o estimador DEA produz uma fronteira côncava e monótona não-decrescente. Embora populares, esses estimadores apresentam algumas características indesejáveis. Do ponto de vista estatístico, a fronteira de produção pode ser vista como a envoltória superior de $\Psi$. Consequentemente, esses estimadores apresentam viés negativo para o estimador da fronteira, visto que a fronteira estimada sempre pertence ao conjunto de produção e, portanto, nunca excede a verdadeira fronteira de produção. Afim de se eliminar o viés, algum estimador que o corrija deverá ser implementado (ver Gijbels et al. (1999) and Park et al. (2000)). Além disso, o estimador FDH produz uma função descontínua (*step function*), enquanto o estimador DEA produz uma função linear por partes (*piecewise linear*).

Dentre os modelos estatísticos e estimadores para fronteiras determinísticas, uma abordagem promissora é apresentada em Martins-Filho and Yao (2007). Esses autores assumem que o produto $Y_i$ é gerado por

$$Y_i = \frac{\sigma(X_i)}{\sigma_R} R_i \text{ for } i = 1, 2, \cdots, n \;, \tag{1.1}$$

onde $R_i$ é uma variável aleatótria não observável representando a eficiência e tomando valores no intervalo $[0, 1]$, $X_i$ é um vetor aleatório representando insumos e tomando valores em $\mathbb{R}^p_+$, $\sigma(x) : \mathbb{R}^p_+ \to (0, \infty)$ é uma função mensurável, $\sigma_R$ é um parâmetro desconhecido e a fronteira de produção é dada por $\rho(x) \equiv \frac{\sigma(x)}{\sigma_R}$. Nesse modelo, quanto maior $R_i$ mais eficiente será a unidade produtiva, pois mais próximo o produto realizado estará da fronteira de produção. Os autores assumem que $E(R_i|X_i = x) \equiv \mu_R$ onde $0 < \mu_R < 1$ e $V(R_i|X_i = x) \equiv \sigma_R^2$. Aqui o parâmetro $\mu_R$ é interpretado como a eficiência média para um dado nível utilizado de insumo e uma dada tecnologia $\Psi$. $\sigma_R$ é um parâmetro de escala para a distribuição condicional de $R_i$ que também posiciona a fronteira de produção. Já o formato da fronteira é capturado por $\sigma(x)$. Essas restrições sobre os momentos condicionais de $R_i$ juntamente com a Eq. (1.1) implicam que $E(Y_i|X_i = x) = \frac{\mu_R}{\sigma_R}\sigma(x)$ e $V(Y_i|X_i = x) = \sigma^2(x)$. Logo, o modelo pode ser reescrito como,

$$Y_i = b\sigma(X_i) + \sigma(X_i)\frac{(R_i - \mu_R)}{\sigma_R} = m(X_i) + \sigma(X_i)\epsilon_i, \tag{1.2}$$

onde $b = \frac{\mu_R}{\sigma_R}$, $\epsilon_i = \frac{R_i - \mu_R}{\sigma_R}$, $m(X_i) = b\sigma(X_i)$, $E(\epsilon_i|X_i = x) = 0$ e $V(\epsilon_i|X_i = x) = 1$.[2]

Martins-Filho e Yao (2007) propõem estimar o modelo (1.2) em três estágios: primeiramente, $m(x)$ é estimado usando o estimador linear local proposto por Fan (1992); no segundo estágio, o quadrado dos resíduos do primeiro estágio são usados para estimar a variância condicional através de estimação linear local, como proposto em Fan e Yao (1998); no terceiro estágio a variância condicional estimada no estágio 2 é usada para estimar $\sigma_R$ baseada na hipótese de que uma firma é eficiente. Esse estimador, originalmente chamado NP, coaduna da flexibilidade da estrutura não-paramétrica, mas também possui propriedades desejáveis extras em comparação aos estimadores DEA e FDH: *i*) O estimador NP é mais robusto a valores extremos; *ii*) o estimador para a fronteira é uma função suave (*smooth*) dos insumos (ou seja, não é descontínua nem linear por partes) e *iii*) embora envelope os dados, o estimador não é

[2] Afim de simplificar a notação, $E(\cdot|X_i = x)$ ou $V(\cdot|X_i = x)$ será denotado simplesmente por $E(\cdot|X_i)$ ou $V(\cdot|X_i)$.

inerentemente viesado como os estimadores DEA e FDH. *iv*) O método de estimação é relativamente simples, visto que é baseado em estimação kernel linear local. Além disso, Martins-Filho e Yao (2007) derivam a normalidade assintótica e a consistência para a fronteira de produção e para os estimadores de eficiência sob hipóteses razoáveis no contexto não-paramétrico.

Apesar das propriedades desejáveis, a implementação do segundo estágio do estimador NP pode ser problemática. Pois da forma como é proposto, o estimador pode produzir valores negativos para a variância condicional. Quando isso ocorre, a implementação do procedimento de estimação não é possível sem que algum critério arbitrário seja adotado para a obtenção de $\sigma(x)$ (raiz quadrada de $\sigma^2(x)$). No capítulo 2 deste trabalho, o modelo e a estratégia de estimação de Martins-Filho e Yao (2007) serão adotados. Contudo, propõe-se modificar esse estimador através da utilização do estimador local exponencial no segundo estágio, como proposto em Ziegelmann (2002). Ou seja, a variância condicional será estimada através do estimador exponencial local, garantindo sua não-negatividade. É importante enfatizar que todas as vantagens do estimador NP citadas anteriormente são mantidas no procedimento proposto. Devido à regressão exponencial, o estimador proposto será chamdo de NPE. Um estudo de Monte Carlo é considerado com o objetivo de avaliar a performance em amostra finita dos estimadores. No estudo considerado, o estimador NPE apresentou melhor performance vis-à-vis o estimador NP. Além disso, o comportamento assintótico do estimador NPE é completamente estabelecido sob hipóteses bastante razoáveis no contexto de estimação não-paramétrica. Os resultados assintóticos obtidos são comparados aos resultados obtidos por Martins-Filho e Yao (2007).

Embora os estimadores NPE e NP possuam vantagens em comparação com os estimadores FDH e DEA, alguns pontos ainda podem ser melhorados. Aqueles estimadores são caracterizados por um procedimento de estimação em três estágios. Os primeiros dois estágios fornecem o formato da fronteira, enquanto que o terceiro é reponsável por posicioná-la. É importante notar ainda que o segundo estágio em ambos estimadores é uma regressão que possui como regressando resíduos (ao quadrado) do primeiro estágio. Essa característica pode ser indesejável, principalmente em aplicações que consideram amostras relativamente pequenas. Contudo, no capítulo 3 deste trabalho será mostrado que pode-se eliminar o segundo estágio dos estimadores NPE e NP, estimando, portanto, a fronteira de produção em apenas dois estágios. Especificamente, o estimador proposto, que chamaremos de NP2S, possui como primeiro estágio exatamente o primeiro estágio dos estimadores NPE e NP. Já o segundo estágio é bastante similar ao terceiro estágio desses estimadores. Portanto, o segundo estágio dos estimadores NPE e NP é de fato eliminado e a fronteira de produção é estimada em apenas dois estágios. A contribuição descrita no capítulo 3 é baseada na idéia de que o formato da fronteira pode ser obtido no primeiro estágio e, então, um estágio adicional é necessário para posicionar a fronteira estimada. Em particular, no modelo (1.2), nota-se que $m(X_i) \equiv \rho(X_i)$. Portanto, ao estimar $m(X_i)$ obtem-se $\hat{m}(x) = \mu_R \hat{\rho}(x)$, visto que $\mu_R$ não depende de $X_i$. Logo, se um estimador para $\mu_R$ estiver disponível, a fronteira poderá ser estimada como $\hat{\rho}(X_i) = \frac{\hat{m}(X_i)}{\hat{\mu}_R}$. Como dito anteriormente, o primeiro estágio consiste em estimar $m(x)$ usando o estimador linear local. Já o segundo estágio, a média condicional estimada no primeiro estágio é usada para estimar $\mu_R$ baseada na hipótese de que uma firma é eficiente. Consequentemente, aliada a simplicidade, o estimador NP2S mantem as vantagens dos estimadores NPE e NP vis-à-vis os estimadores DEA e FDH listadas anteriormente. Um estudo de Monte Carlo é considerado com o objetivo de avaliar a performance em amostra finita do estimador NP2S em comparação com os estimadores NPE e NP. Além disso, o comportamento assintótico do estimador NP2S é completamente estabelecido sob hipóteses bastante razoáveis no contexto de estimação não-paramétrica.

Os estimadores supracitados são estabelecidos considerando-se o caso de múltiplos insumos, i.e., $x \in \mathbb{R}^p$, with $p \geq 1$. Entretanto, por serem estimadores não-paramétricos, todos eles sofrem de perda de qualidade das estimativas à medida que cresce o número de covariáveis consideradas, ou seja, sofrem do amplamente conhecido fenômeno *curse of dimensionality*. A perda de performance de estimadores totalmente não-paramétricos no contexto multivariado limita a aplicabilidade desses estimadores, visto que, em inúmeras situações, o pesquisador estará interessado em estimar eficiência considerando mais de um único insumo. Uma estratégia para mitigar esse problema é a utilização de modelos aditivos. A adoção desses modelos se justifica pela redução da dimensão da componente não-paramétrica, reduzindo o *curse of dimensionality*. No contexto de estimação de fronteira de produção, o estimador NP2S possui uma estrutura conveniente para aplicação de modelos aditivos, visto que assumir que $m(X_i)$ é aditivo é equivalente a assumir que a fronteira de produção, $\rho(X_i)$, é aditiva. É importante ressaltar que a aplicação de modelos aditivos não é tão direta se considerados os estimadores NPE e NP. Nesses casos, modelos aditivos estariam baseados na hipótese de que a função de variância condicional, $\sigma^2(X_i)$, é aditiva. Consequentemente, não está claro, nesse caso, qual estrutura está sendo admitida para a fronteira de produção, $\rho(X_i)$. A extensão do estimador NP2S para o caso de múltiplos insumos sob a hipótese de aditividade da fronteira de produção é objeto do capítulo 4 deste trabalho. Além disso, é estabelecido no capítulo 4, a extensão do estimador NP2S para o caso de múltiplos insumos sob a hipótese de fronteira de produção multiplicativa no contexto de estimação de modelos aditivos. Essa extensão é particularmente importante, visto que a teoria econômica considera amplamente tecnologias em que os insumos se combinam de forma multiplicativa. No que diz respeito à estimação do modelo aditivo supracitado, há quatro estimadores disponíveis na literatura: o *Classical Backfitting estimator* (CBE), proposto por Buja et al. (1989) (ver também Hastie e Tibishirani (1990)); o *Marginal Integration estimator* (MIE), proposto por Newey (1994), Tjøstheim e Auestad (1994) e Linton e Nielsen (1995); um estimador de dois estágios (2SE), que combina um estágio utilizando CBE e outro utilizando MIE, proposto por Linton (1997) e Kim et al. (1999); e o Smooth Backfitting estimator (SBE), proposto por Mammen et al. (1999). O mais tradicional desses estimadores é o CBE. Esse estimador possui destaque em relação aos seus concorrentes por sua performance em simulações e aplicações. Por outro lado, o SBE se destaca, em relação aos seus concorrentes, por sua superioridade no que diz respeito a resultados teóricos. Afim de analisar as propriedades em amostra finita do estimador para fronteira de produção proposto, um estudo de Monte Carlo é considerado no capítulo 4 em que CBE e SBE são os estimadores do modelo aditivo em questão.

# 2 Nonparametric frontier estimation via local exponential regression

Carlos Martins-Filho,[1] Hudson Torrent[2] and Flavio Ziegelmann[3]

January, 2010

**Abstract.** In this paper we consider the estimation of a nonparametric frontier model first proposed in Martins-Filho and Yao (2007). We improve their estimation procedure by adopting a variant of the local exponential smoothing proposed in Ziegelmann (2002). Our estimator is shown to be consistent and asymptotically normal. In addition, due to local exponential smoothing, potential negativity of conditional variance functions that may hinder the use of Martins-Filho and Yao's estimator is avoided. A Monte Carlo study is performed to shed light on the finite sample properties of the estimator and to contrast its performance with that of the estimator proposed in Martins-Filho and Yao (2007). We find that there can be significant improvements in finite sample performance when using exponential smoothing in this context.

**Keywords and phrases.** nonparametric frontier models; local exponential smoothing; local exponential regression.

**JEL Classifications.** C14, C22

[1]Department of Economics, University of Colorado, Boulder, CO 80309-0256, USA, email: carlos.martins@colorado.edu and IFPRI, 2033 K Street NW, Washington, DC 20006-1002, USA, email: c.martins-filho@cgiar.org

[2]Department of Economics, Federal University of Rio Grande do Sul, Porto Alegre, RS 90040-000, Brazil, email: hudsontorrent@ppge.ufrgs.br

[3]Department of Statistics, Federal University of Rio Grande do Sul, Porto Alegre, RS 90040-000, Brazil, email: flavioz@ufrgs.br and Department of Statistics, London School of Economics, London, WC2A 2AE UK.

## 2.1 Introduction

There exists a large and growing literature on the specification and estimation of production frontiers (Simar and Wilson, 2007). Let $\Psi = \{(x, y) \in \Re^{p+1}_{+} : x \text{ can produce } y\}$ be a technology where $x \in \Re^{p}_{+}$ is a vector of inputs used to produce an output $y \in \Re_{+}$. The production frontier associated with $\Psi$ is defined as $\rho(x) = \sup\{y \in \Re_{+} : (x, y) \in \Psi\}$ for all $x \in \Re^{p}_{+}$. Given a sample of $n$ realized production plans (or production units) $\chi_n = \{(X_i, Y_i)\}^{n}_{i=1}$, which share the technology $\Psi$, the main objective of this literature is to estimate $\rho(x)$ for all $x \in \Re^{p}_{+}$. For any given production plan $(X_i, Y_i) \in \Psi$, we define its (inverse) Farrell efficiency as $0 \leq R_i = \frac{Y_i}{\rho(X_i)} \leq 1$. Once an estimate of $\rho$ is available, estimated efficiencies can be readily.

There exists two main statistical approaches for modeling production frontiers. The deterministic approach is based on the assumption that all observed data lie in $\Psi$, i.e., $P((X_i, Y_i) \in \Psi) = 1$ for all $i$, where $P$ is a probability measure. In these models, any deviation of realized output $Y_i$ from $\rho(X_i)$ is attributable to unobserved inefficiencies of the production plan $i$. The stochastic approach allows for random shocks to the production process. As a result observed output $Y_i$ at any input level can be smaller or larger than $\rho(X_i)$. As a result, it may be that $P((X_i, Y_i) \notin \Psi) > 0$ for some $i$. Although more appealing from an econometric perspective, separating inefficiency and random shock in stochastic frontier models requires strong assumptions on the joint density of $(X_i, Y_i)$ (Aigner et al., 1977; Fan et al., 1996; Kumbhakar et al., 2007, Martins-Filho and Yao, 2010). In contrast, deterministic frontier models can be estimated under much milder restrictions on the stochastic process generating $\chi_n$ (Aragon et al., 2005; Martins-Filho and Yao, 2007; Martins-Filho and Yao, 2008; Daouia et al., 2009).

Among the existing statistical models and estimators for deterministic frontiers, a promising approach is that of Martins-Filho and Yao (2007). They assume that output $Y_i$ is generated by

$$Y_i = \frac{\sigma(X_i)}{\sigma_R} R_i \text{ for } i = 1, 2, \cdots, n \tag{2.1}$$

where $R_i$ is an unobserved random variable representing efficiency and taking values in the interval $[0, 1]$, $X_i$ is an observed random vector representing inputs taking values in $\Re^{p}_{+}$, $\sigma(x) : \Re^{p}_{+} \rightarrow (0, \infty)$ is a measurable function, $\sigma_R$ is an unknown parameter and the production frontier is given by $\rho(x) \equiv \frac{\sigma(x)}{\sigma_R}$. In this model $R_i$ has the effect of contracting output from optimal levels that lie on the production frontier. The larger $R_i$ the more efficient the production unit because the closer the realized output is to that on the production frontier. They assume that $E(R_i|X_i = x) \equiv \mu_R$ where $0 < \mu_R < 1$ and

$V(R_i|X_i = x) \equiv \sigma_R^2$. Here, the parameter $\mu_R$ is interpreted as a mean efficiency given input usage and the common technology $\Psi$ and $\sigma_R$ is a scale parameter for the conditional distribution of $R_i$ that also locates the production frontier. Its shape is captured by $\sigma(x)$. These conditional moment restrictions together with equation (1) imply that $E(Y_i|X_i = x) = \frac{\mu_R}{\sigma_R}\sigma(x)$ and $V(Y_i|X_i = x) = \sigma^2(x)$. The model can therefore be rewritten as,

$$Y_i = b\sigma(X_i) + \sigma(X_i)\frac{(R_i - \mu_R)}{\sigma_R} = m(X_i) + \sigma(X_i)\epsilon_i \tag{2.2}$$

where $b = \frac{\mu_R}{\sigma_R}$, $\epsilon_i = \frac{R_i - \mu_R}{\sigma_R}$, $m(X_i) = b\sigma(X_i)$, $E(\epsilon_i|X_i = x) = 0$ and $V(\epsilon_i|X_i = x) = 1$.[1]

Martins-Filho and Yao propose an estimation procedure that consists of three stages: first, $m(x)$ is estimated using the local linear estimator of Fan (1992); second, squared residual from the first stage are used in a local linear procedure to estimate the conditional variance $\sigma^2(x)$; third, the estimated conditional variance from stage 2 is used to estimate $\sigma_R$ based on an anchoring assumption to be discussed in section 2. Their estimator is fairly easy to implement as it involves standard nonparametric procedures. In addition, the frontier estimator has a number of desirable characteristics: first, contrary to the estimators in Aragon et al. (2005), Daouia et al. (2009) and Martins-Filho and Yao (2008), it is a smooth function of input usage; second, although the frontier estimator envelops the data, it is not intrinsically biased as the popular DEA (data envelopment analysis) and FDH (free disposal hull) estimators, therefore no bias correction is needed; third, the estimator is fairly robust to outliers and extreme values. In addition to all of these desirable properties, the estimation procedure leads to a frontier estimator that is consistent and asymptotically normal when suitably centered and normalized.

In spite of these desirable properties, the implementation of the second stage of the estimation can be problematic. The difficulty arises because local linear estimation of the conditional variance function may produce negative estimated conditional variances. When this is the case, implementation of the procedure is not possible unless arbitrary criteria are adopted to provide viable estimated values for $\sigma(x)$. In this paper, we adopt the model and estimation strategy of Martins-Filho and Yao. However, we modify their estimator by adopting local exponential smoothing for the conditional variance as proposed in Ziegelmann (2002). This modification assures positivity of the conditional variance and provides improved performance in finite samples, as will be shown in the Monte Carlo study we conduct. We provide a full characterization of the asymptotic behavior of our proposed frontier estimator under fairly

[1] For simplicity in notation, we will henceforth write $E(\cdot|X_i = x)$ or $V(\cdot|X_i = x)$ simply as $E(\cdot|X_i)$ or $V(\cdot|X_i)$.

mild assumptions and contrast these asymptotic results with those provided by Martins-Filho and Yao (2007). In addition, we conduct a small Monte Carlo experiment which provides both evidence on the estimator's finite sample behavior and its performance relative to the estimator in Martins-Filho and Yao.

Besides this introduction, our paper has four more sections. Section 2 presents the deterministic frontier model under consideration, lists the assumptions on the data generating process and gives a detailed description of the estimator. Section 3 provides the main theorems which characterize the asymptotic behavior of the estimator. Section 4 contains a Monte Carlo simulation and section 5 provides a summary and conclusion.

## 2.2 Statistical model and estimation procedure

In this section we provide a full specification of the statistical model under consideration and give a precise description of the estimation procedure. We start by listing a set of assumptions that mostly coincide with those given in Martins-Filho and Yao. They are sufficient to establish the main asymptotic results in section 3.

Assumption A1. 1. $Z_i = (X_i, R_i)'$ for $i = 1, 2, \cdots, n$ is an independent and identically distributed sequence of random vectors with density $g$. We denote by $g_X(x)$ and $g_R(r)$ the common marginal densities of $X_i$ and $R_i$ respectively, and by $g_{R|X}(r)$ the common conditional density of $R_i$ given $X$. 2. $0 < \underline{B}_{g_X} \leq g_X(x) \leq \bar{B}_{g_X} < \infty$ for all $x \in G$, $G$ a compact subset of $\Theta = \times_{t=1}^{p}(0, \infty)$, which denotes the Cartesian product of the intervals $(0, \infty)$.

Assumption A2. 1. $Y_i = \sigma(X_i)\frac{R_i}{\sigma_R}$. 2. $R_i \in [0, 1]$, $X_i \in \Theta$. 3. $E(R_i|X_i) = \mu_R$, $V(R_i|X_i) = \sigma_R^2$. 4. $0 < \underline{B}_\sigma \leq \sigma(x) \leq \bar{B}_\sigma < \infty$ for all $x \in \Theta$. 5. $\sigma^2(\cdot) : \Theta \to \Re$ is a measurable twice continuously differentiable function in $\Theta$ with first and second derivative denoted by $\sigma^{2(1)}(x)$ and $\sigma^{2(2)}(x)$. 6. $|\sigma^{2(2)}(x)| < \bar{B}_{2\sigma}$ for all $x \in \Theta$. We also assume that $\sigma^2(x) = \exp(f(x))$ where $f(x)$ is everywhere differentiable with first derivative denoted by $f^{(1)}(x)$.

The following assumption is standard in nonparametric estimation and involves only the kernel $K$. We observe that A3 is satisfied by commonly used kernels such as Epanechnikov, Biweight and others.

Assumption A3. $K(x) : \times_{i=1}^{p}[-1, 1] \to \Re$ is a symmetric density function with bounded support satisfying: 1. $\int xK(x)dx = 0$. 2. $\int x^2K(x)dx = \sigma_K^2$. 3. for all $x \in \Re^p$, $|K(x)| < B_K < \infty$. 4. for all $x, x' \in \Re^p$, $|K(x) - K(x')| < m||x - x'||$ for some $0 < m < \infty$, where $||\cdot||$ is the Euclidean norm.

Assumption A4. For all $x, x' \in \Theta$, $|g_X(x) - g_X(x')| < m_g||x - x'||$ for some $0 < m_g < \infty$.

Given that the model can be written as in equation (2), we propose the following three stage estimation procedure. First, for any $x \in \Re^p_+$ we obtain $\hat{m}(x; h_n) \equiv \hat{\alpha}$ where

$$(\hat{\alpha}, \hat{\beta}) = \operatorname{argmin}_{\alpha,\beta} \sum_{i=1}^{n} (Y_i - \alpha - \beta(X_i - x))^2 K\left(\frac{X_i - x}{h_n}\right).$$

The bandwidth $h_n$ satisfies $0 < h_n \to 0$ as $n \to \infty$. This is the local linear kernel estimator of Stone(1977) and Fan(1992) with regressand $Y_i$ and regressors $X_i$. This coincides with the first stage proposed by Martins-Filho and Yao. In the second stage, we follow Ziegelmann (2002) by defining $e_i \equiv (Y_i - \hat{m}(X_i; h_n))^2$ and obtain $\hat{\sigma}_e^2(x; h_n) \equiv \exp(\hat{\theta}_1)$, where

$$(\hat{\theta}_1, \hat{\theta}_2) = \operatorname{argmin}_{\theta_1,\theta_2} \sum_{i=1}^{n} (e_i - \exp(\theta_1 + \theta_2(X_i - x)))^2 K\left(\frac{X_i - x}{h_n}\right).$$

This provides an estimator $\hat{\sigma}(x; h_n) = \left(\hat{\sigma}_e^2(x; h_n)\right)^{1/2}$. In the third stage, an estimator for $\sigma_R$ is obtained by defining

$$s_R(h_n) = \left(\max_{1 \le i \le n} \frac{Y_i}{\hat{\sigma}(X_i; h_n)}\right)^{-1}.$$

As observed in Martins-Filho and Yao, the estimation of $\sigma_R$ by $s_R$ is justified by assuming that there exists *one* observed production unit whose production plan lies on the estimated frontier. This is the anchoring assumption we referred to in the introduction. As a consequence the forecasted value for $R_i$ associated with this unit is identically one. We emphasize that the estimator $s_R$ depends on the bandwidth $h_n$ through $\hat{\sigma}(X_i; h_n)$. Furthermore, in what follows it is desirable to distinguish the bandwidth used in the first two stages of estimation, which we will denote by $h_n$, from that used in defining $s_R$, which we will denote by $g_n$, where $0 < g_n \to 0$ as $n \to \infty$. Therefore, we represent the production frontier estimator at $x \in \Re^p$ by $\hat{\rho}(x; h_n, g_n) = \frac{\hat{\sigma}(x;h_n)}{s_R(g_n)}$. Note that by construction, provided that the chosen kernel $K$ is smooth, $\hat{\rho}(x; h_n, g_n)$ is a smooth estimator that envelops the data (no observed pair $(Y_i, X_i)$ lies above $(\hat{\rho}(X_i; h_n, g_n), X_i))$ but may lie above or below the true frontier $\rho(X_i)$.

## 2.3 Asymptotic characterization of the estimator

Due to the similarity between our proposed estimation strategy and that described in Martins-Filho and Yao, most of our focus will be on establishing the asymptotic properties of the second stage estimation under exponential smoothing. For simplicity, but without loss of generality, all of our results are for the case where is $p = 1$. For the case where $p > 1$, all results hold with appropriate adjustments on the relative speed of $n$, $h_n^p$ and $g_n^p$.

We start by noting that since $\sigma^2(x) = \exp(f(x))$, we have that $\sigma^{2(1)}(x) = \exp(f(x))f^{(1)}(x)$. Hence, consider a local linear approximation for $\sigma^2(X_i)$ given by $L(X_i - x, \theta(x)) = \exp(\theta_1(x) + \theta_2(x)(X_i - x))$, where $\theta(x) = (f(x), f^{(1)}(x))' = (\theta_1(x), \theta_2(x))'$. It is easily verifiable that $L(0, \theta(x)) = \exp(\theta_1(x))$, $L^{(1)}(0, \theta(x)) = \exp(\theta_1(x))\theta_2(x)$ and $L^{(2)}(X_i - x, \theta(x)) = (\theta_2(x))^2 \exp(\theta_1(x) + \theta_2(x)(X_i - x))$. When necessary we will denote by $\theta^0(x) = (\theta_1^0(x), \theta_2^0(x))$ the true values of $f(x)$ and $f^{(1)}(x)$. Let $e_i = (Y_i - \hat{m}(X_i; h_n))^2$ and write the estimator defined in the second stage estimator as

$$\begin{aligned}
\{\hat{\theta}_1(x), \hat{\theta}_2(x)\} &\equiv \text{argmin}_{\theta_1,\theta_2} \frac{1}{n}\sum_{i=1}^{n}(e_i - L(X_i - x, \theta))^2 \frac{1}{h_n} K\left(\frac{X_i - x}{h_n}\right) \\
&\equiv \text{argmin}_{\theta_1,\theta_2} \frac{1}{n}\sum_{i=1}^{n}\Bigg(e_i - \exp(\theta_1) - \theta_2\exp(\theta_1)(X_i - x) \\
&\quad - \frac{1}{2}L^{(2)}(\lambda_i(X_i - x), \theta)(X_i - x)^2\Bigg)^2 \frac{1}{h_n} K\left(\frac{X_i - x}{h_n}\right)
\end{aligned}$$

where $\lambda_i \in [0, 1]$. Now, suppose $\hat{\theta}_1(x)$, and $\hat{\theta}_2(x)$ are uniformly consistent estimator of $\theta_1^0(x)$ and $\theta_2^0(x)$ in a compact set $G$ and put $\hat{\varepsilon}_i = -\frac{1}{2}L^{(2)}(\lambda_i(X_i - x), \hat{\theta}(x))(X_i - x)^2 \frac{1}{h_n} K\left(\frac{X_i - x}{h_n}\right)$. We will first provide the asymptotic properties of the estimator $\gamma_1^*(x)$ defined by

$$\{\gamma_1^*(x), \gamma_2^*(x)\} \equiv \text{argmin}_{\gamma_1,\gamma_2} \frac{1}{n}\sum_{i=1}^{n}(e_i - \hat{\varepsilon}_i - \gamma_1 - \gamma_2(X_i - x))^2 \frac{1}{h_n} K\left(\frac{X_i - x}{h_n}\right), \tag{2.3}$$

where $\gamma_1 = \exp(\theta_1)$ and $\gamma_2 = \theta_2\exp(\theta_1)$. We will need the following auxiliary lemma.

**Lemma 1** *Assume A1-A4. If $h_n \to 0$, $\frac{nh_n^3}{\ln(n)} \to \infty$, then for every $x \in G$ a compact subset of $(0, \infty) \times [0, 1]$ we have*

$$\gamma_1^*(x; h_n) - \sigma^2(x) - \frac{1}{nh_n g_X(x)} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\left(e_i - \hat{\varepsilon}_i - \sigma^2(x) - \sigma^{2(1)}(x)(X_i - x)\right) = O_p(R_{n,1}(x))$$

*uniformly in $G$, with* $R_{n,1}(x) = \frac{1}{n}\Bigg\{\left|\sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)(e_i - \hat{\varepsilon}_i - \sigma^2(x) - \sigma^{2(1)}(x)(X_i - x))\right| +$
$\left|\sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)(e_i - \hat{\varepsilon}_i - \sigma^2(x) - \sigma^{2(1)}(x)(X_i - x))\right|\Bigg\}$.

Lemma 1 reveals that to ascertain the uniform order in probability of

$$\gamma_1^*(x; h_n) - \sigma^2(x) - \frac{1}{nh_n g_X(x)} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\left(e_i - \hat{\varepsilon}_i - \sigma^2(x) - \sigma^{2(1)}(x)(X_i - x)\right)$$

in a compact set $G$, it suffices to investigate the order of the absolute value of the terms

$$c_1(x) = \frac{1}{nh_n g_X(x)} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\left(e_i - \hat{\varepsilon}_i - \sigma^2(x) - \sigma^{2(1)}(x)(X_i - x)\right)$$

and

$$c_2(x) = \frac{1}{nh_n g_X(x)} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)\left(e_i - \hat{\varepsilon}_i - \sigma^2(x) - \sigma^{2(1)}(x)(X_i - x)\right).$$

However, given assumption A3 of compact support for the kernel $K$, it suffices to investigate the order of $|c_1(x)|$.[2] In Theorem 1 we provide the exact order of

$$\gamma_1^*(x;h_n) - \sigma^2(x) - \frac{1}{nh_n g_X(x)}\sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right)\left(e_i - \hat{\varepsilon}_i - \sigma^2(x) - \sigma^{2(1)}(x)(X_i - x)\right)$$

and establish that under suitable normalization and centering $\gamma_1^*(x;h_n)$ is asymptotically normally distributed.

**Theorem 1** *Suppose that assumptions A1-A4 hold. In addition assume that* $E(|\epsilon_i||X_i) = \mu_1(X_i)$ *is a uniformly bounded function of* $X_i \in G$*, a compact subset of* $(0,\infty)$*. If* $h_n \to 0$, $\frac{nh_n^3}{\ln(n)} \to \infty$*, then for every* $x \in G$

*a)*

$$sup_{x\in G}\left|\gamma_1^*(x;h_n) - \sigma^2(x) - \frac{1}{nh_n g_X(x)}\sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right)\left(e_i - \hat{\varepsilon}_i - \sigma^2(x) - \sigma^{2(1)}(x)(X_i - x)\right)\right| =$$

$$O_p(h_n^3) + O_p\left(\left(\frac{h_n \ln(n)}{n}\right)^{1/2}\right)$$

*b) If in addition we assume that* $E(\epsilon_i^4|X_i = x) = \mu_4(x)$ *is continuous in* $(0,\infty)$, $h_n^2\ln(n) \to 0$ *and* $nh_n^5 = O(1)$ *then for every* $x \in G$.

$$\sqrt{nh_n}(\gamma_1^*(x) - \sigma^2(x) - B_{1n}) \xrightarrow{d} N\left(0, \frac{\sigma^4(x)}{g_X(x)}(\mu_4(x) - 1)\int K^2(y)dy\right),$$

*where* $B_{1n} = \frac{h_n^2\sigma_K^2}{2}(\sigma^{2(2)}(x) - L^{(2)}(0,\theta^0(x))) + o_p(h_n^2)$ *and* $\theta^0(x) = (f(x), f(x)^{(1)})$ *is uniquely defined by* $\sigma^{2(i)}(x) = L^{(i)}(0,\theta^0), i = 0,1.$

It is a direct consequence of Theorem 1 and the fact that

$$\sqrt{nh_n}\left(\sqrt{\gamma_1^*(x)} - \sigma(x) - \frac{1}{2\sigma(x)}B_{1n}(x) + \left(\frac{1}{2\sigma(x)} - \frac{1}{2\sigma_b(x)}\right)B_{1n}(x)\right) = \frac{1}{2\sqrt{\sigma_b^2(x)}}\sqrt{nh_n}(\gamma_1^* - \sigma^2(x) - B_{1n})$$

for $\sigma_b^2(x) = \eta\sigma^2(x) + (1-\eta)\gamma_1^*(x)$ and some $\eta \in [0,1]$ that

$$\sqrt{nh_n}\left(\sqrt{\gamma_1^*(x)} - \sigma(x) - B_{2n}\right) \xrightarrow{d} N\left(0, \frac{\sigma^2(x)}{4g_X(x)}(\mu_4(x) - 1)\int K^2(y)dy\right).$$

where $B_{2n} = \frac{h_n^2\sigma_K^2}{4\sigma(x)}(\sigma^{2(2)}(x) - L^{(2)}(0,\theta^0)) + o_p(h_n^2)$.

Theorem 1 relies on the uniform consistency of $\hat{\theta}(x)$. The next theorem establishes the desired uniform consistency.

[2] It should be emphasized that kernels with non-compact support could also be accommodated, provided that their rate of tail decay is sufficiently fast, but this would involve much longer proofs.

**Theorem 2** *Assume that A1-A4 hold and define*

$$(\hat{\theta}_1(x), \hat{\theta}_2(x)) = argmin_{\theta_1,\theta_2} \sum_{i=1}^{n} (e_i - \exp(\theta_1 + \theta_2(X_i - x)))^2 K\left(\frac{X_i - x}{h_n}\right).$$

*Furthermore, assume that the optimand is minimized in the interior of a compact set* $\bar{\Theta}$ *a subset of* $\Re^2$*. Then,* $(\hat{\theta}_1(x), \hat{\theta}_2(x)) - \theta^0(x) = o_p(1)$ *uniformly on a compact set* $G$ *of* $(0, \infty)$*.*

It is a direct consequence of Theorem 2 and the second part of the proof of Theorem 1 in Hall et al. (1999) that $\exp(\hat{\theta}_1(x)) - \gamma_1^*(x) = o_p(h_n^2)$ which combined with Theorem 1 gives,

$$\sqrt{nh_n}\left(\sqrt{\exp(\hat{\theta}_1(x))} - \sigma(x) - B_{2n}\right) \xrightarrow{d} N\left(0, \frac{\sigma^2(x)}{4g_X(x)}(\mu_4(x) - 1)\int K^2(y)dy\right).$$

where $B_{2n} = \frac{h_n^2 \sigma_K^2}{4\sigma(x)}(\sigma^{2(2)}(x) - L^{(2)}(0, \theta^0)) + o_p(h_n^2)$. The results in theorems 1 and 2 refer to the estimator $\hat{\sigma}(x; h_n) = \sqrt{\exp(\hat{\theta}_1(x))}$, but since our main interest lies on $\hat{\rho}(x; h_n, g_n) \equiv \frac{\hat{\sigma}(x;h_n)}{s_R(g_n)}$, a complete characterization of the asymptotic behavior of the frontier estimator requires a characterization of the asymptotic behavior of $s_R(g_n)$, and how it combines with the results obtained from Theorem 1 for $\hat{\sigma}(x; h_n)$. The following theorem is present without proof, as it can be obtained directly form Martins-Filho and Yao (2007). Part a) of Theorem 3 is a general result regarding the order in probability of $s_R(g_n) - \sigma_R$. It states that if the estimator $\hat{\sigma}(x; g_n)$ used to obtain $s_R$ is $O_p(L_n)$, where $L_n$ is an arbitrary nonstochastic sequence such that $0 < L_n \to 0$ as $n \to \infty$, and if $1 - max_{1 \le t \le n} R_t = O_p(L_n)$, then $s_R(g_n) - \sigma_R = O_p(L_n)$. The result is useful in that from part a) of Theorem 1, if $\frac{ng_n^5}{ln(n)} \to \infty$, then $\hat{\sigma}(x; g_n) - \sigma(x) = O_p(g_n^2)$. Hence, together with the assumption that $1 - max_{1 \le i \le n} R_i = O_p(g_n^2)$ we obtain $s_R(g_n) - \sigma_R = O_p(g_n^2)$. It should be noted that the required boundedness in probability of $1 - max_{1 \le i \le n} R_i$ is not necessary to establish the consistency of $s_R(g_n)$, which results directly from part a) of Theorem 1. Its use is confined to part b) of Theorem 3, where we use the result on the order of $s_R(g_n)$ to obtain the asymptotic normality of $\hat{\rho}(x; h_n, g_n)$ under a suitable normalization.

**Theorem 3** *Let* $L_n$ *be a nonstochastic sequence such that* $0 < L_n \to 0$ *as* $n \to \infty$ *and suppose that (1)* $\hat{\sigma}(x; g_n) - \sigma(x) = O_p(L_n)$ *uniformly in* $G$*, and (2)* $1 - max_{1 \le i \le n} R_i = O_p(L_n)$*. Then,*

*a)* $s_R(g_n) - \sigma_R = O_p(L_n)$*,*

*b) Under the assumptions in Theorem 1 part b), if* $\frac{ng_n^5}{ln(n)} \to \infty$*,* $nh_n^5 = o(1)$*, and* $nh_n g_n^4 = O(1)$ *then*

$$\sqrt{nh_n}\left(\frac{\hat{\sigma}(x; h_n)}{s_R(g_n)} - \frac{\sigma(x)}{\sigma_R} - B_{2n}\right) \xrightarrow{d} N\left(0, \frac{\sigma^2(x)}{4\sigma_R^2 g_X(x)}(\mu_4(x) - 1)\int K^2(y)dy\right)$$

*where* $B_{2n} = O_p(g_n^2)$*.*

The conditions on the order of the bandwidths $h_n$ and $g_n$ are also crucial for asymptotic normality of the estimated frontier. In particular, they imply that the bandwidth $h_n$, used in the first and second stages of the estimation, must satisfy $nh_n^5 = o(1)$, which represents an undersmoothing in the estimation $\hat{\sigma}(x, h_n)$. In addition, the bandwidth $g_n$ used to obtain $s_R$in the third stage must converge to zero slower than $h_n$. The requirement $ng_n^5 \to \infty$ in the estimation of $s_R$ is necessary only in that it provides a convenient order for $B_{2n}$.

A sharper result on the bias term $B_{2n}$ can be obtained by assuming that $1 - \max_{1\leq i\leq n} R_i = o_p(g_n^2)$. In this case part (b) of Theorem 2 can be extended to give

$$\sqrt{nh_n}\left(\frac{\hat{\sigma}(x;h_n)}{s_R(g_n)} - \frac{\sigma(x)}{\sigma_R} - B_{3n}\right) \xrightarrow{d} N\left(0, \frac{\sigma^2(x)}{4\sigma_R^2 g_X(x)}(\mu_4(x)-1)\int K^2(y)dy\right)$$

where $B_{3n} = \frac{g_n^2\sigma(x)\sigma_k^2}{4\sigma_R}\sup_{x\in G, R\in[0,1]}\left(-\frac{[\sigma^{2(2)}(x)-L^{(2)}(0,\theta^0)]R}{\sigma^2(x)}\right)+o_p(g_n^2)$. We note that this increased precision in the expression of the bias in unnecessary for inference purposes, since it is normally conducted under the assumption that $nh_ng_n^4 \to 0$, in which case $\sqrt{nh_n}B_{3n} \to 0$ as $n \to \infty$. If we compare the preceding result to that obtained from Theorem 2 in Martins-Filho and Yao, we can see that the two estimators have exactly the same asymptotic variance (resulting in the same efficiency) but a different bias. The difference is governed by the term $L^{(2)}(0,\theta^0)$. As mentioned in Ziegelmann (2002), since $L^{(2)}(0,\theta^0)$ is a nonnegative quantity, we conclude that the bias of the estimator we propose can be smaller than that of the local linear estimator if $\sigma^{2(2)}(x)$ is nonnegative and greater than $L^{(2)}(0,\theta^0)$.Our results show that a local exponential estimator can be incorporated into the second stage estimation replacing the local linear estimator without loss of consistency or asymptotic normality, previously established under the assumptions of Martins-Filho and Yao (2007).

## 2.4 Monte Carlo Study

In this section we investigate some of the finite sample properties of our estimator, henceforth referred to as NPE, via a Monte Carlo study. For comparison purposes, we also include in the study the local linear frontier estimator proposed in Martins-Filho and Yao (2007), referred to as NP. Our simulations are based on model (1), i.e., $Y_i = \frac{\sigma(X_i)R_i}{\sigma_R}$, with $p = 1$. We generate data with the following characteristics. The $X_i$ are pseudorandom variables from a uniform distribution with support given by $[a_l, b_u]$. $R_i = \exp(-Z_i)$, where $Z_i$ are pseudorandom variables from an exponential distribution with parameter $\beta > 0$, therefore

$R_i$ has support on $(0,1]$. We consider two specifications for $\sigma(x)$:

$$\sigma_1(x) = \sqrt{x},\ \text{with } x \in [a_l, b_u] = [10, 100] \text{ and}$$

$$\sigma_2(x) = 3(x-1.5)^3 + 0.25x + 1.125,\ \text{with } x \in [a_l, b_u] = [1, 2],$$

which are associated with convex and non-convex production technologies. Three parameters for the exponential distribution are considered: $\beta_1 = 3$, $\beta_2 = 1$ and $\beta_3 = 1/3$. These choices of parameters produce, respectively, the following values for the parameters of $g_{R|X} : (\mu_R, \sigma_R^2) = (0.25, 0.08), (0.5, 0.08)$ and $(0.75, 0.04)$. Three sample sizes $n = 200, 300, 400$ were used. We evaluate the frontiers at $x_1 = 32.5$, $x_2 = 55$ and $x_3 = 77.5$ for $\sigma_1(x)$ and at $x_1 = 1.25$, $x_2 = 1.5$ and $x_3 = 1.75$ for $\sigma_2(x)$. These values of $X$ correspond to the 25th, 50th and 75th percentile of its support.

An important aspect in the implementation of our frontier estimator is bandwidth selection. We consider the following rule-of-thumb bandwidth.

$$\hat{h}_{ROT} = \left( \frac{\int K^2(\phi)d\phi(\hat{\mu}_4(\lambda_n) - 1) \int \dot{\sigma}^2(x)dx}{(\sigma_K^2)^2 \left( \max_{1\le i\le n} \left( \frac{(\dot{\sigma}^{2(2)}(x_i) - \dot{\beta}^2 e^{\dot{\alpha}})\dot{R}_i}{\dot{\sigma}^2(x_i)} \right) \right)^2 \frac{1}{n}\sum_{i=1}^n \dot{\sigma}^2(x_i)} \right)^{1/5} n^{-(1+4\gamma)/5}$$

where $\gamma$ is set to be 0.11 in all experiments, which satisfies the requirements in Theorem 3, $\hat{g}_{ROT} = n^\gamma \hat{h}_{ROT}$. The sequence $\{\dot{\sigma}^2(X_i)\}_{i=1}^n$ is estimated with an ordinary least square quartic regression of $\{\hat{\epsilon}_i^2\}_{i=1}^n$ on $\{X_i\}_{i=1}^n$, with $\hat{\epsilon}_i = Y_i - \hat{m}(X_i)$, where $\hat{m}(X_i)$ is estimated via local linear regression with a rule-of-thumb bandwidth as in Ruppert et al. (1995). $\{\dot{\sigma}^2(X_i)\}_{i=1}^n$ is then used to construct $\int \dot{\sigma}^2(x)dx$, $\max_{1\le i\le n}\left( \frac{(\dot{\sigma}^{2(2)}(x_i) - \dot{\beta}^2 e^{\dot{\alpha}})\dot{R}_i}{\dot{\sigma}^2(x_i)} \right)$ and $\frac{1}{n}\sum_{i=1}^n \dot{\sigma}^2(x_i)$. $\dot{\alpha}$ and $\dot{\beta}$ are estimated via a local exponential regression with the same bandwidth as used to estimate $\hat{m}(X_i)$. $\hat{\mu}_4(\lambda_n) = \frac{1}{n}\sum_{i=1}^n \left( \frac{Y_i}{\hat{\sigma}(X_i, \lambda_n)} - \hat{b} \right)^4$, where $\hat{b} = \frac{\sum_{i=1}^n \hat{\sigma}(X_i,\lambda_n) Y_i}{\sum_{i=1}^n \hat{\sigma}(X_i,\lambda_n)}$ is an estimator for $b = \mu_R/\sigma_R$. $\{\hat{\sigma}^2(X_i, \lambda_n)\}_{i=1}^n$ in $\hat{\mu}_4$ is estimated via local linear regression of $\{\hat{\epsilon}_i^2\}_{i=1}^n$ on $\{X_i\}_{i=1}^n$, with a rule-of-thumb bandwidth $\lambda_n$ as in Ruppert et al. (1995) and Fan and Yao (1998).

We evaluate the overall performance of the efficiency estimator based on three different measures. First, we consider the correlation between the efficiency rankings produced by the estimator and the true efficiency rankings:

$$R_{rank} = \frac{cov(rank(\hat{R}_i), rank(R_i))}{\sqrt{var(rank(\hat{R}_i))var(rank(R_i))}}$$

where $rank(R_i)$ gives the ranking index according to the magnitude of $R_i$. The closer $R_{rank}$ for $\hat{R}_i$ is to 1, the higher the correlation between the true $R_i$ and $\hat{R}_i$, thus the better the estimator $\hat{R}_i$. The second measure we consider is $R_{mag} = \frac{1}{n}\sum_{i=1}^{n}(\hat{R}_i - R_i)^2$ which is simply the squared Euclidean distance between the estimated vector of efficiencies and the true vector of efficiencies. The third measure we use is $R_{rel} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{R}_i}{\hat{R}_t} - \frac{R_i}{R_t}\right|$, where $t$ is the position index for $R_t = \max_{1\leq i\leq n} R_i$, and $\hat{R}_t$ is the $t$th corresponding element in $\{\hat{R}_i\}_{i=1}^{n}$, which may not be the maximum of $\hat{R}_i$. Hence $R_{rank}$, $R_{mag}$ summarize the performance of the estimator $\hat{R}_i$ in rankings and calculating the magnitude of efficiency. $R_{rel}$ captures the relative efficiency.

The results of our simulations are summarized in tables 1-6 and figures 1-13. Whenever negative estimates for $\sigma^2(\cdot)$ occur in the case of NP, the sample is discarded. In this case, another sample is generated until 1000 *valid* repetitions are obtained. Figures 1-12 give boxplots of MSE for the frontier estimator ($\hat{\rho}(\cdot)$), efficiency estimator ($\hat{R}_i$), location parameter estimator ($s_R$) and frontier shape estimator ($\hat{\sigma}(\cdot)$). Each boxplot is constructed from 1000 points (repetitions), where each point corresponds to a sample draw and is calculated as the squared Euclidean distance between the estimate and true value of $\rho(\cdot)$, $R_i$, $\sigma_R$ and $\sigma(\cdot)$. The thick horizontal line inside the rectangle in each boxplot corresponds to the median of the distribution, and the rectangle height corresponds to interquartile range. Consequently 50% of data is represented by the rectangle. The two thin horizontal lines below and above the rectangle are the whiskers. The whiskers extend to the most extreme data point which is no more than 1.5 times the interquartile range. Tables 1-6 are constructed based on the data points between the whiskers in the boxplots. Tables 1-2 provide the bias and MSE of $s_R$ and $\hat{\sigma}(x)$ at three different values of $x$. Tables 3-4 give the bias and MSE of the frontier estimators and Figure 13 shows kernel density estimates for the two frontier estimators around the true value at $x_2 = 55$ based on 1000 simulations, $\mu_R = 0.5$ and $\sigma_1(x)$, for $n = 200$ and $n = 400$. Tables 5-6 give the overall performance of the efficiency estimators according to the measures described above.

### 2.4.1 General regularities

As expected from the asymptotic results of section 3, as the sample size $n$ increases, the boxplots show that MSE decreases for the vast majority of simulations for all estimators and values for $\mu_R$ considered. The bias and MSE for $s_R$, $\hat{\sigma}(x)$, and the frontier estimator based on NPE and NP, presented on Tables 1-4, generally decrease, with some exceptions when it comes to the bias. Regarding the measures of overall performance for efficiency estimators mentioned above, NPE and NP perform better as $n$ increases. The asymptotics of both estimators seem to be confirmed in general terms as their performance improve with large $n$.

We now turn to the impact of different values of $\mu_R$ on the performance of NPE and NP. Regarding $s_R$, $\hat{\sigma}(x)$, and the frontier estimator, the best performance in terms of MSE occurs when $\mu_R = 0.5$, and the worst performance occurs when $\mu_R = 0.75$. The relative diminished performance when $\mu_R = 0.75$ is most likely explained by the fact that for this DGP $\sigma_R^2$ is half of its value in other DGPs, contributing to it has higher variance as suggested in Theorem 2. Regarding $s_R$, the best performance in terms of bias is generally for $\mu_R = 0.5$, with some exceptions, and the worst performance occurs for the DGP with smallest value for $\sigma_R$, that is, when $\mu_R = 0.75$. The bias of $\hat{\sigma}_1(x, h_n)$ is negative for most experiments considered, which is in accordance to the asymptotic results due to presence of $\sigma^{2(2)}(x) - \theta_2^2 e^{\theta_1}$ and $\sigma^{2(2)}(x)$ in the case of NPE and NP respectively. The bias of $\hat{\sigma}_2(x, h_n)$ for NPE is negative in most experiments at the points $x_1 = 1.25$ and $x_2 = 1.5$, but the bias is positive when $x_3 = 1.75$. For NP, the bias of $\hat{\sigma}_2(x, h_n)$ is positive at $x_3 = 1.75$ and at $x_2 = 1.5$ when $\mu_R = 0.25$ and $\mu_R = 0.5$. Following the asymptotic results in Theorem 2, the bias of the frontier estimator of NPE as well as of NP is generally positive, except for small $\mu_R$. The bias seems to increase with $\mu_R$.

Regarding the measures of overall performance for the efficient estimator described above, NPE and NP estimator seem to perform worse when $\mu_R$ is large for $R_{rank}$, $R_{mag}$ and $R_{rel}$.

### 2.4.2 Relative performance of estimators

The main differences occur when DGP uses $\sigma_1(x)$. In that case, on estimating the production frontier (Table 3 and figures 1-6) there seems to be evidence that NPE dominates NP in terms of MSE in all cases considered. In terms of bias, NPE dominates NP when $\mu_R = 0.5$ and $\mu_R = 0.75$. This better performance of NPE against NP may be understood looking at Table 1 and Figs. 1-6. The gain of NPE seems to be on estimating $\sigma_R$, since that NPE outperforms NP on estimating $\sigma_R$, while NP outperforms NPE on estimating $\sigma_1(x)$. When the different measures of overall performance we considered are analyzed (Table 5), we observe that NPE outperforms NP in terms of $R_{rel}$. Regarding $R_{rank}$, NPE and NP present the same performance when $\mu_R = 0.25$ and $\mu_R = 0.5$. In the case of $\mu_R = 0.75$, NP performs better than NPE. When DGP uses $\sigma_2(x)$ the performance of NPE and NP are more similar. Nevertheless, we see in Figs. 7-12 that NPE generally presents a better performance in estimating $\sigma_R$ and NP usually performs better estimating $\sigma_2(x)$. Such conclusion may be in part confirmed looking at Table 2, noting that NPE outperforms NP in terms of MSE regarding the estimation of $\sigma_R$. On estimating $\sigma(x)$, NP outperforms

NPE in terms of MSE for $\hat{\sigma}_2(x_1)$, but on estimating $\sigma_2(x_2)$ and $\sigma_2(x_3)$ NPE performs better in some cases, while NP performs better in others. Regarding the frontier estimator, Figs. 7-12 show that NPE and NP perform very similar and Table 4 show that NPE usually performs better than NP in terms of MSE. When the different measures of overall performance we considered are analyzed (Table 6), we observe that NP usually performs better than NPE for all measures considered. This is more evident in the case of $\mu_R = 0.75$. Fig. 13 shows kernel density estimates for the frontier around the true value evaluated at $x_2 = 55$ for NPE and NP based on 1000 simulations, $\mu_R = 0.5$ and $\sigma(x) = \sqrt{x}$, for $n = 200$ and $n = 400$. The kernel density estimates are calculated using an Epanechnikov kernel and bandwidths are selected using rule-of-thumb of Silverman (1986). For both estimators we observe the familiar symmetric bell shape and that the NPE is more tightly centered around the true frontier. Fig. 13 also shows that the estimated densities become tighter with more acute spikes as the sample size increases, as expected from the available asymptotic results.[3]

## 2.5 Real Data Example

We illustrate our methodology analyzing USA crime data. The goal is to estimate a production frontier and efficiency for 294 USA Law Enforcement agencies using data for the year 2000. Data sources are FBI's Uniform Crime Reports and LEMAS (Law Enforcement Management and Administrative Statistics) survey. All data used is available on the Internet in the site of Bureau of Justice Statistics (http://bjsdata.ojp.usdoj.gov).

In order to measure crime we consider crime trend data from FBI's Uniform Crime Reports for large agencies in USA (population coverage $\geq 80,000$). This data was used to construct the output, which is defined as population per total crime, where total crime is number of violent crimes plus number of property crimes[4]. This output measure is consistent with our methodology and also is considered in Gorman and Ruggiero (2008).

To measure resources invested in police force we consider data from LEMAS (Law Enforcement Management and Administrative Statistics) survey. This data was used to construct a measure of input, which is in fact an index composed by base annual starting salaries for three categories: chief executive (*chief*), sergeant (*sergeant*) and entry-level officer (*entry*). In order to simplify our analysis and avoid

[3] The densities for $\mu_R = 0.25$ and $\mu_R = 0.75$ show a similar pattern.

[4] Violent crimes are murder and non-negligent manslaughter, forcible rape, robbery and aggravated assault. Property crimes are burglary, larceny-theft and motor vehicle theft.

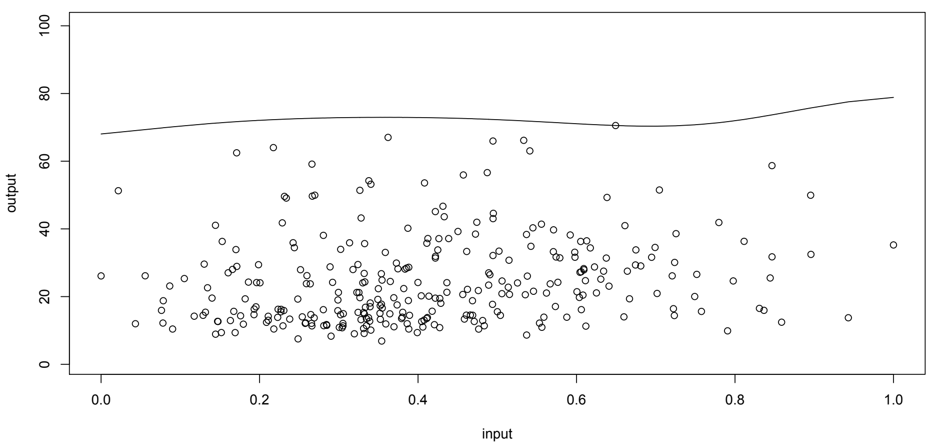
any *curse of dimensionality*, we reduce our problem dimension via principal components analysis of those three inputs. Proceeding in this manner, we can see that the first orthogonal component is responsible by 79.18% of the total variability. Therefore, we decide to use a single component, which is the following linear combination of the three original variables

$$input = 0.958(chief - \overline{chief}) + 0.242(sergeant - \overline{sergeant}) + 0.157(entry - \overline{entry}),$$

where $\bar{x}$ denotes the sample mean of a variable $x$. In order to avoid negative values and therefore ease interpretability of our index, we rescale the above index such that its minimum value is zero. Moreover, we project the above index into the interval [0,1] in order to facilitate the bandwidth choice.

The estimated frontier is displayed in Fig. 2.1. We can notice it is a smooth function, where the point lying on the estimated curve reflects the model anchoring assumption, corresponding to San Jose Police Department-CA. Efficiency rank and efficiency scores are shown on Table 2.7. Using the output and input measures mentioned above, we see that the estimated five most efficient agencies are San Jose Police Department-CA, Shreveport Police Department-LA, Fulton County Police Department-GA, El Dorado County Sheriff Department-CA and Harford County Sheriff Office-MD, in decreasing order. Moreover four californian agencies are in the top ten efficiency rank. On the other hand, the estimated five least efficient agencies are Tallahassee Police Department-FL, Chattanooga Police Department-TN, Orlando Police Department-FL, Atlanta Police Department-GA and St Joseph County Sheriff Department-IN, where the last is the least efficient.

Figura 2.1: NPE Frontier Estimation



## 2.6 Summary and conclusions

In this paper we use the idea of local exponential smoothing to improve the nonparametric frontier estimator proposed by Martins-Filho and Yao (2007). Their estimation strategy suffered from the undesirable property of potentially generating negative estimated conditional variances. Local exponential smoothing prevents this problem. In addition, there seems to be finite sample gains in adopting exponential smoothing. These gains are particularly large in the estimation of the location parameter in the frontier model. Our simulation results confirm and give added support those in Ziegelmann (2002).

## 2.7 Appendix 1: Tables and Graphics

Tabela 2.1: Frontier I - Bias and MSE for $S_R$ and $\hat{\sigma}(x)$

| $\sigma_1(x)$ | $n$ | | $S_R(\times 10^{-2})$ | | $\hat{\sigma}_1(x_1)$ | | $\hat{\sigma}_1(x_2)$ | | $\hat{\sigma}_1(x_3)$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | NPE | NP | NPE | NP | NPE | NP | NPE | NP |
| $\mu_R = 0.25$ | 200 | Bias | -0.313 | -0.665 | -0.177 | -0.047 | -0.045 | 0.060 | -0.271 | -0.155 |
| | | MSE | 0.019 | 0.030 | 0.207 | 0.146 | 0.307 | 0.252 | 0.526 | 0.427 |
| | 300 | Bias | -0.497 | -1.012 | -0.189 | -0.045 | -0.006 | 0.087 | -0.243 | -0.132 |
| | | MSE | 0.016 | 0.035 | 0.169 | 0.103 | 0.211 | 0.171 | 0.369 | 0.290 |
| | 400 | Bias | -0.620 | -0.884 | -0.197 | -0.062 | -0.003 | 0.071 | -0.195 | -0.104 |
| | | MSE | 0.014 | 0.025 | 0.153 | 0.092 | 0.182 | 0.140 | 0.291 | 0.224 |
| $\mu_R = 0.50$ | 200 | Bias | -0.357 | -0.719 | -0.093 | 0.026 | 0.099 | 0.167 | -0.118 | -0.076 |
| | | MSE | 0.010 | 0.022 | 0.096 | 0.075 | 0.129 | 0.143 | 0.203 | 0.188 |
| | 300 | Bias | -0.429 | -0.671 | -0.083 | 0.029 | 0.131 | 0.182 | -0.055 | -0.025 |
| | | MSE | 0.009 | 0.016 | 0.074 | 0.056 | 0.119 | 0.130 | 0.142 | 0.135 |
| | 400 | Bias | -0.405 | -0.473 | -0.059 | 0.038 | 0.148 | 0.191 | -0.044 | -0.009 |
| | | MSE | 0.007 | 0.010 | 0.059 | 0.048 | 0.108 | 0.117 | 0.122 | 0.111 |
| $\mu_R = 0.75$ | 200 | Bias | -2.138 | -2.652 | -0.378 | -0.252 | -0.310 | -0.212 | -0.593 | -0.462 |
| | | MSE | 0.062 | 0.097 | 0.322 | 0.205 | 0.373 | 0.283 | 0.776 | 0.592 |
| | 300 | Bias | -1.974 | -2.389 | -0.391 | -0.256 | -0.233 | -0.164 | -0.482 | -0.404 |
| | | MSE | 0.050 | 0.077 | 0.282 | 0.176 | 0.269 | 0.210 | 0.542 | 0.434 |
| | 400 | Bias | -1.833 | -2.252 | -0.367 | -0.237 | -0.200 | -0.140 | -0.460 | -0.392 |
| | | MSE | 0.042 | 0.068 | 0.246 | 0.148 | 0.207 | 0.164 | 0.476 | 0.381 |

Tabela 2.2: Frontier II - Bias and MSE for $S_R$ and $\hat{\sigma}(x)$

| $\sigma_2(x)$ | $n$ | | $S_R(\times 10^{-2})$ | | $\hat{\sigma}_2(x_1)$ | | $\hat{\sigma}_2(x_2)$ | | $\hat{\sigma}_2(x_3)$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | NPE | NP | NPE | NP | NPE | NP | NPE | NP |
| $\mu_R = 0.25$ | 200 | Bias | -0.404 | -0.383 | -0.059 | -0.054 | -0.015 | 0.014 | 0.047 | 0.059 |
| | | MSE | 0.017 | 0.018 | 0.011 | 0.010 | 0.006 | 0.006 | 0.014 | 0.015 |
| | 300 | Bias | -0.668 | -0.710 | -0.064 | -0.058 | -0.017 | 0.010 | 0.048 | 0.058 |
| | | MSE | 0.018 | 0.019 | 0.010 | 0.008 | 0.004 | 0.004 | 0.010 | 0.012 |
| | 400 | Bias | -0.780 | -0.805 | -0.059 | -0.055 | -0.012 | 0.011 | 0.050 | 0.059 |
| | | MSE | 0.017 | 0.017 | 0.008 | 0.007 | 0.003 | 0.003 | 0.008 | 0.009 |
| $\mu_R = 0.50$ | 200 | Bias | -0.603 | -0.631 | -0.033 | -0.028 | 0.009 | 0.031 | 0.070 | 0.070 |
| | | MSE | 0.013 | 0.014 | 0.005 | 0.004 | 0.003 | 0.004 | 0.012 | 0.011 |
| | 300 | Bias | -0.639 | -0.696 | -0.029 | -0.025 | 0.013 | 0.034 | 0.072 | 0.072 |
| | | MSE | 0.010 | 0.011 | 0.004 | 0.004 | 0.002 | 0.003 | 0.011 | 0.010 |
| | 400 | Bias | -0.669 | -0.789 | -0.026 | -0.022 | 0.015 | 0.033 | 0.069 | 0.075 |
| | | MSE | 0.010 | 0.012 | 0.003 | 0.002 | 0.002 | 0.003 | 0.008 | 0.010 |
| $\mu_R = 0.75$ | 200 | Bias | -2.244 | -2.264 | -0.100 | -0.090 | -0.060 | -0.033 | 0.003 | 0.012 |
| | | MSE | 0.064 | 0.066 | 0.019 | 0.016 | 0.011 | 0.011 | 0.014 | 0.015 |
| | 300 | Bias | -2.030 | -2.077 | -0.089 | -0.082 | -0.055 | -0.032 | 0.004 | 0.011 |
| | | MSE | 0.050 | 0.053 | 0.014 | 0.012 | 0.008 | 0.006 | 0.009 | 0.010 |
| | 400 | Bias | -1.917 | -1.999 | -0.084 | -0.076 | -0.052 | -0.031 | 0.003 | 0.012 |
| | | MSE | 0.045 | 0.048 | 0.012 | 0.010 | 0.007 | 0.006 | 0.008 | 0.008 |

Tabela 2.3: Frontier I - Bias and MSE of frontier estimators

| $\sigma_1(x)$ | $n$ | | $x_1 = 32.5$ | | $x_2 = 55$ | | $x_3 = 77.5$ | |
|---|---|---|---|---|---|---|---|---|
| | | | NPE | NP | NPE | NP | NPE | NP |
| $\mu_R = 0.25$ | 200 | Bias | -0.303 | -0.027 | 0.215 | 0.421 | -0.587 | -0.550 |
| | | MSE | 2.250 | 3.069 | 2.965 | 6.915 | 4.636 | 10.442 |
| | 300 | Bias | -0.154 | 0.159 | 0.616 | 0.914 | -0.189 | -0.058 |
| | | MSE | 1.764 | 2.020 | 2.659 | 5.049 | 3.331 | 6.403 |
| | 400 | Bias | -0.142 | 0.063 | 0.683 | 0.847 | 0.069 | 0.020 |
| | | MSE | 1.386 | 1.717 | 2.423 | 4.117 | 2.729 | 4.434 |
| $\mu_R = 0.50$ | 200 | Bias | 0.041 | 0.288 | 0.788 | 0.996 | 0.055 | 0.261 |
| | | MSE | 1.017 | 1.607 | 1.935 | 3.929 | 1.656 | 3.734 |
| | 300 | Bias | 0.092 | 0.288 | 0.947 | 1.086 | 0.350 | 0.538 |
| | | MSE | 0.833 | 1.284 | 2.006 | 3.575 | 1.765 | 3.120 |
| | 400 | Bias | 0.155 | 0.181 | 0.995 | 0.955 | 0.386 | 0.359 |
| | | MSE | 0.705 | 0.940 | 2.031 | 2.651 | 1.526 | 2.111 |
| $\mu_R = 0.75$ | 200 | Bias | 1.402 | 2.099 | 2.713 | 3.475 | 1.856 | 2.467 |
| | | MSE | 8.170 | 14.777 | 16.466 | 31.507 | 13.399 | 32.398 |
| | 300 | Bias | 0.950 | 1.629 | 2.671 | 3.366 | 1.932 | 2.677 |
| | | MSE | 5.023 | 10.351 | 13.795 | 28.208 | 11.644 | 27.862 |
| | 400 | Bias | 0.938 | 1.525 | 2.682 | 3.303 | 1.907 | 2.601 |
| | | MSE | 4.566 | 9.115 | 12.851 | 24.609 | 11.321 | 22.928 |

Tabela 2.4: Frontier II - Bias and MSE of frontier estimators

| $\sigma_2(x)$ | $n$ | | $x_1 = 1.25$ | | $x_2 = 1.5$ | | $x_3 = 1.75$ | |
|---|---|---|---|---|---|---|---|---|
| | | | NPE | NP | NPE | NP | NPE | NP |
| $\mu_R = 0.25$ | 200 | Bias | -0.109 | -0.107 | 0.043 | 0.041 | 0.246 | 0.242 |
| | | MSE | 0.074 | 0.077 | 0.044 | 0.048 | 0.194 | 0.193 |
| | 300 | Bias | -0.072 | -0.072 | 0.093 | 0.092 | 0.317 | 0.317 |
| | | MSE | 0.050 | 0.053 | 0.042 | 0.042 | 0.213 | 0.213 |
| | 400 | Bias | -0.053 | -0.047 | 0.114 | 0.123 | 0.329 | 0.346 |
| | | MSE | 0.039 | 0.043 | 0.036 | 0.040 | 0.184 | 0.198 |
| $\mu_R = 0.50$ | 200 | Bias | 0.007 | 0.016 | 0.164 | 0.175 | 0.389 | 0.403 |
| | | MSE | 0.034 | 0.041 | 0.052 | 0.060 | 0.227 | 0.239 |
| | 300 | Bias | 0.031 | 0.035 | 0.194 | 0.199 | 0.417 | 0.431 |
| | | MSE | 0.032 | 0.038 | 0.061 | 0.066 | 0.238 | 0.248 |
| | 400 | Bias | 0.039 | 0.057 | 0.196 | 0.221 | 0.401 | 0.439 |
| | | MSE | 0.025 | 0.033 | 0.057 | 0.071 | 0.215 | 0.250 |
| $\mu_R = 0.75$ | 200 | Bias | 0.320 | 0.326 | 0.586 | 0.577 | 0.992 | 0.964 |
| | | MSE | 0.387 | 0.418 | 0.512 | 0.514 | 1.341 | 1.214 |
| | 300 | Bias | 0.272 | 0.283 | 0.520 | 0.530 | 0.905 | 0.910 |
| | | MSE | 0.277 | 0.311 | 0.404 | 0.420 | 1.037 | 1.007 |
| | 400 | Bias | 0.276 | 0.291 | 0.491 | 0.517 | 0.842 | 0.876 |
| | | MSE | 0.252 | 0.278 | 0.368 | 0.399 | 0.914 | 0.959 |

Tabela 2.5: Frontier I - Overall Efficiency Measures by NPE and NP

| $\sigma_1(x)$ | $n$ | $R_{rank}$ | | $R_{mag}(\times 10^{-2})$ | | $R_{rel}$ | |
|---|---|---|---|---|---|---|---|
| | | NPE | NP | NPE | NP | NPE | NP |
| $\mu_R = 0.25$ | 200 | 0.998 | 0.998 | 0.118 | 0.130 | 0.021 | 0.019 |
| | 300 | 0.999 | 0.999 | 0.084 | 0.108 | 0.018 | 0.016 |
| | 400 | 0.999 | 0.999 | 0.071 | 0.081 | 0.017 | 0.015 |
| $\mu_R = 0.50$ | 200 | 0.994 | 0.994 | 0.115 | 0.176 | 0.028 | 0.028 |
| | 300 | 0.995 | 0.995 | 0.101 | 0.140 | 0.027 | 0.024 |
| | 400 | 0.996 | 0.996 | 0.087 | 0.108 | 0.024 | 0.022 |
| $\mu_R = 0.75$ | 200 | 0.932 | 0.934 | 0.576 | 0.948 | 0.060 | 0.059 |
| | 300 | 0.948 | 0.950 | 0.423 | 0.764 | 0.053 | 0.050 |
| | 400 | 0.950 | 0.957 | 0.403 | 0.676 | 0.050 | 0.045 |

Tabela 2.6: Frontier II - Overall Efficiency Measures by NPE and NP

| $\sigma_2(x)$ | $n$ | $R_{rank}$ | | $R_{mag}(\times 10^{-2})$ | | $R_{rel}$ | |
|---|---|---|---|---|---|---|---|
| | | NPE | NP | NPE | NP | NPE | NP |
| $\mu_R = 0.25$ | 200 | 0.999 | 0.999 | 0.076 | 0.074 | 0.016 | 0.016 |
| | 300 | 0.999 | 0.999 | 0.071 | 0.065 | 0.015 | 0.014 |
| | 400 | 0.999 | 0.999 | 0.062 | 0.058 | 0.014 | 0.013 |
| $\mu_R = 0.50$ | 200 | 0.994 | 0.995 | 0.148 | 0.147 | 0.027 | 0.026 |
| | 300 | 0.995 | 0.996 | 0.146 | 0.141 | 0.025 | 0.024 |
| | 400 | 0.996 | 0.996 | 0.134 | 0.140 | 0.024 | 0.023 |
| $\mu_R = 0.75$ | 200 | 0.941 | 0.947 | 0.689 | 0.678 | 0.055 | 0.051 |
| | 300 | 0.953 | 0.957 | 0.556 | 0.565 | 0.048 | 0.045 |
| | 400 | 0.958 | 0.960 | 0.489 | 0.535 | 0.046 | 0.044 |

Figura 2.2: Frontier I - Boxplot of Estimators - $n = 200$ - $\mu_r = 0.25$

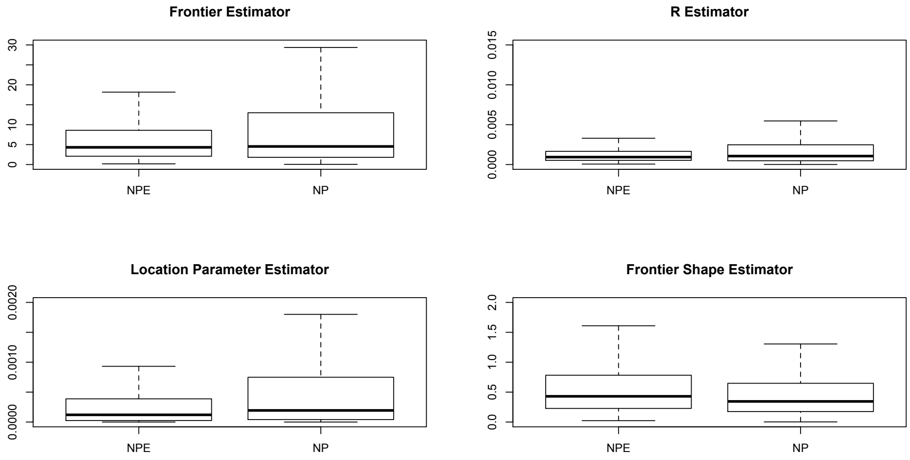



Figura 2.3: Frontier I - Boxplot of Estimators - $n = 400$ - $\mu_r = 0.25$

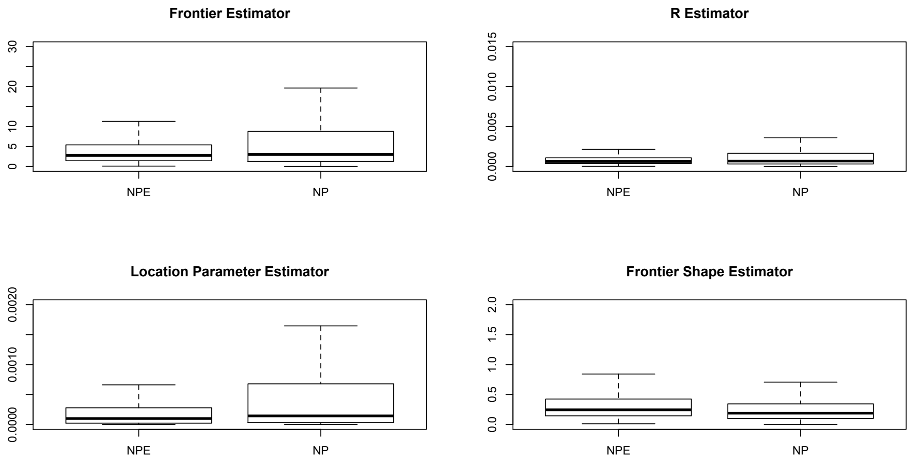


Figura 2.4: Frontier I - Boxplot of Estimators - $n = 200$ - $\mu_r = 0.5$

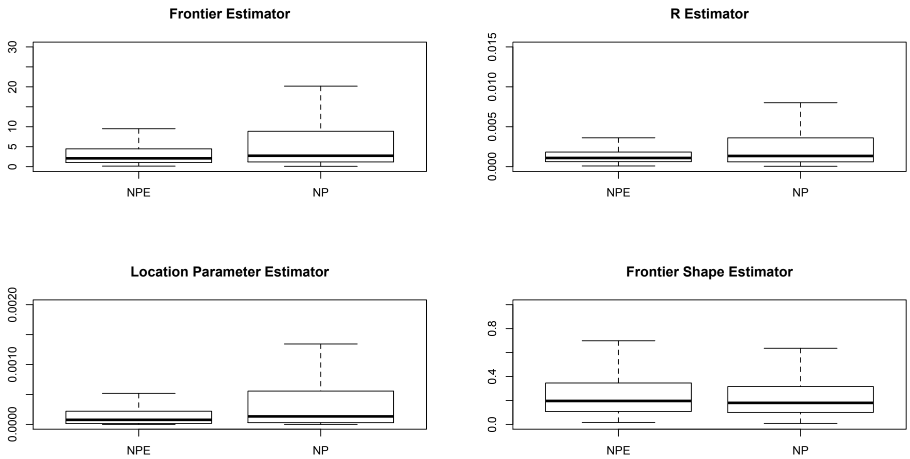



Figura 2.5: Frontier I - Boxplot of Estimators - $n = 400$ - $\mu_r = 0.5$

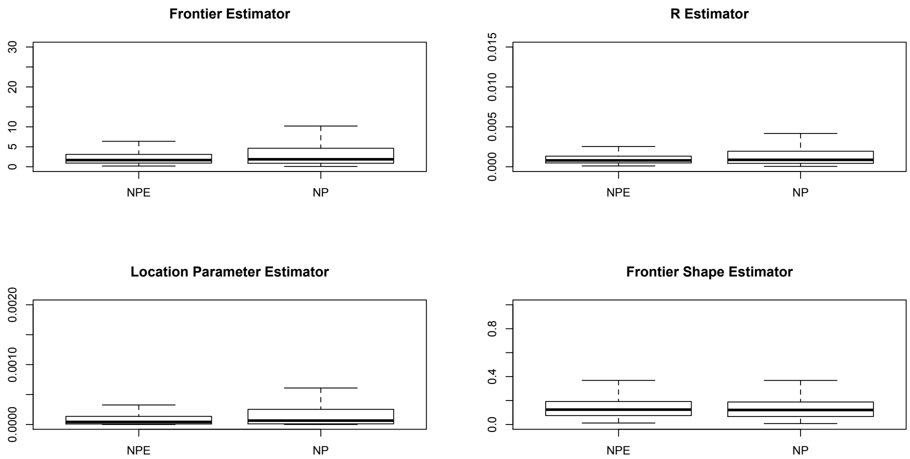


Figura 2.6: Frontier I - Boxplot of Estimators - $n = 200$ - $\mu_r = 0.75$

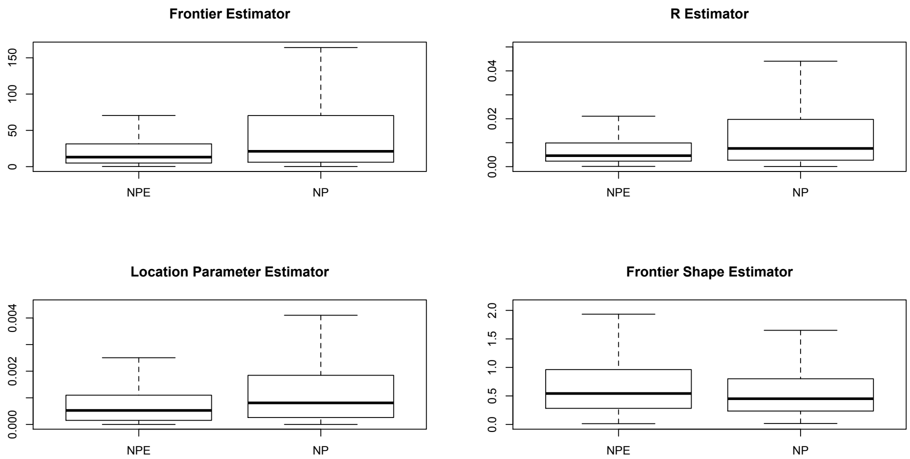



Figura 2.7: Frontier I - Boxplot of Estimators - $n = 400$ - $\mu_r = 0.75$

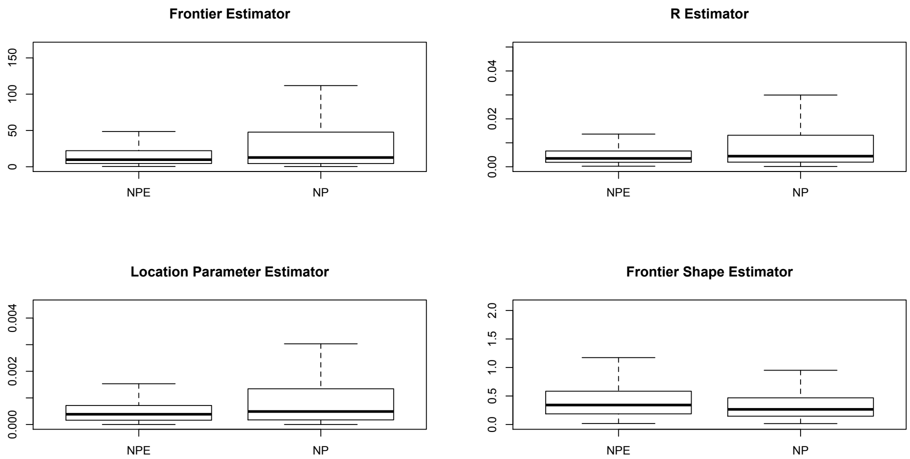


Figura 2.8: Frontier II - Boxplot of Estimators - $n = 200$ - $\mu_r = 0.25$

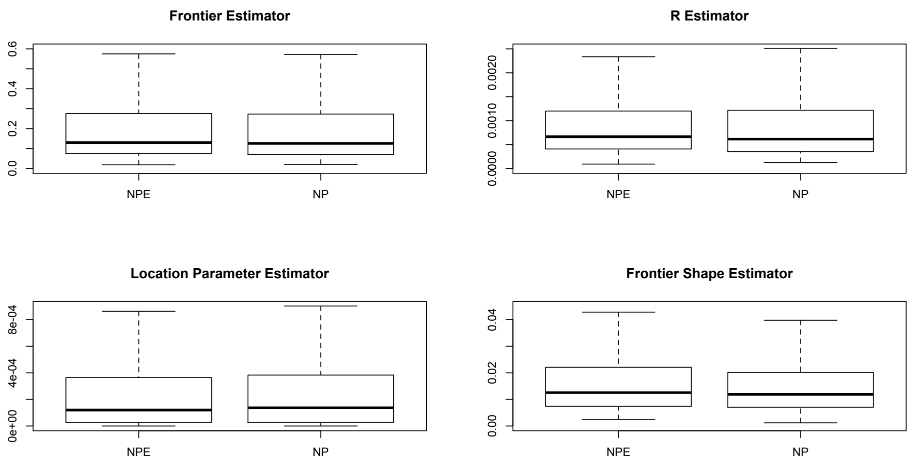



Figura 2.9: Frontier II - Boxplot of Estimators - $n = 400$- $\mu_r = 0.25$

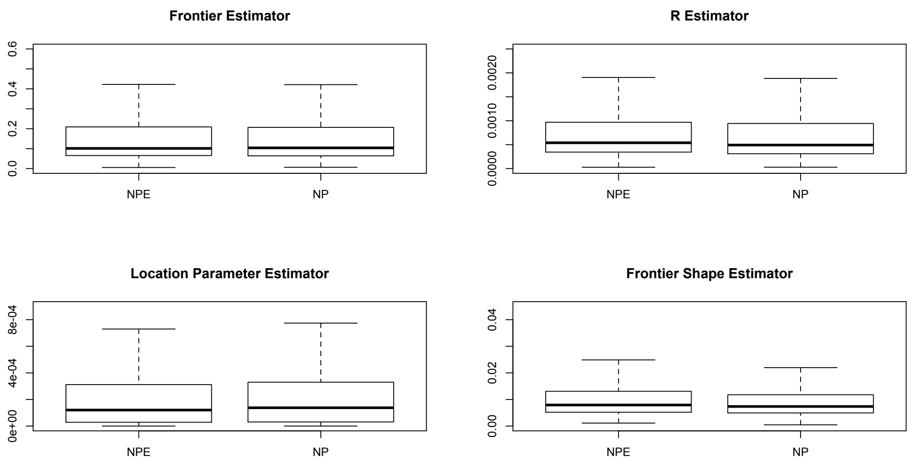


Figura 2.10: Frontier II - Boxplot of Estimators - $n = 200$ - $\mu_r = 0.5$

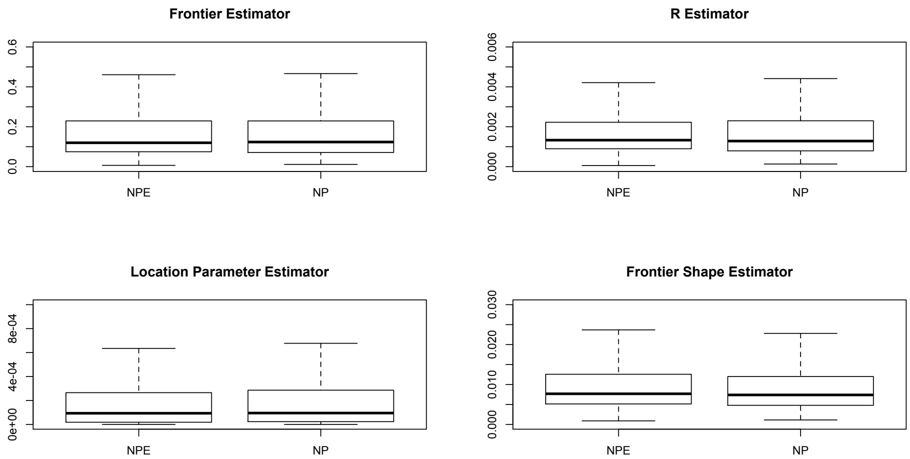



Figura 2.11: Frontier II - Boxplot of Estimators - $n = 400$ - $\mu_r = 0.5$

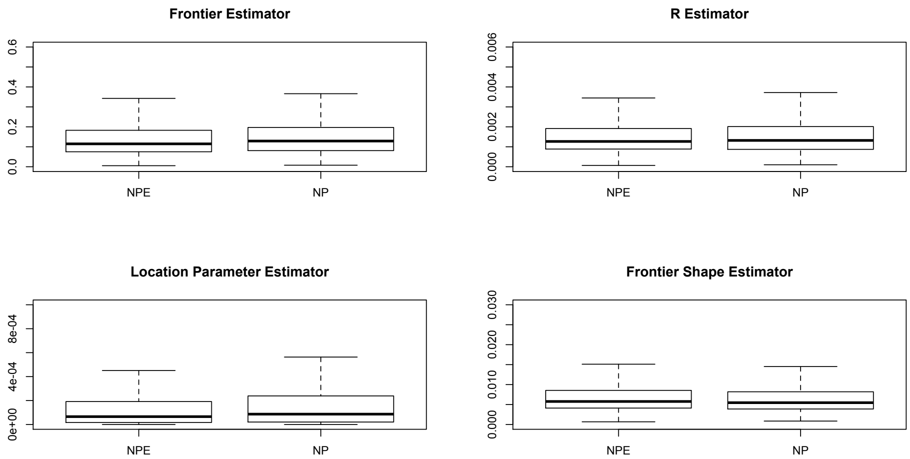


Figura 2.12: Frontier II - Boxplot of Estimators - $n = 200$ - $\mu_r = 0.75$

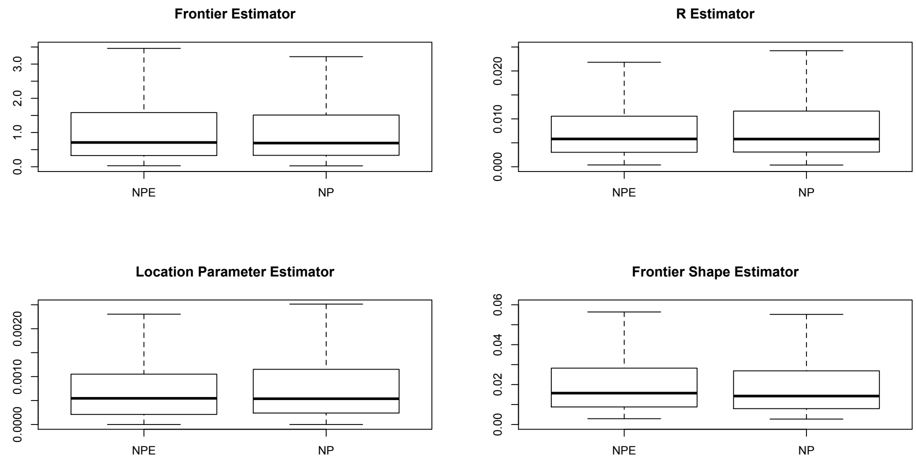



Figura 2.13: Frontier II - Boxplot of Estimators - $n = 400$ - $\mu_r = 0.75$

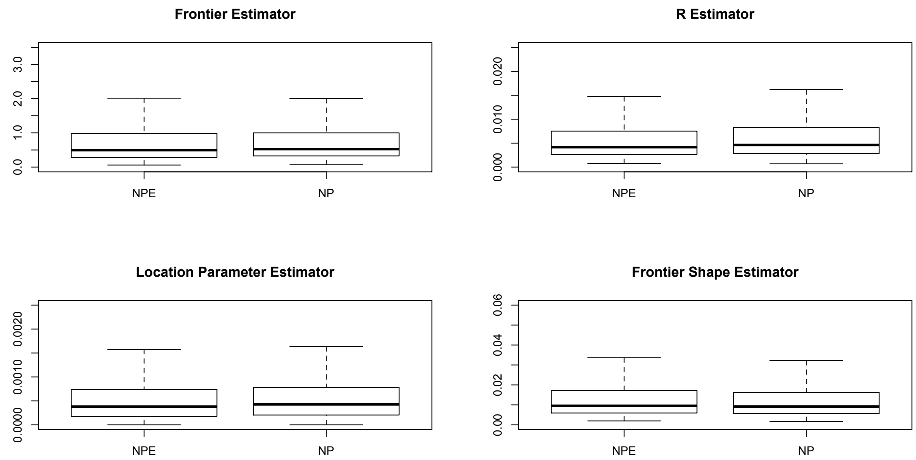



Figura 2.14: Frontier I - Density estimates for NPE and NP estimators - $\mu_r = 0.5$.

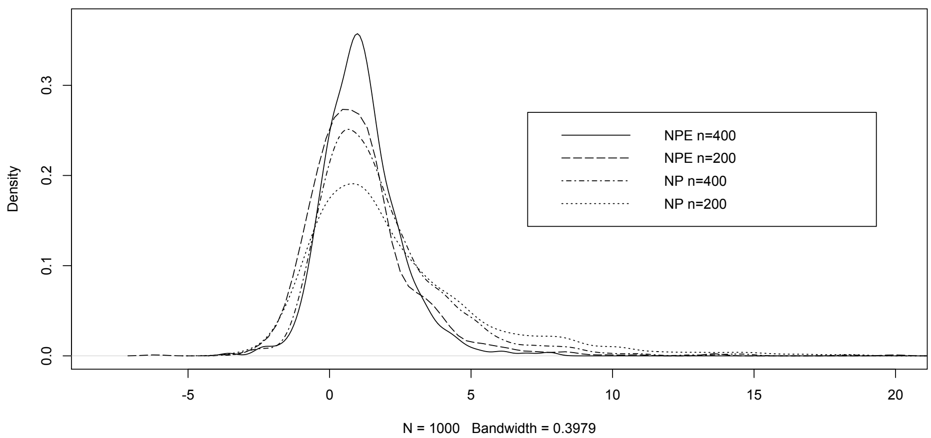


Tabela 2.7: Real Data Example - NPE estimated efficiency

| Rank | Agency | State | $\hat{R}$ | Rank | Agency | State | $\hat{R}$ |
|---|---|---|---|---|---|---|---|
| 1 | San Jose Police Dept | CA | 1.00 | 81 | Fullerton Police Dept | CA | 0.43 |
| 2 | Shreveport Police Dept | LA | 0.92 | 82 | Hidalgo County Sheriff Department | TX | 0.42 |
| 3 | Fulton County Police Department | GA | 0.92 | 83 | Fresno County Sheriff Department | CA | 0.42 |
| 4 | El Dorado County Sheriff Department | CA | 0.91 | 84 | Harris County Sheriff Office | TX | 0.41 |
| 5 | Harford County Sheriff Office | MD | 0.89 | 85 | Mobile County Sheriff Department | AL | 0.41 |
| 6 | Oakland County Sheriff Office | MI | 0.88 | 86 | Green Bay Police Dept | WI | 0.41 |
| 7 | Springfield Police Dept MO | MO | 0.87 | 87 | Sacramento County Sheriff Department | CA | 0.41 |
| 8 | Jefferson County Sheriff Department CO | CO | 0.81 | 88 | Bexar County Sheriff Office | TX | 0.40 |
| 9 | San Diego County Sheriff Department | CA | 0.80 | 89 | Simi Valley Police Dept | CA | 0.40 |
| 10 | Santa Barbara County Sheriff Department | CA | 0.78 | 90 | Santa Rosa Police Dept | CA | 0.40 |
| 11 | Naperville Police Dept | IL | 0.77 | 91 | Midland Police Dept | TX | 0.40 |
| 12 | Jefferson County Sheriff Department MO | MO | 0.75 | 92 | Santa Cruz County Sheriff Department | CA | 0.39 |
| 13 | Allen County Sheriff Department | IN | 0.74 | 93 | Ann Arbor Police Dept | MI | 0.39 |
| 14 | Amherst Town Police Dept | NY | 0.74 | 94 | Forsyth County Sheriff Department | NC | 0.39 |
| 15 | Snohomish County Sheriff Office | WA | 0.73 | 95 | Oceanside | CA | 0.39 |
| 16 | Washington Metropolitan Police Dept | DC | 0.73 | 96 | Jefferson County Sheriff Department AL | AL | 0.39 |
| 17 | New Castle County Police Department | DE | 0.70 | 97 | Brevard County Sheriff Department | FL | 0.39 |
| 18 | Monterey County Sheriff Department | CA | 0.70 | 98 | Plano Police Dept | TX | 0.39 |
| 19 | El Paso | CO | 0.69 | 99 | Kitsap County Sheriff Office | WA | 0.39 |
| 20 | Arapahoe County Sheriff Department | CO | 0.69 | 100 | Pasco County Sheriff Department | FL | 0.38 |
| 21 | Fort Bend County Sheriff Department | TX | 0.68 | 101 | Pima County Sheriff Department | AZ | 0.38 |
| 22 | Onondaga County Sheriff Department | NY | 0.68 | 102 | Lexington County Sheriff Department | SC | 0.38 |
| 23 | Stockton Police Dept | CA | 0.66 | 103 | Hampton Police Dept | VA | 0.38 |
| 24 | Waco Police Dept | TX | 0.64 | 104 | Kern County Sheriff Department | CA | 0.38 |
| 25 | Butte County Sheriff Department | CA | 0.62 | 105 | Erie City Police Dept | PA | 0.38 |
| 26 | Irvine Police Police | CA | 0.62 | 106 | Anderson County Sheriff Department | SC | 0.38 |
| 27 | St Charles County Sheriff Department | MO | 0.60 | 107 | Pinellas County Sheriff Department | FL | 0.38 |
| 28 | Charlotte County Sheriff Department | FL | 0.59 | 108 | Clay County Sheriff Department | FL | 0.37 |
| 29 | St Petersburg Police Dept | FL | 0.59 | 109 | Long Beach Police Dept | CA | 0.37 |
| 30 | Huntington Beach Police Dept | CA | 0.59 | 110 | Richland County Sheriff Department | SC | 0.37 |
| 31 | Monroe County Sheriff Office | NY | 0.58 | 111 | Charles County Sheriff Office | MD | 0.37 |
| 32 | Sioux Falls Police Dept | SD | 0.58 | 112 | Hernando County Sheriff Department | FL | 0.37 |
| 33 | Kent County Sheriff Office | MI | 0.58 | 113 | Madison City Police Dept | WI | 0.36 |
| 34 | Stanislaus County Sheriff Department | CA | 0.58 | 114 | Escondido Police Dept | CA | 0.36 |
| 35 | Buncombe County Sheriff Department | NC | 0.58 | 115 | Collier County Sheriff Department | FL | 0.36 |
| 36 | Warren Police Dept | MI | 0.56 | 116 | Lee County Sheriff Department | FL | 0.36 |
| 37 | Glendale Police Dept CA | CA | 0.56 | 117 | Hayward Police Dept | CA | 0.35 |
| 38 | Tucson Police Dept | AZ | 0.55 | 118 | Pomona Police Dept | CA | 0.35 |
| 39 | San Bernardino County Sheriff Department | CA | 0.55 | 119 | Chula Vista Police Dept | CA | 0.34 |
| 40 | San Joaquin County Sheriff Department | CA | 0.54 | 120 | Garland Police Dept | TX | 0.34 |
| 41 | St Tammany Parish Sheriff Department | LA | 0.54 | 121 | Concord Police Dept | CA | 0.34 |
| 42 | Okaloosa County Sheriff Department | FL | 0.53 | 122 | Waterbury Police Dept | CT | 0.34 |
| 43 | Orange Police Dept | CA | 0.53 | 123 | Anne Arundel County Police Department | MD | 0.34 |
| 44 | Arlington County Police Department | VA | 0.52 | 124 | Spartanburg County Sheriff Office | SC | 0.34 |
| 45 | Contra Costa County Sheriff Department | CA | 0.51 | 125 | Greenville County Sheriff Office | SC | 0.34 |
| 46 | Seattle Police Dept | WA | 0.51 | 126 | Chesapeake Police Dept | VA | 0.33 |
| 47 | Guilford County Sheriff Office | NC | 0.51 | 127 | Bakersfield Police Dept | CA | 0.33 |
| 48 | Clark County Sheriff Department | WA | 0.51 | 128 | Fort Collins Police Dept | CO | 0.33 |
| 49 | Torrance Police Dept | CA | 0.51 | 129 | Alexandria Police Dept | VA | 0.33 |
| 50 | Montgomery County Sheriff Department | TX | 0.51 | 130 | Pasadena Police Dept CA | CA | 0.33 |
| 51 | Riverside County Sheriff Department | CA | 0.50 | 131 | Abilene Police Dept | TX | 0.33 |
| 52 | Prince William County Police Department | VA | 0.50 | 132 | Inglewood Police Dept | CA | 0.33 |
| 53 | Tempe Police Dept | AZ | 0.49 | 133 | Bernalillo County Sheriff Department | NM | 0.33 |
| 54 | Knox County Sheriff Office | TN | 0.49 | 134 | Aurora Police Dept IL | IL | 0.33 |
| 55 | Worcester Police Dept | MA | 0.49 | 135 | Henrico County Police Dept | VA | 0.33 |
| 56 | Akron City Police Dept | OH | 0.49 | 136 | Vallejo Police Dept | CA | 0.33 |
| 57 | Downey Police Dept | CA | 0.49 | 137 | Cambridge Police Dept | MA | 0.33 |
| 58 | San Francisco Police Dept | CA | 0.49 | 138 | Salinas Police Dept | CA | 0.32 |
| 59 | King County Sheriff Office | WA | 0.48 | 139 | South Bend Police Dept | IN | 0.32 |
| 60 | Stark County Sheriff Office | OH | 0.48 | 140 | Cumberland County Sheriff Office | NC | 0.32 |
| 61 | Cobb County Police Department | GA | 0.47 | 141 | Las Vegas Metropolitan Police Department | NV | 0.31 |
| 62 | Chesterfield County Police Department | VA | 0.47 | 142 | Pasadena Police Dept TX | TX | 0.31 |
| 63 | Marion County Sheriff Department | FL | 0.47 | 143 | Boise Police Dept | ID | 0.31 |
| 64 | Anaheim Police Dept | CA | 0.47 | 144 | Escambia County Sheriff Department | FL | 0.30 |
| 65 | Wake County Sheriff Department | NC | 0.46 | 145 | Polk County Sheriff Department | FL | 0.30 |
| 66 | Norwalk | CA | 0.46 | 146 | Savannah Police Dept | GA | 0.30 |
| 67 | Gwinnett County Police Department | GA | 0.46 | 147 | Denver Police Dept | CO | 0.30 |
| 68 | Livonia Police Dept | MI | 0.45 | 148 | Stamford Police Dept | CT | 0.30 |
| 69 | El Monte Police Dept | CA | 0.45 | 149 | Baltimore County Police Department | MD | 0.29 |
| 70 | Alameda County Sheriff Department | CA | 0.45 | 150 | Chandler Police Dept | AZ | 0.29 |
| 71 | Burbank Police Dept | CA | 0.44 | 151 | Indianapolis Police Dept | IN | 0.29 |
| 72 | Garden Grove Police Dept | CA | 0.44 | 152 | Irving Police Dept | TX | 0.29 |
| 73 | Virginia Beach Police Dept | VA | 0.44 | 153 | Pierce County Sheriff Department | WA | 0.29 |
| 74 | Fontana Police Dept | CA | 0.44 | 154 | Riverside Police Dept | CA | 0.29 |
| 75 | Hamilton County Sheriff Department | OH | 0.44 | 155 | Manatee County Sheriff Department | FL | 0.29 |
| 76 | Howard County Police Department | MD | 0.44 | 156 | Ontario Police Dept | CA | 0.29 |
| 77 | Henderson Police Dept | NV | 0.43 | 157 | Mesquite Police Dept | TX | 0.29 |
| 78 | Topeka Police Dept | KS | 0.43 | 158 | Clackamas County Sheriff Department | OR | 0.28 |
| 79 | Montgomery County Police Department | MD | 0.43 | 159 | Palm Beach County Sheriff Department | FL | 0.28 |
| 80 | Oxnard Police Dept | CA | 0.43 | 160 | Jersey City Police Dept | NJ | 0.28 |

| Rank | Agency (cont.) | State | $\hat{R}$ | Rank | Agency | State | $\hat{R}$ |
|---|---|---|---|---|---|---|---|
| 161 | Anchorage Police Dept | AK | 0.28 | 228 | Corpus Christi Police Dept | TX | 0.19 |
| 162 | Colorado Springs | CO | 0.28 | 229 | Eugene Police Dept | OR | 0.19 |
| 163 | Evansville Police Dept | IN | 0.27 | 230 | Beaumont Police Dept | TX | 0.19 |
| 164 | Hillsborough County Sheriff Department | FL | 0.27 | 231 | Lubbock Police Dept | TX | 0.19 |
| 165 | Dekalb County Public Safety Department | GA | 0.27 | 232 | Salem Police Dept | OR | 0.19 |
| 166 | Pueblo Police Dept | CO | 0.27 | 233 | Milwaukee Police Dept | WI | 0.19 |
| 167 | Allentown City Police Dept | PA | 0.27 | 234 | Providence Police Dept | RI | 0.19 |
| 168 | Clearwater Police Dept | FL | 0.27 | 235 | Albany Police Dept | NY | 0.18 |
| 169 | Grand Prairie Police Dept | TX | 0.27 | 236 | Lafayette Police Dept | LA | 0.18 |
| 170 | Winston-Salem Police Dept | NC | 0.27 | 237 | Berkeley Police Dept | CA | 0.18 |
| 171 | Lakewood | CO | 0.27 | 238 | Spokane Police Dept | WA | 0.18 |
| 172 | Cedar Rapids Police Dept | IA | 0.26 | 239 | Laredo | TX | 0.18 |
| 173 | Salt Lake County Sheriff Office | UT | 0.26 | 240 | Huntsville Police Dept | AL | 0.18 |
| 174 | Honolulu Police Dept | HI | 0.26 | 241 | Fresno Police Dept | CA | 0.18 |
| 175 | Charleston County Sheriff Department | SC | 0.26 | 242 | Thurston County Sheriff Department | WA | 0.18 |
| 176 | Reno Police Dept | NV | 0.25 | 243 | Portland Police Dept | OR | 0.18 |
| 177 | Newport News Police Dept | VA | 0.25 | 244 | Sonoma County Sheriff Department | CA | 0.18 |
| 178 | Modesto Police Dept | CA | 0.24 | 245 | Amarillo Police Dept | TX | 0.18 |
| 179 | Aurora Police Dept CO | CO | 0.24 | 246 | Washington County Sheriff Office | OR | 0.18 |
| 180 | Pittsburgh Bureau Of Police | PA | 0.24 | 247 | Rochester Police Dept | NY | 0.17 |
| 181 | Elizabeth Police Dept | NJ | 0.24 | 248 | Charlotte-Mecklenburg Police Department | NC | 0.17 |
| 182 | Lansing City Police Dept | MI | 0.24 | 249 | Volusia County Sheriff Department | FL | 0.17 |
| 183 | Fort Wayne Police Dept | IN | 0.24 | 250 | Tacoma Police Dept | WA | 0.17 |
| 184 | Louisville Police Dept | KY | 0.23 | 251 | New Haven Police Dept | CT | 0.17 |
| 185 | Boston Police Dept | MA | 0.23 | 252 | Montgomery Police Dept | AL | 0.17 |
| 186 | Austin Police Dept | TX | 0.23 | 253 | St Paul Police Dept | MN | 0.17 |
| 187 | Lincoln Police Dept | NE | 0.23 | 254 | Brownsville Police Dept | TX | 0.17 |
| 188 | West Covina Police Dept | CA | 0.23 | 255 | Scottsdale Police Dept | AZ | 0.17 |
| 189 | Knoxville Police Dept | TN | 0.23 | 256 | East Baton Rouge Parish Sheriff Dept | LA | 0.17 |
| 190 | Gary Police Dept | IN | 0.23 | 257 | Rockford Police Dept | IL | 0.17 |
| 191 | Tulsa Police Dept | OK | 0.23 | 258 | Springfield Police Dept IL | IL | 0.17 |
| 192 | El Paso Police Dept | TX | 0.22 | 259 | Sarasota County Sheriff Department | FL | 0.17 |
| 193 | Hialeah Police Dept | FL | 0.22 | 260 | Spokane County Sheriff Department | WA | 0.16 |
| 194 | Sunnyvale Dept Of Public Safety | CA | 0.22 | 261 | Fort Lauderdale Police Dept | FL | 0.16 |
| 195 | San Bernardino Police Dept | CA | 0.22 | 262 | Seminole County Sheriff Department | FL | 0.16 |
| 196 | Wichita Falls Police Dept | TX | 0.22 | 263 | Birmingham Police Dept | AL | 0.16 |
| 197 | Columbus Police Dept GA | GA | 0.22 | 264 | Columbus Police Dept OH | OH | 0.16 |
| 198 | Oakland Police Dept | CA | 0.22 | 265 | Columbia Police Dept | SC | 0.16 |
| 199 | Prince Georges County Police Department | MD | 0.22 | 266 | Albuquerque Police Dept | NM | 0.16 |
| 200 | Portsmouth Police Dept | VA | 0.22 | 267 | Springfield Police Dept MA | MA | 0.16 |
| 201 | Bridgeport Police Dept | CT | 0.22 | 268 | Durham Police Dept | NC | 0.16 |
| 202 | Mesa Police Dept | AZ | 0.21 | 269 | Richmond (City) Bureau Of Police | VA | 0.16 |
| 203 | Jefferson Parish Sheriff Department | LA | 0.21 | 270 | Flint City Police Dept | MI | 0.16 |
| 204 | Washtenaw County Sheriff Department | MI | 0.21 | 271 | Nashville-Davidson Metro Police Dept | TN | 0.16 |
| 205 | St Louis Police Dept | MO | 0.21 | 272 | Travis County Sheriff Department | TX | 0.15 |
| 206 | Grand Rapids Police Dept | MI | 0.21 | 273 | Wichita Police Dept | KS | 0.15 |
| 207 | Glendale Police Dept AZ | AZ | 0.21 | 274 | Peoria Police Dept | IL | 0.15 |
| 208 | Greensboro Police Dept | NC | 0.21 | 275 | Memphis Police Dept | TN | 0.15 |
| 209 | Jacksonville | FL | 0.20 | 276 | Hartford Police Dept | CT | 0.15 |
| 210 | Cincinnati Police Dept | OH | 0.20 | 277 | Baton Rouge Police Dept | LA | 0.15 |
| 211 | Sacramento Police Dept | CA | 0.20 | 278 | Salt Lake City Police Dept | UT | 0.15 |
| 212 | Mobile Police Dept | AL | 0.20 | 279 | Oklahoma City Police Dept | OK | 0.15 |
| 213 | Norfolk Police Dept | VA | 0.20 | 280 | Little Rock Police Dept | AR | 0.14 |
| 214 | Cleveland | OH | 0.20 | 281 | Syracuse Police Dept | NY | 0.14 |
| 215 | Raleigh Police Dept | NC | 0.20 | 282 | Tampa Police Dept | FL | 0.14 |
| 216 | Tulare County Sheriff Department | CA | 0.20 | 283 | Dayton | OH | 0.14 |
| 217 | Omaha Police Dept | NE | 0.20 | 284 | Baltimore City Police Dept | MD | 0.14 |
| 218 | Buffalo Police Dept | NY | 0.20 | 285 | Toledo Police Dept | OH | 0.13 |
| 219 | Des Moines Police Dept | IA | 0.20 | 286 | Jackson Police Dept | MS | 0.13 |
| 220 | Hollywood Police Dept | FL | 0.20 | 287 | Kansas City Police Dept | MO | 0.13 |
| 221 | New Orleans Police Dept | LA | 0.20 | 288 | Miami Police Dept | FL | 0.13 |
| 222 | Metro-Dade Police Department | FL | 0.20 | 289 | Macon Police Dept | GA | 0.12 |
| 223 | Sterling Heights Police Dept | MI | 0.20 | 290 | Tallahassee Police Dept | FL | 0.12 |
| 224 | Independence Police Dept | MO | 0.20 | 291 | Chattanooga Police Dept | TN | 0.12 |
| 225 | Newark Police Dept | NJ | 0.20 | 292 | Orlando Police Dept | FL | 0.11 |
| 226 | Minneapolis Police Dept | MN | 0.19 | 293 | Atlanta Police Dept | GA | 0.10 |
| 227 | Fort Worth Police Dept | TX | 0.19 | 294 | St Joseph County Sheriff Department | IN | 0.09 |

## 2.8 Appendix 2: Proofs

*Proof of Lemma 1.* Given the algebraic structure of the optimand in equation (3) we can write

$$A_n \equiv \gamma_1^*(x) - \sigma^2(x) - \frac{1}{nh_n g_X(x)} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\left(e_i - \hat{\varepsilon}_i - \sigma^2(x) - \sigma^{2(1)}(x)(X_i - x)\right).$$

Letting $S(x) = \begin{pmatrix} g_X(x) & 0 \\ 0 & g_X(x)\sigma_K^2 \end{pmatrix}$ we have by the Cauchy-Schwarz inequality that

$$|A_n| \le \frac{1}{h_n}\left((1,0)(S_n^{-1}(x) - S^{-1}(x))^2(1,0)'\right)^{1/2} R_{n,1}(x).$$

From part (b) of Lemma 1 in Martins-Filho and Yao (2007),

$$B_n(x) \equiv h_n^{-1}\left((1,0)(S_n^{-1}(x) - S^{-1}(x))^2(1,0)'\right)^{1/2} = O_p(1)$$

uniformly in $G$, therefore completing the proof.

*Proof of Theorem 1.* a) Given the comments following Lemma, it suffices to investigate the order of $|c_1(x)|$. After substituting $e_i$, we write $c_1(x) = I_{1n}(x) + I_{2n}(x) + I_{3n}(x) + I_{4n}(x) - I_{5n}(x)$, where

$$\begin{aligned}
I_{1n}(x) &= \frac{1}{nh_n g_X(x)} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\left(\sigma^2(X_i) - \sigma^2(x) - \sigma^{2(1)}(x)(X_i - x)\right) \\
I_{2n}(x) &= \frac{1}{nh_n g_X(x)} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)(\epsilon_i^2 - 1)\sigma^2(X_i) \\
I_{3n}(x) &= \frac{2}{nh_n g_X(x)} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\sigma(X_i)\epsilon_i(\hat{m}(X_i; h_n) - m(X_i)) \\
I_{4n}(x) &= \frac{1}{nh_n g_X(x)} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)(\hat{m}(X_i; h_n) - m(X_i))^2 \\
I_{5n}(x) &= \frac{1}{nh_n g_X(x)} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\hat{\varepsilon}_i
\end{aligned}$$

The uniform order in probability of $I_{jn}(x)$ for $j = 2, 3, 4$ on the set $G$ is given in Martins-Filho and Yao (2007) Theorem 1, part (a). Here we study the order of $I_{1n}(x) - I_{5n}(x)$. Note that

$$\begin{aligned}
I_{1n}(x) - I_{5n}(x) &= \frac{1}{nh_n g_X(x)} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\Big(\sigma^2(X_i) - \sigma^2(x) - \sigma^{2(1)}(x)(X_i - x) \\
&\quad - \frac{1}{2}L^{(2)}(\lambda_i(X_i - x), \hat{\theta}(x))(X_i - x)^2\Big) \\
&= \frac{1}{nh_n g_X(x)} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\Big(\frac{1}{2}\sigma^{2(2)}(\lambda_i'(X_i - x) + x)(X_i - x)^2 \\
&\quad - \frac{1}{2}L^{(2)}(\lambda_i(X_i - x), \hat{\theta}(x))(X_i - x)^2\Big) \\
&= \frac{1}{nh_n g_X(x)} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right) R_i(X_i - x)^2 \\
&= \frac{h_n}{n g_X(x)} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)^2 R_i.
\end{aligned}$$

where $R_i = \frac{1}{2}\bigg(\sigma^{2(2)}(\lambda_i'(X_i - x) + x) - L^{(2)}(\lambda_i(X_i - x), \hat{\theta}(x))\bigg)$ and $\lambda_i' \in [0, 1]$. Now, consider rewriting $R_i$ as

$$\begin{aligned}
R_i &= \frac{1}{2}\bigg(\sigma^{2(2)}(\lambda_i^{'}(X_i - x) + x) - L^{(2)}(\lambda_i(X_i - x), \hat{\theta}(x)) + L^{(2)}(0, \theta^0) - L^{(2)}(0, \theta^0)\bigg) \\
&= \frac{1}{2}\bigg(\sigma^{2(2)}(\lambda_i^{'}(X_i - x) + x) - L^{(2)}(0, \theta^0)\bigg) + \frac{1}{2}\bigg(L^{(2)}(0, \theta^0) - L^{(2)}(\lambda_i(X_i - x), \hat{\theta}(x))\bigg).
\end{aligned}$$

Then,

$$\begin{aligned}
I_{1n}(x) - I_{5n}(x) &= \frac{h_n}{2ng_X(x)} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)^2 \bigg(\sigma^{2(2)}(\lambda_i^{'}(X_i - x) + x) - L^{(2)}(0, \theta^0)\bigg) \\
&\quad + \frac{h_n}{2ng_X(x)} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)^2 \bigg(L^{(2)}(0, \theta^0) - L^{(2)}(\lambda_i(X_i - x), \hat{\theta}(x))\bigg) \\
&= J_{1n}(x) + J_{2n}(x).
\end{aligned}$$

Recall that $L^{(2)}(0, \theta) = \exp(\theta_1(x))(\theta_2(x))^2 = \sigma^2(x)(f^{(1)}(x))^2$ and $L^{(2)}(0, \theta) < C$ an arbitrary bound provided $|f^{(1)}(x)| < B_f$, given that $0 < \sigma^2(x) < \bar{B}_\sigma^2$. Also, since $|\sigma^{2(2)}(x)| < \bar{B}_{2\sigma}$ for all $x$ by Martins-Filho and Yao (2007, p. 306), we have that $\sup_{x\in G}|J_{1n}(x)| \leq O_p(h_n^2)$. Now,

$$\begin{aligned}
J_{2n}(x) &= -\frac{h_n}{2ng_X(x)} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)^2 \bigg((\hat{\theta}_2(x)^2 \exp(\hat{\theta}_1(x) - \theta_2(x)^2 \exp(\theta_1(x))) \\
&\qquad \times \exp(\hat{\theta}_2(x)(X_i - x)\lambda_i) + \theta_2^2(x) \exp(\theta_1(x))(\exp(\hat{\theta}_2(x)(X_i - x)\lambda_i) - 1)\bigg) \\
&= (\hat{\theta}_2(x)^2 \exp(\hat{\theta}_1(x)) - \theta_2(x)^2 \exp(\theta_1(x)) \frac{-h_n}{2ng_X(x)} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)^2 \exp(\hat{\theta}_2(x)(X_i - x)\lambda_i) \\
&\qquad - \theta_2^2(x) \exp(\theta_1(x)) \frac{h_n}{2ng_X(x)} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)^2 (\exp(\hat{\theta}_2(x)(X_i - x)\lambda_i) - 1).
\end{aligned}$$

Note that whenever $\left|\frac{X_i - x}{h_n}\right| > 1$ we have $K\left(\frac{X_i - x}{h_n}\right) = 0$. Hence, consider

$M_n(x) = \frac{-h_n}{ng_X(x)} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)^2 \exp(\hat{\theta}_2(x)(X_i - x)\lambda_i)$.

All terms in the sum are positive, and since the exponential function is everywhere increasing $\exp(\hat{\theta}_2(x)(X_i-$

$x)\lambda_i) \le \exp(|\hat{\theta}_2(x)|h_n)$ since $0 < \lambda_i < 1$ and $|X_i - x| \le h_n$, otherwise $K\left(\frac{X_i - x}{h_n}\right) = 0$. Therefore,

$$\begin{aligned}|M_n(x)| &\le \frac{h_n e^{|\hat{\theta}_2(x)|h_n}}{n g_X(x)} \sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)^2 \\ &\le \underline{\mathrm{B}}_{gX}^{-1} e^{|\hat{\theta}_2(x)|h_n} \frac{h_n^2}{n h_n} \sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)^2 \\ &= \underline{\mathrm{B}}_{gX}^{-1} e^{|\hat{\theta}_2(x)|h_n} h_n^2 \Bigg\{ \frac{1}{n h_n} \sum_{i=1}^n \Bigg[ K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)^2 \\ &\quad - \frac{1}{h_n} E\Bigg( K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)^2 \Bigg)\Bigg] + \frac{1}{h_n} E\Bigg( K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)^2 \Bigg)\Bigg\}\end{aligned}$$

and consequently

$$\begin{aligned}\sup_{x\in G} |M_n(x)| &\le \underline{\mathrm{B}}_{gX}^{-1} \sup_{x\in G} e^{|\hat{\theta}_2(x)|h_n} h_n^2 \Bigg\{ \sup_{x\in G} \Bigg| \frac{1}{n h_n} \sum_{i=1}^n \Bigg[ K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)^2 \\ &\quad - \frac{1}{h_n} E\Bigg( K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)^2 \Bigg)\Bigg]\Bigg| \\ &\quad + \sup_{x\in G} \frac{1}{h_n} E\Bigg( K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)^2 \Bigg)\Bigg\}.\end{aligned}$$

From Martins-Filho and Yao (2007, p. 306),

$$\sup_{x\in G} \Bigg| \frac{1}{n h_n} \sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)^2 - \frac{1}{h_n} E\Bigg( K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)^2 \Bigg)\Bigg| = o_p(h_n),$$

and $\sup_{x\in G} \frac{1}{h_n} E\Big( K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)^2 \Big) = O(1)$. Furthermore, given that $\hat{\theta}_2(x)$ is an uniformly consistent estimator for $\theta_2(x)$, we have $e^{|\hat{\theta}_2(x)|} = e^{|\hat{\theta}_2(x) - \theta_2(x) + \theta_2(x)|h_n} \le e^{|\hat{\theta}_2(x) - \theta_2(x)|h_n + |\theta_2(x)|h_n} \xrightarrow{p} 1$ uniformly in $G$. Hence,

$$\begin{aligned}\sup_{x\in G} |M_n(x)| &\le \underline{\mathrm{B}}_{gX}^{-1} h_n^2 (h_n o_p(1) + O(1)) \\ &= \underline{\mathrm{B}}_{gX}^{-1} h_n^3 o_p(1) + \underline{\mathrm{B}}_{gX}^{-1} O(h_n^2) \\ &= O_p(h_n^2).\end{aligned}$$

Since $(\hat{\theta}(x) - \theta_(x)) = o_p(1)$ uniformly in $G$, we have by Slutsky Theorem that $\hat{\theta}_2(x)^2 e^{\hat{\theta}_1(x)} - \theta_2^2(x) e^{\theta_1(x)} = o_p(1)$. Similarly,

$$-\theta_2^2(x) e^{\theta_1(x)} \frac{h_n}{2 n g_X(x)} \sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)^2 (e^{\hat{\theta}_2(x)(X_i - x)\lambda_i} - 1) = o_p(h_n^2).$$

Hence, $\sup_{x\in G} |J_{2n}(x)| = o_p(h_n^2)$. In all, $\sup_{x\in G} |I_{1n}(x) - I_{5n}(x)| = O_p(h_n^2)$, and using the results in Martins-Filho and Yao (2007) for $I_{2n}(x)$, $I_{3n}(x)$ and $I_{4n}(x)$ we have

$$sup_{x\in G} \Bigg| \gamma_1^*(x) - \sigma^2(x) - \frac{1}{n h_n g_X(x)} \sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right)\Big(e_i - \hat{\varepsilon}_i - \sigma^2(x) - \sigma^{2(1)}(x)(X_i - x)\Big)\Bigg| \le$$

$$O_p(h_n^3) + O_p\Bigg(\left(\frac{h_n \ln(n)}{n}\right)^{1/2}\Bigg),$$

which completes the proof of part a).

b) A direct consequence of a) is the fact that

$$\sqrt{nh_n}(\gamma_1^*(x)-\sigma^2(x)) - \frac{1}{\sqrt{nh_n}g_X(x)}\sum_{i=1}^n K\left(\frac{X_i-x}{h_n}\right)\left(e_i-\hat{\varepsilon}_i-\sigma^2(x)-\sigma^{2(1)}(x)(X_i-x)\right)$$
$$= \sqrt{nh_n^7}O_p(1)+(h_n^2\ln(n))^{1/2}O_p(1).$$

Hence, provided $h_n^2\ln(n)\to 0$ as $n\to\infty$, the asymptotic distribution of $\sqrt{nh_n}(\gamma_1^*(x)-\sigma^2(x))$ is the same as that of

$$\frac{1}{\sqrt{nh_n}g_X(x)}\sum_{i=1}^n K\left(\frac{X_i-x}{h_n}\right)\left(e_i-\hat{\varepsilon}_i-\sigma^2(x)-\sigma^{2(1)}(x)(X_i-x)\right) = \sqrt{nh_n}c_1(x),$$

which can be written as

$$\sqrt{nh_n}c_1(x) = \sqrt{nh_n}(I_{1n}(x)-I_{5n}(x)+I_{2n}(x)+I_{3n}(x)+I_{4n}(x)).$$

From Martins-Filho and Yao (2007) we have that

$$\sqrt{nh_n}I_{2n}(x) \xrightarrow{d} N\left(0, \frac{\sigma^4(x)}{g_X(x)}(\mu_4(x)-1)\int K^2(y)dy\right),$$

Also, $\sqrt{nh_n}I_{3n}(x) = \sqrt{nh_n}\left(\frac{1}{\sqrt{n}}o_p(1)+h_n^2 o_p(1)\right) = \sqrt{h_n}o_p(1)+\sqrt{nh_n^5}o_p(1)$. Hence, provided that $nh_n^5 = O(1)$, $\sqrt{nh_n}I_{3n}(x) = o_p(1)$. Moreover, $\sqrt{nh_n}I_{4n}(x) = \sqrt{nh_n}(h_n^2 o_p(1)) = \sqrt{nh_n^5}o_p(1) = o_p(1)$. We now focus on $\sqrt{nh_n}(I_{1n}(x)-I_{5n}(x)) = \sqrt{nh_n}\frac{1}{nh_n g_X(x)}\sum_{i=1}^n K\left(\frac{X_i-x}{h_n}\right)\left(\frac{X_i-x}{h_n}\right)^2 R_i = \sqrt{nh_n}B_n(x)$. Note that we can write,

$$\sqrt{nh_n}B_n(x) = \sqrt{nh_n}\left(\frac{h_n}{n}\frac{1}{2g_X(x)}\sum_{i=1}^n K\left(\frac{X_i-x}{h_n}\right)\left(\frac{X_i-x}{h_n}\right)^2(\sigma^{2(2)}(\lambda_i'(X_i-x)+x)-L^{(2)}(0,\theta^0))\right)$$
$$+\sqrt{nh_n}\left(\frac{h_n}{n}\frac{1}{2g_X(x)}\sum_{i=1}^n K\left(\frac{X_i-x}{h_n}\right)\left(\frac{X_i-x}{h_n}\right)^2(L^{(2)}(0,\theta^0))-L^{(2)}(\lambda_i(X_i-x),\hat{\theta}(x))\right).$$

But using the definition of $J_{2n}(x)$ given in part a) as well as its order in probability we have that the last term is $\sqrt{nh_n}J_{2n}(x) = \sqrt{nh_n}(h_n^2 o_p(1)) = \sqrt{nh_n^5}o_p(1) = o_p(1)$ provided $nh_n^5 = O(1)$. Now,

$$E\left(\frac{1}{h_n^2}\frac{h_n}{n}\frac{1}{2g_X(x)}\sum_{i=1}^n K\left(\frac{X_i-x}{h_n}\right)\left(\frac{X_i-x}{h_n}\right)^2[\sigma^{2(2)}(\lambda_i'(X_i-x)+x)-L^{(2)}(0,\theta^0)]\right)$$
$$= \frac{1}{h_n^2}\left(\frac{h_n}{n}\frac{1}{2g_X(x)}n\int K(\phi)\phi^2[\sigma^{2(2)}(\lambda_i'(X_i-x)+x)-L^{(2)}(0,\theta^0)]g_X(x+h_n\phi)h_n d\phi\right)$$
$$= \frac{1}{2g_X(x)}\left(\int K(\phi)\phi^2 h^{(2)}(x+\lambda_i' h_n\phi)g(x+h_n\phi)h_n d\phi\right.$$
$$\left.-L^{(2)}(0,\theta^0)\int K(\phi)\phi^2 g_X(x+h_n\phi)d\phi\right) \to \frac{1}{2}h^{(2)}(x)\sigma_K^2 - \frac{1}{2}L^{(2)}(0,\theta^0)\sigma_K^2.$$

as $n \to \infty$ by Lebesgue's dominated convergence Theorem. Also,

$$
\begin{aligned}
&V\Bigg(\frac{1}{h_n^2}\frac{h_n}{n}\frac{1}{2g_X(x)}\sum_{i=1}^n K\left(\frac{X_i-x}{h_n}\right)\left(\frac{X_i-x}{h_n}\right)^2[\sigma^{2(2)}(\lambda_i^{'}(X_i-x)+x)-L^{(2)}(0,\theta^0)]\Bigg)\\
&=\frac{1}{4g_X^2(x)}\frac{1}{n^2h_n^2}nVK\left(\frac{X_i-x}{h_n}\right)\left(\frac{X_i-x}{h_n}\right)^2[\sigma^{2(2)}(\lambda_i^{'}(X_i-x)+x)-L^{(2)}(0,\theta^0)]\\
&=\frac{1}{4g_X^2(x)}\frac{1}{nh_n^2}\Bigg\{E\Bigg(K^2\left(\frac{X_i-x}{h_n}\right)\left(\frac{X_i-x}{h_n}\right)^4[\sigma^{2(2)}(\lambda_i^{'}(X_i-x)+x)-L^{(2)}(0,\theta^0)]^2\Bigg)\\
&\quad-\Bigg(E\Bigg(K\left(\frac{X_i-x}{h_n}\right)\left(\frac{X_i-x}{h_n}\right)^2[\sigma^{2(2)}(\lambda_i^{'}(X_i-x)+x)-L^{(2)}(0,\theta^0)]\Bigg)\Bigg)^2\Bigg\}\\
&=\frac{1}{h_n^2}\Bigg(\frac{h_n}{n}\frac{1}{2g_X(x)}n\int K(\phi)\phi^2[\sigma^{2(2)}(\lambda_i^{'}(X_i-x)+x)-L^{(2)}(0,\theta^0)]g_X(x+h_n\phi)h_nd\phi\Bigg)\\
&=\frac{1}{4g_X^2(x)}\Bigg\{\frac{1}{nh_n^2}\int K^2(\phi)\phi^4[h^{(2)}(x+\lambda_i^{'}h_n\phi)-L^{(2)}(0,\theta^0)]^2g(x+h_n\phi)h_nd\phi\\
&\quad-\frac{1}{n}\Bigg(\frac{1}{h_n}\int K(\phi)\phi[h^{(2)}(x+\lambda_i^{'}h_n\phi)-L^{(2)}(0,\theta^0)]g_X(x+h_n\phi)h_nd\phi\Bigg)^2\Bigg\}.
\end{aligned}
$$

Observe that

$$
\frac{-1}{4g_X^2(x)}\frac{1}{n}\Bigg(\int K(\phi)\phi[h^{(2)}(x+\lambda_i^{'}h_n\phi)-L^{(2)}(0,\theta^0)]g_X(x+h_n\phi)d\phi\Bigg)^2\to 0
$$

as $n \to \infty$ and

$$
\frac{1}{4g_X^2(x)}\frac{1}{nh_n^2}\int K^2(\phi)\phi^4[h^{(2)}(x+\lambda_i^{'}h_n\phi)-L^{(2)}(0,\theta^0)]^2g(x+h_n\phi)d\phi\to 0
$$

provided that $nh_n \to \infty$ and given that $\int K^2(\phi)\phi^4 d\phi < C$, an arbitrary bound. In all,

$$
\begin{aligned}
&\frac{h_n}{n}\frac{1}{2g_X(x)}\sum_{i=1}^n K\left(\frac{X_i-x}{h_n}\right)\left(\frac{X_i-x}{h_n}\right)^2[\sigma^{2(2)}(\lambda_i^{'}(X_i-x)+x)-L^{(2)}(0,\theta^0)]\\
&\qquad=\frac{1}{2}h_n^2(h^{(2)}(x)-L^{(2)}(0,\theta^0))\sigma_K^2+o_p(h_n^2),
\end{aligned}
$$

and

$$
\frac{h_n}{n}\frac{1}{2g_X(x)}\sum_{i=1}^n K\left(\frac{X_i-x}{h_n}\right)\left(\frac{X_i-x}{h_n}\right)^2[L^{(2)}(0,\theta^0)-L^{(2)}(\lambda_i(X_i-x),\hat{\theta}(x))]=o_p(h_n^2).
$$

Hence, we conclude that

$$
\sqrt{nh_n}\Bigg(\gamma_1^*(x)-\sigma^2(x)-\frac{1}{2}h_n^2\sigma_K^2(\sigma^{2(2)}(x)-L^{(2)}(0,\theta^0))+o_p(h_n^2)\Bigg)\xrightarrow{d}N\Bigg(0,\frac{\sigma^4(x)}{g_X(x)}(\mu_4(x)-1)\int K^2(y)dy\Bigg).
$$

which completes the proof.

*Proof of Theorem 2.* Suppose $\hat{\theta}_1(x)$, $\hat{\theta}_2(x)$ minimizes $\frac{1}{nh_n}\sum_{i=1}^n(e_i-L(X_i-x,\theta))^2K\left(\frac{X_i-x}{h_n}\right)$. Under the assumption that $\hat{\theta}(x)\equiv\begin{bmatrix}\hat{\theta}_1(x)\\ \hat{\theta}_2(x)\end{bmatrix}$ is in the interior of some compact subset $\bar{\Theta}$ of $\Re^2$, we have that

$$
\begin{bmatrix}\frac{1}{nh_n}\sum_{i=1}^n(e_i-L(X_i-x,\hat{\theta}(x)))L(X_i-x,\hat{\theta}(x))K\left(\frac{X_i-x}{h_n}\right)\\ \frac{1}{nh_n}\sum_{i=1}^n(e_i-L(X_i-x,\hat{\theta}(x)))L(X_i-x,\hat{\theta}(x))(X_i-x)K\left(\frac{X_i-x}{h_n}\right)\end{bmatrix}=0.
$$

We focus on the first element of this vector, since given assumption A3, all orders obtained for the first element will hold for the second element of the vector. Given that $e_i = (Y_i - \hat{m}(X_i;h_n))^2$ we can write

$$\frac{1}{nh_n}\sum_{i=1}^{n}(\sigma^2(X_i)\epsilon_i^2 - \sigma^2(X_i) + \sigma^2(X_i) - L(X_i - x, \hat{\theta}(x)))L(X_i - x, \hat{\theta}(x))K\left(\frac{X_i - x}{h_n}\right)$$
$$+ \frac{1}{nh_n}\sum_{i=1}^{n}(m(X_i) - \hat{m}(X_i;h_n))^2 L(X_i - x, \hat{\theta}(x))K\left(\frac{X_i - x}{h_n}\right)$$
$$- \frac{2}{nh_n}\sum_{i=1}^{n}(\sigma^2(X_i))^{1/2}\epsilon_i L(X_i - x, \hat{\theta}(x))K\left(\frac{X_i - x}{h_n}\right)(\hat{m}(X_i;h_n) - m(X_i)) = 0.$$

or $D_n(x, \hat{\theta}) \equiv I_{1n}^*(x) + I_{2n}^*(x) - I_{3n}^*(x) + I_{4n}^*(x) = 0$ where

$$I_{1n}^*(x) = \frac{1}{nh_n}\sum_{i=1}^{n}(\sigma^2(X_i) - L(X_i - x, \hat{\theta}(x)))L(X_i - x, \hat{\theta}(x))K\left(\frac{X_i - x}{h_n}\right)$$
$$I_{2n}^*(x) = \frac{1}{nh_n}\sum_{i=1}^{n}\sigma^2(X_i)(\epsilon_i^2 - 1)L(X_i - x, \hat{\theta}(x))K\left(\frac{X_i - x}{h_n}\right)$$
$$I_{3n}^*(x) = \frac{2}{nh_n}\sum_{i=1}^{n}(\sigma^2(X_i))^{1/2}\epsilon_i L(X_i - x, \hat{\theta}(x))K\left(\frac{X_i - x}{h_n}\right)(\hat{m}(X_i) - m(X_i))$$
$$I_{4n}^*(x) = \frac{1}{nh_n}\sum_{i=1}^{n}(m(X_i) - \hat{m}(X_i;h_n))^2 L(X_i - x, \hat{\theta}(x))K\left(\frac{X_i - x}{h_n}\right)..$$

Let $D_n(x,\theta)$ be the left-hand side of the equality with $\hat{\theta}(x)$ substituted by $\theta$. We will show that $\sup_{x\in G}\sup_{\theta\in\bar{\Theta}}\left|D_n(x,\theta) - g_X(x)(\sigma^2(x) - L(0,\theta))L(0,\theta)\right| = o_p(1)$, and since $g_X(x)(\sigma^2(x) - L(0,\theta^0))L(0,\theta^0) = 0$, it follows directly from the arguments in Hall et al. (1999) that $\hat{\theta}(x)$ is a uniformly consistent estimator for $\theta^0$. We start by considering $I_{2n}^*(x)$.

$$I_{2n}^*(x) = \frac{1}{nh_n}\sum_{i=1}^{n}\sigma^2(X_i)(\epsilon_i^2 - 1)[L(0,\theta) + L^{(1)}(0,\theta)(X_i - x)\exp(\theta_2\lambda_i(X_i - x))]K\left(\frac{X_i - x}{h_n}\right)$$
$$= L(0,\theta)\frac{1}{nh_n}\sum_{i=1}^{n}\sigma^2(X_i)(\epsilon_i^2 - 1)K\left(\frac{X_i - x}{h_n}\right) + L^{(1)}(0,\theta)\frac{h_n}{nh_n}\sum_{i=1}^{n}\sigma^2(X_i)$$
$$\times\left(\frac{X_i - x}{h_n}\right)\exp(\theta_2\lambda_i(X_i - x))K\left(\frac{X_i - x}{h_n}\right)$$
$$= I_{21,n}^*(x) + I_{22,n}^*(x)$$

Now,

$$\sup_{x\in G}|I_{21,n}^*(x)| \le |L(0,\theta)|\bar{B}_{gX}\sup_{x\in G}\left|\frac{h_n}{nh_n}\sum_{i=1}^{n}\sigma^2(X_i)(\epsilon_i^2 - 1)K\left(\frac{X_i - x}{h_n}\right)\right|$$
$$= |L(0,\theta)|\bar{B}_{gX}O_p\left(\left(\frac{\ln(n)}{nh_n}\right)^{1/2}\right),$$

and $\sup_{\theta\in\Theta}\sup_{x\in G}|I_{21,n}^*(x)| \le \sup_{\theta\in\Theta}|L(0,\theta)|\bar{B}_{gX}O_p\left(\left(\frac{\ln(n)}{nh_n}\right)^{1/2}\right)$. For $I_{22,n}^*(x)$ we have

$$|I_{22,n}^*(x)| \le h_n|L^{(1)}(0,\theta)|\bar{B}_{gX}\frac{1}{nh_ng_X(x)}\sum_{i=1}^{n}\sigma^2(X_i)|\epsilon_i^2 - 1|\left|\frac{X_i - x}{h_n}\right|\exp(\theta_2\lambda_i(X_i - x))K\left(\frac{X_i - x}{h_n}\right)$$
$$\le h_n|L^{(1)}(0,\theta)|\bar{B}_{gX}\frac{1}{nh_ng_X(x)}\sum_{i=1}^{n}\sigma^2(X_i)|\epsilon_i^2 - 1|\left|\frac{X_i - x}{h_n}\right|\exp(|\theta_2|h_n)K\left(\frac{X_i - x}{h_n}\right)$$

and therefore

$$\sup_{x\in G}|I^*_{22,n}(x)| \leq h_n|L^{(1)}(0,\theta)|\bar{B}_{gX}e^{|\theta_2|h_n}\sup_{x\in G}\frac{1}{nh_ng_X(x)}\sum_{i=1}^{n}\sigma^2(X_i)|\epsilon_i^2-1|\left|\frac{X_i-x}{h_n}\right|\exp(|\theta_2|h_n)K\left(\frac{X_i-x}{h_n}\right).$$

Now, note that $\sigma^2(X_i)|\epsilon_i^2-1| = |(R_i-\mu_R)^2-\sigma_R^2|\sigma_R^{-2}\sigma^2(X_i) \leq C$ given that $R_i \in [0,1]$, and

$$\begin{aligned}E\left(\frac{1}{h_n}K\left(\frac{X_i-x}{h_n}\right)h(X_i)|\epsilon_i^2-1|\right) &= \frac{1}{h_n}\int K\left(\frac{X_i-x}{h_n}\right)\sigma^2(X_i)\frac{|(R_i-\mu_R)^2-\sigma_R^2|}{\sigma_R^2}g_X(X_i)\\ &\quad g_R(R_i|X_i)dX_idR_i\\ &= \frac{1}{\sigma_R^2h_n}\int K\left(\frac{X_i-x}{h_n}\right)\sigma^2(X_i)\int|(R_i-\mu_R)^2-\sigma_R^2|\\ &\quad g_R(R_i|X_i)dR_ig_X(X_i)dX_i.\end{aligned}$$

Now, since $\int|(R_i-\mu_R)^2-\sigma_R^2|g_R(R_i|X_i)dR_i \leq \int|(R_i-\mu_R)^2|g_R(R_i|X_i)dR_i+\sigma_R^2 \leq 2\sigma_R^2$, we will denote this integral by $\eta(X_i)$. So,

$$E\left(\frac{1}{h_n}K\left(\frac{X_i-x}{h_n}\right)\sigma^2(X_i)|\epsilon_i^2-1|\right)=\frac{1}{h_n\sigma_R^2}\int K(\phi)\sigma^2(x+h_n\phi)\eta(x+h_n\phi)g_X(x+h_n\phi)h_nd\phi$$

with

$$\sup_{x\in G}E\left(\frac{1}{h_n}K\left(\frac{X_i-x}{h_n}\right)\sigma^2(X_i)|\epsilon_i^2-1|\right)=\frac{1}{\sigma_R^2}\int K(\phi)d\phi\sup_{x\in G}\sigma^2(x)\sup_{x\in G}\eta(x)\sup_{x\in G}g_X(x)\leq C.$$

where $C$ is an arbitrary constant. By Lemma 1 - part (a) in Martins-Filho and Yao (2007)

$$\sup_{x\in G}\frac{1}{nh_ng_X(x)}\sum_{i=1}^{n}\sigma^2(X_i)|\epsilon_i^2-1|K\left(\frac{X_i-x}{h_n}\right)=O_p\left(\left(\frac{\ln(n)}{nh_n}\right)^{1/2}\right)+O(1),$$

and consequently

$$\sup_{x\in G}|I^*_{22,n}(x)| \leq h_n|L^{(1)}(0,\theta)|\bar{B}_{gX}\exp(|\theta_2|h_n)\left[O_p\left(\left(\frac{\ln(n)}{nh_n}\right)^{1/2}\right)+O(1)\right]$$
$$\sup_{\theta\in\Theta}\sup_{x\in G}|I^*_{22,n}(x)| \leq \sup_{\theta\in\Theta}h_n|L^{(1)}(0,\theta)|\bar{B}_{gX}\exp(|\theta_2|h_n)\left[O_p\left(\left(\frac{\ln(n)}{nh_n}\right)^{1/2}\right)+O(1)\right].$$

We now turn our attention to $I^*_{3n}(x)$, which can be written as

$$\begin{aligned}I^*_{3n}(x) &= \frac{2}{nh_n}\sum_{i=1}^{n}(\sigma^2(X_i))^{1/2}\epsilon_iK\left(\frac{X_i-x}{h_n}\right)(m(X_i)-\hat{m}(X_i;h_n))\Big[L(0,\theta)+L^{(1)}(\lambda_i(X_i-x),\theta)\Big]\\ &= \frac{2}{nh_n}\sum_{i=1}^{n}(\sigma^2(X_i))^{1/2}\epsilon_iK\left(\frac{X_i-x}{h_n}\right)(m(X_i)-\hat{m}(X_i;h_n)\Big[L(0,\theta)\\ &\quad +L^{(1)}(0,\theta)(X_i-x)\exp(\theta_2\lambda_i(X_i-x))\Big]\\ &= L(0,\theta)\frac{2}{nh_n}\sum_{i=1}^{n}(m(X_i)-\hat{m}(X_i;h_n))(\sigma^2(X_i))^{1/2}\epsilon_iK\left(\frac{X_i-x}{h_n}\right)\\ &\quad +L^{(1)}(0,\theta)\frac{2}{nh_n}\sum_{i=1}^{n}(m(X_i)-\hat{m}(X_i;h_n))(\sigma^2(X_i))^{1/2}\epsilon_iK\left(\frac{X_i-x}{h_n}\right)\\ &\quad\times(X_i-x)\exp(\theta_2\lambda_i)(X_i-x))\\ &= I^*_{31,n}(x)+I^*_{32,n}(x).\end{aligned}$$

We write $I^*_{31,n}(x) = L(0,\theta)g_X(x)\frac{2}{nh_n g_X(x)}\sum_{i=1}^n (m(X_i) - \hat{m}(X_i;h_n))(\sigma^2(X_i))^{1/2}\epsilon_i K\left(\frac{X_i - x}{h_n}\right)$. From Martins-Filho and Yao (2007) we have that

$$\sup_{x\in G}\left|L(0,\theta)\frac{2}{nh_n}\sum_{i=1}^n (m(X_i) - \hat{m}(X_i;h_n))(\sigma^2(X_i))^{1/2}\epsilon_i K\left(\frac{X_i - x}{h_n}\right)\right| = O_p(h_n^2) + O_p\left(\left(\frac{\ln(n)}{nh_n}\right)^{1/2}\right).$$

Hence,

$$sup_{x\in G}|I^*_{31,n}(x)| \leq L(0,\theta)\bar{B}_{g_X}\left[O_p(h_n^2) + O_p\left(\left(\frac{\ln(n)}{nh_n}\right)^{1/2}\right)\right],$$

and

$$\sup_{\theta\in\bar{\Theta}}\sup_{x\in G}|I^*_{31,n}(x)| \leq C\bar{B}_{g_X}\left[O_p(h_n^2) + O_p\left(\left(\frac{\ln(n)}{nh_n}\right)^{1/2}\right)\right],$$

since $\sup_{\theta\in\bar{\Theta}} L(0,\theta) = C$, given that $\bar{\Theta}$ is compact.

$I^*_{32,n}(x)$ can be written as,

$$\begin{aligned} I^*_{32,n}(x) &= 2L^{(1)}(0,\theta)g_X(x)\frac{h_n}{nh_n g_X(x)}\sum_{i=1}^n (m(X_i) - \hat{m}(X_i;h_n))(\sigma^2(X_i))^{1/2}\epsilon_i K\left(\frac{X_i - x}{h_n}\right) \\ &\quad \times \left(\frac{X_i - x}{h_n}\right)\exp(\theta_2\lambda_i(X_i - x)) \\ &= 2L^{(1)}(0,\theta)g_X(x)h_n I^*_{321,n}(x) \end{aligned}$$

$$|I^*_{32,n}(x)| = 2|L^{(1)}(0,\theta)|g_X(x)h_n|I^*_{321,n}(x)|.$$

$$\begin{aligned} |I^*_{321,n}(x)| &\leq \frac{1}{nh_n g_X(x)}\sum_{i=1}^n |m(X_i) - \hat{m}(X_i,h_n)|(\sigma^2(X_i))^{1/2}|\epsilon_i|K\left(\frac{X_i - x}{h_n}\right) \\ &\quad \times \left|\frac{X_i - x}{h_n}\right|\exp(\theta_2\lambda_i(X_i - x)). \end{aligned}$$

Since, we have that if $\left|\frac{X_i - x}{h_n}\right| > 1$ then $K\left(\frac{X_i - x}{h_n}\right) = 0$ we can write that

$$|I^*_{321,n}(x)| \leq \frac{1}{nh_n g_X(x)}\sum_{i=1}^n |m(X_i) - \hat{m}(X_i;h_n)|(\sigma^2(X_i))^{1/2}|\epsilon_i|K\left(\frac{X_i - x}{h_n}\right)\left|\frac{X_i - x}{h_n}\right|e^{|\theta_2|h_n}$$

since $e^{\theta_2\lambda_i(X_i - x)} \leq e^{|\theta_2|h_n}$ given that $0 \leq \lambda_i \leq 1$. Now, note that from Fan and Yao (1998) and arguments similar to those used to establish Lemma 1, we have that for a bandwidth $h_1$ used in the first stage estimation, we have

$$\hat{m}(X_i;h_1) - m(X_i) = \frac{1}{nh_1 g_X(X_i)}\sum_{t=1}^n K\left(\frac{X_t - X_i}{h_1}\right)\left[Y_t - m(X_i) - m^{(1)}(X_i)(X_t - X_i)\right] + O_p(R_{n,2}(X_i))$$

where

$$\begin{aligned} R_{n,2}(X_i) &= \frac{1}{n}\left|\sum_{t=1}^n K\left(\frac{X_t - X_i}{h_1}\right)Y_t^*\right| \\ &+ \frac{1}{n}\left|\sum_{t=1}^n K\left(\frac{X_t - X_i}{h_1}\right)\left(\frac{X_t - X_i}{h_1}\right)Y_t^*\right| \end{aligned}$$

and $Y_t^* = Y_t - m(x) - m^{(1)}(x)(X_t - x) = \frac{1}{2}m^{(2)}(X_{ti})(X_t - X_i)^2 + (\sigma^2(X_t))^{1/2}\epsilon_t$, for $X_{ti} = \phi X_t + (1-\phi)X_i$ for some $\phi \in [0,1]$. Thus, we can write

$$\begin{aligned}\hat{m}(X_i; h_1) - m(X_i) = &\frac{h_1^2}{nh_1 g_X(X_i)} \sum_{t=1}^{n} K\left(\frac{X_t - X_i}{h_1}\right)\left(\frac{X_t - X_i}{h_1}\right)^2 \frac{1}{2} m^{(2)}(X_{ti}) \\ &+ \frac{1}{nh_1 g_X(X_i)} \sum_{t=1}^{n} K\left(\frac{X_t - X_i}{h_1}\right)(h(X_t))^{1/2}\epsilon_t + O_p(R_{n,2}(X_i))\end{aligned}$$

and

$$\begin{aligned}|\hat{m}(X_i, h_1) - m(X_i)| \leq &\frac{h_1^2}{nh_1 g_X(X_i)} \sum_{t=1}^{n} K\left(\frac{X_t - X_i}{h_1}\right)\left(\frac{X_t - X_i}{h_1}\right)^2 \frac{1}{2} |m^{(2)}(X_{ti})| \\ &+ \frac{1}{nh_1 g_X(X_i)} \sum_{t=1}^{n} K\left(\frac{X_t - X_i}{h_1}\right)(h(X_t))^{1/2}|\epsilon_t| + O_p(R_{n,2}(X_i)).\end{aligned}$$

Consequently, we have

$$\begin{aligned}|I^*_{321,n}(x)| \leq e^{|\theta_2|h_n} &\frac{1}{nh_n g_X(x)} \sum_{i=1}^{n} \left[\frac{h_1^2}{nh_1 g_X(X_i)} \sum_{t=1}^{n} K\left(\frac{X_t - X_i}{h_1}\right)\left(\frac{X_t - X_i}{h_1}\right)^2 \frac{1}{2} |m^{(2)}(X_{ti})|\right. \\ &\left. + \frac{1}{nh_1 g_X(X_i)} \sum_{t=1}^{n} K\left(\frac{X_t - X_i}{h_1}\right)(h(X_t))^{1/2}|\epsilon_t| + O_p(R_{n,2}(X_i))\right] \\ &\times (h(X_t))^{1/2}|\epsilon_t| K\left(\frac{X_i - x}{h_n}\right) \\ = e^{|\theta|h_n} &\frac{1}{nh_n g_X(x)} \frac{h_1^2}{nh_1} \sum_{i=1}^{n}\sum_{t=1}^{n} \frac{1}{g_X(X_i)} K\left(\frac{X_t - X_i}{h_1}\right)\left(\frac{X_t - X_i}{h_1}\right)^2 K\left(\frac{X_i - x}{h_n}\right) \\ &\times (\sigma^2(X_t))^{1/2}|\epsilon_t| \frac{1}{2}|m^{(2)}(X_{ti})| \\ &+ e^{|\theta|h_n} \frac{1}{nh_n g_X(x)} \frac{1}{nh_1} \sum_{i=1}^{n}\sum_{t=1}^{n} \frac{1}{g_X(X_i)} K\left(\frac{X_t - X_i}{h_1}\right)(\sigma^2(X_t))^{1/2}|\epsilon_t| \\ &\times (h(X_i))^{1/2}|\epsilon_i| K\left(\frac{X_i - x}{h_n}\right) \\ &+ e^{|\theta|h_n} \frac{1}{nh_n g_X(x)} \sum_{i=1}^{n} (\sigma^2(X_i))^{1/2}|\epsilon_i| K\left(\frac{X_i - x}{h_n}\right) O_p(R_{n,2}(X_i)) \\ = L_{n,1}&(x) + L_{n,2}(x) + L_{n,3}(x).\end{aligned}$$

We investigate each of these terms separately. First, we write

$$\begin{aligned}L_{n,1}(x) &\leq e^{|\theta_2|h_n} \underline{\mathrm{B}}_{g_X}^{-1} \frac{1}{nh_n} \frac{h_1^2}{nh_1} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right) \frac{\sigma^2(X_i)^{1/2}}{g_X(X_i)} |\epsilon_i| \frac{1}{nh_1} K\left(\frac{X_t - X_i}{h_1}\right)\left(\frac{X_t - X_i}{h_1}\right)^2 \\ &\leq e^{|\theta_2|h_n} \underline{\mathrm{B}}_{g_X}^{-1} \frac{h_1^2}{nh_n} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right) \frac{\sigma^2(X_i)^{1/2}}{g_X(X_i)} |\epsilon_i| \frac{1}{h_1} \sup_{x \in G} \frac{1}{n} \sum_{t=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)^2.\end{aligned}$$

By Martins-Filho and Yao (2007, p. 307) we have that

$$\frac{1}{n}\sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right) \frac{\sigma^2(X_i)^{1/2}}{g_X(X_i)} |\epsilon_i| = O_p(h_n)$$

and

$$\sup_{x \in G} \frac{1}{n} \sum_{t=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)^2 = O_p(h_1).$$

Hence, $\sup_{\theta\in\bar{\Theta}}\sup_{x\in G} L_{n,1}(x) \leq \sup_{\theta\in\bar{\Theta}} e^{|\theta_2|h_n}\underline{\mathrm{B}}_{g_X}^{-1}h_1^2 O_p(1)$. Second, we write

$$\begin{aligned}
L_{n,2}(x) &= e^{|\theta|h_n}\frac{1}{nh_n g_X(x)}\sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\frac{\sigma^2(X_i)^{1/2}}{g_X(X_i)}|\epsilon_i| \\
&\quad \times \sum_{t=1}^{n}\frac{1}{nh_1}K\left(\frac{X_t - X_i}{h_1}\right)\sigma^2(X_t)^{1/2}|\epsilon_t| \\
&\leq e^{|\theta_2|h_n}\underline{\mathrm{B}}_{g_X}^{-1}\frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\frac{\sigma^2(X_i)^{1/2}}{g_X(X_i)}|\epsilon_i| \\
&\quad \times \sup_{x\in G}\frac{1}{nh_1}\sum_{t=1}^{n}K\left(\frac{X_t - x}{h_1}\right)\sigma^2(X_t)^{1/2}|\epsilon_t|.
\end{aligned}$$

Now, observe that

$$\begin{aligned}
E\left(\frac{1}{h_1}K\left(\frac{X_t - x}{h_1}\right)(\sigma^2(X_t))^{1/2}|\epsilon_t|\right) &= \frac{1}{h_1}\int K\left(\frac{X_t - x}{h_1}\right)(\sigma^2(X_t))^{1/2}\frac{1}{\sigma_R}|R_t - \mu_R| g_X(X_t) \\
&\quad \times g_{R|X}(R_t)dX_t dR_t \\
&= \frac{1}{\sigma_R}\frac{1}{h_1}\int K\left(\frac{X_t - x}{h_1}\right)(\sigma^2(X_t))^{1/2}\int |R_t - \mu_R| \\
&\quad \times g_{R|X}(R_t)dR_t g_X(X_t)dX_t \\
&= \frac{1}{\sigma_R}\frac{1}{h_1}\int K\left(\frac{X_t - x}{h_1}\right)(\sigma^2(X_t))^{1/2}\mu_1(X_t)g_X(X_t)dX_t \\
&= \frac{1}{\sigma_R}\frac{1}{h_1}\int K(\phi)(\sigma^2(x + h_1\phi))^{1/2}\mu_1(x + h_1\phi)h_1 g_X(x + h_1\phi)d\phi.
\end{aligned}$$

Hence,

$$\sup_{x\in G} E\left(\frac{1}{h_1}K\left(\frac{X_t - x}{h_1}\right)(\sigma^2(X_t))^{1/2}|\epsilon_t|\right) \leq \frac{1}{\sigma_R}\int K(\phi)d\phi \sup_{x\in G}\sigma^2(x))^{1/2}\sup_{x\in G} g_X(x)\sup_{x\in G}\mu_1(x) \leq C,$$

since $(\sigma^2(X_t))^{1/2}|\epsilon_t| = \sigma^2(X_t)^{1/2}\frac{1}{\sigma_R}|R_t - \mu_R| < C$, by Lemma 1 in Martins-Filho and Yao (2007), part (a), if $nh_n^3 \to \infty$. Therefore, we can write

$$\begin{aligned}
\sup_{x\in G}\frac{1}{nh_1}\sum_{t=1}^{n}K\left(\frac{X_t - x}{h_1}\right)(\sigma^2(X_t))^{1/2}|\epsilon_t| &\leq \sup_{x\in G}\left|\frac{1}{nh_1}\sum_{t=1}^{n}K\left(\frac{X_t - x}{h_1}\right)(\sigma^2(X_t))^{1/2}|\epsilon_t|\right. \\
&\quad \left. - E\left(\frac{1}{h_1}K\left(\frac{X_t - x}{h_1}\right)(\sigma^2(X_t))^{1/2}|\epsilon_t|\right)\right| \\
&\quad + \sup_{x\in G}\frac{1}{h_1}E\left(K\left(\frac{X_t - x}{h_1}\right)(\sigma^2(X_t))^{1/2}|\epsilon_t|\right) \\
&= O_p\left(\left(\frac{\ln(n)}{nh_1}\right)^{1/2}\right) + O(1).
\end{aligned}$$

Hence, $\sup_{\theta\in\bar{\Theta}}\sup_{x\in G} L_{n,2}(x) \leq \sup_{\theta\in\Theta} e^{|\theta_2|h_n}\underline{\mathrm{B}}_{g_X}^{-1}O_p(1)$. Lastly, by Martins-Filho and Yao (2007, p. 308) $\sup_{\theta\in\Theta}\sup_{x\in G} L_{n,3}(x) \leq \sup_{\theta\in\bar{\Theta}} e^{|\theta_2|h_n}\underline{\mathrm{B}}_{g_X}^{-1}h_n^2 o_p(1)$. Combining the results on $L_{n,1}(x)$, $L_{n,2}(x)$ and

$L_{n,3}(x)$ we have, together with compactness of $\bar{\Theta}$, that $\sup_{\theta\in\bar{\Theta}}\sup_{x\in G}|I^*_{32,n}(x)| = o_p(1)$. Now, consider

$$\begin{aligned}
I^*_{4n}(x) &= \frac{1}{nh_n}\sum_{i=1}^{n}(m(X_i)-\hat{m}(X_i;h_1))^2\left[L(0,\theta)+L^{(1)}(0,\theta)(X_i-x)e^{\theta_2\lambda_i(X_i-x)}\right]K\left(\frac{X_i-x}{h_n}\right)\\
&= L(0,\theta)\frac{1}{nh_n}\sum_{i=1}^{n}(m(X_i)-\hat{m}(X_i;h_1))^2K\left(\frac{X_i-x}{h_n}\right)\\
&\quad + L^{(1)}(0,\theta)\frac{1}{nh_n}\sum_{i=1}^{n}(m(X_i)-\hat{m}(X_i;h_1))^2(X_i-x)e^{\theta_2\lambda_i(X_i-x)}K\left(\frac{X_i-x}{h_n}\right)\\
&= I^*_{41,n}(x)+I^*_{42,n}(x).
\end{aligned}$$

We consider each of these terms separately.

$$I^*_{41,n}(x) = L(0,\theta)g_X(x)\frac{1}{nh_ng_X(x)}\sum_{i=1}^{n}(m(X_i)-\hat{m}(X_i;h_n))^2K\left(\frac{X_i-x}{h_n}\right) \text{ which gives}$$

$$\sup_{x\in G}|I^*_{41,n}(x)| \leq |L(0,\theta)|\bar{B}_{g_X}\sup_{x\in G}\left|\frac{1}{nh_ng_X(x)}\sum_{i=1}^{n}(m(X_i)-\hat{m}(X_i;h_n))^2K\left(\frac{X_i-x}{h_n}\right)\right|$$

By Martins-Filho and Yao (2007, p. 310)

$$\sup_{x\in G}\left|\frac{1}{nh_ng_X(x)}\sum_{i=1}^{n}(m(X_i)-\hat{m}(X_i;h_n))^2K\left(\frac{X_i-x}{h_n}\right)\right| = o_p(h_n^2) \text{ and consequently .}$$

$\sup_{\theta\in\bar{\Theta}}\sup_{x\in G}|I^*_{41,n}(x)| \leq \sup_{\theta\in\bar{\Theta}}|L(0,\theta)|\bar{B}_{g_X}h_n^2o_p(1)$. Now,

$$I^*_{42,n}(x) = L^{(1)}(0,\theta)h_n\frac{1}{nh_n}\sum_{i=1}^{n}(m(X_i)-\hat{m}(X_i;h_1))^2K\left(\frac{X_i-x}{h_n}\right)\left(\frac{X_i-x}{h_n}\right)e^{\theta_2\lambda_i(X_i-x)} \text{ and}$$

$$|I^*_{42,n}(x)| \leq |L^{(1)}(0,\theta)|h_n\frac{1}{nh_n}\sum_{i=1}^{n}(m(X_i)-\hat{m}(X_i;h_1))^2K\left(\frac{X_i-x}{h_n}\right)\left|\frac{X_i-x}{h_n}\right|e^{\theta_2\lambda_i(X_i-x)}$$

Again, by compactness of the support of $K$, we have

$$|I^*_{42,n}(x)| \leq |L^{(1)}(0,\theta)|h_n\frac{1}{nh_n}\sum_{i=1}^{n}(m(X_i)-\hat{m}(X_i;h_1))^2K\left(\frac{X_i-x}{h_n}\right)\left|\frac{X_i-x}{h_n}\right|e^{|\theta_2|h_n} \text{ and}$$

$$\sup_{x\in G}|I^*_{42,n}(x)| \leq h_n\bar{B}_{g_X}|L^{(1)}(0,\theta)|e^{|\theta_2|h_n}o_p(h_n^2).$$

Consequently, $\sup_{\theta\in\bar{\Theta}}\sup_{x\in G}|I^*_{42,n}(x)| \leq h_n^3\bar{B}_{g_X}\sup_{\theta\in\Theta}|L^{(1)}(0,\theta)|e^{|\theta_2|h_n}$. We now examine $I^*_{1n}(x)$. We start by noting that

$$\begin{aligned}
\sigma^2(X_i)-L(X_i-x,\hat{\theta}(x)) &= \sigma^2(X_i)-L(0,\theta)-L^{(1)}(0,\theta)(X_i-x)e^{\theta_2\lambda_i(X_i-x)}\\
&= \sigma^2(x)+\sigma^{2(1)}(\lambda_i'(X_i-x)+x)(X_i-x)-L(0,\theta)\\
&\quad - L^{(1)}(0,\theta)(X_i-x)e^{\theta_2\lambda_i(X_i-x)}.
\end{aligned}$$

Therefore,

$$
\begin{aligned}
I_{1n}^{*}(x) &= g_X(x)\frac{1}{nh_n g_X(x)}\sum_{i=1}^{n}\Big[\sigma^2(x) - L(0,\theta) + \sigma^{2(1)}(\lambda_i^{'}(X_i - x) + x)(X_i - x) \\
&\quad - L^{(1)}(0,\theta)(X_i - x)e^{\theta_2\lambda_i(X_i-x)}\Big]K\left(\frac{X_i - x}{h_n}\right)L(X_i - x,\theta) \\
&= g_X(x)\Big\{\frac{1}{nh_n g_X(x)}(\sigma^2(x) - L(0,\theta))\sum_{i=1}^{n}K\left(\frac{X_i - x}{h_n}\right)L(X_i - x,\theta) \\
&\quad + \frac{1}{nh_n g_X(x)}\sum_{i=1}^{n}K\left(\frac{X_i - x}{h_n}\right)\Big[\sigma^{2(1)}(\lambda_i^{'}(X_i - x) + x)(X_i - x) \\
&\quad - L^{(1)}(0,\theta)(X_i - x)e^{\theta_2\lambda_i(X_i-x)}\Big]L(X_i - x,\theta)\Big\} \\
&= g_X(x)(I_{11,n}^{*}(x) + I_{12,n}^{*}(x)).
\end{aligned}
$$

We now look at $I_{11,n}^{*}(x)$, and $I_{12,n}^{*}(x)$ in isolation.

$$
\begin{aligned}
I_{11,n}^{*}(x) &= [\sigma^2(x) - L(0,\theta)]\frac{1}{nh_n}\sum_{i=1}^{n}K\left(\frac{X_i - x}{h_n}\right)L(X_i - x,\theta) \\
&= [\sigma^2(x) - L(0,\theta)]\frac{1}{nh_n}\sum_{i=1}^{n}K\left(\frac{X_i - x}{h_n}\right)\Big(L(0,\theta) \\
&\quad - L^{(1)}(0,\theta)(X_i - x)e^{\theta_2\lambda_i(X_i-x)}\Big) \\
&= [\sigma^2(x) - L(0,\theta)]\Big\{L(0,\theta)\frac{1}{nh_n}\sum_{i=1}^{n}K\left(\frac{X_i - x}{h_n}\right) - h_n L^{(1)}(0,\theta) \\
&\quad \times \frac{1}{nh_n}\sum_{i=1}^{n}K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)e^{\theta_2\lambda_i(X_i-x)}\Big\}.
\end{aligned}
$$

From Lemma 1 in Martins-Filho and Yao (2007) we have that $\frac{1}{nh_n}\sum_{i=1}^{n}K\left(\frac{X_i-x}{h_n}\right)$ converges uniformly to $g_X(x)$ on a compact set $G$, hence the first term converges uniformly to $g_X(x)L(0,\theta)$ on $G$. By arguments made earlier in the proof we have that $\frac{1}{nh_n}\sum_{i=1}^{n}K\left(\frac{X_i-x}{h_n}\right)\left(\frac{X_i-x}{h_n}\right)\times e^{\theta_2\lambda_i(X_i-x)}$ is uniformly bounded in probability on $G$, hence the last term is $o_p(1)$ uniformly in $G$. Hence, we gave $I_{11,n}^{*}(x) \xrightarrow{p}$

$g_X(x)[h(x) - L(0,\theta)]L(0,\theta)$. Now we treat $I^*_{12,n}(x)$. Note that,

$$
\begin{aligned}
I^*_{12,n}(x) &= \frac{h_n}{nh_n}\sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)\Bigg(\sigma^{2(1)}(\lambda_i^{'}(X_i - x) + x) - L^{(1)}(0,\theta) \\
&\quad \times e^{\theta_2\lambda_i(X_i - x)}\Bigg)\Bigg(L(0,\theta) - h_n L^{(1)}(0,\theta)\left(\frac{X_i - x}{h_n}\right)e^{\theta_2\lambda_i(X_i-x)}\Bigg) \\
&= h_n\frac{1}{nh_n}L(0,\theta)\sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)\Bigg(\sigma^{2(1)}(\lambda_i^{'}(X_i - x) + x) \\
&\quad - L^{(1)}(0,\theta)e^{\theta_2\lambda_i(X_i-x)}\Bigg) - h_n^2\frac{1}{nh_n}L^{(1)}(0,\theta)\sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)^2 \\
&\quad \times e^{\theta_2\lambda_i(X_i-x)}\Bigg(\sigma^{2(1)}(\lambda_i^{'}(X_i - x) + x) - L^{(1)}(0,\theta)e^{\theta_2\lambda_i(X_i-x)}\Bigg) \\
&= h_n\frac{1}{nh_n}L(0,\theta)\sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)\Bigg(\sigma^{2(1)}(\lambda_i^{'}(X_i - x) + x) \\
&\quad - L^{(1)}(\lambda_i(X_i - x),\theta)\Bigg) - h_n^2\frac{1}{nh_n}L^{(1)}(0,\theta)\sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)^2 \\
&\quad \times e^{\theta_2\lambda_i(X_i-x)}\Bigg(\sigma^{2(1)}(\lambda_i^{'}(X_i - x) + x) - L^{(1)}(\lambda_i(X_i - x),\theta)\Bigg) \\
&= I^*_{121,n}(x) - I^*_{122,n}(x).
\end{aligned}
$$

Observe that

$$
\begin{aligned}
I^*_{121,n}(x) &= h_n\frac{1}{nh_n}L(0,\theta)\sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)[\sigma^{2(1)}(\lambda_i^{'}(X_i - x) + x) - L^{(1)}(0,\theta)] \\
&\quad + h_n\frac{1}{nh_n}L(0,\theta)\sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)[L^{(1)}(0,\theta) - L^{(1)}(\lambda_i(X_i - x),\theta)],
\end{aligned}
$$

where $L^{(1)}(0,\theta) = \sigma^2(x) < C$. Also, $|\sigma^{2(1)}(x)| \leq \sigma^2(x)|f^{(1)}(x)| < C$ provided $|f^{(1)}(x)| < B_f$. Hence,

$$
\begin{aligned}
|I^*_{121,n}(x)| &\leq h_n\underline{\mathrm{B}}_{g_X}^{-1}C\frac{1}{nh_n}\sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right)\left|\frac{X_i - x}{h_n}\right| \\
&\quad + h_n\underline{\mathrm{B}}_{g_X}^{-1}C\frac{1}{nh_n}\sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right)\left|\frac{X_i - x}{h_n}\right|[L^{(1)}(0,\theta) - L^{(1)}(\lambda_i(X_i - x),\theta)].
\end{aligned}
$$

Since, $K\left(\frac{X_i - x}{h_n}\right) = 0$ whenever $\left|\frac{X_i - x}{h_n}\right| < 1$,

$$
\begin{aligned}
|I^*_{121,n}(x)| &\leq h_n\underline{\mathrm{B}}_{g_X}^{-1}C\frac{1}{nh_n}\sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right) \\
&\quad + h_n\underline{\mathrm{B}}_{g_X}^{-1}C\frac{1}{nh_n}\sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right)|L^{(1)}(0,\theta) - L^{(1)}(\lambda_i(X_i - x),\theta)|.
\end{aligned}
$$

Now, $|L^{(1)}(0,\theta) - L^{(1)}(\lambda_i(X_i - x),\theta)| = \sigma^2(x)|1 - e^{\theta_2\lambda_i(X_i-x)}| \leq \sigma^2(x)(1 + e^{h_n|\theta_2|})$. Hence, $|I^*_{121,n}(x)| \leq h_n\underline{\mathrm{B}}_{g_X}^{-1}C\frac{1}{nh_n}\sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right) + h_n\sigma^2(x)(1 + e^{h_n|\theta_2|})\underline{\mathrm{B}}_{g_X}^{-1}C\frac{1}{nh_n}\sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right)$. Again, using Lemma 1 in Martins-Filho and Yao (2007),

$$
\begin{aligned}
\sup_{x\in G}|I^*_{121,n}(x)| &\leq h_n\underline{\mathrm{B}}_{g_X}^{-1}B_hC[g_X(x) + O_p(h_n)] \\
&\quad + h_nh(x)(1 + e^{h_n|\theta_2|})\underline{\mathrm{B}}_{g_X}^{-1}B_h[g_X(x) + O_p(h_n)],
\end{aligned}
$$

which gives $\sup_{\theta\in\bar{\Theta}}\sup_{x\in G}|I^*_{121,n}(x)| = o_p(1)$. Similar arguments show that $\sup_{\theta\in\Theta}\sup_{x\in G}|I^*_{122,n}(x)| = o_p(1)$, which completes the proof.

## 2.9 References


1. Aigner, D., C.A.K. Lovell and P. Schmidt, Formulation and estimation of stochastic frontiers production function models. **Journal of Econometrics**, 6, 21-37, 1977.

2. Aragon, Y., A. Daouia, C. Thomas-Agnan, Nonparametric frontier estimation: a conditional quantile-based approach. **Econometric Theory**, 21, 358-389, 2005.

3. Daouia, A., Gardes, L. and S. Girard, 2009, Large Sample Approximation of the Distribution for Smoothed Monotone Frontier Estimators . Working paper.

4. Fan, J., Design adaptive nonparametric regression. **Journal of the American Statistical Association**, 87, 998-1004, 1992.

5. Fan, J. and I. Gijbels, Data driven bandwidth selection in local polynomial fitting: variable bandwidth and spatial adaptation. **Journal of the Royal Statistical Society B**, 57, 371-394, 1995.

6. Fan, Y., Q. Li and A. Weersink, Semiparametric estimation of stochastic production frontier models. **Journal of Business and Economic Statistics**, 14, 460-468, 1996.

7. Fan, J., and Q. Yao, Efficient estimation of conditional variance functions in stochastic regression. **Biometrika**, 85, 645-660, 1998.

8. Farrell, M., The measurement of productive efficiency. **Journal of the Royal Statistical Society A**, 120, 253-290, 1957.

9. Gorman, M.F., Ruggiero, J., Evaluating US state police performance using data envelopment analysis. **Int. J. Production Economics**, 113, 1031-1037, 2008.

10. Hall, P., Wolff, R. and Q. Yao, Methods for estimating a conditional distribution function. **Journal of the American Statistical Association**, 94, 154-163, 1999.

11. Kumbhakar, S. C., B. U. Park, L. Simar and E. Tsionas, Nonparametric stochastic frontiers: a local maximum likelihood approach. **Journal of Econometrics**, 137, 1-27, 2007.

12. Martins-Filho, C. and F. Yao, Nonparametric frontier estimation via local linear regression. Journal of Econometrics, 141, 283-319, 2007.

13. Martins-Filho, C. and F. Yao, A Smooth Nonparametric Conditional Quantile Frontier Estimator. **Journal of Econometrics**, 143, 317-333, 2008.

14. Martins-Filho, C. and F. Yao, 2010, Nonparametric stochastic frontier estimation via profile likelihood. Working paper, University of Colorado, Boulder.

15. Ruppert D., S. Sheather, M. Wand, An effective bandwidth selector for least squares regression. **Journal of the American Statistical Association**, 90, 1257-1270, 1995.

16. Silverman, B.W., **Density estimation for statistics and data analysis**. Chapman and Hall, London, 1986.

17. Simar, L. and P. Wilson, Statistical inference in nonparametric frontier models: recent developments and perspectives, in: H. Fried, C.A.K. Lovell, and S.S. Schmidt, (Eds.), **The Measurement of Productive Efficiency**, 2nd edition. Oxford University Press, Oxford, 2007.

18. Stone, C. J., Consistent nonparametric regression. **Annals of Statistics** 5, 595-620, 1977.

19. Ziegelmann, F., Nonparametric estimation of volatility functions: the local exponential approach. **Econometric Theory**, 18, 985-991, 2002.

# 3 Nonparametric Frontier Estimation in Two Steps

Hudson Torrent[1] and Flavio Ziegelmann[2]

**Abstract.** In this paper we propose a novel method of estimation of production frontiers using nonparametric local linear Kernel regression in two stages, with some advantages over traditional frontier estimators as DEA and FDH. Whereas the first stage gives the shape of the frontier, the second stage is responsible for locating it. Our method may be viewed as a modification of those there proposed by Martins-Filho and Yao (2007) and Martins-Filho, Torrent and Ziegelmann (2010), who estimate the same frontier model in three stages under different versions for the second stage. In these studies the first two stages are responsible for estimating the frontier shape, while the third stage finds the location frontier. Here we show that we can eliminate the second stage of Martins-Filho and Yao, obtaining the frontier in only two stages. We study asymptotic properties showing consistency and $\sqrt{nh_n}$ asymptotic normality of our proposed estimator under standard assumptions. We also perform a simulation study comparing our estimator with two other: the estimator of Martins-Filho and Yao (2007); and an improved estimator proposed by Martins-Filho, Torrent and Ziegelmann (2010). In our numerical implementations our proposed estimator outperforms its competitors for mean efficiency above 50% in all sample sizes considered. Finally we present an empirical analysis to USA crime data.

**Keywords and phrases.** nonparametric frontier models; local linear smoothing; local linear regression.

**JEL Classifications.** C14, C22

[1]Department of Economics, Federal University of Rio Grande do Sul, Porto Alegre, RS 90040-000, Brazil, email: hudsontorrent@ppge.ufrgs.br

[2]Department of Statistics, Federal University of Rio Grande do Sul, Porto Alegre, RS 90040-000, Brazil, email: flavioz@ufrgs.br and Department of Statistics, London School of Economics, London, WC2A 2AE UK.

## 3.1 Introduction

Estimation of production frontiers and therefore efficiency (and inefficiency) of production processes have been the subject of a vast and growing literature since Farrell (1957). The problem can be stated as follows. Let $x \in \mathbb{R}_+^p$ be a set of inputs used to produce a set of outputs $y \in \mathbb{R}_+^q$. So, there is a technological or production set defined as $\Psi = \{(x, y) \in \mathbb{R}_+^{p+q} \mid x \text{ can produce } y\}$. A production frontier associated with $\Psi$ is defined as $\rho(x) = \sup\{y \in \mathbb{R}_+^q \mid (y, x) \in \Psi\}$ for all $x \in \mathbb{R}_+^p$. Thus, for given $(x_0, y_0) \in \Psi$, efficiency is measured by the distance between $y_0$ and $\rho(x_0)$. In this context we aim to estimate from a given random sample $\chi = \{(x_i, y_i), i = 1, ..., n\}$ the associated production frontier $\rho(\cdot)$ and efficiency measures for the observed production units.

To solve these problems we can find in the literature two traditional nonparametric estimation procedures[1], the Free Disposal Hull (FDH) estimator, introduced by Deprins et al. (1984) and the Data Envelopment Analysis (DEA) estimator, represented by Charnes et al. (1978). The idea is to estimate a production set from an observed random sample without assuming any restrictive parametric structure either on the production frontier $\rho(\cdot)$ or on the joint density of $(X_i, Y_i)$. There is a vast literature on these methodologies. Gijbels et al. (1999) and Park et al. (2000) obtained asymptotic distributions for DEA and FDH estimators, respectively. However, these estimators have some undesirable characteristics. Both estimators are based on the assumption that all observed data lie inside the technological set, so the estimators consist of obtaining the smallest set that envelops all data. Therefore the estimated set never exceeds the true frontier; hence the frontier estimators are inherently downwards biased. Moreover, FDH produces a discontinuous function that envelops the data and DEA produces a piecewise linear function.

Martins-Filho and Yao (2007) propose a deterministic production frontier model and a nonparametric production frontier estimator, called NP3S[2]. They derive the asymptotic normality and consistency of both production frontier and efficiency estimators under reasonable assumptions in the nonparametric context. Their estimator shares the flexible nonparametric structure but has some extra desirable properties if compared to FDH and DEA estimators: *i*) NP3S estimator is demonstrated to be more robust to extreme values; *ii*) the frontier estimator is a smooth function of input usage (not discontinuous neither piecewise linear) and *iii*) although the estimator envelops the data, it is not inherently biased as FDH

[1]We consider here only the deterministic approach for the efficiency estimation problem. This approach lies on the assumption that all the observations lie in the technological set. That is, one does not consider the problem where there is noise on the data.

[2]In this paper, we call the estimator proposed by Martins-Filho and Yao as NP3S, contrasting with our proposed estimator, which will be called NP2S

and DEA estimators. Another gain from this procedure is that the estimation method is fairly simple since it is based on local linear Kernel estimation. Martins-Filho and Yao (2007) propose an estimation process consisting of three steps. First step is estimating a conditional mean using local linear Kernel estimation. The second step follows Fan and Yao (1998), i.e., a local linear Kernel method is used to estimate a conditional variance function. The third and final step is a novel estimator that is related to their proposed production frontier model.

However, an undesirable result may be emerging in the second step of NP3S estimator since this estimation procedure allows for a negative variance estimate. To overcome this problem, Martins-Filho, Torrent and Ziegelmann (2010) propose to use the local exponential Kernel estimator of Ziegelmann (2002) to estimate the conditional variance functions, ensuring its nonnegativity. They also derive the asymptotic normality and consistency of production frontier under reasonable assumptions in the nonparametric context.This estimator consists of using an exponential functional at the minimization problem that characterizes the local Kernel regressions for estimating a nonnegative conditional variance. Thus, the authors use the local exponential Kernel estimator at the second step of the proposed method by Martins-Filho and Yao (2007). We call this frontier estimator as NPE.

Although NP3S and NPE have advantages in comparison with FDH and DEA estimators, we further pursue some improvements. The former estimators are characterized by an estimation procedure in three steps. The first two steps give the shape of the frontier and the third step is responsible for locating it. It is important to emphasize that the second step of both estimators is a regression that has as regressands squared residuals from the first stage. This feature is sometimes undesirable, specially for practitioners working with relatively small sample size. Here we propose the elimination of the second step of NP3S and NPE estimators, estimating the frontier in just two steps. Furthermore, our proposed estimator has as first step exactly the same first step presented in NP3S and NPE estimators, whereas our second step is very similar to the third step used for those. Therefore, we in fact eliminate the second step of NP3S and NPE and get the frontier in a simpler estimation procedure. Our contribution lies in perceiving that we actually get the frontier shape from the first step, and thus we need only one additional step to locate the estimated frontier. Hence, besides the simpler fashion of our estimator, which will be called NP2S, it maintains the advantages over FDH and DEA listed above.

This paper is composed as follows. In the second section we present the model originally proposed by Martins-Filho and Yao (2007). We then present our estimation process and compare it with NP3S and NPE. Section 3 discusses the asymptotic properties of our proposed estimator. In section 4 a Monte Carlo study is presented comparing NP2S, NPE and NP3S estimators. In section 5 we illustrate the application of our proposed estimator by an empirical analysis to USA crime data. Finally, in Section 6 conclusions and final comments are stated.

## 3.2 Nonparametric Frontier Estimation via Local Kernel Regression

### 3.2.1 The Model

In this section we present the model developed by Martins-Filho and Yao (2007). The problem may be viewed considering a firm that makes only one product from $k$ inputs, that is, $(x, y) \in \mathbb{R}^p_+ \times \mathbb{R}_+$, where $x$ describes $p$ inputs used for production and $y$ describes the output (one-output case) of a production unit. The production set is defined as previously. In a unique product case we have the following:

$$\Psi = \{(x, y) \in \mathbb{R}^{p+1}_+ | x \text{ can produce } y\}.$$

The production function or frontier associated with $\Psi$ is

$$\rho(x) = \sup\{y \in \mathbb{R}^q_+ \mid (y, x) \in \Psi\} \text{ for all } x \in \mathbb{R}^p_+.$$

In practice $\Psi$ and its frontier are unknown, so our main interest is in estimating this frontier from a set of observed firms, i.e., given a random sample of production units $\{(X_i, Y_i)\}_{i=1}^n$ that share a technology $\Psi$, obtaining estimates of $\rho(\cdot)$. By extension we are interested in constructing efficiency ranks and relative performance of production units. To see this, let $(x_0, y_0) \in \Psi$ characterize the performance of a production unit and define $0 \le R_0 \equiv \frac{y_0}{\rho(x_0)} \le 1$ to be this unit's (inverse) Farrell output efficiency measure.[3] From estimates of $\rho$ we can obtain estimates of $R_0$.

The construction of our frontier regression model is inspired by DGP for multiplicative regression. Primitive assumptions take place on $(X_i, R_i)$ and the properties of $Y_i$ arise from a suitable regression function. We assume that $Z_i \equiv (X_i, R_i)'$ is a $p + 1-$dimensional random vector with common density $g$ for all $i \in \{1, 2, ...\}$ and $\{Z_i\}$ forms an independently distributed sequence.

If there are observations on a random variable $Y_i$, the suitable regression model is defined as

$$Y_i = \rho(X_i)R_i = \frac{\sigma(X_i)}{\sigma_R}R_i \tag{3.1}$$

[3]Note that if the production level $y_0$ associated with $x_0$ lies on the frontier function we have $y_0 = \rho(x_0)$. The production process is then efficient and $R_0 = 1$.

where $R_i$ is an unobserved random variable and $X_i$ is an observed random vector in $\mathbb{R}^p_+$. In this context $Y_i$ is the output and $\rho(\cdot) = \frac{\sigma(\cdot)}{\sigma_R}$ is the production frontier, where $\sigma(\cdot) : \mathbb{R}^p_+ \to (0, \infty)$ is a measurable function and $\sigma_R$ is an unknown parameter. $X_i$ are the inputs and $R_i$ is the efficiency with values in $[0, 1]$. The closer $R_i$ is to 1 the closer are observed output and frontier. For an observed $(x_i, y_i)$, if $y_i$ is far from $\rho(x_i)$ it means low efficiency and so a small value for $R_i$. There is no specification of $R_i$ density, however two moment restrictions on $R_i$ must be assumed:

$$E(R_i|X_i = x) \equiv \mu_R, \text{ where } 0 < \mu_R < 1; \tag{3.2}$$

$$V(R_i|X_i = x) \equiv \sigma^2_R. \tag{3.3}$$

Note that $R_i \in [0, 1]$ and $0 < \mu_R < 1$ imply by construction that $0 < \sigma^2_R < \mu_R < 1$. The unknown parameter $\sigma_R$ locates the production frontier. For example, if a random sample of a population is far from the true frontier, estimated efficiency is low hence $\hat{\mu}_R$ and $\hat{\sigma}_R$ are small. In this case, DEA or FDH estimators will produce a sub-estimated production frontier. Due to presence of $\sigma_R$ in NP3S and NPE models, the estimated frontier is shifted to a higher level when compared to DEA or FDH. In next subsection, we present the original estimation procedure for this model and propose a new one, eliminating the second step.

### 3.2.2 The Estimation Procedure

In this section we characterize our estimator (NP2S). We can rewrite equation (3.1) as follows

$$Y_i = \frac{\sigma(X_i)}{\sigma_R} R_i = \frac{\mu_R}{\sigma_R}\sigma(X_i) + \sigma(X_i)\frac{(R_i - \mu_R)}{\sigma_R}.$$

Hence,

$$Y_i = m(X_i) + \sigma(X_i)\epsilon_i, \tag{3.4}$$

where $\epsilon_i = \frac{(R_i - \mu_R)}{\sigma_R}$ and $m(X_i) = \frac{\mu_R}{\sigma_R}\sigma(X_i)$. Given the conditional moment restrictions (3.2) and (3.3) on $R_i$, we have that $E(\varepsilon_i|X_i = x) = 0$ and $V(\varepsilon_i|X_i = x) = 1$. Hence, $E(Y_i|X_i = x) = m(X_i)$ and $V(Y_i|X_i = x) = \sigma^2(x)$.

First, we note that $m(X_i) \equiv \mu_R\rho(X_i)$. Therefore, estimating $m(X_i)$ gives us $\hat{m}(x) = \mu_R\hat{\rho}(x)$, since $\mu_R$ does not depend on $X_i$. We thus get from $\hat{m}(x)$ an estimation of $\rho(x)$, but in a wrong position, i.e., up to the scale. Then, if we have an estimator for $\mu_R$ we can propose to estimate the frontier as $\hat{\rho}(X_i) = \frac{\hat{m}(X_i)}{\hat{\mu}_R}$. With this in mind, we propose to estimate $\rho(X_i)$ in two simple steps. The first is simply the local linear Kernel estimator of Fan (1992) with regressand $Y_i$ and regressors $X_i$. That is, for any $x \in \mathbb{R}^p_+$ we obtain $\hat{m}(x) \equiv \hat{\alpha}$, where

$$(\hat{\alpha}, \hat{\beta}) = \arg\min_{\alpha,\beta} \sum_{i=1}^{n} (Y_i - \alpha - \beta(X_i - x))^2 K_{h_n}(X_i - x), \tag{3.5}$$

with $K(\cdot) : \mathbb{R}^p \to \mathbb{R}$ a symmetric probability density function, $K_h(u) = (1/h)K(u/h)$ and $0 < h_n \to 0$ as $n \to \infty$ a bandwidth. This first step gives us the frontier multiplied by $\mu_R$. Then, in the second stage we propose an estimator for $\mu_R$,

$$\hat{\mu}_r = \left( \max_{1 \leqslant i \leqslant n} \frac{Y_i}{\hat{m}(X_i; h_n)} \right)^{-1}.$$

In the second step, we use the idea originally proposed by Martins-Filho and Yao (2007); we assume that there exists at least one observed production unit that is efficient, i.e. there is at least one $R_i$ identically equal to one. The idea behind the estimator proposed above is to establish $Y_i = \rho(X_i)R_i = \frac{m(X_i)}{\mu_R} R_i$ and then set one firm to be efficient. Therefore, after these two steps the proposed estimator for the frontier at $x \in \mathbb{R}^p$ is given by $\hat{\rho}(X_i) = \frac{\hat{m}(X_i, h_n)}{\hat{\mu}_R}$. We note that $\hat{\rho}(X_i)$ is a smooth estimator that envelops the data but may lie above or below the true frontier $\rho(X_i)$.

It is worth to emphasize that $\hat{\mu}_R$ depends on $h_n$ trough $\hat{m}(X_i, h_n)$. Moreover, in section 3 it will be convenient to distinguish the bandwidth used in the first step from that used in the second step, which we will denote by $g_n$. Thus, in this new notation, the production frontier estimator at $x \in \mathbb{R}^p$ is given by $\hat{\rho}(X_i, h_n, g_n) = \frac{\hat{m}(X_i, h_n)}{\hat{\mu}_R(g_n)}$.

**Comparing NP2S Estimation Procedure with NP3S and NPE**

One of the goals of this paper is to propose an alternative estimation procedure for the model in equation (3.4). Thus, in this subsection we outline the estimation procedures proposed by Martins-Filho and Yao (2007) (NP3S estimator) and Martins-Filho et. al (2009) (NPE estimator) and compare some features of those estimators with the one proposed in this paper (NP2S estimator).

NP3S and NPE estimation methods for the model described in equation (3.4) are composed of three steps. The first two steps are responsible for estimating the frontier shape, $\sigma(\cdot)$, while the third step gives an estimative of the frontiers position, $\sigma_R$. The first step is the same as the first step for NP2S as in equation (5). The second step consists of implementing again a local kernel regression, but now for the conditional volatility function. The idea is using $\hat{m}(\cdot)$ from the first step and define $e_i \equiv (Y_i - \hat{m}(X_i))^2$

to obtain $\hat{\sigma}^2(x)$ as follows:

$$(\hat{\alpha}_1, \hat{\beta}_1) = \arg\min_{\alpha_1,\beta_1} \sum_{i=1}^{n} (e_i - \psi(\alpha_1 - \beta_1(X_i - x)))^2 K_{h_n}(X_i - x)\,. \tag{3.6}$$

For NP3S estimator, we have $\psi(x) \equiv x$ and the estimator for conditional volatility function is given by $\hat{\sigma}^2_l(x) = \hat{\alpha}$. This is the local linear kernel estimator for the variance as defined by Fan and Yao (1998). For NPE estimator, the functional has the form $\psi(x) \equiv \exp(x)$ and the variance estimator is defined as $\hat{\sigma}^2_e(x) = \exp(\hat{\alpha})$, as proposed in Ziegelmann (2002). After that, the frontier shape is estimated in NP3S model as $\hat{\sigma}_l(X_i) = \sqrt{\hat{\sigma}^2_l(X_i)}$ and in NPE model as $\hat{\sigma}_e(X_i) = \sqrt{\hat{\sigma}^2_e(X_i)}$.

After obtaining an estimate for frontier shape, $\hat{\sigma}^2(\cdot)$, a third step is proposed to estimate the frontier position, $\sigma_R$. The proposed estimator is

$$s_R = \left( \max_{1 \le i \le n} \frac{Y_i}{\hat{\sigma}(X_i)} \right)^{-1}, \tag{3.7}$$

where $\hat{\sigma}(X_i)$ is the estimate from the previous second step. We use $s^l_R$ to represent the location estimator for NP3S, and for NPE we use $s^e_R$. As pointed out earlier, the intuition behind these estimators is to assume that there exists one observed production unit that is efficient. Hence, a production frontier estimator at $x \in \mathbb{R}^p$ is given by $\hat{\rho}_l(\cdot) = \frac{\hat{\sigma}_l(\cdot)}{s^l_R}$ for NP3S case and by $\hat{\rho}_e(\cdot) = \frac{\hat{\sigma}_e(\cdot)}{s^e_R}$ for NPE.

Comparing NP2S, NPE and NP3S, we see that the first step is exactly the same in all cases. Furthermore, the step responsible to locate the frontier - second step in NP2S and third step in NP3s and NPE - is built over the same idea. Note however, that NP2S eliminates one step, and therefore, does not require estimation of a conditional volatility function; and thus NP2S eliminates the need of estimating a regression that has as depended variable residuals of a previous regression. In other words, our contribution is to get the frontier shape from the first step, via $m(X_i) = \mu_R \rho(X_i)$, then positioning it at the second step through an estimative for $\mu_R$. On the other hand, in NP3S and NPE cases, frontier shapes are captured after two steps, but miss positioned by a factor $\sigma_R$, since $\sigma(X_i) = \sigma_R \rho(X_i)$, then a third step is necessary to correct its position, using an estimate for $\sigma_R$.

## 3.3 Asymptotic Characterization

In this section we discuss the asymptotic properties of the proposed estimator. The following assumptions are assumed:

**Assumption A1.** 1. $Z_i = (X_i, R_i)'$ for $i = 1, 2, ..., n$ is an independent and identically distributed sequence of random vectors with density $f$, where $f_X(x)$ and $f_R(r)$ denote the common marginal densities of $X_i$ and $R_i$ respectively, and $f_{R|X}(r; X)$ denotes the common density of $R_i$ given $X$. 2. $0 \leq \underline{B}_{f_X} \leq f_X(x) \leq \bar{B}_{f_X} < \infty$ for all $x \in G$, $G$ a compact subset of $\Theta = \times_{i=1}^{p}(0, \infty)$, which denotes the Cartesian product of the intervals $(0, \infty)$ .

**Assumption A2.** 1. $Y_i = \sigma(X_i)\frac{R_i}{\sigma_R}$. 2. $R_i \in [0, 1], X_i \in \Theta$. 3. $E(R_i|X_i) = \mu_R, V(R_i|X_i) = \sigma_R^2$. 4. The regression function $m(x)$ has a bounded and continuous second derivative for all $x \in \Theta$, which will be denoted by $m^{(2)}(x)$. 5. $0 < \underline{B}_\sigma \leq \sigma(x) \leq \bar{B}_\sigma < \infty$ for all $x \in \Theta$.

**Assumption A3.** $K(x) : S_p \to \mathbb{R}$ is a symmetric density function with bounded support $S_p \to \mathbb{R}^p$ satisfying: 1. $\int xK(x)dx = 0$. 2. $\int x^2 K(x)dx = \sigma_p^2$. 3. For all $x \in \imath, |K(x)| < B_p < \infty$. 4. For all $x, x' \in \mathbb{R}_p, |K(x) - K(x')| < m\|x - x'\|$ for some $0 < m < \infty$.

**Assumption A4.** For all $x, x' \in \Theta, |g_X(x) - g_X(x')| < m_g\|x - x'\|$ for some $0 < m < \infty$.

Assumptions $A1.1$ and $A2$ imply that $\{Y_i, X_i\}_{i=1}^n$ forms an iid sequence of random variables with joint probability density $\phi(y, x)$. Comparing with Martins-Filho and Yao (2007) and Martins-Filho, Torrent and Ziegelmann (2010), we do not have to assume anything about the second derivative of $\sigma^2(x)$. Furthermore, we do not have to deal with regressands that are themselves residuals from a first stage nonparametric regression, due to elimination of the second step in the estimation procedure. Therefore, asymptotic properties are much easier to obtain. The uniform consistency and asymptotic normality of the frontier estimator are presented in the following theorems.

**Theorem 4** *Suppose that Assumptions A1-A4 hold. In addition assume that* $E(|\epsilon_i|^{2+\delta}|X_i) < C_1 < \infty$. *Then for every* $x \in G$

$$\sqrt{nh_n}\left(\hat{m}(x, h_n) - m(x) - B_{1n}\right) \xrightarrow{d} N\left(0, \frac{\sigma^2(x)}{f_X(x)} \int K^2(\phi)d\phi\right),$$

*where* $B_{1n} = \frac{h_n^2 m^{(2)}(x)\sigma_k^2}{2} + o_p(h_n^2)$.

**Theorem 5** *Let* $L_n$ *be a non-stochastic sequence such that* $0 < L_n \to 0$ *as* $n \to \infty$ *and suppose that* (*i*) $\hat{m}(x, g_n) - m(x) = O_p(L_n)$ *uniformly in* $G$ *and* (*ii*) $1 - \max_{1 \leqslant i \leqslant n} R_t = O_p(L_n)$. *Then,*

*a)* $\hat{\mu}_R(g_n) - \mu_R = O_p(L_n)$;

*b) Under the assumptions* $A1 - A4$, *if* $\frac{ng_n^5}{ln(n)} \to \infty$, $nh_n^5 = o(1)$, *and* $nh_n g_n^4 = O(1)$, *then:*

$$\sqrt{nh_n}\left(\frac{\hat{m}(x, h_n)}{\hat{\mu}_R(g_n)} - \frac{m(x)}{\mu_R} - B_{2n}\right) \xrightarrow{d} N\left(0, \frac{\sigma^2(x)}{\mu_R^2 f_X(x)} \int K^2(\phi)d\phi\right),$$

*where* $B_{2n} = O_p(g_n^2)$.

## 3.4 Monte Carlo Study

In this section we investigate some of the finite sample properties of our estimator, NP2S, via a Monte Carlo study. For comparison purposes, we also include in the study the local exponential frontier estimator proposed in Martins-Filho, Torrent and Ziegelmann (2010), NPE, and the local linear frontier estimator proposed in Martins-Filho and Yao (2007), NP3S. Our simulations are based on model (1), i.e., $Y_i = \frac{\sigma(X_i)R_i}{\sigma_R}$, with $p = 1$. We generate data with the following characteristics. The $X_i$'s are pseudorandom variables with a uniform distribution whose support is given on $[a_l, b_u]$. $R_i = \exp(-Z_i)$, where $Z_i$ are pseudorandom variables with exponential distribution (parameter $\beta > 0$), therefore $R_i$ has support on $(0, 1]$. We consider two specifications for $\sigma(x)$:

$$\sigma_1(x) = \sqrt{x}, \text{ with } x \in [a_l, b_u] = [10, 100] \text{ and}$$

$$\sigma_2(x) = 3(x - 1.5)^3 + 0.25x + 1.125, \text{ with } x \in [a_l, b_u] = [1, 2],$$

which are associated with convex and non-convex production technologies. Four parameters for the exponential distribution are considered: $\beta_1 = 3$, $\beta_2 = 1$, $\beta_3 = 2/3$ and $\beta_4 = 1/3$. These choices of parameters produce, respectively, the following values for the parameters of $g_{R|X} : (\mu_R, \sigma_R^2) = (0.25, 0.08), (0.5, 0.08)$, $(0.6, 0.07)$ and $(0.75, 0.04)$. Three sample sizes $n = 200, 300, 400$ are used. We evaluate the frontiers at $x_1 = 32.5$, $x_2 = 55$ and $x_3 = 77.5$ for $\sigma_1(x)$ and at $x_1 = 1.25$, $x_2 = 1.5$ and $x_3 = 1.75$ for $\sigma_2(x)$. These values of $X$ correspond to the 25th, 50th and 75th percentiles.

An important aspect in the implementation of our frontier estimator is bandwidth selection. We consider the following rule-of-thumb bandwidth:

$$\hat{h}_{ROT} = \left( \frac{\int K^2(\phi)d\phi \int \dot{\sigma}^2(x)dx}{(\sigma_K^2)^2 \left( \max_{1 \le i \le n} \left( \frac{\dot{m}^{(2)}(x_i)\dot{R}_i}{\dot{m}(x_i)} \right) \right)^2 \frac{1}{n} \sum_{i=1}^n \dot{m}(x_i)} \right)^{1/5} n^{-(1+4\gamma)/5}$$

where $\gamma$ is set to be 0.11 in all experiments, which satisfies the requirements on relative speed of the bandwidths $h_n$ and $g_n$, as stated in Theorem 5. Thus, $\hat{g}_{ROT} = n^{\gamma}\hat{h}_{ROT}$. The sequence $\{\dot{m}(X_i)\}_{i=1}^n$ is estimated with an ordinary least square quartic regression of $\{Y_i\}_{i=1}^n$ on $\{X_i\}_{i=1}^n$. $\{\dot{m}(X_i)\}_{i=1}^n$ is then used to construct $\max_{1 \le i \le n} \left( \frac{\dot{m}^{(2)}(x_i)\dot{R}_i}{\dot{m}(x_i)} \right)$ and $\frac{1}{n} \sum_{i=1}^n \dot{m}(x_i)$. $\int \dot{\sigma}^2(x)dx$ is estimated by sum of squared residuals of that regression.

We evaluate the overall performance of the efficiency estimator based on three different measures. First, we consider the correlation between the efficiency ranks produced by the estimator and the true efficiency ranks, given by:

$$R_{rank} = \frac{cov(rank(\hat{R}_i), rank(R_i))}{\sqrt{var(rank(\hat{R}_i))var(rank(R_i))}},$$

where $rank(R_i)$ is the ranking index according to the magnitude of $R_i$. The closer $R_{rank}$ for $\hat{R}_i$ is to 1, the higher the correlation between the true $R_i$ and $\hat{R}_i$, thus the better the estimator $\hat{R}_i$. The second measure we consider is $R_{mag} = \frac{1}{n}\sum_{i=1}^{n}(\hat{R}_i - R_i)^2$ which is simply the squared Euclidean distance between the estimated vector of efficiencies and the true vector of efficiencies. The third measure we use is $R_{rel} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{R}_i}{\hat{R}_t} - \frac{R_i}{R_t}\right|$, where $t$ is the position index for $R_t = \max_{1\leq i\leq n} R_i$, and $\hat{R}_t$ is the $t$th corresponding element in $\{\hat{R}_i\}_{i=1}^{n}$, which may not be the maximum of $\hat{R}_i$. Hence $R_{rank}$ and $R_{mag}$ summarize the performance of the estimator $\hat{R}_i$ in rankings and in calculating the magnitude of efficiency. On the other hand $R_{rel}$ captures the relative efficiency.

The results of our simulations are summarized in tables 1-4 and figures 1-16. Figures 1-16 are boxplots of MSE for the frontier estimator ($\hat{\rho}(\cdot)$) and efficiency estimator ($\hat{R}_i$). Each boxplot is constructed from 1000 points (repetitions), where each point corresponds to a sample draw and is calculated as the squared Euclidean distance between the estimate and true value of $\rho(\cdot)$ and $R_i$. The whiskers extend to the most extreme data point which is no more than 1.5 times the interquartile range. Tables 1-4 are constructed based on the data points between the whiskers in the boxplots. Tables 1-2 provide the bias and MSE of the frontier estimators at three different values of $x$. Tables 3-4 give the overall performance of the efficiency estimators according to the measures described above.

### 3.4.1 General regularities

As expected from the asymptotic results of section 3, as the sample size $n$ increases, the boxplots show that MSE decreases for all estimators and values for $\mu_R$ considered. The bias and MSE for the frontier estimator based on NP2S, NPE and NP3S, presented on Tables 1-2, generally decrease, with some exceptions when it comes to the bias. Regarding the measures of overall performance for efficiency estimators mentioned above, all estimators perform better as $n$ increases. The asymptotics of the three estimators seem to be confirmed in general terms as their performances improve with large $n$.

We now turn to the impact of different values of $\mu_R$ on the performance of NP2S, NPE and NP3S.

For NP2S, regarding the frontier estimator, the performance in terms of MSE improves as the value of $\mu_R$ increases. This pattern is most likely explained by the fact its variance is inversely proportional to $\mu_R$ value, as stated in Theorem 5. Following the asymptotic results in Theorem 2, the bias of the frontier estimator of NP2S is positive for all $\mu_R$. No clear pattern is discerned from the impact of $\mu_R$ on the bias. Regarding the measures of overall performance for the efficient estimator described above, NP2S estimator seems to perform worse when $\mu_R$ is large for $R_{rank}$. On the other hand, for $R_{mag}$ NP2S performs better as $\mu_R$ increases. Besides no clear pattern is discerned from the impact of a larger $\mu_R$ on $R_{rel}$. Regarding the frontier estimator for NPE and NP3S, the best performance in terms of MSE occurs when $\mu_R = 0.5$, and the worst performance occurs when $\mu_R = 0.75$. The relative diminished performance when $\mu_R = 0.75$ is most likely explained by the fact that for this DGP $\sigma_R^2$ is roughly half of its value in other DGPs, contributing to it to have higher variance as suggested by asymptotic theory[4]. The bias of the frontier estimator of NPE as well as of NP is generally positive, except for small $\mu_R$. The bias seems to increase with $\mu_R$. Regarding the measures of overall performance for the efficient estimator described above, NPE and NP estimator seem to perform worse when $\mu_R$ is large for $R_{rank}$, $R_{mag}$ and $R_{rel}$.

### 3.4.2 Relative performance of estimators

On estimating the production frontier (Tables 1-2 and Figs. 1-16) there seems to be evidence that NP2S dominates NPE and NP3S in terms of MSE when $\mu_R = 0.6$ and $0.75$, while NPE and NP3S dominate NP2S for $\mu_R = 0.25$. For $\mu_R = 0.5$ NPE and NP3S perform better than NP2S in case where $\sigma(x) = \sigma_2(x)$. When $\sigma(x) = \sigma_1(x)$ NP2S performs better than its competitors with few exceptions for $n = 200$. When different measures of overall performance we considered are analyzed (Tables 3-4), we observe that NP2S outperforms NPE and NP3S when $\mu_R = 0.6$ and $0.75$ for all measures considered. For $\mu_R = 0.25$ the estimators present a very similar performance regarding $R_{rank}$, however NPE and NP3S generally perform better than NP2S when considering $R_{mag}$ and $Rrel$. When $\mu_R = 0.5$ and $\sigma(x) = \sigma_1(x)$ all estimators perform similar for $R_{rank}$. For $R_{mag}$ NP2S outperforms NP3S, while compared with NPE, NP2S performs better only when $n = 400$. Regarding $R_{rel}$ NP2S performs better than its competitors for $n = 300$ and $n = 400$. When $\mu_R = 0.5$ and $\sigma(x) = \sigma_2(x)$ NPE and NP3S generally perform better than NP2S. Based on these results, it seems reasonable to conclude higher the mean efficiency is the better NP2S performance is.

[4]See Martins-Filho, Torrent and Ziegelmann (2010) for NPE and Martins-Filho and Yao (2007) for NP3S.

## 3.5 Real Data Example

We illustrate our methodology analyzing USA crime data. The goal is to estimate a production frontier and efficiency for 294 USA Law Enforcement agencies using data for the year 2000. Data sources are FBI's Uniform Crime Reports and LEMAS (Law Enforcement Management and Administrative Statistics) survey. All data used is available on the Internet in the site of Bureau of Justice Statistics (http://bjsdata.ojp.usdoj.gov).

In order to measure crime we consider crime trend data from FBI's Uniform Crime Reports for large agencies in USA (population coverage $\geq 80,000$). This data was used to construct the output, which is defined as population per total crime, where total crime is number of violent crimes plus number of property crimes[5]. This output measure is consistent with our methodology and also is considered in Gorman and Ruggiero (2008).

To measure resources invested in police force we consider data from LEMAS (Law Enforcement Management and Administrative Statistics) survey. This data was used to construct a measure of input, which is in fact an index composed by base annual starting salaries for three categories: chief executive (*chief*), sergeant (*sergeant*) and entry-level officer (*entry*). In order to simplify our analysis and avoid any *curse of dimensionality*, we reduce our problem dimension via principal component analysis of those three inputs. Proceeding in this manner, we can see that the first orthogonal component is responsible by 79.18% of the total variability. Therefore, we decide to use a single component, which is the following linear combination of the three original variables

$$input = 0.958(chief - \overline{chief}) + 0.242(sergeant - \overline{sergeant}) + 0.157(entry - \overline{entry}),$$

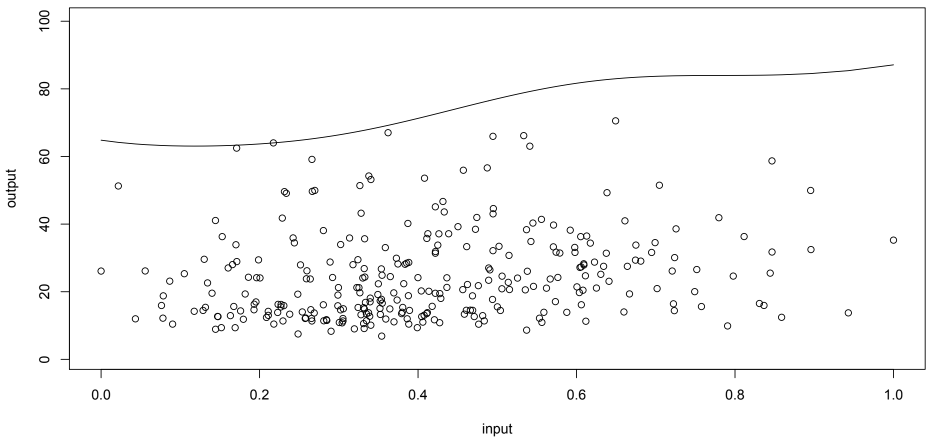
where $\bar{x}$ denotes the sample mean of a variable $x$. In order to avoid negative values and therefore ease interpretability of our index, we rescale the above index such that its minimum value is zero. Moreover, we project the above index into the interval [0,1] in order to facilitate the bandwidth choice.

The estimated frontier is displayed in Fig. 3.1. We can notice it is a smooth function, where the point lying on the estimated curve reflects the model anchoring assumption, corresponding to Harford County Sheriff Office-MD. Efficiency rank and efficiency scores are shown on Table 3.5. Using the output and input measures mentioned above, we see that the estimated five most efficient agencies are Harford County

[5] Violent crimes are murder and non-negligent manslaughter, forcible rape, robbery and aggravated assault. Property crimes are burglary, larceny-theft and motor vehicle theft.

Sheriff Office-MD, Springfield Police Department-MO, Fulton County Police Department-GA, Jefferson County Sheriff Department-CO and El Dorado County Sheriff Department-CA, in decreasing order. On the other hand, the estimated five least efficient agencies are Orlando Police Department-FL, Baltimore City Police Department-MD, Atlanta Police Department-GA, Chattanooga Police Department-TN and St Joseph County Sheriff Department-IN, where the last is the least efficient.

Figura 3.1: NP2S Frontier Estimation




## 3.6 Summary and conclusions

In this paper we propose a novel approach of estimating the production frontier model developed by Martins-Filho and Yao (2007). There they use Kernel regression for estimating production frontier and therefore efficiency for production units with significant advantages when compared to DEA and FDH estimators. However, their estimation process is made in three stages. We here propose a modification on the estimation procedure, eliminating the need of the second step. The result is a simpler estimation procedure that retains all inherent advantages present in the original estimator. A Monte Carlo study was performed comparing three estimators: our estimator, called NP2S; NP3S from Martins-Filho and Yao (2007); and NPE, presented in Martins-Filho, Torrent and Ziegelmann (2010). The results show that NP2S outperforms its competitors for mean efficiency above 50%, i.e., $\mu_R = 0.60$ and $\mu_R = 0.75$ in the simulations considered.

## 3.7 Appendix 1: Tables and Graphics

Tabela 3.1: Frontier I - Bias and MSE of frontier estimators

| $\sigma_1(x)$ | $n$ | | $x_1 = 32.5$ | | | $x_2 = 55$ | | | $x_3 = 77.5$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | NP2S | NPE | NP3S | NP2S | NPE | NP3S | NP2S | NPE | NP3S |
| $\mu_R = 0.25$ | 200 | Bias | 0.410 | -0.303 | -0.027 | 0.260 | 0.215 | 0.421 | 0.462 | -0.587 | -0.550 |
| | | MSE | 3.346 | 2.250 | 3.069 | 2.873 | 2.965 | 6.915 | 5.377 | 4.636 | 10.442 |
| | 300 | Bias | 0.339 | -0.154 | 0.159 | 0.294 | 0.616 | 0.914 | 0.645 | -0.189 | -0.058 |
| | | MSE | 1.873 | 1.764 | 2.020 | 1.626 | 2.659 | 5.049 | 4.069 | 3.331 | 6.403 |
| | 400 | Bias | 0.298 | -0.142 | 0.063 | 0.313 | 0.683 | 0.847 | 0.790 | 0.069 | 0.020 |
| | | MSE | 1.558 | 1.386 | 1.717 | 1.409 | 2.423 | 4.117 | 3.366 | 2.729 | 4.434 |
| $\mu_R = 0.50$ | 200 | Bias | 0.441 | 0.041 | 0.288 | 0.230 | 0.788 | 0.996 | 0.874 | 0.055 | 0.261 |
| | | MSE | 1.305 | 1.017 | 1.607 | 0.941 | 1.935 | 3.929 | 2.634 | 1.656 | 3.734 |
| | 300 | Bias | 0.346 | 0.092 | 0.288 | 0.279 | 0.947 | 1.086 | 0.959 | 0.350 | 0.538 |
| | | MSE | 0.766 | 0.833 | 1.284 | 0.618 | 2.006 | 3.575 | 2.293 | 1.765 | 3.120 |
| | 400 | Bias | 0.312 | 0.155 | 0.181 | 0.243 | 0.995 | 0.955 | 0.871 | 0.386 | 0.359 |
| | | MSE | 0.549 | 0.705 | 0.940 | 0.439 | 2.031 | 2.651 | 1.681 | 1.526 | 2.111 |
| $\mu_R = 0.60$ | 200 | Bias | 0.413 | 0.295 | 0.812 | 0.197 | 1.114 | 1.695 | 0.916 | 0.534 | 1.233 |
| | | MSE | 0.913 | 1.510 | 3.536 | 0.648 | 3.065 | 8.807 | 2.171 | 2.762 | 8.919 |
| | 300 | Bias | 0.334 | 0.213 | 0.587 | 0.242 | 1.209 | 1.577 | 0.940 | 0.682 | 1.176 |
| | | MSE | 0.520 | 1.114 | 2.357 | 0.386 | 2.823 | 6.676 | 1.798 | 2.591 | 6.430 |
| | 400 | Bias | 0.310 | 0.311 | 0.516 | 0.254 | 1.281 | 1.489 | 0.857 | 0.798 | 1.089 |
| | | MSE | 0.370 | 1.080 | 1.842 | 0.312 | 3.062 | 5.178 | 1.358 | 2.832 | 4.709 |
| $\mu_R = 0.75$ | 200 | Bias | 0.364 | 1.402 | 2.099 | 0.278 | 2.713 | 3.475 | 1.099 | 1.856 | 2.467 |
| | | MSE | 0.481 | 8.170 | 14.777 | 0.436 | 16.466 | 31.507 | 2.076 | 13.399 | 32.398 |
| | 300 | Bias | 0.344 | 0.950 | 1.629 | 0.331 | 2.671 | 3.366 | 1.140 | 1.932 | 2.677 |
| | | MSE | 0.308 | 5.023 | 10.351 | 0.310 | 13.795 | 28.208 | 1.931 | 11.644 | 27.862 |
| | 400 | Bias | 0.333 | 0.938 | 1.525 | 0.380 | 2.682 | 3.303 | 1.126 | 1.907 | 2.601 |
| | | MSE | 0.251 | 4.566 | 9.115 | 0.299 | 12.851 | 24.609 | 1.721 | 11.321 | 22.928 |

Tabela 3.2: Frontier II - Bias and MSE of frontier estimators

| $\sigma_2(x)$ | $n$ | | $x_1 = 1.25$ | | | $x_2 = 1.5$ | | | $x_3 = 1.75$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | NP2S | NPE | NP3S | NP2S | NPE | NP3S | NP2S | NPE | NP3S |
| $\mu_R = 0.25$ | 200 | Bias | 0.106 | -0.109 | -0.107 | 0.237 | 0.043 | 0.041 | 0.427 | 0.246 | 0.242 |
| | | MSE | 0.266 | 0.074 | 0.077 | 0.304 | 0.044 | 0.048 | 0.510 | 0.194 | 0.193 |
| | 300 | Bias | 0.108 | -0.072 | -0.072 | 0.267 | 0.093 | 0.092 | 0.474 | 0.317 | 0.317 |
| | | MSE | 0.179 | 0.050 | 0.053 | 0.240 | 0.042 | 0.042 | 0.468 | 0.213 | 0.213 |
| | 400 | Bias | 0.095 | -0.053 | -0.047 | 0.269 | 0.114 | 0.123 | 0.491 | 0.329 | 0.346 |
| | | MSE | 0.141 | 0.039 | 0.043 | 0.200 | 0.036 | 0.040 | 0.416 | 0.184 | 0.198 |
| $\mu_R = 0.50$ | 200 | Bias | 0.136 | 0.007 | 0.016 | 0.220 | 0.164 | 0.175 | 0.460 | 0.389 | 0.403 |
| | | MSE | 0.111 | 0.034 | 0.041 | 0.140 | 0.052 | 0.060 | 0.314 | 0.227 | 0.239 |
| | 300 | Bias | 0.129 | 0.031 | 0.035 | 0.239 | 0.194 | 0.199 | 0.484 | 0.417 | 0.431 |
| | | MSE | 0.076 | 0.032 | 0.038 | 0.120 | 0.061 | 0.066 | 0.308 | 0.238 | 0.248 |
| | 400 | Bias | 0.116 | 0.039 | 0.057 | 0.241 | 0.196 | 0.221 | 0.470 | 0.401 | 0.439 |
| | | MSE | 0.059 | 0.025 | 0.033 | 0.107 | 0.057 | 0.071 | 0.274 | 0.215 | 0.250 |
| $\mu_R = 0.60$ | 200 | Bias | 0.134 | 0.087 | 0.092 | 0.218 | 0.267 | 0.277 | 0.464 | 0.533 | 0.548 |
| | | MSE | 0.077 | 0.075 | 0.087 | 0.115 | 0.117 | 0.130 | 0.288 | 0.388 | 0.399 |
| | 300 | Bias | 0.138 | 0.086 | 0.094 | 0.242 | 0.270 | 0.282 | 0.490 | 0.514 | 0.535 |
| | | MSE | 0.062 | 0.058 | 0.064 | 0.106 | 0.107 | 0.117 | 0.294 | 0.346 | 0.364 |
| | 400 | Bias | 0.126 | 0.096 | 0.115 | 0.247 | 0.263 | 0.289 | 0.485 | 0.490 | 0.526 |
| | | MSE | 0.049 | 0.049 | 0.061 | 0.099 | 0.100 | 0.120 | 0.276 | 0.316 | 0.352 |
| $\mu_R = 0.75$ | 200 | Bias | 0.147 | 0.320 | 0.326 | 0.309 | 0.586 | 0.577 | 0.591 | 0.992 | 0.964 |
| | | MSE | 0.062 | 0.387 | 0.418 | 0.155 | 0.512 | 0.514 | 0.410 | 1.341 | 1.214 |
| | 300 | Bias | 0.153 | 0.272 | 0.283 | 0.329 | 0.520 | 0.530 | 0.606 | 0.905 | 0.910 |
| | | MSE | 0.053 | 0.277 | 0.311 | 0.147 | 0.404 | 0.420 | 0.410 | 1.037 | 1.007 |
| | 400 | Bias | 0.164 | 0.276 | 0.291 | 0.339 | 0.491 | 0.517 | 0.606 | 0.842 | 0.876 |
| | | MSE | 0.053 | 0.252 | 0.278 | 0.152 | 0.368 | 0.399 | 0.403 | 0.914 | 0.959 |

Tabela 3.3: Frontier I - Overall Efficiency Measures

| $\sigma_1(x)$ | $n$ | $R_{rank}$ | | | $R_{mag}(\times 10^{-2})$ | | | $R_{rel}$ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | NP2S | NPE | NP3S | NP2S | NPE | NP3S | NP2S | NPE | NP3S |
| $\mu_R = 0.25$ | 200 | 0.998 | 0.998 | 0.998 | 0.157 | 0.118 | 0.130 | 0.024 | 0.021 | 0.019 |
| | 300 | 0.998 | 0.999 | 0.999 | 0.105 | 0.084 | 0.108 | 0.019 | 0.018 | 0.016 |
| | 400 | 0.999 | 0.999 | 0.999 | 0.087 | 0.071 | 0.081 | 0.017 | 0.017 | 0.015 |
| $\mu_R = 0.50$ | 200 | 0.994 | 0.994 | 0.994 | 0.140 | 0.115 | 0.176 | 0.029 | 0.028 | 0.028 |
| | 300 | 0.995 | 0.995 | 0.995 | 0.104 | 0.101 | 0.140 | 0.023 | 0.027 | 0.024 |
| | 400 | 0.996 | 0.996 | 0.996 | 0.086 | 0.087 | 0.108 | 0.022 | 0.024 | 0.022 |
| $\mu_R = 0.60$ | 200 | 0.991 | 0.986 | 0.985 | 0.121 | 0.185 | 0.362 | 0.028 | 0.036 | 0.037 |
| | 300 | 0.993 | 0.989 | 0.989 | 0.095 | 0.152 | 0.273 | 0.023 | 0.033 | 0.032 |
| | 400 | 0.995 | 0.991 | 0.991 | 0.082 | 0.144 | 0.210 | 0.021 | 0.030 | 0.028 |
| $\mu_R = 0.75$ | 200 | 0.985 | 0.932 | 0.934 | 0.088 | 0.576 | 0.948 | 0.024 | 0.060 | 0.059 |
| | 300 | 0.987 | 0.948 | 0.950 | 0.083 | 0.423 | 0.764 | 0.021 | 0.053 | 0.050 |
| | 400 | 0.989 | 0.950 | 0.957 | 0.073 | 0.403 | 0.676 | 0.019 | 0.050 | 0.045 |

Tabela 3.4: Frontier II - Overall Efficiency Measures

| $\sigma_2(x)$ | $n$ | $R_{rank}$ | | | $R_{mag}(\times 10^{-2})$ | | | $R_{rel}$ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | NP2S | NPE | NP3S | NP2S | NPE | NP3S | NP2S | NPE | NP3S |
| $\mu_R = 0.25$ | 200 | 0.997 | 0.999 | 0.999 | 0.222 | 0.076 | 0.074 | 0.027 | 0.016 | 0.016 |
| | 300 | 0.998 | 0.999 | 0.999 | 0.164 | 0.071 | 0.065 | 0.022 | 0.015 | 0.014 |
| | 400 | 0.999 | 0.999 | 0.999 | 0.137 | 0.062 | 0.058 | 0.020 | 0.014 | 0.013 |
| $\mu_R = 0.50$ | 200 | 0.992 | 0.994 | 0.995 | 0.223 | 0.148 | 0.147 | 0.032 | 0.027 | 0.026 |
| | 300 | 0.995 | 0.995 | 0.996 | 0.184 | 0.146 | 0.141 | 0.027 | 0.025 | 0.024 |
| | 400 | 0.996 | 0.996 | 0.996 | 0.159 | 0.134 | 0.140 | 0.024 | 0.024 | 0.023 |
| $\mu_R = 0.60$ | 200 | 0.990 | 0.986 | 0.988 | 0.201 | 0.271 | 0.268 | 0.031 | 0.036 | 0.034 |
| | 300 | 0.992 | 0.989 | 0.990 | 0.185 | 0.232 | 0.230 | 0.026 | 0.032 | 0.030 |
| | 400 | 0.994 | 0.991 | 0.991 | 0.167 | 0.216 | 0.227 | 0.024 | 0.030 | 0.029 |
| $\mu_R = 0.75$ | 200 | 0.982 | 0.941 | 0.947 | 0.193 | 0.689 | 0.678 | 0.027 | 0.055 | 0.051 |
| | 300 | 0.985 | 0.953 | 0.957 | 0.184 | 0.556 | 0.565 | 0.024 | 0.048 | 0.045 |
| | 400 | 0.987 | 0.958 | 0.960 | 0.182 | 0.489 | 0.535 | 0.022 | 0.046 | 0.044 |

Figura 3.2: Frontier I - Boxplot of Estimators - $n = 200$ - $\mu_r = 0.25$

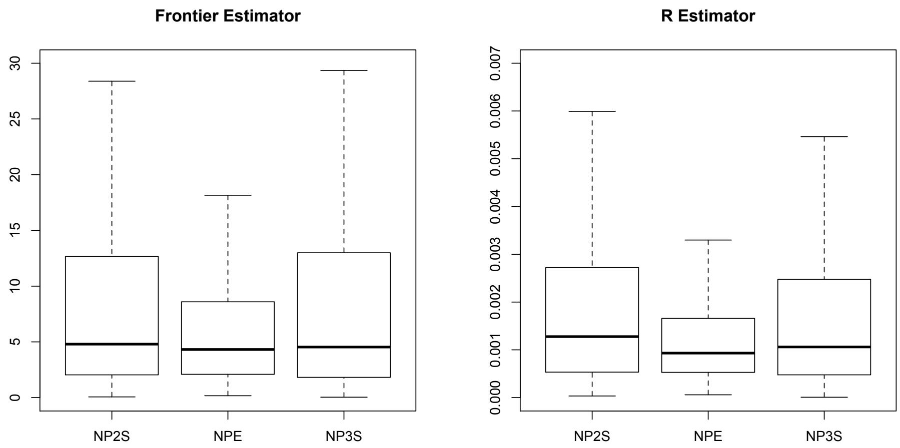



Figura 3.3: Frontier I - Boxplot of Estimators - $n = 400$ - $\mu_r = 0.25$

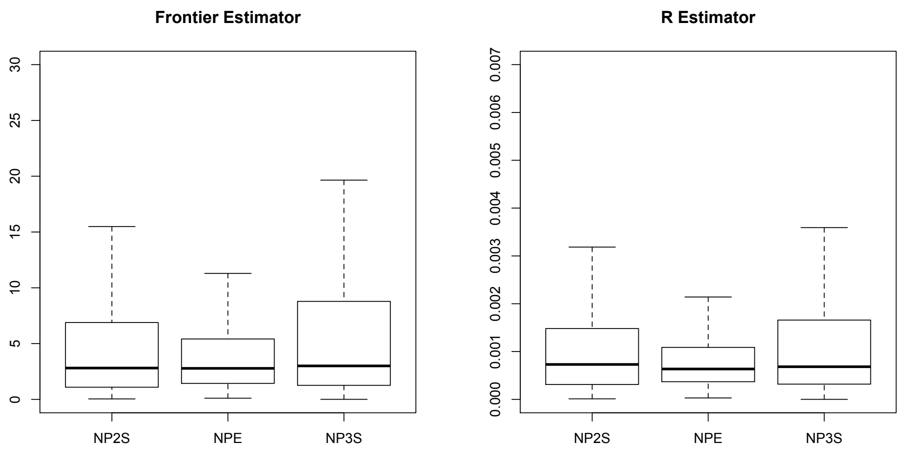


Figura 3.4: Frontier I - Boxplot of Estimators - $n = 200$ - $\mu_r = 0.5$

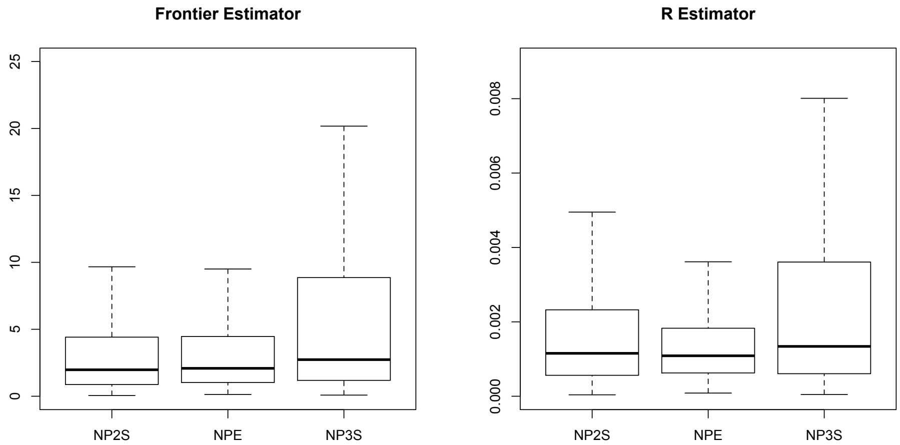



Figura 3.5: Frontier I - Boxplot of Estimators - $n = 400$ - $\mu_r = 0.5$

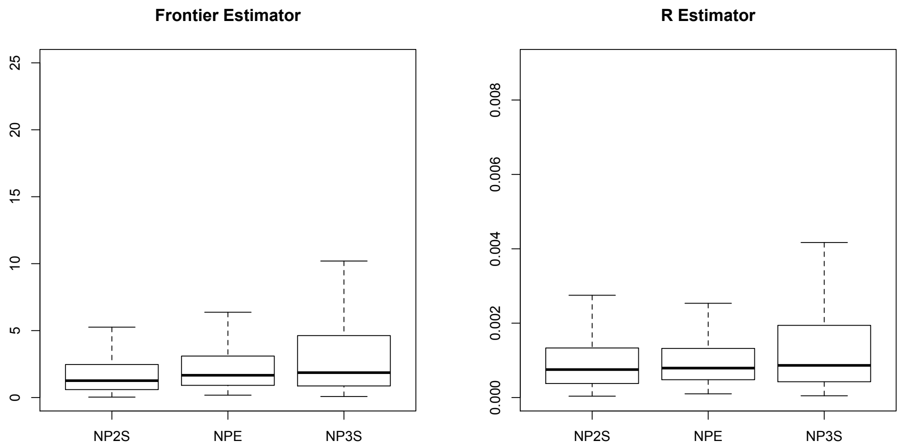


Figura 3.6: Frontier I - Boxplot of Estimators - $n = 200$ - $\mu_r = 0.6$

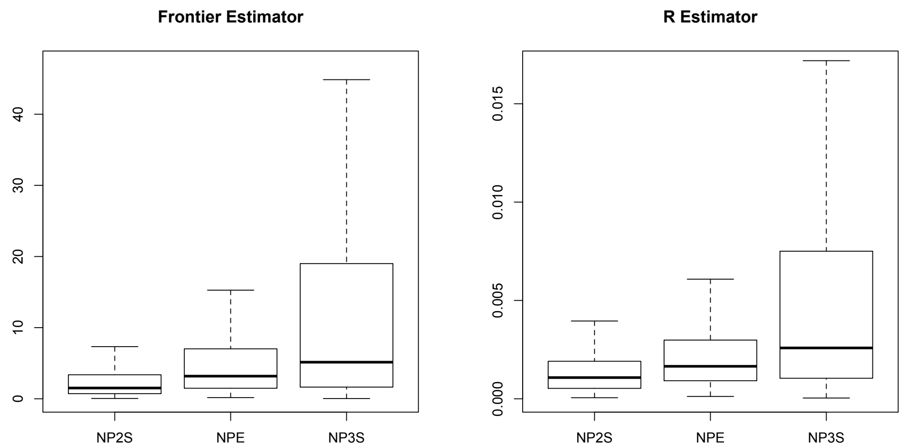



Figura 3.7: Frontier I - Boxplot of Estimators - $n = 400$ - $\mu_r = 0.6$

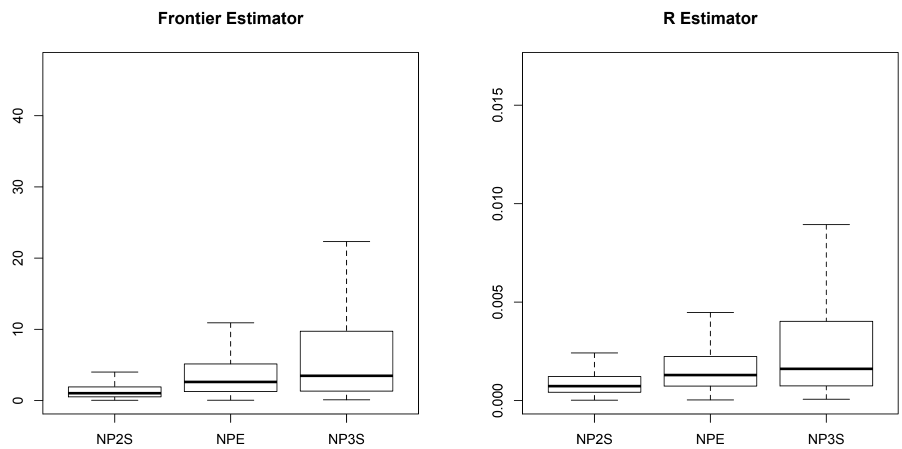


Figura 3.8: Frontier I - Boxplot of Estimators - $n = 200$ - $\mu_r = 0.75$

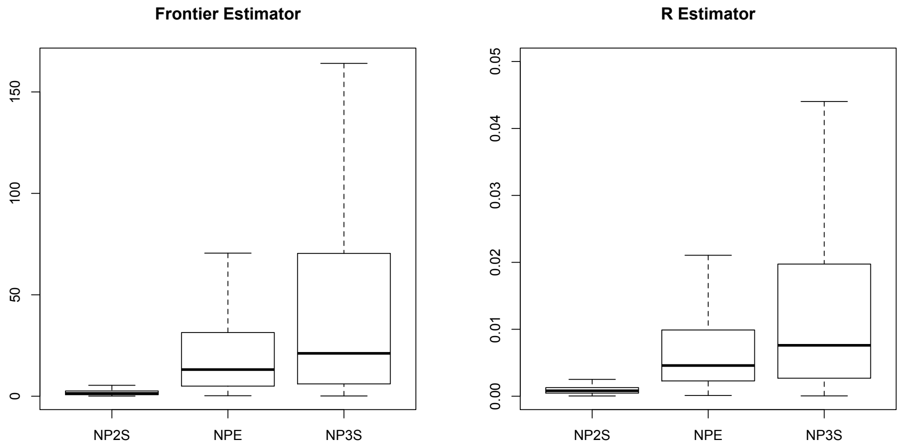



Figura 3.9: Frontier I - Boxplot of Estimators - $n = 400$ - $\mu_r = 0.75$

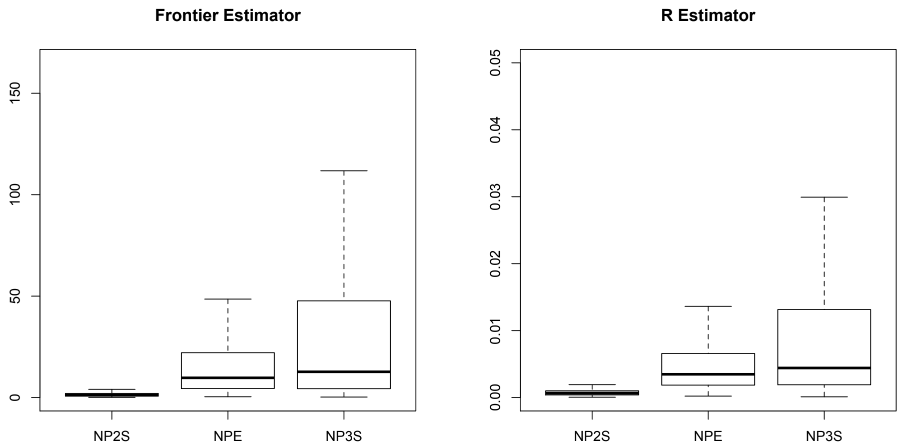


Figura 3.10: Frontier II - Boxplot of Estimators - $n = 200$ - $\mu_r = 0.25$

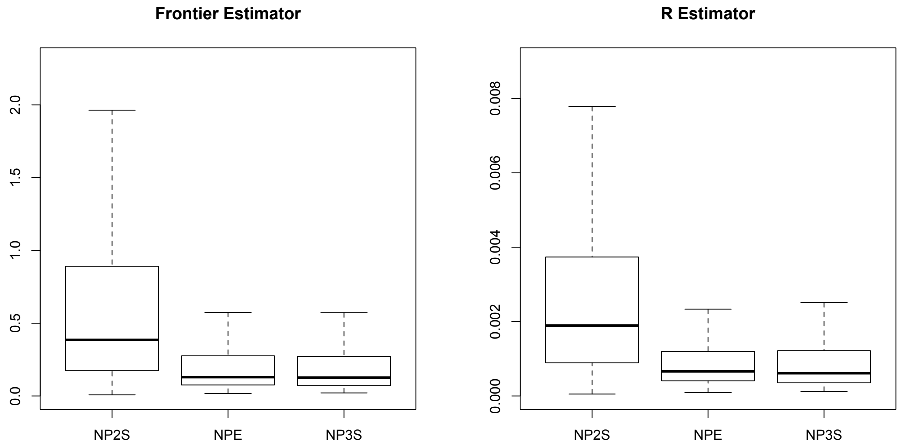



Figura 3.11: Frontier II - Boxplot of Estimators - $n = 400$- $\mu_r = 0.25$

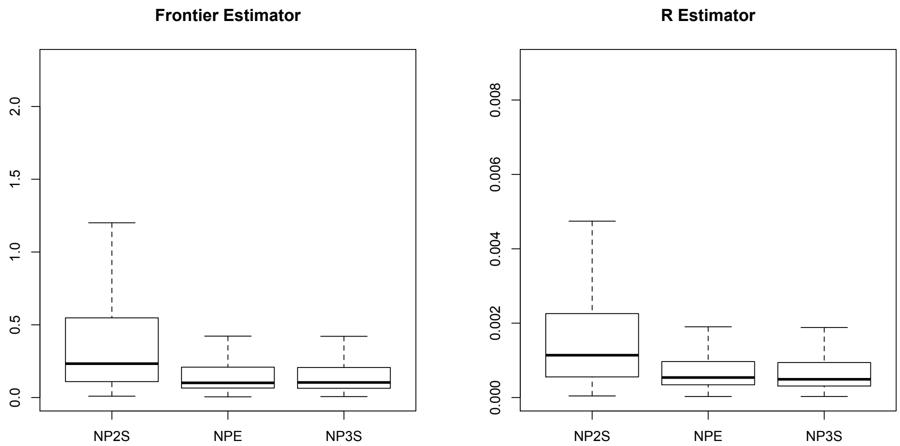


Figura 3.12: Frontier II - Boxplot of Estimators - $n = 200$ - $\mu_r = 0.5$

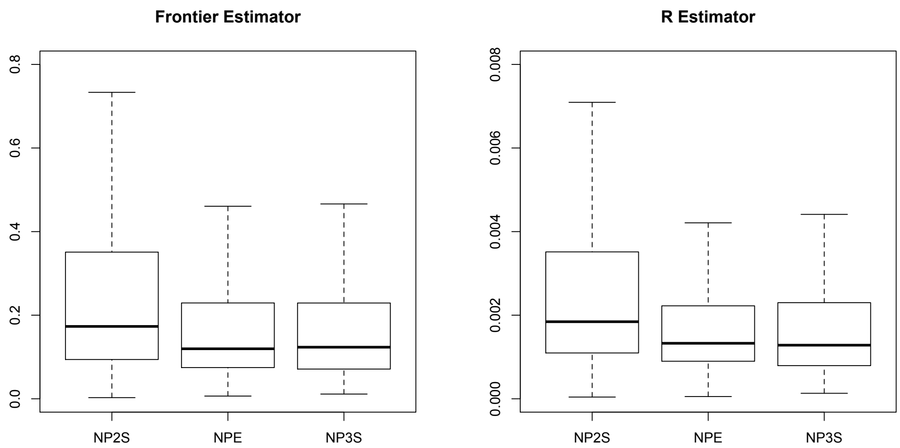



Figura 3.13: Frontier II - Boxplot of Estimators - $n = 400$- $\mu_r = 0.5$

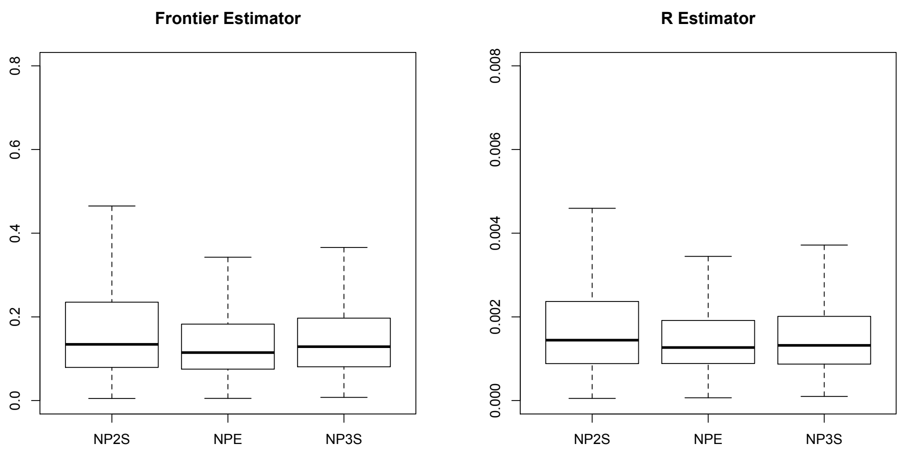


Figura 3.14: Frontier II - Boxplot of Estimators - $n = 200$ - $\mu_r = 0.6$

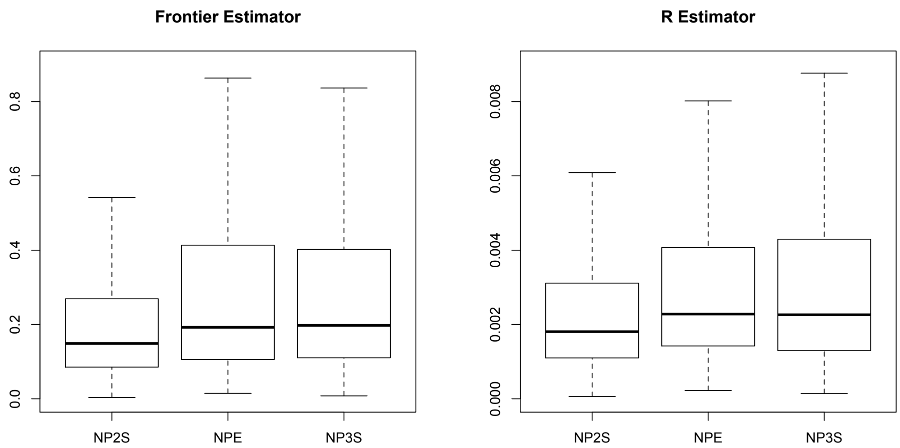



Figura 3.15: Frontier II - Boxplot of Estimators - $n = 400$ - $\mu_r = 0.6$

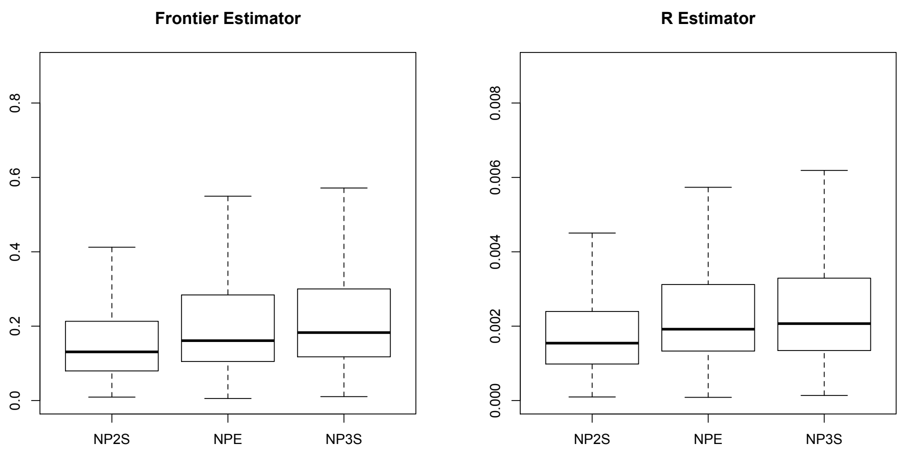


Figura 3.16: Frontier II - Boxplot of Estimators - $n = 200$ - $\mu_r = 0.75$

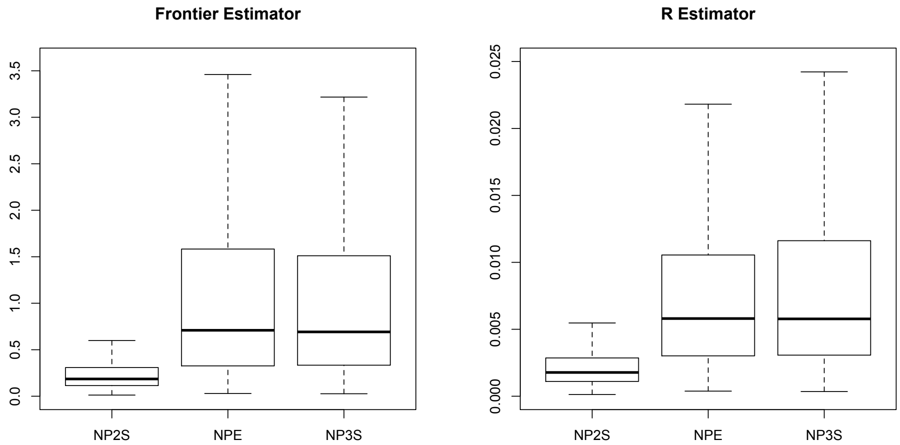



Figura 3.17: Frontier II - Boxplot of Estimators - $n = 400$ - $\mu_r = 0.75$

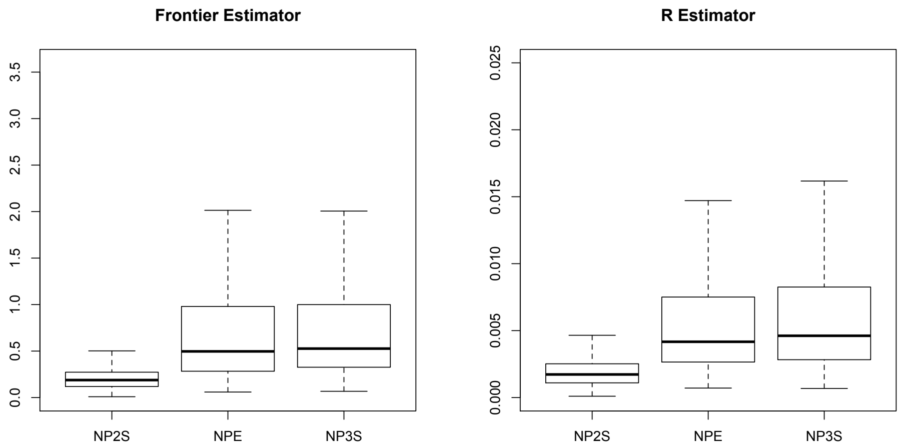


Tabela 3.5: Real Data Example - NP2S estimated efficiency

| Rank | Agency | State | $\hat{R}$ | Rank | Agency | State | $\hat{R}$ |
|---|---|---|---|---|---|---|---|
| 1 | Harford County Sheriff Office | MD | 1.00 | 81 | Ann Arbor Police Dept | MI | 0.41 |
| 2 | Springfield Police Dept MO | MO | 0.99 | 82 | Anaheim Police Dept | CA | 0.41 |
| 3 | Fulton County Police Department | GA | 0.97 | 83 | King County Sheriff Office | WA | 0.41 |
| 4 | Jefferson County Sheriff Department CO | CO | 0.91 | 84 | Alameda County Sheriff Department | CA | 0.40 |
| 5 | El Dorado County Sheriff Department | CA | 0.86 | 85 | Plano Police Dept | TX | 0.40 |
| 6 | San Jose Police Dept | CA | 0.85 | 86 | Jefferson County Sheriff Department AL | AL | 0.40 |
| 7 | Shreveport Police Dept | LA | 0.84 | 87 | Hampton Police Dept | VA | 0.40 |
| 8 | Jefferson County Sheriff Department MO | MO | 0.80 | 88 | Escondido Police Dept | CA | 0.40 |
| 9 | Allen County Sheriff Department | IN | 0.80 | 89 | Collier County Sheriff Department | FL | 0.40 |
| 10 | Oakland County Sheriff Office | MI | 0.80 | 90 | Kitsap County Sheriff Office | WA | 0.40 |
| 11 | Washington Metropolitan Police Dept | DC | 0.78 | 91 | Richland County Sheriff Department | SC | 0.40 |
| 12 | Arapahoe County Sheriff Department | CO | 0.77 | 92 | Fullerton Police Dept | CA | 0.39 |
| 13 | El Paso | CO | 0.76 | 93 | Fontana Police Dept | CA | 0.39 |
| 14 | Onondaga County Sheriff Department | NY | 0.76 | 94 | Charles County Sheriff Office | MD | 0.39 |
| 15 | Fort Bend County Sheriff Department | TX | 0.76 | 95 | Hamilton County Sheriff Department | OH | 0.39 |
| 16 | New Castle County Police Department | DE | 0.76 | 96 | Burbank Police Dept | CA | 0.39 |
| 17 | Naperville Police Dept | IL | 0.75 | 97 | Montgomery County Police Department | MD | 0.38 |
| 18 | Amherst Town Police Dept | NY | 0.75 | 98 | Anne Arundel County Police Department | MD | 0.38 |
| 19 | Santa Barbara County Sheriff Department | CA | 0.74 | 99 | Greenville County Sheriff Office | SC | 0.38 |
| 20 | San Diego County Sheriff Department | CA | 0.70 | 100 | Virginia Beach Police Dept | VA | 0.38 |
| 21 | Buncombe County Sheriff Department | NC | 0.65 | 101 | El Monte Police Dept | CA | 0.38 |
| 22 | Stanislaus County Sheriff Department | CA | 0.65 | 102 | Chesapeake Police Dept | VA | 0.38 |
| 23 | St Petersburg Police Dept | FL | 0.64 | 103 | Topeka Police Dept | KS | 0.38 |
| 24 | Waco Police Dept | TX | 0.64 | 104 | Henrico County Police Dept | VA | 0.37 |
| 25 | Butte County Sheriff Department | CA | 0.62 | 105 | Alexandria Police Dept | VA | 0.37 |
| 26 | Snohomish County Sheriff Office | WA | 0.61 | 106 | Bernalillo County Sheriff Department | NM | 0.37 |
| 27 | Monterey County Sheriff Department | CA | 0.60 | 107 | Vallejo Police Dept | CA | 0.36 |
| 28 | St Charles County Sheriff Department | MO | 0.60 | 108 | Garland Police Dept | TX | 0.36 |
| 29 | Stockton Police Dept | CA | 0.59 | 109 | Fort Collins Police Dept | CO | 0.36 |
| 30 | Irvine Police Police | CA | 0.58 | 110 | Oxnard Police Dept | CA | 0.36 |
| 31 | Arlington County Police Department | VA | 0.58 | 111 | South Bend Police Dept | IN | 0.36 |
| 32 | Montgomery County Sheriff Department | TX | 0.57 | 112 | Aurora Police Dept IL | IL | 0.36 |
| 33 | Tucson Police Dept | AZ | 0.57 | 113 | Clay County Sheriff Department | FL | 0.35 |
| 34 | Charlotte County Sheriff Department | FL | 0.56 | 114 | Spartanburg County Sheriff Office | SC | 0.35 |
| 35 | Prince William County Police Department | VA | 0.56 | 115 | Fresno County Sheriff Department | CA | 0.35 |
| 36 | Sioux Falls Police Dept | SD | 0.55 | 116 | Sacramento County Sheriff Department | CA | 0.35 |
| 37 | Tempe Police Dept | AZ | 0.54 | 117 | Harris County Sheriff Office | TX | 0.35 |
| 38 | Cobb County Police Department | GA | 0.54 | 118 | Hernando County Sheriff Department | FL | 0.35 |
| 39 | Stark County Sheriff Office | OH | 0.53 | 119 | Santa Rosa Police Dept | CA | 0.34 |
| 40 | San Joaquin County Sheriff Department | CA | 0.53 | 120 | Santa Cruz County Sheriff Department | CA | 0.34 |
| 41 | Akron City Police Dept | OH | 0.53 | 121 | Abilene Police Dept | TX | 0.34 |
| 42 | Kent County Sheriff Office | MI | 0.52 | 122 | Pasco County Sheriff Department | FL | 0.33 |
| 43 | Torrance Police Dept | CA | 0.52 | 123 | Brevard County Sheriff Department | FL | 0.33 |
| 44 | Chesterfield County Police Department | VA | 0.51 | 124 | Kern County Sheriff Department | CA | 0.33 |
| 45 | Clark County Sheriff Department | WA | 0.51 | 125 | Oceanside | CA | 0.33 |
| 46 | Orange Police Dept | CA | 0.51 | 126 | Madison City Police Dept | WI | 0.33 |
| 47 | Warren Police Dept | MI | 0.51 | 127 | Inglewood Police Dept | CA | 0.33 |
| 48 | Guilford County Sheriff Office | NC | 0.51 | 128 | Pasadena Police Dept TX | TX | 0.33 |
| 49 | Huntington Beach Police Dept | CA | 0.50 | 129 | Indianapolis Police Dept | IN | 0.32 |
| 50 | Knox County Sheriff Office | TN | 0.50 | 130 | Las Vegas Metropolitan Police Department | NV | 0.32 |
| 51 | Glendale Police Dept CA | CA | 0.49 | 131 | Concord Police Dept | CA | 0.32 |
| 52 | Monroe County Sheriff Office | NY | 0.49 | 132 | Pinellas County Sheriff Department | FL | 0.32 |
| 53 | Okaloosa County Sheriff Department | FL | 0.49 | 133 | Irving Police Dept | TX | 0.32 |
| 54 | Livonia Police Dept | MI | 0.48 | 134 | Pierce County Sheriff Department | WA | 0.31 |
| 55 | St Tammany Parish Sheriff Department | LA | 0.47 | 135 | Long Beach Police Dept | CA | 0.31 |
| 56 | Hidalgo County Sheriff Department | TX | 0.47 | 136 | Evansville Police Dept | IN | 0.31 |
| 57 | Wake County Sheriff Department | NC | 0.46 | 137 | Bakersfield Police Dept | CA | 0.31 |
| 58 | Green Bay Police Dept | WI | 0.46 | 138 | Salinas Police Dept | CA | 0.31 |
| 59 | San Bernardino County Sheriff Department | CA | 0.46 | 139 | Lee County Sheriff Department | FL | 0.30 |
| 60 | Simi Valley Police Dept | CA | 0.46 | 140 | Allentown City Police Dept | PA | 0.30 |
| 61 | Gwinnett County Police Department | GA | 0.45 | 141 | Pomona Police Dept | CA | 0.30 |
| 62 | Contra Costa County Sheriff Department | CA | 0.44 | 142 | Hayward Police Dept | CA | 0.30 |
| 63 | Seattle Police Dept | WA | 0.44 | 143 | Waterbury Police Dept | CT | 0.30 |
| 64 | Forsyth County Sheriff Department | NC | 0.44 | 144 | Lakewood | CO | 0.30 |
| 65 | Downey Police Dept | CA | 0.44 | 145 | Clearwater Police Dept | FL | 0.30 |
| 66 | Howard County Police Department | MD | 0.44 | 146 | Pasadena Police Dept CA | CA | 0.30 |
| 67 | Bexar County Sheriff Office | TX | 0.44 | 147 | Boise Police Dept | ID | 0.30 |
| 68 | Midland Police Dept | TX | 0.44 | 148 | Chula Vista Police Dept | CA | 0.29 |
| 69 | Norwalk | CA | 0.43 | 149 | Cumberland County Sheriff Office | NC | 0.29 |
| 70 | Henderson Police Dept | NV | 0.43 | 150 | Dekalb County Public Safety Department | GA | 0.29 |
| 71 | Riverside County Sheriff Department | CA | 0.43 | 151 | Chandler Police Dept | AZ | 0.29 |
| 72 | Pima County Sheriff Department | AZ | 0.43 | 152 | Savannah Police Dept | GA | 0.29 |
| 73 | Mobile County Sheriff Department | AL | 0.43 | 153 | Salt Lake County Sheriff Office | UT | 0.29 |
| 74 | Erie City Police Dept | PA | 0.43 | 154 | Anchorage Police Dept | AK | 0.28 |
| 75 | Garden Grove Police Dept | CA | 0.42 | 155 | Pueblo Police Dept | CO | 0.28 |
| 76 | San Francisco Police Dept | CA | 0.42 | 156 | Cedar Rapids Police Dept | IA | 0.28 |
| 77 | Lexington County Sheriff Department | SC | 0.42 | 157 | Colorado Springs | CO | 0.28 |
| 78 | Marion County Sheriff Department | FL | 0.41 | 158 | Cambridge Police Dept | MA | 0.28 |
| 79 | Worcester Police Dept | MA | 0.41 | 159 | Clackamas County Sheriff Department | OR | 0.28 |
| 80 | Anderson County Sheriff Department | SC | 0.41 | 160 | Polk County Sheriff Department | FL | 0.27 |

| Rank | Agency (cont.) | State | $\hat{R}$ | Rank | Agency | State | $\hat{R}$ |
|---|---|---|---|---|---|---|---|
| 161 | Riverside Police Dept | CA | 0.27 | 228 | Laredo | TX | 0.20 |
| 162 | Grand Prairie Police Dept | TX | 0.27 | 229 | Montgomery Police Dept | AL | 0.19 |
| 163 | Fort Wayne Police Dept | IN | 0.27 | 230 | Des Moines Police Dept | IA | 0.19 |
| 164 | Winston-Salem Police Dept | NC | 0.27 | 231 | Omaha Police Dept | NE | 0.19 |
| 165 | Charleston County Sheriff Department | SC | 0.27 | 232 | Spokane Police Dept | WA | 0.19 |
| 166 | Reno Police Dept | NV | 0.27 | 233 | Providence Police Dept | RI | 0.19 |
| 167 | Ontario Police Dept | CA | 0.26 | 234 | Volusia County Sheriff Department | FL | 0.19 |
| 168 | Baltimore County Police Department | MD | 0.26 | 235 | Hollywood Police Dept | FL | 0.19 |
| 169 | Escambia County Sheriff Department | FL | 0.26 | 236 | Beaumont Police Dept | TX | 0.19 |
| 170 | Mesquite Police Dept | TX | 0.26 | 237 | Oakland Police Dept | CA | 0.19 |
| 171 | Aurora Police Dept CO | CO | 0.26 | 238 | Milwaukee Police Dept | WI | 0.19 |
| 172 | Gary Police Dept | IN | 0.26 | 239 | Brownsville Police Dept | TX | 0.19 |
| 173 | Denver Police Dept | CO | 0.26 | 240 | Tacoma Police Dept | WA | 0.19 |
| 174 | El Paso Police Dept | TX | 0.25 | 241 | Rochester Police Dept | NY | 0.19 |
| 175 | Hialeah Police Dept | FL | 0.25 | 242 | Buffalo Police Dept | NY | 0.19 |
| 176 | Lansing City Police Dept | MI | 0.25 | 243 | Sarasota County Sheriff Department | FL | 0.19 |
| 177 | Lincoln Police Dept | NE | 0.25 | 244 | Rockford Police Dept | IL | 0.19 |
| 178 | Pittsburgh Bureau Of Police | PA | 0.25 | 245 | San Bernardino Police Dept | CA | 0.19 |
| 179 | Stamford Police Dept | CT | 0.25 | 246 | East Baton Rouge Parish Sheriff Dept | LA | 0.19 |
| 180 | Louisville Police Dept | KY | 0.25 | 247 | Thurston County Sheriff Department | WA | 0.18 |
| 181 | Manatee County Sheriff Department | FL | 0.25 | 248 | Springfield Police Dept IL | IL | 0.18 |
| 182 | Honolulu Police Dept | HI | 0.25 | 249 | Berkeley Police Dept | CA | 0.18 |
| 183 | Austin Police Dept | TX | 0.25 | 250 | Seminole County Sheriff Department | FL | 0.18 |
| 184 | Wichita Falls Police Dept | TX | 0.25 | 251 | Springfield Police Dept MA | MA | 0.18 |
| 185 | Newport News Police Dept | VA | 0.25 | 252 | New Haven Police Dept | CT | 0.18 |
| 186 | Columbus Police Dept GA | GA | 0.25 | 253 | Columbia Police Dept | SC | 0.17 |
| 187 | Sunnyvale Dept Of Public Safety | CA | 0.25 | 254 | Flint City Police Dept | MI | 0.17 |
| 188 | Portsmouth Police Dept | VA | 0.24 | 255 | Minneapolis Police Dept | MN | 0.17 |
| 189 | Jersey City Police Dept | NJ | 0.24 | 256 | Durham Police Dept | NC | 0.17 |
| 190 | Jefferson Parish Sheriff Department | LA | 0.24 | 257 | Fort Lauderdale Police Dept | FL | 0.17 |
| 191 | West Covina Police Dept | CA | 0.24 | 258 | Jacksonville | FL | 0.17 |
| 192 | Prince Georges County Police Department | MD | 0.24 | 259 | Newark Police Dept | NJ | 0.17 |
| 193 | Palm Beach County Sheriff Department | FL | 0.24 | 260 | Nashville-Davidson Metro Police Dept | TN | 0.17 |
| 194 | Hillsborough County Sheriff Department | FL | 0.23 | 261 | Spokane County Sheriff Department | WA | 0.17 |
| 195 | Modesto Police Dept | CA | 0.23 | 262 | Portland Police Dept | OR | 0.17 |
| 196 | Norfolk Police Dept | VA | 0.23 | 263 | Metro-Dade Police Department | FL | 0.17 |
| 197 | Mobile Police Dept | AL | 0.23 | 264 | Richmond (City) Bureau Of Police | VA | 0.17 |
| 198 | St Louis Police Dept | MO | 0.23 | 265 | Charlotte-Mecklenburg Police Department | NC | 0.17 |
| 199 | New Orleans Police Dept | LA | 0.23 | 266 | Baton Rouge Police Dept | LA | 0.17 |
| 200 | Grand Rapids Police Dept | MI | 0.23 | 267 | Memphis Police Dept | TN | 0.16 |
| 201 | Raleigh Police Dept | NC | 0.23 | 268 | Little Rock Police Dept | AR | 0.16 |
| 202 | Cleveland | OH | 0.23 | 269 | Salt Lake City Police Dept | UT | 0.16 |
| 203 | Greensboro Police Dept | NC | 0.22 | 270 | Birmingham Police Dept | AL | 0.16 |
| 204 | Independence Police Dept | MO | 0.22 | 271 | Washington County Sheriff Office | OR | 0.16 |
| 205 | Tulare County Sheriff Department | CA | 0.22 | 272 | Wichita Police Dept | KS | 0.16 |
| 206 | Sacramento Police Dept | CA | 0.22 | 273 | Oklahoma City Police Dept | OK | 0.16 |
| 207 | Glendale Police Dept AZ | AZ | 0.22 | 274 | Peoria Police Dept | IL | 0.15 |
| 208 | Washtenaw County Sheriff Department | MI | 0.22 | 275 | St Paul Police Dept | MN | 0.15 |
| 209 | Bridgeport Police Dept | CT | 0.22 | 276 | Albuquerque Police Dept | NM | 0.15 |
| 210 | Corpus Christi Police Dept | TX | 0.22 | 277 | Hartford Police Dept | CT | 0.15 |
| 211 | Fort Worth Police Dept | TX | 0.22 | 278 | Toledo Police Dept | OH | 0.15 |
| 212 | Cincinnati Police Dept | OH | 0.22 | 279 | Dayton | OH | 0.15 |
| 213 | Elizabeth Police Dept | NJ | 0.21 | 280 | Jackson Police Dept | MS | 0.15 |
| 214 | Salem Police Dept | OR | 0.21 | 281 | Scottsdale Police Dept | AZ | 0.15 |
| 215 | Lafayette Police Dept | LA | 0.21 | 282 | Tampa Police Dept | FL | 0.15 |
| 216 | Sterling Heights Police Dept | MI | 0.20 | 283 | Macon Police Dept | GA | 0.14 |
| 217 | Huntsville Police Dept | AL | 0.20 | 284 | Columbus Police Dept OH | OH | 0.14 |
| 218 | Fresno Police Dept | CA | 0.20 | 285 | Syracuse Police Dept | NY | 0.14 |
| 219 | Lubbock Police Dept | TX | 0.20 | 286 | Travis County Sheriff Department | TX | 0.14 |
| 220 | Mesa Police Dept | AZ | 0.20 | 287 | Miami Police Dept | FL | 0.13 |
| 221 | Sonoma County Sheriff Department | CA | 0.20 | 288 | Tallahassee Police Dept | FL | 0.13 |
| 222 | Amarillo Police Dept | TX | 0.20 | 289 | Kansas City Police Dept | MO | 0.13 |
| 223 | Albany Police Dept | NY | 0.20 | 290 | Orlando Police Dept | FL | 0.13 |
| 224 | Eugene Police Dept | OR | 0.20 | 291 | Baltimore City Police Dept | MD | 0.12 |
| 225 | Knoxville Police Dept | TN | 0.20 | 292 | Atlanta Police Dept | GA | 0.12 |
| 226 | Tulsa Police Dept | OK | 0.20 | 293 | Chattanooga Police Dept | TN | 0.11 |
| 227 | Boston Police Dept | MA | 0.20 | 294 | St Joseph County Sheriff Department | IN | 0.10 |

## 3.8 Appendix 2: Proofs

In this section we prove Theorems 1 and 2. Starting with a lemma, which will be useful in establishing the asymptotic normality and uniform consistency of the estimator $\hat{m}(X_t)$. This result is the local linear non-parametric regression with $V(Y_i|X_i = x) = \sigma^2(x)$ of Fan (1992), a well known result in non-parametric literature, thus we state without proof.

**Lemma 2** *Suppose the assumptions A1-A3 are holding and*

$$s_j(x) = \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)^j,\ j = 0, 1, 2.$$

*Then, if* $nh_n^3 \to \infty$, *we have* $\sup_{x\in G} |s_j(x) - E(s_j(x))| = o_p(h_n^2)$.

**Proof of Lemma 1:** See Martins-Filho and Yao (2007) pg. 303-304.

**Proof of Theorem 1:** Let

$$S_n(x) = (nh_n)^{-1}\begin{pmatrix} \sum_{i=1}^n K\left(\frac{X_i-x}{h_n}\right) & \sum_{i=1}^n K\left(\frac{X_i-x}{h_n}\right)\left(\frac{X_i-x}{h_n}\right) \\ \sum_{i=1}^n K\left(\frac{X_i-x}{h_n}\right)\left(\frac{X_i-x}{h_n}\right) & \sum_{i=1}^n K\left(\frac{X_i-x}{h_n}\right)\left(\frac{X_i-x}{h_n}\right)^2 \end{pmatrix},$$

and

$$S(x) = \begin{pmatrix} f_X(x) & 0 \\ 0 & f_X(x)\sigma_k^2 \end{pmatrix}.$$

Now, $\hat{m}(x, h_n) = \frac{1}{nh_n}\sum_{i=1}^n W_n\left(\frac{X_i-x}{h_n}, x\right)Y_i$, where $W_n(z, x) = (1, 0)S_n^{-1}(x)(1, z)'K(z)$. Note also that $m(x) = \frac{1}{nh_n}\sum_{i=1}^n W_n\left(\frac{X_i-x}{h_n}, x\right)(m(x) + m^{(1)}(x)(X_i - x))$. Therefore,

$$\hat{m}(x, h_n) - m(x) = \frac{1}{nh_n}\sum_{i=1}^n W_n\left(\frac{X_i - x}{h_n}, x\right)Y_i^*,$$

where $Y_i^* = Y_i - m(x) - m^{(1)}(x)(X_i - x)$.

Let $D_n(x) \equiv \hat{m}(x, h_n) - m(x) - \frac{1}{nh_n f_X(x)}\sum_{i=1}^n K\left(\frac{X_i-x}{h_n}\right)Y_i^*$, then

$$\begin{aligned}
|D_n(x)| &= \frac{1}{nh_n}\left|\sum_{i=1}^n\left(W_n\left(\frac{X_i - x}{h_n}, x\right) - \frac{1}{f_X(x)}K\left(\frac{X_i - x}{h_n}\right)\right)Y_i^*\right| \\
&= \frac{1}{nh_n}\left|(1, 0)(S_n^{-1} - S^{-1}(x))\begin{pmatrix} \sum_{i=1}^n K\left(\frac{X_i-x}{h_n}\right)Y_i^* \\ \sum_{i=1}^n K\left(\frac{X_i-x}{h_n}\right)\left(\frac{X_i-x}{h_n}\right)Y_i^* \end{pmatrix}\right| \\
&\le \frac{1}{h_n}((1, 0)(S_n^{-1}(x) - S^{-1}(x))^2(1, 0)')^{1/2}\frac{1}{n}\left(\left|\sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right)Y_i^*\right|\right. \\
&\left. + \left|\sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right)\left(\frac{X_i - x}{h_n}\right)Y_i^*\right|\right),
\end{aligned}$$

where the inequality above follows from the Cauchy-Schwarz inequality and the fact that for a set $a_i$, $i = 1, ..., n$ of positive numbers $\sum_{i=1}^n a_i^2 \leq (\sum_{i=1}^n a_i)^2$.

Our approach to establishing the asymptotic properties of $\hat{m}(x, h_n)$ will be to first establish that the right hand side of the last inequality is $o_p(1)$. Then, by the asymptotic equivalence theorem we conclude that $\hat{m}(x, h_n) - m(x)$ will inherit the asymptotic properties of $\frac{1}{nh_n f_X(x)} \sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right) Y_i^*$. Thus, from Lemma 1 one can show that $h_n^{-1}(S_0(x) - f_X(x)) = O_p(1)$, $h_n^{-1} S_1(x) = O_p(1)$, and $h_n^{-1}(S_2(x) - \sigma_k^2 f_X(x)) = O_p(1)$ uniformly in $G$. Hence, we can conclude that $B_n(x) \equiv h_n^{-1}((1,0)(S_n^{-1}(x) - S^{-1}(x))^2(1,0)')^{1/2} = O_p(1)$ uniformly in $G$.

Given the upperbound $\bar{B}_{f_X}$,

$$
\begin{aligned}
|D_n(x)| \leq \bar{B}_{f_X} B_n(x) h_n \Bigg( \frac{1}{nh_n f_X(x)} \Bigg( \left| \sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right) Y_i^* \right| \\
+ \left| \sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right) \left(\frac{X_i - x}{h_n}\right) Y_i^* \right| \Bigg) \Bigg) \\
= \bar{B}_{f_X} B_n(x) h_n (|c_1(x)| + |c_2(x)|).
\end{aligned} \tag{3.8}
$$

Since $B_n(x) = O_p(1)$ uniformly in $G$, it suffices to investigate the order in probability of $|c_1(x)|$ and $|c_2(x)|$. If we can prove that $c_1(x) = o_p(1)$ and $c_2(x) = o_p(1)$ we can concentrate on

$$
\frac{1}{\sqrt{nh_n} f_X(x)} \left( \left| \sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right) Y_i^* \right| \right)
$$

to obtain the asymptotic distribution of $\sqrt{nh_n}(\hat{m}(x, h_n) - m(x))$. Here we establish a result for $c_1(x)$ noting that the proof for $c_2(x)$ follows a similar argument given assumption A.3.

Now, we write $c_1(x) \equiv I_{1n}(x) + I_{2n}(x)$, where

$$
\begin{aligned}
I_{1n}(x) &= \frac{1}{nh_n f_X(x)} \sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right) (m(X_i) - m(x) - m^{(1)}(x)(X_i - x)), \text{ and} \\
I_{2n}(x) &= \frac{1}{nh_n f_X(x)} \sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right) \sigma(X_i) \xi_i.
\end{aligned}
$$

Now, we analyze each term separately. For $I_{1n}(x)$, by Taylor's Theorem there exists $X_{ib} = x + c_i(X_i - x)$ for some $c_i \in [0, 1]$ such that $I_{1n}(x) = \frac{1}{2nh_n f_X(x)} \sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right) (X_i - x)^2 m^{(2)}(X_{ib})$. Now define $Z_{1i} \equiv \frac{1}{2h_n f_X(x)} K\left(\frac{X_i - x}{h_n}\right) (X_i - x)^2 m^{(2)}(x + c_i(X_i - x))$. Let $i$ being arbitrary we have

$$
\begin{aligned}
E\left(\frac{Z_{1i}}{h_n^2}\right) &= E\left( \frac{1}{2h_n f_X(x)} K\left(\frac{X_i - x}{h_n}\right) \left(\frac{X_i - x}{h_n}\right)^2 m^{(2)}(x + c_i(X_i - x)) \right) \\
&= \frac{1}{2h_n f_X(x)} \int K\left(\frac{\psi - x}{h_n}\right) \left(\frac{\psi - x}{h_n}\right)^2 m^{(2)}(x + c_i(X_i - x)) f_X(\psi) d\psi.
\end{aligned}
$$

Given $i$, $c_i$ is a constant in $[0,1]$. Let $c_i = c_0$ and $\psi - x = \phi h_n$. Changing variables, we have

$$E\left(\frac{Z_{1i}}{h_n^2}\right) = \frac{1}{2f_X(x)} \int \phi^2 K(\phi) m^{(2)}(x + c_0 \phi h_n) f_X(x + \phi h_n) d\phi.$$

Since $k(\cdot)$ has compact support, and given assumptions A.1(2) and A.2(4) we are able to apply Lebesgue's Dominated Convergence Theorem, hence

$$E\left(\frac{Z_{1i}}{h_n^2}\right) = \frac{m^{(2)}(x)}{2} f_X(x) \int \phi^2 K(\phi) d\phi + o(1).$$

By Kolmogorov's LLN,

$$I_{1n}(x) = \frac{h_n^2 m^{(2)}(x) \sigma_k^2 f_X(x)}{2} + h_n^2 o_p(1) \tag{3.9}$$

Given assumptions A.2(4), A.1(2) and A.3(3) we conclude that

$$I_{1n}(x) = h_n^2 O(1) + h_n^2 o_p(1).$$

Hence, $I_{1n}(x) = o_p(1)$.

For $I_{2n}(x)$, let $Z_{2i} = \frac{1}{nh_n} K\left(\frac{X_i - x}{h_n}\right) \sigma(X_i)\epsilon_i$. Given that $E(\epsilon_i|X_i) = 0$ we have

$$E(Z_{2i}) = E\left(\frac{1}{nh_n} K\left(\frac{X_i - x}{h_n}\right) \sigma(X_i) E(\epsilon_i|X_i\right) = 0.$$

Furthermore, $E(\epsilon_i^2|X_i) = 1$, thus

$$\begin{aligned} V(Z_{2i}) = E(Z_{2i}^2) &= \frac{1}{(nh_n)^2} E\left(K^2\left(\frac{X_i - x}{h_n}\right)\sigma^2(X_i)\right) \\ &= \frac{1}{(nh_n)^2} \int K^2\left(\frac{\psi - x}{h_n}\right) \sigma^2(\psi) f_X(\psi) d\psi. \end{aligned}$$

Now, let $S_n^2 = \sum_{i=1}^n E(Z_{2i}^2) = (nh_n)^{-1} \frac{1}{h_n} \int K^2(h_n^{-1}(\psi - x))\sigma^2(\psi) f_X(\psi) d\psi$ and $W_{in} = \frac{Z_{2i}}{S_n}$. By Liapounov's central limit theorem $\sum_{i=1}^n W_{in} \xrightarrow{d} N(0,1)$ provided that

$$\lim_{n\to\infty} E(|W_{in}|^{2+\delta}) = 0$$

for some $\delta > 0$. Note that

$$|W_{in}| = \frac{|K(h_n^{-1}(\psi - x))||\sigma(X_i)||\epsilon_i|}{(nh_n)^{1/2} c_n^{1/2}},$$

where $c_n = h_n^{-1} \int K^2(h_n^{-1}(\psi - x))\sigma^2(\psi) f_X(\psi) d\psi$. Now,

$$|W_{in}|^{2+\delta} = \frac{|K(h_n^{-1}(\psi - x))|^{2+\delta} |\sigma(X_i)|^{2+\delta} |\epsilon_i|^{2+\delta}}{(nh_n)^{1+\delta/2} c_n^{1+\delta/2}}.$$

Since $c_n$ is non-stochastic, it follows that

$$E(|W_{in}|^{2+\delta}) = (nh_nc_n)^{-1-\delta/2}E(|K(h_n^{-1}(\psi - x))|^{2+\delta}|\sigma(X_i)|^{2+\delta}|\epsilon_i|^{2+\delta}),$$

and

$$\sum_{i=1}^{n} E(|W_{in}|^{2+\delta}) = (nh_nc_n)^{-1-\delta/2}\sum_{i=1}^{n} E(|K(h_n^{-1}(\psi - x))|^{2+\delta}|\sigma(X_i)|^{2+\delta}|\epsilon_i|^{2+\delta}).$$

Now, using law of iterated expectations and the assumption that $E(|\epsilon_i|^{2+\delta}|X) < C_1 < \infty$,

$$\begin{aligned} E(|K(h_n^{-1}(\psi - x))|^{2+\delta}|\sigma(X_i)|^{2+\delta}|\epsilon_i|^{2+\delta}) &= E(|K(\cdot)|^{2+\delta}|\sigma(X_i)|^{2+\delta}E(|\epsilon_i|^{2+\delta}|X_i)) \\ &\le C_1 \int \left|K\left(\frac{\psi - x}{h_n}\right)\right|^{2+\delta} |\sigma(\psi)|^{2+\delta}|f_X(\psi)d\psi. \end{aligned}$$

Hence,

$$\sum_{i=1}^{n} E(|W_{in}|^{2+\delta}) \le (nh_nc_n)^{-1-\delta/2}nC_1 \int K\left(\frac{\psi - x}{h_n}\right)^{2+\delta} \sigma(\psi)^{2+\delta}f_X(\psi)d\psi.$$

Let $\psi - x = \phi h_n$ and changing variables,

$$\sum_{i=1}^{n} E(|W_{in}|^{2+\delta}) \le (nh_n)^{-\delta/2}(c_n)^{-1-\delta/2}C_1 \int K(\phi)^{2+\delta}\sigma(x + h_n\phi)^{2+\delta}f_X(x + h_n\phi)d\phi.$$

By assumptions A.1(2), A.2(5) and A.3(3) we are able to apply Lebesgue's Dominated Convergence Theorem. Hence, as $n \to \infty$, $c_n \to \sigma^2(x)f_X(x) \int K^2(\phi)d\phi$ and

$$\int K^{2+\delta}(\phi)f_X(x + h_n\phi)d\phi \to \sigma^{2+\delta}f_X(x) \int K^{2+\delta}(\phi)d\phi.$$

Since $nh_n \to \infty$, we have that

$$\lim_{n\to\infty} \sum_{i=1}^{n} E(|W_{in}|^{2+\delta}) = 0.$$

Therefore,

$$\frac{(nh_n)^{1/2}(nh_n)^{-1}\sum_{i=1}^{n} K(h_n^{-1}(X_i - x))\sigma(X_i)\epsilon_i}{c_n^{1/2}} \xrightarrow{d} N(0, 1);$$

which implies that

$$\sqrt{nh_n}I_{2n}(x) = \frac{1}{\sqrt{nh_n}}\sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\sigma(X_i)\epsilon_i \xrightarrow{d} N\left(0, \sigma^2(x)f_X(x)\int K^2(\phi)d\phi\right). \tag{3.10}$$

Therefore, $I_{2n}(x) = (\sqrt{nh_n})^{-1}O_p(1) = o_p(1)$ since $nh_n \to \infty$.

Now, $I_{1n}(x) = o_p(1)$ and $I_{2n}(x) = o_p(1)$ which imply that $D_n(x) = o_p(1)$. Thus, it suffices to examine $\frac{1}{nh_nf_X(x)}\sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)Y_i^* = I_{1n}(x) + I_{2n}(x)$ to uncover asymptotic behavior of $\hat{m}(x, h_n) - m(x)$. Moreover, $D_n(x) = o_p(1)$ combined with $I_{1n}(x) = o_p(1)$ and $I_{2n}(x) = o_p(1)$ imply that $\hat{m}(x, h_n) - m(x) = o_p(1)$, thus we have consistency of $\hat{m}(x)$.

Now, combining previous results in equations (11) and (12) we have

$$\sqrt{nh_n}(\hat{m}(x,h_n) - m(x) - B_{1n}(x)) \xrightarrow{d} N\Big(0, \frac{\sigma^2(x)}{f_X(x)} \int K^2(\phi)d\phi\Big), \tag{3.11}$$

where $B1n(x) = \frac{h_n^2 m^{(2)}(x)\sigma_k^2}{2} + o_p(h_n^2)$. □

**Proof of Theorem 2:** To prove Theorem 2 we first note that Martins-Filho and Yao (2007) get after two steps $\hat{\sigma}(X_t,h_n)$ which is in fact $\sigma_R\hat{\rho}(X_t,h_n)$. Then, to obtain asymptotic normality of the estimated frontier, $\hat{\rho}(.)$, they divided $\hat{\sigma}(x,h_n)$ by $s_R(g_n)$ and combine their Theorem 1 and their Theorem 2 part $(a)$ to achieve the desired result. In our case, after one step, we get $\hat{m}(X_i,h_n)$ which is in fact $\mu_R\hat{\rho}(X_i,h_n)$. Therefore, to obtain the result claimed in our Theorem 2 part (b), we just need to combine the results from our Theorem 1 and our Theorem 2 part (a).

To prove Theorem 2 part $(a)$, we use the same argument presented in the proof of Theorem 2 part $(a)$ of Martins-Filho and Yao (2007); but substituting in their proof $\sigma(X_t)$ by $m(X_i)$ as well as $\hat{\sigma}(X_t,g_n)$ by $\hat{m}(X_i,g_n)$, and $\sigma_R$ by $\mu_R$ as well as $s_R(g_n)$ by $\hat{\mu}_R(g_n)$.

For a proof of part $(b)$, we note that

$$\sqrt{nh_n}\left(\frac{\hat{m}(x,h_n)}{\mu_R} - \frac{m(x)}{\mu_R} - \frac{B_{1n}}{\mu_R}\right) \equiv \sqrt{nh_n}\left(\frac{\hat{m}(x,h_n)}{\hat{\mu}_R(g_n)} - \frac{m(x)}{\mu_R} - \hat{m}(x,h_n)\left(\frac{1}{\hat{\mu}(g_n)} - \frac{1}{\mu_R}\right) - \frac{B_{1n}}{\mu_R}\right).$$

From Theorem 1 we have

$$\sqrt{nh_n}\left(\frac{\hat{m}(x,h_n)}{\mu_R} - \frac{m(x)}{\mu_R} - \frac{B_{1n}}{\mu_R}\right) \xrightarrow{d} N\left(0, \frac{\sigma^2(x)}{\mu_R^2 f_X(x)} \int K^2(\phi)d\phi\right),$$

and from Theorem 2 part $(a)$, provided that $\frac{ng_n^5}{ln(n)} \to \infty$ we have that $\hat{\mu}_R(x,h_n)(\hat{\mu}_R(g_n)^{-1} - {\mu_R}^{-1}) = O_p(g_n^2)$. Hence, given that $nh_n^5 \to 0$ and $nh_n g_n^4 = O(1)$

$$\sqrt{nh_n}\left(\frac{\hat{m}(x,h_n)}{\hat{\mu}_R(g_n)} - \frac{m(x)}{\mu_R} - B_{2n}\right) \xrightarrow{d} N\left(0, \frac{\sigma^2(x)}{\mu_R^2 f_X(x)} \int K^2(\phi)d\phi\right),$$

where, $B_{2n} = O_p(g_n^2)$. □

## 3.9 References


1. Aigner, D., C.A.K. Lovell and P. Schmidt, Formulation and estimation of stochastic frontiers production function models. **Journal of Econometrics**, 6, 21-37, 1977.

2. Aragon, Y., A. Daouia, C. Thomas-Agnan, Nonparametric frontier estimation: a conditional quantile-based approach. **Econometric Theory**, 21, 358-389, 2005.

3. Daouia, A., Gardes, L. and S. Girard, 2009, Large Sample Approximation of the Distribution for Smoothed Monotone Frontier Estimators . Working paper.

4. Fan, J., Design adaptive nonparametric regression. **Journal of the American Statistical Association**, 87, 998-1004, 1992.

5. Fan, J. and I. Gijbels, Data driven bandwidth selection in local polynomial fitting: variable bandwidth and spatial adaptation. **Journal of the Royal Statistical Society B**, 57, 371-394, 1995.

6. Fan, Y., Q. Li and A. Weersink, Semiparametric estimation of stochastic production frontier models. **Journal of Business and Economic Statistics**, 14, 460-468, 1996.

7. Fan, J., and Q. Yao, Efficient estimation of conditional variance functions in stochastic regression. **Biometrika**, 85, 645-660, 1998.

8. Farrell, M., The measurement of productive efficiency. **Journal of the Royal Statistical Society A**, 120, 253-290, 1957.

9. Gorman, M.F., Ruggiero, J., Evaluating US state police performance using data envelopment analysis. **Int. J. Production Economics**, 113, 1031-1037, 2008.

10. Hall, P., Wolff, R. and Q. Yao, Methods for estimating a conditional distribution function. **Journal of the American Statistical Association**, 94, 154-163, 1999.

11. Kumbhakar, S. C., B. U. Park, L. Simar and E. Tsionas, Nonparametric stochastic frontiers: a local maximum likelihood approach. **Journal of Econometrics**, 137, 1-27, 2007.

12. Martins-Filho, C., Torrent, H., Ziegelmann, F., 2010, Nonparametric Frontier Estimation: Using Local Exponential Regression for Conditional Variance. Submitted paper.

13. Martins-Filho, C. and F. Yao, Nonparametric frontier estimation via local linear regression. Journal of Econometrics, 141, 283-319, 2007.

14. Martins-Filho, C. and F. Yao, A Smooth Nonparametric Conditional Quantile Frontier Estimator. **Journal of Econometrics**, 143, 317-333, 2008.

15. Martins-Filho, C. and F. Yao, 2010, Nonparametric stochastic frontier estimation via profile likelihood. Working paper, University of Colorado, Boulder.

16. Ruppert D., S. Sheather, M. Wand, An effective bandwidth selector for least squares regression. **Journal of the American Statistical Association**, 90, 1257-1270, 1995.

17. Silverman, B.W., **Density estimation for statistics and data analysis**. Chapman and Hall, London, 1986.

18. Simar, L. and P. Wilson, Statistical inference in nonparametric frontier models: recent developments and perspectives, in: H. Fried, C.A.K. Lovell, and S.S. Schmidt, (Eds.), **The Measurement of Productive Efficiency**, 2nd edition. Oxford University Press, Oxford, 2007.

19. Stone, C. J., Consistent nonparametric regression. **Annals of Statistics** 5, 595-620, 1977.

20. Ziegelmann, F., Nonparametric estimation of volatility functions: the local exponential approach. **Econometric Theory**, 18, 985-991, 2002.

# 4 Nonparametric Frontier Estimation: Using Additive Models in a Multivariate input case

Hudson Torrent[1] and Flavio Ziegelmann[2]

**Abstract.** In this paper we propose a semiparametric variation of the frontier model studied by Torrent and Ziegelmann (2010). We rewrite their model allowing for estimating the production frontier and efficiency of production units in a multiple input context without suffering the *curse of dimensionality.* Our approach places their model within the framework of additive models based on assumptions regarding the way inputs combine in production. In particular, we consider the cases of additive and multiplicative inputs, which are widely considered in economic theory and applications. We conduct Monte Carlo simulation examples to compare the finite sample performances of the classical backfitting estimator of Hastie and Tibishirani (1990) and the smooth backfitting estimator of Mammen et al. (1999) and Nielsen and Sperlich (2005) as estimators of the additive model. Furthermore a real data study is carried out, from which we rank efficiency within a sample of USA Law Enforcement agencies using USA crime data.

**Keywords and phrases.** nonparametric frontier models; additive models; semiparametric regression; Classical Backfitting; Smooth Backfitting.

**JEL Classifications.** C14, C22

[1]Department of Economics, Federal University of Rio Grande do Sul, Porto Alegre, RS 90040-000, Brazil, email: hudsontorrent@ppge.ufrgs.br

[2]Department of Statistics, Federal University of Rio Grande do Sul, Porto Alegre, RS 90040-000, Brazil, email: flavioz@ufrgs.br and Department of Statistics, London School of Economics, London, WC2A 2AE UK.

## 4.1 Introduction

There exists a large and growing literature on the specification and estimation of production frontiers (Simar and Wilson, 2007). Let $\Psi = \{(x, y) \in \mathbb{R}_+^{p+1} : x \text{ can produce } y\}$ be a technology where $x \in \mathbb{R}_+^p$ is a vector of inputs used to produce an output $y \in \mathbb{R}_+$. The production frontier associated with $\Psi$ is defined as $\rho(x) = \sup\{y \in \mathbb{R}_+ : (x, y) \in \Psi\}$ for all $x \in \mathbb{R}_+^p$. Given a sample of $n$ realized production plans (or production units) $\chi_n = \{(X_i, Y_i)\}_{i=1}^n$, which share the technology $\Psi$, the main objective of this literature is to estimate $\rho(x)$ for all $x \in \mathbb{R}_+^p$. For any given production plan $(X_i, Y_i) \in \Psi$, we define its (inverse) Farrell efficiency as $0 \leq R_i = \frac{Y_i}{\rho(X_i)} \leq 1$. Once an estimate of $\rho$ is available, estimated efficiencies can be readily.

There exist two main statistical approaches for modeling production frontiers. The deterministic approach is based on the assumption that all observed data lie in $\Psi$, i.e., $P((X_i, Y_i) \in \Psi) = 1$ for all $i$, where $P$ is a probability measure. In these models, any deviation of realized output $Y_i$ from $\rho(X_i)$ is attributable to unobserved inefficiencies of the production plan $i$. On the other hand the stochastic approach allows for random shocks to the production process. As a result observed output $Y_i$ at any input level can be smaller or larger than $\rho(X_i)$. As a result, it may be that $P((X_i, Y_i) \notin \Psi) > 0$ for some $i$. Although more appealing from an econometric perspective, separating inefficiency and random shock in stochastic frontier models requires strong assumptions on the joint density of $(X_i, Y_i)$ (Aigner et al., 1977; Fan et al., 1996; Kumbhakar et al., 2007, Martins-Filho and Yao, 2010). In contrast, deterministic frontier models can be estimated under much milder restrictions on the stochastic process generating $\chi_n$ (Aragon et al., 2005; Martins-Filho and Yao, 2007; Martins-Filho and Yao, 2008; Daouia et al., 2009).

Among the existing statistical models and estimators for deterministic frontiers, a very interesting approach is that of Martins-Filho and Yao (2007). They assume that output $Y_i$ is generated by

$$Y_i = \rho(X_i)R_i = \frac{\sigma(X_i)}{\sigma_R}R_i \text{ for } i = 1, 2, \cdots, n \ , \tag{4.1}$$

where $R_i$ is an unobserved random variable representing efficiency and taking values in the interval $[0, 1]$, $X_i$ is an observed random vector representing inputs taking values in $\mathbb{R}_+^p$, $\sigma(x) : \mathbb{R}_+^p \to (0, \infty)$ is a measurable function, $\sigma_R$ is an unknown parameter and the production frontier is given by $\rho(x) \equiv \frac{\sigma(x)}{\sigma_R}$. In this model $R_i$ has the effect of contracting output from optimal levels that lie on the production frontier. The larger $R_i$ the more efficient the production unit because the closer the realized output is to that on the production frontier. They assume that $E(R_i|X_i = x) \equiv \mu_R$ where $0 < \mu_R < 1$ and

$V(R_i|X_i = x) \equiv \sigma_R^2$. Here, the parameter $\mu_R$ is interpreted as a mean efficiency given input usage and the common technology $\Psi$ and $\sigma_R$ is a scale parameter for the conditional distribution of $R_i$ that also locates the production frontier. Its shape is captured by $\sigma(x)$. These conditional moment restrictions along with equation (4.1) imply that $E(Y_i|X_i = x) = \frac{\mu_R}{\sigma_R}\sigma(x)$ and $V(Y_i|X_i = x) = \sigma^2(x)$. The model can therefore be rewritten as,

$$Y_i = b\sigma(X_i) + \sigma(X_i)\frac{(R_i - \mu_R)}{\sigma_R} = m(X_i) + \sigma(X_i)\epsilon_i \tag{4.2}$$

where $b = \frac{\mu_R}{\sigma_R}$, $\epsilon_i = \frac{R_i - \mu_R}{\sigma_R}$, $m(X_i) = b\sigma(X_i)$, $E(\epsilon_i|X_i = x) = 0$ and $V(\epsilon_i|X_i = x) = 1$.

Martins-Filho and Yao propose an estimation procedure that consists of three stages: i) $m(x)$ is estimated using the local linear estimator of Fan (1992); ii) squared residuals from the first stage are used in a local linear procedure to estimate the conditional variance $\sigma^2(x)$; iii) the estimated conditional variance from stage 2 is used to estimate $\sigma_R$ based on a simple anchoring assumption. Another estimation procedure was proposed for that model in Martins-Filho, Torrent and Ziegelmann (2010). They propose to use the local exponential Kernel estimator of Ziegelmann (2002) applied to estimate the conditional volatility function in stage two, ensuring its nonnegativity. This feature is important since we are interested in the square root of estimated conditional variance.

Both estimators above are fairly easy to implement as it involves standard nonparametric procedures. In addition, those frontier estimators have a number of desirable characteristics: first, contrary to the estimators in Aragon et al. (2005), Daouia et al. (2009) and Martins-Filho and Yao (2008), they are a smooth function of input usage; second, although the frontier estimator envelops the data, they are not intrinsically biased as the popular DEA (data envelopment analysis) and FDH (free disposal hull) estimators, therefore no bias correction is needed; third, both estimators are fairly robust to outliers and extreme values. In addition to all of these desirable properties, those estimation procedures lead to frontier estimators that are consistent and asymptotically normal when suitably centered and normalized.

An undesirable characteristic of Martins-Filho and Yao (2007) and Martins-Filho, Torrent and Ziegelmann (2010) estimators is that they are characterized by an estimation procedure in three steps. Furthermore, the second step of both estimators is a regression that has as regressands squared residuals from stage 1. This feature is sometimes undesirable, specially for practitioners working with relatively small sample size. Torrent and Ziegelmann (2010) overcome this problem getting the frontier in a simpler estimation procedure. Their contribution lies in perceiving that frontier shape can be get from the first step, and thus an additional step to locate the estimated frontier is implemented. In particular, they notice that in eq. (4.2) we have $m(X_i) \equiv \mu_R \rho(X_i)$. Therefore, estimating $m(X_i)$ gives $\hat{m}(x) = \mu_R \hat{\rho}(x)$, since $\mu_R$ does not depend on $X_i$. Then, if an estimator for $\mu_R$ is available the frontier can be estimated as $\hat{\rho}(X_i) = \frac{\hat{m}(X_i)}{\hat{\mu}_R}$. Torrent and Ziegelmann (2010) propose an estimation procedure that consists of two stages: first, $m(x)$ is estimated using the local linear estimator of Fan (1992); second, the estimated conditional mean from stage 1 is used to estimate $\mu_R$ based on a simple anchoring assumption. Hence, besides the simpler fashion of their estimator, it maintains the advantages over FDH and DEA listed above.

All aforementioned estimators are established considering a multivariate input case, i.e., $x \in \mathbb{R}^p$. However, due to its nonparametric nature those estimators suffer from the well known *curse of dimensionality*. This poor performance of fully nonparametric models in a multivariate setup limits their applicability, since in many practical situations researchers may be interested in estimating efficiency considering more than one single input. One strategy to overcome this problem is to use additive models. The rationale for using these models is to reduce the dimension of the nonparametric component in order to mitigate the curse of dimensionality. In frontier estimation context, the estimator proposed by Torrent and Ziegelmann (2010) pursue a convenient structure for applicability of additive models, since assuming that $m(X_i)$ is additive is equivalent to assume that $\rho(X_i)$ is additive. We note that the implementation of additive models is not so simple if one considers the estimators proposed by Martins-Filho and Yao (2007) or Martins-Filho et. al. (2010). In those cases additive models would be based on the assumption that the conditional variance function ($\sigma^2(X_i)$) is additive. As a consequence, it is not clear which structure is being assumed for $\rho(X_i)$. Therefore, in our paper we present a semiparametric version of the deterministic frontier model proposed by Torrent and Ziegelmann (2010). This allows for additive models to address the problem of estimating production frontiers with multiples inputs, relying on the assumption that the production frontier is additive. Furthermore, we present a model for estimating a multiplicative frontier in the framework of additive models. This last extension is particularly important since economic theory mostly considers the case of multiplicative inputs.

Besides this introduction, our paper has four more sections. Section 2 presents the deterministic frontier models under the assumptions of additive frontier and multiplicative frontier. It also describes how to estimate those models within the framework of additive models. Section 3 contains a Monte Carlo simulation that sheds some light on the finite sample properties of the proposed estimators. In Section 4 a real data example is analyzed, and Section 5 provides a summary and a conclusion.

## 4.2 Model and Estimation

In this section we characterize our model and estimator. We focus here on the estimation of a possibly high dimensional regression model $E(Y|X_i = x) = m(x), Y \in \mathbb{R}_+,\ X \in \mathbb{R}_+^p$, given the random sample $\chi_n = \{(X_{1i}, ..., X_{pi}, Y_i)\}_{i=1}^n$. As we noted before, considering fully nonparametric estimation procedures leads to problems related to the curse of dimensionality. Our strategy to overcome this issue is to use additive models. That is, the aim is to model $m(\cdot)$ additively, i.e.,

$$m(x) = m_0 + \sum_{j=1}^{p} m_j(x_j), \tag{4.3}$$

where $m_0 = E(Y)$, and $E(m_j(X_j)) = 0$ for all $j = 1, \cdots, p$ for identification. In order to rewrite the frontier model in the framework of additive models we have to impose some structure on the frontier function. In this paper we consider two cases: the additive frontier and the multiplicative frontier functions. It is important to emphasize that assumptions of additive frontier as well as multiplicative frontier are supported by Economic theory.

### 4.2.1 Additive Frontier

As established in equation (4.2),

$$Y_i = m(X_i) + \sigma(X_i)\epsilon_i \ ,$$

where $\epsilon_i = \frac{(R_i - \mu_R)}{\sigma_R}$ and $m(X_i) = \frac{\mu_R}{\sigma_R}\sigma(X_i)$.

First, we note that $m(X_i) \equiv \mu_R \rho(X_i)$. Therefore, estimating $m(X_i)$ gives to us $\hat{m}(x) = \mu_R \hat{\rho}(x)$, since $\mu_R$ does not depend on $X_i$. We thus get from $\hat{m}(x)$ an estimate of $\rho(x)$, but in a wrong position (only up to the scale). Then, if we have an estimator for $\mu_R$ we can propose to estimate the frontier as $\hat{\rho}(X_i) = \frac{\hat{m}(X_i)}{\hat{\mu}_R}$. Therefore, we propose an estimation procedure in two stages. Firstly, regarding the estimation of $m(\cdot)$, if we assume that $\rho(\cdot)$ is additive it follows that $m(\cdot)$ is additive as well, i.e.,

$$m(x) = m_0 + \sum_{j=1}^{p} m_j(x_j) \ ,$$

where $m_0 = E(Y)$. Hence, we can estimate the conditional mean in the framework of additive models. In subsection 4.2.3 we discuss the estimation approach for this first stage. In the second stage, we estimate $\mu_R$ in the same way that is proposed in Torrent and Ziegelmann (2010). That is, we assume that there exists at least *one* observed production unit that is efficient, i.e., there is at least one $R_i$ identically equal to one. Thus, the proposed estimator for $\mu_R$ is

$$\hat{\mu}_r = \left( \max_{1 \leqslant i \leqslant n} \frac{Y_i}{\hat{m}(X_i; h_n)} \right)^{-1} .$$

To understand the idea behind the estimator proposed above one should note that $Y_i = \rho(X_i)R_i = \frac{m(X_i)}{\mu_R}R_i$, and then set one firm to be efficient. Therefore, after these two stages, the proposed estimator for the frontier at $x \in \mathbb{R}^p$ is given by $\hat{\rho}(X_i) = \frac{\hat{m}(X_i, h_n)}{\hat{\mu}_R}$.

### 4.2.2 Multiplicative Frontier

The model is based on the following multiplicative regression (eq. (4.1)):

$$Y_i = \rho(X_i)R_i \ .$$

Since $Y_i$, $\rho(X_i)$ and $R_i$ are nonnegative (and assuming that $R_i \in (0, 1]$) for $i = 1, ..., n$, we are able to write

$$\log Y_i = \log \rho(X_i) + \log R_i \ . \tag{4.4}$$

Similar to the previous case, two important conditional moment restrictions on $\log R_i$ must be assumed: $E(\log R_i|X_i = x) \equiv m_R$, where $m_R \in (-\infty, 0)$, and $V(\log R_i|X_i = x) \equiv \gamma_R^2$, where $\gamma_R^2 \in (0, \infty)$. These conditional moment restrictions together with eq. (4.4) imply that $E(\log Y_i|X_i = x) = \log \rho(x) + m_R$ and $V(\log Y_i|X_i = x) = \gamma_R^2$. The model can therefore be rewritten as

$$\log Y_i = (m_R + \log \rho(X_i)) + (\log R_i - m_R) = g(X_i) + \epsilon_i \ , \tag{4.5}$$

where $g(X_i) = m_R + \log \rho(X_i)$, $\epsilon_i = \log R_i - m_R$, $E(\epsilon_i|X_i = x) = 0$ and $V(\epsilon_i|X_i = x) = \gamma_R^2$.

Now, if we assume that the production frontier, $\rho(\cdot)$, is multiplicative, it follows that $g(\cdot)$ is additive,

$$g(x) = \alpha + \sum_{j=1}^{p} g_j(x_j) \tag{4.6}$$

where $\alpha = m_R + \log \rho_0 = E(\log(Y))$ and $g_j(x_j) = \log \rho_j(x_j)$. Hence, we are again in the framework of additive models. If we were able to identify $m_R$ and $\log \rho_0$ in the estimation of the above model, the estimated frontier would be $\hat{\rho}(X_i) = \exp(\hat{g}_0 + \sum_{j=1}^{p} \hat{g}_j(X_{ji}))$ where $g_0 = \log \rho_0$. Unfortunately we are not able to identify either $m_R$ or $g_0$. Hence we still need another step to get an estimate for the frontier function. As before, we get an estimate for frontier in the first step but in a wrong position. In particular the frontier is positioned below the true value since $m_R$ is negative. To circumvent it we once more estimate the frontier in two stages. First, we estimate eq. (4.6) using some estimation technique for additive models as we discuss in subsection 4.2.3. In the second stage, we estimate $m_R$ based on the assumption that there exists *one* efficient firm ($\log R_i \equiv 0$ for some $i = 1, ..., n$). Therefore, the proposed estimator of $m_R$ is

$$\hat{m}_R = - \max_{1 \leq i \leq n} \bigg( \log Y_i - \hat{g}(X_i) \bigg) \tag{4.7}$$

leading to a frontier estimator given by

$$\hat{\rho}(X_i) = \exp \bigg( \hat{g}(X_i) - \hat{m}_R \bigg) \tag{4.8}$$

### 4.2.3 Estimation of Additive Models

We now discuss how to estimate the additive model (4.3). Four estimators emerge as viable alternatives to estimate that model: the Classical Backfitting estimator (CBE), proposed by Buja et al. (1989) and Hastie and Tibishirani (1990); the Marginal Integration estimator (MIE), proposed by Newey (1994), Tjøstheim and Auestad (1994) and Linton and Nielsen (1995); a two-stage estimator (2SE), proposed by Linton (1997) and Kim et al. (1999); and the Smooth Backfitting estimator (SBE), proposed by Mammen et al. (1999).

The CBE is the most used estimator to estimate model (4.3). It has proved to be very useful in applications with real data and simulation studies. However, due to its iterative nature, it is difficult to analyze its statistical properties. Nevertheless, Opsomer and Ruppert (1997) and Opsomer (2000) addressed the algorithmic and statistical properties of CBE. They show, using local polynomial smoothers, that the asymptotic variance is equal to the oracle bound, but the bias is not oracle, except when the covariates are mutually independent. Furthermore, to establish those results they assume that the amount of dependency in the covariates is strictly limited. Compared with CBE, the MIE is computationally more expensive, but it reaches the oracle efficiency bounds. However, the MIE becomes less efficient as the correlation among regressors increases. The 2SE proposed by Kim et al. (1999) reduces asymptotic variance by combining the MIE with a one-step backfitting. More recently, Mammen et al. (1999) propose a smooth backfitting procedure that is based on an interpretation of both Nadaraya-Watson estimator and local linear estimator as projections in an appropriate Hilbert space, which is first provided by Mammen et al. (2001). The SBE does not have the drawbacks of CBE, MIE or 2SE. It achieves the same bias and variance as the oracle estimator. In addition, to establish the asymptotic result for SBE, no restriction on the amount of dependency between the covariates is imposed in any way.

Although SBE pursue more interesting asymptotic properties when compared to its competitors, it is not obvious that it performs better in finite sample estimation. With this in mind we decided to consider CBE and SBE in a Monte Carlo study that we conduct in next section. We choose CBE due to its well known usefulness and success in simulations and applications with real data, and SBE due to its theoretical advantages and hence an expectation regarding its performance in finite sample. Now we briefly define both estimators and its algorithms.

**Classical Backfitting estimator - CBE**

The CBE is motivated by solving an empirical version of the following set of equations

$$m_j(x_j) = E(Y|X_j = x_j) - m_0 - \sum_{k \neq j}^{p} E(m_k(X_k)|X_j = x_j), \text{ for } j = 1, \cdots, p. \tag{4.9}$$

An empirical version of the above set of equations is solved via an iterative procedure. Let $\hat{m}_j^{[0]}(x_j)$ be some initial estimator for $m_j(x_j)$, for instance, $\hat{m}_j^{[0]}(x_j) = 0$, also set $\hat{m}_0 = n^{-1} \sum_{i=1}^{n} Y_i$. Thus, the iterative procedure is such that for $j = 1, \cdots, p$ and for $l = 1, 2, \cdots$, compute the $l^{th}$ step $\hat{m}_j^{[l]}(x_j)$ by

$$\hat{m}_j^{[l]}(x_j) = \hat{E}(Y|X_j = x_j) - \hat{m}_0 - \sum_{k=1}^{j-1} \hat{E}(\hat{m}_k^{[l]}(X_k)|X_j = x_j) - \sum_{k=j+1}^{p} \hat{E}(\hat{m}_k^{[l-1]}(X_k)|X_j = x_j),$$

where, for a random variable $A$, $\hat{E}(A|X_j = x_j)$ is the univariate kernel estimator of $E(A|X_j = x_j)$. $\hat{E}(A|X_j = x_j)$ may be the Nadaraya-Watson estimator or the local linear estimator. We consider both cases in the Monte Carlo study presented in the next section.

Iteration ends when a pre-specified convergence criterion is reached. We consider the following criterion in our simulations,

$$\frac{\sum_{i=1}^{n} [\hat{m}_j^{[l]}(x_{ij}) - \hat{m}_j^{[l-1]}(x_{ij})]^2}{\sum_{i=1}^{n} [\hat{m}_j^{[l-1]}(x_{ij})]^2 + 0.0001} < 0.0001, \text{ for all } j = 1, \cdots, p. \tag{4.10}$$

**Smooth Backfitting estimator - SBE**

Here we present briefly the Nadaraya-Watson smooth backfitting estimator in its simpler version as established in Nielsen and Sperlich (2005). Consider first the Nadaraya-Watson regression smoother with product kernels

$$\tilde{m}(x) = \frac{\sum_{i=1}^{n} \prod_{j=1}^{p} K_h(X_{ij} - x_j) Y_i}{\sum_{i=1}^{n} \prod_{j=1}^{p} K_h(X_{ij} - x_j)}$$

and the multidimensional kernel density estimate

$$\tilde{f}(x) = n^{-1} \sum_{i=1}^{n} \prod_{j=1}^{p} K_h(X_{ij} - x_j)$$

where $K_h(u) = (1/h)K(u/h)$. $\tilde{m}(\cdot)$. Then SBEs ($\hat{m}_j(\cdot)$, $j = 0, \cdots, p$) are defined as the minimizers of the criterion

$$\int (\tilde{m}(x) - m_0 - m_1(x_1) - \cdots - m_p(x_p))^2 \tilde{f}(x)dx. \tag{4.11}$$

The solution to the minimization (4.11) is characterized by the following system of equations ($j = 1, \cdots, p$):

$$\hat{m}_j(x_j) = \int \tilde{m}(x)\frac{\tilde{f}(x)}{\tilde{f}_j(x_j)}dx_{\underline{j}} - \sum_{k\neq j}^{p} \int \hat{m}(x_k)\frac{\tilde{f}(x)}{\tilde{f}_j(x_j)}dx_{\underline{j}} - \bar{Y} \tag{4.12}$$

$$0 = \int \hat{m}_j(x_j)\tilde{f}_j(x_j)dx_j \tag{4.13}$$

where $\tilde{f}_j(x_j) = \int \tilde{f}(x)dx_{\underline{j}}$ is the marginal of the density estimate $\tilde{f}(x)$ and $\bar{Y} = n^{-1}\sum_{i=1}^n Y_i$. Furthermore, note that

$$\int \tilde{m}(x)\frac{\tilde{f}(x)}{\tilde{f}_j(x_j)}dx_{\underline{j}} = \frac{n^{-1}\sum_{i=1}^n K_h(X_{ij} - x_j)Y_i}{\tilde{f}_j(x_j)} \equiv \tilde{m}_j(x_j), \tag{4.14}$$

i.e. $\tilde{m}_j(x_j)$ is the univariate Nadaraya-Watson estimator. Furthermore, $\hat{m}_0 = n^{-1}\sum_{i=1}^n Y_i = \bar{Y}$. All in all we can rewrite equation (4.12) as

$$\hat{m}_j(x_j) = \tilde{m}_j(x_j) - \sum_{k\neq j} \int \hat{m}_k(x_k)\frac{\tilde{f}_{jk}(x_j, x_k)}{\tilde{f}_k(x_j)}dx_k - \bar{Y} \tag{4.15}$$

where $\tilde{f}_{jk}(x_j, x_k) = n^{-1}\sum_{i=1}^n K_h(X_{ij} - x_j)K_h(X_{ik} - x_k)$ is the two-dimensional marginals of the full dimensional kernel density estimate $\tilde{f}(x)$. Equations (4.12), (4.14) and (4.15) lead to the following iterative estimation procedure (for all $j = 1, \cdots, p$),

$$\hat{m}_j^{[l+1]}(x_j) = \tilde{m}_j(x_j) - \sum_{k=1}^{j-1} \int \hat{m}_k^{[l+1]}(x_k)\frac{\tilde{f}_{j,k}(x_j, x_k)}{\tilde{f}_j(x_j)}dx_k - \sum_{k=j+1}^{p} \int \hat{m}_k^{[l]}(x_k)\frac{\tilde{f}_{j,k}(x_j, x_k)}{\tilde{f}_j(x_j)}dx_k, \tag{4.16}$$

where $l = 1, 2, \cdots$ denotes the number of iterations. Iteration ends when a pre-specified convergence criterion is reached. We consider the following criterion in our simulations,

$$\frac{\sum_{i=1}^n [\hat{m}_j^{[l]}(x_{ij}) - \hat{m}_j^{[l-1]}(x_{ij})]^2}{\sum_{i=1}^n [\hat{m}_j^{[l-1]}(x_{ij})]^2 + 0.0001} < 0.0001, \text{ for all } j = 1, \cdots, p. \tag{4.17}$$

## 4.3 Monte Carlo Study

In this section we investigate some of the finite sample properties of our estimator via a Monte Carlo study. We consider two formulations of CBE and two formulations of SBE for estimating the conditional mean, i.e., stage 1 of our frontier estimator. For CBE, we consider the classical backfitting Nadaraya-Watson estimator (CBE NW) and the classical backfitting local linear estimator (CBE LL) to solve the system (4.9), as defined in Hastie and Tibishirani (1990). Regarding SBE, we consider the smooth backfitting Nadaraya-Watson estimator as stated in Nielsen and Sperlich (2005), referred to as SBE NW. This means that $\tilde{m}_j(\cdot)$ in (4.16) is estimated by the univariate Nadaraya-Watson estimator, for all $j = 1, \cdots, p$. Furthermore, we consider a naive extension of SBE NW, referred to as SBE NLL, consisting of estimating $\tilde{m}_j(\cdot)$ in (4.16) by the univariate local linear estimator, for all $j = 1, \cdots, p$.[3] The bandwidth for both estimators are selected by a cross-validation procedure as suggested in Nielsen and Sperlich (2005). Our simulations are based on model (1), i.e., $Y_i = \frac{\sigma(X_i)R_i}{\sigma_R}$, with $p = 2$. We consider three models and two specifications for $\sigma(\cdot)$. In model I we consider additive frontier with independent regressors. Model II is characterized by additive frontier and correlated regressors and model III has multiplicative frontier with independent regressors. We generate data with the following characteristics. For models I and III, the $X_{ji},\ j = 1, 2$ are pseudorandom variables from a uniform distribution with support given by $[1, 2]$. $R_i = \exp(-Z_i)$, where $Z_i$ are pseudorandom variables from an exponential distribution with parameter $\beta > 0$, therefore $R_i$ has support on $(0, 1]$. Regarding model II, we firstly consider $S_{ji},\ j = 1, 2$ which are pseudorandom variables from a uniform distribution with support given by $[1, 2]$, then we construct $X_{1i} = S_{1i}$ and $X_{2i} = 0.3S_{1i} + 0.7S_{2i}$ resulting in a correlation between $X_{1i}$ and $X_{2i}$ approximately equal to 0.4. As before, $R_i = \exp(-Z_i)$, where $Z_i$ are pseudorandom variables from an exponential distribution with parameter $\beta > 0$, therefore $R_i$ has support on $(0, 1]$. The specifications for $\sigma(x)$ we consider are:

$$\sigma_1(x) = x_1^{0.5} + x_2^{0.4}, \text{ and}$$
$$\sigma_2(x) = x_1^{0.5} x_2^{0.4}.$$

Two parameters for the exponential distribution are considered: $\beta_1 = 3$ and $\beta_2 = 1/3$. These choices of parameters produce, respectively, the following values for the parameters of $g_{R|X} : (\mu_R, \sigma_R^2) = (0.25, 0.08)$[4] and $(0.75, 0.04)$. Two sample sizes $n = 200, 400$ were used.

We evaluate the overall performance of the efficiency estimator based on three different measures. First, we consider the correlation between the efficiency rankings produced by the estimator and the true

[3] Although intuitive, we note that this is not the correct extension of the SBE NW to the smooth backfitting local linear estimator. For the correct extension see Nielsen and Sperlich (2005) and Mammen et al. (1999).

[4] We truncate the distribution for $\beta_1 = 3$ so that $R_i \geq 0.005$, $i = 1, \cdots, n$ in order to avoid estimation problems related to small values of $\log(R_i)$.

efficiency rankings:

$$R_{rank} = \frac{cov(rank(\hat{R}_i), rank(R_i))}{\sqrt{var(rank(\hat{R}_i))var(rank(R_i))}}$$

where $rank(R_i)$ gives the ranking index according to the magnitude of $R_i$. The closer $R_{rank}$ for $\hat{R}_i$ is to 1, the higher the correlation between the true $R_i$ and $\hat{R}_i$, thus the better the estimator $\hat{R}_i$. The second measure we consider is $R_{mag} = \frac{1}{n}\sum_{i=1}^{n}(\hat{R}_i - R_i)^2$ which is simply the squared Euclidean distance between the estimated vector of efficiencies and the true vector of efficiencies. The third measure we use is $R_{rel} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{R}_i}{\hat{R}_t} - \frac{R_i}{R_t}\right|$, where $t$ is the position index for $R_t = \max_{1\leq i\leq n} R_i$, and $\hat{R}_t$ is the $t$th corresponding element in $\{\hat{R}_i\}_{i=1}^{n}$, which may not be the maximum of $\hat{R}_i$. Hence $R_{rank}$, $R_{mag}$ summarize the performance of the estimator $\hat{R}_i$ in rankings and calculating the magnitude of efficiency. $R_{rel}$ captures the relative efficiency.

The results of our simulations are summarized in tables 4.1-4.3 and figures 4.1-4.12. Tables 4.1-4.3 provide the overall performance of the efficiency estimators according to the measures described above. Figures 4.1-4.12 give boxplots of MSE for the frontier estimator ($\hat{\rho}(\cdot)$) and efficiency estimator ($\hat{R}_i$). Each boxplot is constructed from 500 points (repetitions), where each point corresponds to a sample draw and is calculated as the squared Euclidean distance between the estimate and true value of $\rho(\cdot)$ and $R_i$.

**General regularities**

As expected, as the sample size $n$ increases, the boxplots show that MSE decreases for the vast majority of simulations for all estimators and values for $\mu_R$ considered. The only exception occurs in the case of SBE NW for $\mu_R = 0.75$ in the model with correlated inputs (see Figs. 4.7 and 4.8). Moreover, regarding the measures of overall performance for efficiency estimators mentioned above, all estimators perform better as $n$ increases. We now turn to the impact of different values of $\mu_R$ on the performance of the estimators. Regarding the frontier estimator, the best performance in terms of MSE occurs when $\mu_R = 0.75$ for all estimators considered. When the different measures of overall performance we considered are analyzed, we note that for $R_{rank}$ all estimators seem to perform worse when $\mu_R = 0.75$. Concerning $R_{mag}$ and $R_{rel}$ we note that almost all estimators perform better when $\mu_R = 0.75$. The exceptions occur for CBE NW and SBE NW in the model with correlated inputs for $R_{rel}$ and also for SBE NW in the same model for $R_{mag}$ (Tables 4.2 and 4.3).

**Relative performance of estimators**

Each Nadaraya-Watson type estimator performs worse than its respective extension to local linear type estimator. Furthermore, CBE NW and CBE LL perform very similar to SBE NW and SBE NLL, respectively. This pattern occurs in term of MSE (Figs. 4.1-4.12) as well as in terms of the measures of overall efficiency we considered (Tables 4.1-4.3). However, an important exception in term of MSE occurs when considering the model with correlated inputs. In this case, CBE NW outperforms SBE NW, specially when $\mu_R = 0.75$ (see Figs. 4.5-4.8).

## 4.4 Real Data Example

We illustrate our methodology analyzing USA crime data. The goal is to estimate a production frontier and efficiency for 294 USA Law Enforcement agencies using data for the year 2000. Data sources are FBI's Uniform Crime Reports and LEMAS (Law Enforcement Management and Administrative Statistics) survey. All data used is available on the Internet in the site of Bureau of Justice Statistics (http://bjsdata.ojp.usdoj.gov).

In order to measure crime we consider crime trend data from FBI's Uniform Crime Reports for large agencies in USA (population coverage $\geq 80,000$). This data was used to construct the output, which is defined as population per total crime, where total crime is number of violent crimes plus number of property crimes[5]. This output measure is consistent with our methodology and also is considered in Gorman and Ruggiero (2008).

To measure resources invested in police force we consider data from LEMAS (Law Enforcement Management and Administrative Statistics) survey. This data was used to construct three inputs, which are base annual starting salary for three categories: chief executive, sergeant and entry-level officer. We suppose that those inputs combine additively and use CBE to estimate the efficiency scores.

Efficiency rank and efficiency scores are shown on Table 4.4. Using the output and input measures mentioned above, we see that the estimated five most efficient agencies are Fulton County Police Department-GA, Harford County Sheriff Office-MD, Jefferson County Sheriff Department-MO, Springfield Police Department-MO and Jefferson County Sheriff Department-CO, in decreasing order. On the other hand, the estimated five least efficient agencies are Miami Police Department-FL, Baltimore City

[5]Violent crimes are murder and non-negligent manslaughter, forcible rape, robbery and aggravated assault. Property crimes are burglary, larceny-theft and motor vehicle theft.

Police Department-MD, Atlanta Police Department-GA, Chattanooga Police Department-TN and St Joseph County Sheriff Department-IN, where the last is the least efficient.

## 4.5 Summary and conclusions

In this paper we propose a novel extension of the production frontier model presented by Torrent and Ziegelmann (2010). This extension allows for estimating production frontier and efficiency of production units in a multiple inputs context without suffering the *curse of dimensionality*. To overcome this problem we rewrite that model in the framework of additive models based on assumptions regarding how the inputs combine in production. In particular, we consider the cases of additive and multiplicative inputs, which are in accordance with economic theory. In the Monte Carlo study we conduct the classical backfitting estimator and smooth backfitting estimator are considered as estimators of the additive model.

## 4.6 Appendix: Tables and Garphics

Tabela 4.1: Overall Efficiency Measures - $R_{rank}$

| Model | $\mu_R$ | $n$ | $R_{rank}$ | | | |
|---|---|---|---|---|---|---|
| | | | CBE NW | CBE LL | SBE NW | SBE NLL |
| I | 0.25 | 200 | 0.992 | 0.995 | 0.992 | 0.995 |
| | | 400 | 0.996 | 0.997 | 0.995 | 0.997 |
| | 0.75 | 200 | 0.968 | 0.985 | 0.966 | 0.984 |
| | | 400 | 0.979 | 0.992 | 0.978 | 0.992 |
| II | 0.25 | 200 | 0.992 | 0.995 | 0.991 | 0.995 |
| | | 400 | 0.995 | 0.997 | 0.995 | 0.997 |
| | 0.75 | 200 | 0.963 | 0.985 | 0.955 | 0.983 |
| | | 400 | 0.976 | 0.992 | 0.968 | 0.991 |
| III | 0.25 | 200 | 0.983 | 0.990 | 0.982 | 0.990 |
| | | 400 | 0.989 | 0.995 | 0.989 | 0.995 |
| | 0.75 | 200 | 0.937 | 0.976 | 0.937 | 0.975 |
| | | 400 | 0.958 | 0.987 | 0.957 | 0.986 |

Model I: additive frontier and independent inputs;

Model II: additive frontier and correlated inputs;

Model III: multiplicative frontier and independent inputs.

Tabela 4.2: Overall Efficiency Measures - $R_{mag}$

| Model | $\mu_R$ | $n$ | $R_{mag}(\times 10^{-2})$ | | | |
|---|---|---|---|---|---|---|
| | | | CBE NW | CBE LL | SBE NW | SBE NLL |
| I | 0.25 | 200 | 0.589 | 0.329 | 0.577 | 0.326 |
| | | 400 | 0.422 | 0.215 | 0.443 | 0.214 |
| | 0.75 | 200 | 0.430 | 0.152 | 0.455 | 0.128 |
| | | 400 | 0.335 | 0.065 | 0.335 | 0.066 |
| II | 0.25 | 200 | 0.591 | 0.340 | 0.801 | 0.333 |
| | | 400 | 0.463 | 0.217 | 0.605 | 0.208 |
| | 0.75 | 200 | 0.507 | 0.127 | 0.826 | 0.164 |
| | | 400 | 0.451 | 0.065 | 0.723 | 0.098 |
| III | 0.25 | 200 | 1.387 | 0.883 | 1.423 | 0.868 |
| | | 400 | 1.129 | 0.563 | 1.151 | 0.567 |
| | 0.75 | 200 | 1.042 | 0.205 | 1.110 | 0.212 |
| | | 400 | 0.870 | 0.111 | 0.849 | 0.119 |

Model I: additive frontier and independent inputs;

Model II: additive frontier and correlated inputs;

Model III: multiplicative frontier and independent inputs.

Tabela 4.3: Overall Efficiency Measures - $R_{rel}$

| Model | $\mu_R$ | $n$ | $R_{rel}(\times 10^{-1})$ | | | |
|---|---|---|---|---|---|---|
| | | | CBE NW | CBE LL | SBE NW | SBE NLL |
| I | 0.25 | 200 | 0.414 | 0.347 | 0.413 | 0.347 |
| | | 400 | 0.318 | 0.262 | 0.325 | 0.263 |
| | 0.75 | 200 | 0.383 | 0.246 | 0.391 | 0.247 |
| | | 400 | 0.319 | 0.168 | 0.324 | 0.171 |
| II | 0.25 | 200 | 0.406 | 0.358 | 0.430 | 0.361 |
| | | 400 | 0.312 | 0.261 | 0.335 | 0.256 |
| | 0.75 | 200 | 0.410 | 0.240 | 0.451 | 0.260 |
| | | 400 | 0.336 | 0.168 | 0.389 | 0.182 |
| III | 0.25 | 200 | 0.679 | 0.519 | 0.695 | 0.516 |
| | | 400 | 0.582 | 0.365 | 0.588 | 0.368 |
| | 0.75 | 200 | 0.581 | 0.328 | 0.584 | 0.336 |
| | | 400 | 0.466 | 0.234 | 0.480 | 0.239 |

Model I: additive frontier and independent inputs;

Model II: additive frontier and correlated inputs;

Model III: multiplicative frontier and independent inputs.

Figura 4.1: Model I - Boxplot of Estimators - $n = 200$ - $\mu_r = 0.25$

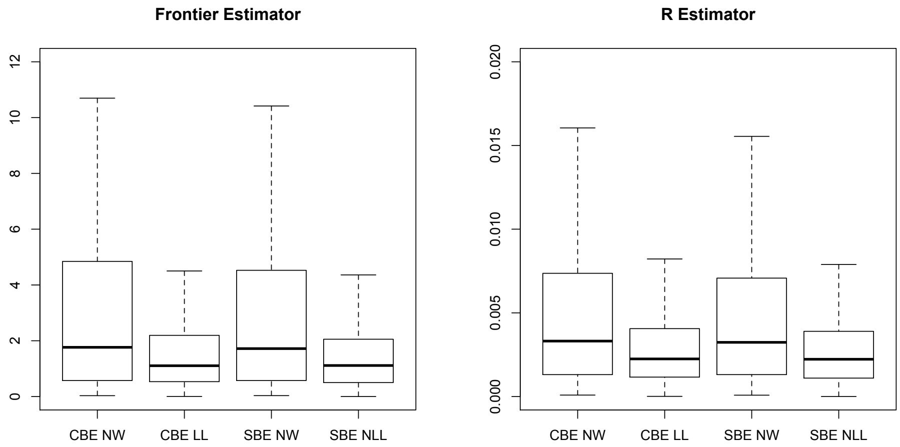



Figura 4.2: Model I - Boxplot of Estimators - $n = 400$ - $\mu_r = 0.25$

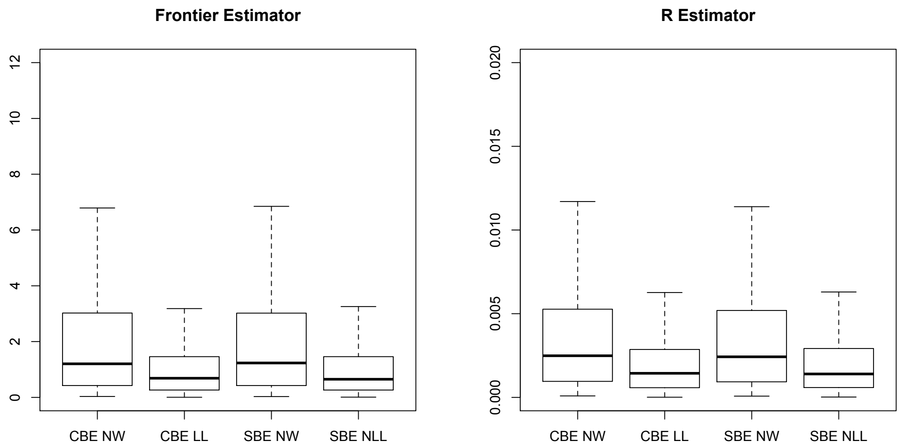


Figura 4.3: Model I - Boxplot of Estimators - $n = 200$ - $\mu_r = 0.75$

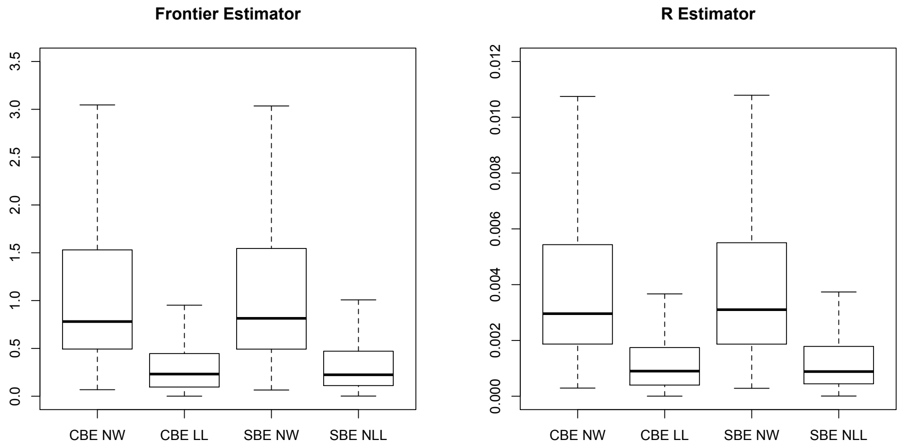



Figura 4.4: Model I - Boxplot of Estimators - $n = 400$ - $\mu_r = 0.75$

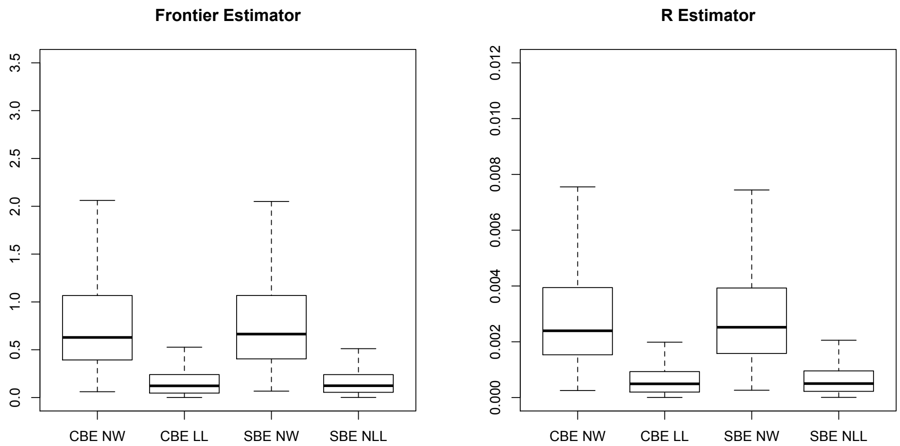


Figura 4.5: Model II - Boxplot of Estimators - $n = 200$ - $\mu_r = 0.25$

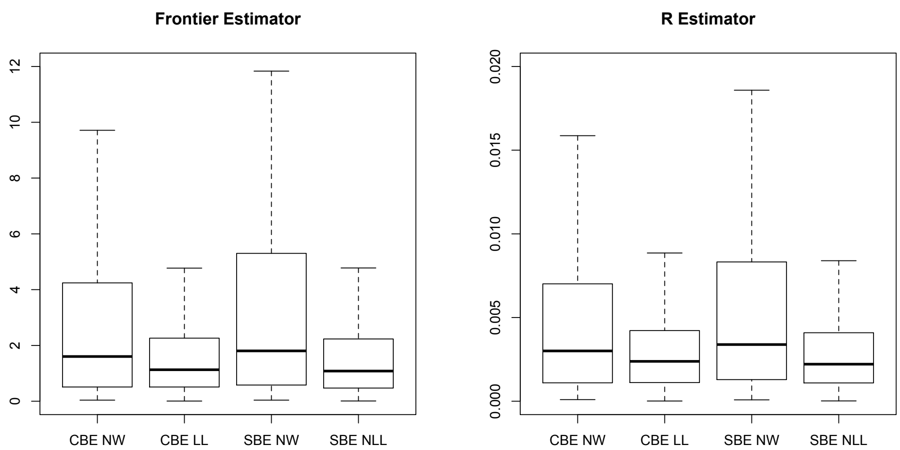



Figura 4.6: Model II - Boxplot of Estimators - $n = 400$ - $\mu_r = 0.25$

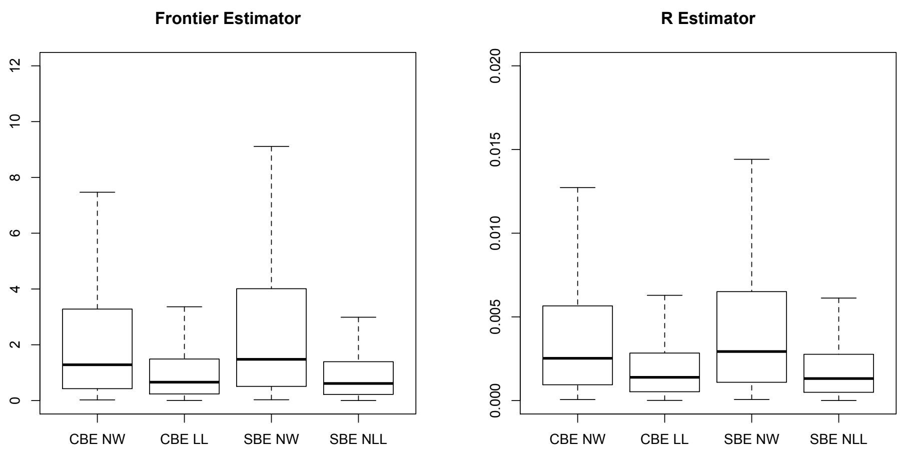


Figura 4.7: Model II - Boxplot of Estimators - $n = 200$ - $\mu_r = 0.75$

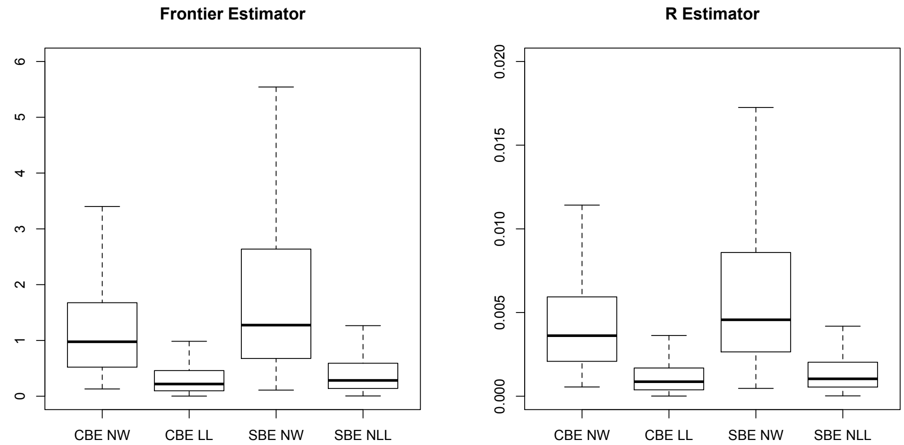



Figura 4.8: Model II - Boxplot of Estimators - $n = 400$ - $\mu_r = 0.75$

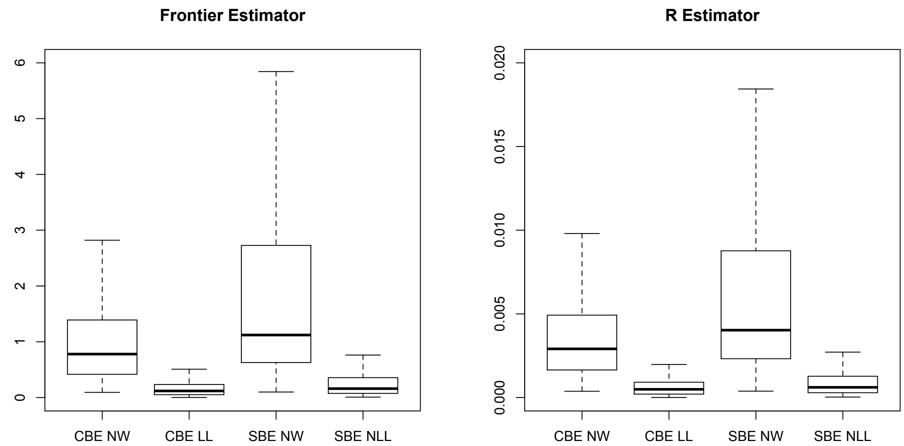


Figura 4.9: Model III - Boxplot of Estimators - $n = 200$ - $\mu_r = 0.25$

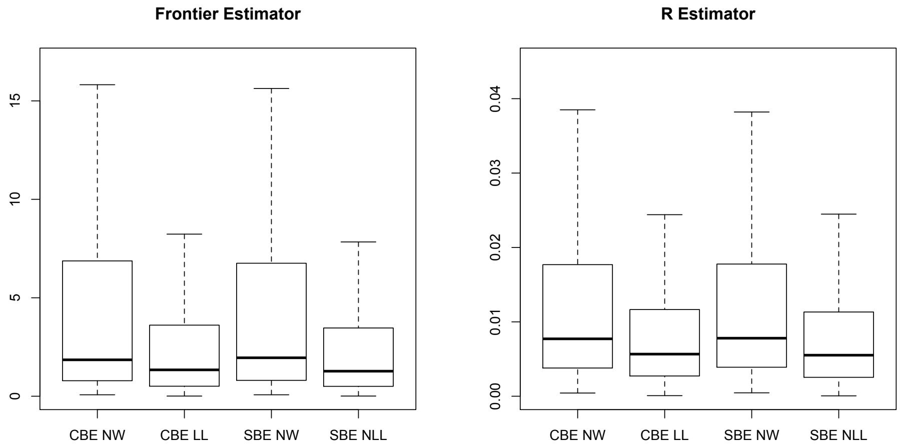



Figura 4.10: Model III - Boxplot of Estimators - $n = 400$ - $\mu_r = 0.25$

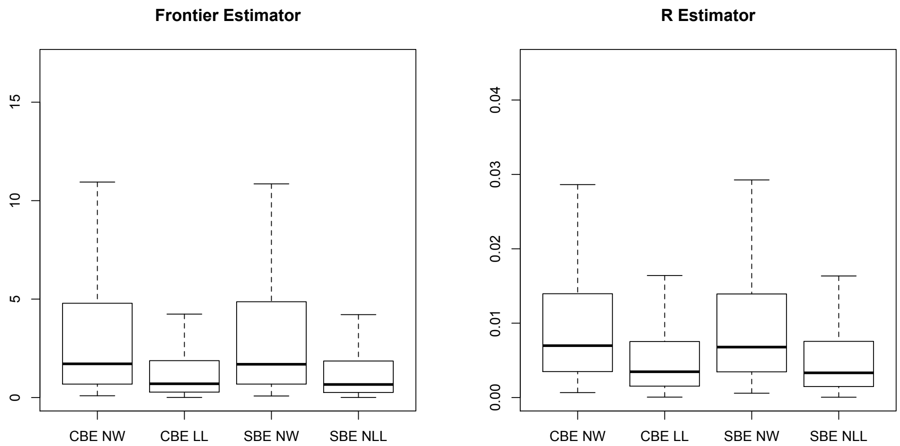


Figura 4.11: Model III - Boxplot of Estimators - $n = 200$ - $\mu_r = 0.75$

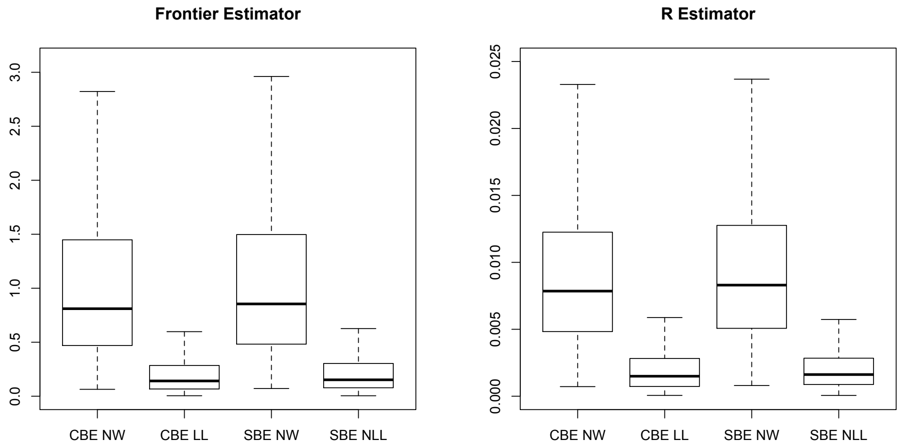



Figura 4.12: Model III - Boxplot of Estimators - $n = 400$ - $\mu_r = 0.75$

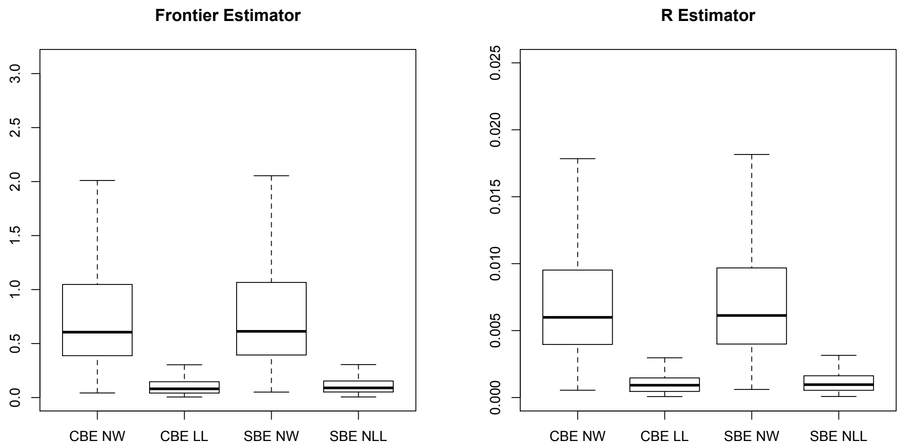


Tabela 4.4: Real Data Example - Estimated efficiency

| Rank | Agency | State | $\hat{R}$ | Rank | Agency | State | $\hat{R}$ |
|---|---|---|---|---|---|---|---|
| 1 | Fulton County Police Department | GA | 1.00 | 81 | Riverside County Sheriff Department | CA | 0.39 |
| 2 | Harford County Sheriff Office | MD | 0.89 | 82 | Hernando County Sheriff Department | FL | 0.39 |
| 3 | Jefferson County Sheriff Department MO | MO | 0.88 | 83 | Lexington County Sheriff Department | SC | 0.38 |
| 4 | Springfield Police Dept MO | MO | 0.86 | 84 | Hampton Police Dept | VA | 0.38 |
| 5 | Jefferson County Sheriff Department CO | CO | 0.81 | 85 | San Francisco Police Dept | CA | 0.38 |
| 6 | Oakland County Sheriff Office | MI | 0.81 | 86 | Kitsap County Sheriff Office | WA | 0.38 |
| 7 | Amherst Town Police Dept | NY | 0.80 | 87 | Midland Police Dept | TX | 0.38 |
| 8 | El Dorado County Sheriff Department | CA | 0.80 | 88 | Henderson Police Dept | NV | 0.38 |
| 9 | Arapahoe County Sheriff Department | CO | 0.80 | 89 | Anne Arundel County Police Department | MD | 0.38 |
| 10 | Shreveport Police Dept | LA | 0.79 | 90 | Collier County Sheriff Department | FL | 0.37 |
| 11 | Washington Metropolitan Police Dept | DC | 0.78 | 91 | Plano Police Dept | TX | 0.37 |
| 12 | San Jose Police Dept | CA | 0.77 | 92 | Garden Grove Police Dept | CA | 0.37 |
| 13 | El Paso | CO | 0.76 | 93 | Kern County Sheriff Department | CA | 0.37 |
| 14 | New Castle County Police Department | DE | 0.75 | 94 | Clay County Sheriff Department | FL | 0.37 |
| 15 | Onondaga County Sheriff Department | NY | 0.75 | 95 | Charles County Sheriff Office | MD | 0.37 |
| 16 | Naperville Police Dept | IL | 0.74 | 96 | Virginia Beach Police Dept | VA | 0.36 |
| 17 | San Diego County Sheriff Department | CA | 0.72 | 97 | Escondido Police Dept | CA | 0.36 |
| 18 | Allen County Sheriff Department | IN | 0.67 | 98 | Garland Police Dept | TX | 0.36 |
| 19 | Fort Bend County Sheriff Department | TX | 0.65 | 99 | Alameda County Sheriff Department | CA | 0.36 |
| 20 | Santa Barbara County Sheriff Department | CA | 0.64 | 100 | Fontana Police Dept | CA | 0.36 |
| 21 | Stark County Sheriff Office | OH | 0.61 | 101 | Spartanburg County Sheriff Office | SC | 0.36 |
| 22 | Stanislaus County Sheriff Department | CA | 0.61 | 102 | Oceanside | CA | 0.36 |
| 23 | Buncombe County Sheriff Department | NC | 0.60 | 103 | Fullerton Police Dept | CA | 0.36 |
| 24 | Waco Police Dept | TX | 0.59 | 104 | Harris County Sheriff Office | TX | 0.35 |
| 25 | Sioux Falls Police Dept | SD | 0.56 | 105 | Montgomery County Police Department | MD | 0.35 |
| 26 | Charlotte County Sheriff Department | FL | 0.56 | 106 | Richland County Sheriff Department | SC | 0.35 |
| 27 | Butte County Sheriff Department | CA | 0.56 | 107 | Pinellas County Sheriff Department | FL | 0.35 |
| 28 | St Charles County Sheriff Department | MO | 0.55 | 108 | Greenville County Sheriff Office | SC | 0.35 |
| 29 | Monterey County Sheriff Department | CA | 0.55 | 109 | Burbank Police Dept | CA | 0.35 |
| 30 | Snohomish County Sheriff Office | WA | 0.55 | 110 | Boise Police Dept | ID | 0.35 |
| 31 | Montgomery County Sheriff Department | TX | 0.55 | 111 | South Bend Police Dept | IN | 0.35 |
| 32 | Kent County Sheriff Office | MI | 0.55 | 112 | Vallejo Police Dept | CA | 0.35 |
| 33 | St Petersburg Police Dept | FL | 0.54 | 113 | Cedar Rapids Police Dept | IA | 0.35 |
| 34 | Tucson Police Dept | AZ | 0.54 | 114 | Henrico County Police Dept | VA | 0.34 |
| 35 | Arlington County Police Department | VA | 0.54 | 115 | Contra Costa County Sheriff Department | CA | 0.34 |
| 36 | Torrance Police Dept | CA | 0.54 | 116 | Chesapeake Police Dept | VA | 0.34 |
| 37 | Irvine Police Police | CA | 0.52 | 117 | Abilene Police Dept | TX | 0.34 |
| 38 | Tempe Police Dept | AZ | 0.51 | 118 | Topeka Police Dept | KS | 0.34 |
| 39 | Green Bay Police Dept | WI | 0.50 | 119 | Brevard County Sheriff Department | FL | 0.34 |
| 40 | Wake County Sheriff Department | NC | 0.50 | 120 | Colorado Springs | CO | 0.34 |
| 41 | Erie City Police Dept | PA | 0.50 | 121 | Fresno County Sheriff Department | CA | 0.34 |
| 42 | Warren Police Dept | MI | 0.49 | 122 | Santa Rosa Police Dept | CA | 0.33 |
| 43 | Prince William County Police Department | VA | 0.49 | 123 | Alexandria Police Dept | VA | 0.33 |
| 44 | Cobb County Police Department | GA | 0.48 | 124 | Pasco County Sheriff Department | FL | 0.33 |
| 45 | Livonia Police Dept | MI | 0.48 | 125 | El Monte Police Dept | CA | 0.33 |
| 46 | Chesterfield County Police Department | VA | 0.48 | 126 | Waterbury Police Dept | CT | 0.33 |
| 47 | Seattle Police Dept | WA | 0.47 | 127 | Madison City Police Dept | WI | 0.33 |
| 48 | Guilford County Sheriff Office | NC | 0.47 | 128 | Lee County Sheriff Department | FL | 0.33 |
| 49 | Akron City Police Dept | OH | 0.46 | 129 | Fort Collins Police Dept | CO | 0.32 |
| 50 | San Joaquin County Sheriff Department | CA | 0.46 | 130 | Pasadena Police Dept TX | TX | 0.32 |
| 51 | St Tammany Parish Sheriff Department | LA | 0.46 | 131 | Clearwater Police Dept | FL | 0.32 |
| 52 | Ann Arbor Police Dept | MI | 0.46 | 132 | Hialeah Police Dept | FL | 0.31 |
| 53 | Worcester Police Dept | MA | 0.46 | 133 | Oxnard Police Dept | CA | 0.31 |
| 54 | Okaloosa County Sheriff Department | FL | 0.45 | 134 | Cambridge Police Dept | MA | 0.31 |
| 55 | Monroe County Sheriff Office | NY | 0.45 | 135 | Manatee County Sheriff Department | FL | 0.31 |
| 56 | Glendale Police Dept CA | CA | 0.45 | 136 | Sacramento County Sheriff Department | CA | 0.31 |
| 57 | Orange Police Dept | CA | 0.45 | 137 | Santa Cruz County Sheriff Department | CA | 0.31 |
| 58 | Clark County Sheriff Department | WA | 0.45 | 138 | Jersey City Police Dept | NJ | 0.30 |
| 59 | Hamilton County Sheriff Department | OH | 0.45 | 139 | Las Vegas Metropolitan Police Department | NV | 0.30 |
| 60 | Simi Valley Police Dept | CA | 0.44 | 140 | Pierce County Sheriff Department | WA | 0.30 |
| 61 | Knox County Sheriff Office | TN | 0.44 | 141 | Lakewood | CO | 0.30 |
| 62 | Stockton Police Dept | CA | 0.44 | 142 | Denver Police Dept | CO | 0.30 |
| 63 | San Bernardino County Sheriff Department | CA | 0.44 | 143 | Aurora Police Dept CO | CO | 0.29 |
| 64 | Huntington Beach Police Dept | CA | 0.44 | 144 | Salt Lake County Sheriff Office | UT | 0.29 |
| 65 | Anaheim Police Dept | CA | 0.43 | 145 | Savannah Police Dept | GA | 0.29 |
| 66 | Bexar County Sheriff Office | TX | 0.43 | 146 | Bakersfield Police Dept | CA | 0.29 |
| 67 | Hidalgo County Sheriff Department | TX | 0.43 | 147 | Evansville Police Dept | IN | 0.29 |
| 68 | Howard County Police Department | MD | 0.42 | 148 | Dekalb County Public Safety Department | GA | 0.29 |
| 69 | Gwinnett County Police Department | GA | 0.42 | 149 | Pomona Police Dept | CA | 0.28 |
| 70 | King County Sheriff Office | WA | 0.41 | 150 | Inglewood Police Dept | CA | 0.28 |
| 71 | Marion County Sheriff Department | FL | 0.41 | 151 | Salinas Police Dept | CA | 0.28 |
| 72 | Aurora Police Dept IL | IL | 0.41 | 152 | Reno Police Dept | NV | 0.28 |
| 73 | Bernalillo County Sheriff Department | NM | 0.41 | 153 | Irving Police Dept | TX | 0.28 |
| 74 | Norwalk | CA | 0.41 | 154 | Concord Police Dept | CA | 0.28 |
| 75 | Forsyth County Sheriff Department | NC | 0.41 | 155 | Grand Prairie Police Dept | TX | 0.28 |
| 76 | Pima County Sheriff Department | AZ | 0.40 | 156 | Escambia County Sheriff Department | FL | 0.28 |
| 77 | Jefferson County Sheriff Department AL | AL | 0.40 | 157 | Long Beach Police Dept | CA | 0.27 |
| 78 | Downey Police Dept | CA | 0.40 | 158 | Clackamas County Sheriff Department | OR | 0.27 |
| 79 | Mobile County Sheriff Department | AL | 0.40 | 159 | Anchorage Police Dept | AK | 0.27 |
| 80 | Anderson County Sheriff Department | SC | 0.39 | 160 | Chandler Police Dept | AZ | 0.27 |

| Rank | Agency (cont.) | State | $\hat{R}$ | Rank | Agency | State | $\hat{R}$ |
|---|---|---|---|---|---|---|---|
| 161 | Louisville Police Dept | KY | 0.27 | 228 | St Louis Police Dept | MO | 0.20 |
| 162 | Lincoln Police Dept | NE | 0.27 | 229 | Glendale Police Dept AZ | AZ | 0.19 |
| 163 | Sunnyvale Dept Of Public Safety | CA | 0.27 | 230 | Minneapolis Police Dept | MN | 0.19 |
| 164 | Cumberland County Sheriff Office | NC | 0.27 | 231 | New Haven Police Dept | CT | 0.19 |
| 165 | Pittsburgh Bureau Of Police | PA | 0.27 | 232 | Albany Police Dept | NY | 0.19 |
| 166 | El Paso Police Dept | TX | 0.27 | 233 | Boston Police Dept | MA | 0.19 |
| 167 | Indianapolis Police Dept | IN | 0.27 | 234 | Lubbock Police Dept | TX | 0.19 |
| 168 | Pasadena Police Dept CA | CA | 0.26 | 235 | Beaumont Police Dept | TX | 0.19 |
| 169 | Riverside Police Dept | CA | 0.26 | 236 | Huntsville Police Dept | AL | 0.19 |
| 170 | Polk County Sheriff Department | FL | 0.26 | 237 | Raleigh Police Dept | NC | 0.19 |
| 171 | Newport News Police Dept | VA | 0.26 | 238 | Knoxville Police Dept | TN | 0.19 |
| 172 | Allentown City Police Dept | PA | 0.26 | 239 | Eugene Police Dept | OR | 0.19 |
| 173 | Elizabeth Police Dept | NJ | 0.26 | 240 | Laredo | TX | 0.19 |
| 174 | Chula Vista Police Dept | CA | 0.26 | 241 | Montgomery Police Dept | AL | 0.19 |
| 175 | Austin Police Dept | TX | 0.26 | 242 | Portland Police Dept | OR | 0.19 |
| 176 | Baltimore County Police Department | MD | 0.26 | 243 | Fresno Police Dept | CA | 0.19 |
| 177 | Hayward Police Dept | CA | 0.26 | 244 | Sterling Heights Police Dept | MI | 0.19 |
| 178 | Lansing City Police Dept | MI | 0.25 | 245 | Milwaukee Police Dept | WI | 0.18 |
| 179 | Mesquite Police Dept | TX | 0.25 | 246 | Sonoma County Sheriff Department | CA | 0.18 |
| 180 | Charleston County Sheriff Department | SC | 0.25 | 247 | Jacksonville | FL | 0.18 |
| 181 | Ontario Police Dept | CA | 0.25 | 248 | Springfield Police Dept IL | IL | 0.18 |
| 182 | Grand Rapids Police Dept | MI | 0.24 | 249 | Sarasota County Sheriff Department | FL | 0.18 |
| 183 | Winston-Salem Police Dept | NC | 0.24 | 250 | Flint City Police Dept | MI | 0.18 |
| 184 | Fort Wayne Police Dept | IN | 0.24 | 251 | Oklahoma City Police Dept | OK | 0.18 |
| 185 | Newark Police Dept | NJ | 0.24 | 252 | Toledo Police Dept | OH | 0.17 |
| 186 | Palm Beach County Sheriff Department | FL | 0.24 | 253 | Little Rock Police Dept | AR | 0.17 |
| 187 | Honolulu Police Dept | HI | 0.23 | 254 | Salt Lake City Police Dept | UT | 0.17 |
| 188 | Portsmouth Police Dept | VA | 0.23 | 255 | Wichita Police Dept | KS | 0.17 |
| 189 | Independence Police Dept | MO | 0.23 | 256 | Fort Lauderdale Police Dept | FL | 0.17 |
| 190 | Columbus Police Dept GA | GA | 0.23 | 257 | Thurston County Sheriff Department | WA | 0.17 |
| 191 | Lafayette Police Dept | LA | 0.23 | 258 | Metro-Dade Police Department | FL | 0.17 |
| 192 | Greensboro Police Dept | NC | 0.23 | 259 | St Paul Police Dept | MN | 0.16 |
| 193 | Tulare County Sheriff Department | CA | 0.23 | 260 | Oakland Police Dept | CA | 0.16 |
| 194 | Spokane Police Dept | WA | 0.22 | 261 | Durham Police Dept | NC | 0.16 |
| 195 | Hillsborough County Sheriff Department | FL | 0.22 | 262 | Providence Police Dept | RI | 0.16 |
| 196 | Stamford Police Dept | CT | 0.22 | 263 | San Bernardino Police Dept | CA | 0.16 |
| 197 | Prince Georges County Police Department | MD | 0.22 | 264 | Hartford Police Dept | CT | 0.16 |
| 198 | Rochester Police Dept | NY | 0.22 | 265 | Columbus Police Dept OH | OH | 0.16 |
| 199 | Norfolk Police Dept | VA | 0.22 | 266 | Richmond (City) Bureau Of Police | VA | 0.16 |
| 200 | Gary Police Dept | IN | 0.22 | 267 | Seminole County Sheriff Department | FL | 0.16 |
| 201 | Mobile Police Dept | AL | 0.22 | 268 | Columbia Police Dept | SC | 0.16 |
| 202 | Amarillo Police Dept | TX | 0.22 | 269 | Springfield Police Dept MA | MA | 0.16 |
| 203 | Modesto Police Dept | CA | 0.22 | 270 | East Baton Rouge Parish Sheriff Dept | LA | 0.15 |
| 204 | New Orleans Police Dept | LA | 0.22 | 271 | Tulsa Police Dept | OK | 0.15 |
| 205 | Bridgeport Police Dept | CT | 0.22 | 272 | Nashville-Davidson Metro Police Dept | TN | 0.15 |
| 206 | Fort Worth Police Dept | TX | 0.21 | 273 | Kansas City Police Dept | MO | 0.15 |
| 207 | Pueblo Police Dept | CO | 0.21 | 274 | Berkeley Police Dept | CA | 0.15 |
| 208 | Jefferson Parish Sheriff Department | LA | 0.21 | 275 | Jackson Police Dept | MS | 0.15 |
| 209 | Corpus Christi Police Dept | TX | 0.21 | 276 | Albuquerque Police Dept | NM | 0.15 |
| 210 | Volusia County Sheriff Department | FL | 0.21 | 277 | Baton Rouge Police Dept | LA | 0.15 |
| 211 | Omaha Police Dept | NE | 0.21 | 278 | Birmingham Police Dept | AL | 0.14 |
| 212 | Washtenaw County Sheriff Department | MI | 0.21 | 279 | Peoria Police Dept | IL | 0.14 |
| 213 | Wichita Falls Police Dept | TX | 0.21 | 280 | Washington County Sheriff Office | OR | 0.14 |
| 214 | West Covina Police Dept | CA | 0.21 | 281 | Tampa Police Dept | FL | 0.14 |
| 215 | Salem Police Dept | OR | 0.21 | 282 | Travis County Sheriff Department | TX | 0.14 |
| 216 | Cleveland | OH | 0.21 | 283 | Tallahassee Police Dept | FL | 0.14 |
| 217 | Charlotte-Mecklenburg Police Department | NC | 0.20 | 284 | Syracuse Police Dept | NY | 0.14 |
| 218 | Cincinnati Police Dept | OH | 0.20 | 285 | Memphis Police Dept | TN | 0.14 |
| 219 | Brownsville Police Dept | TX | 0.20 | 286 | Scottsdale Police Dept | AZ | 0.13 |
| 220 | Spokane County Sheriff Department | WA | 0.20 | 287 | Dayton | OH | 0.13 |
| 221 | Des Moines Police Dept | IA | 0.20 | 288 | Orlando Police Dept | FL | 0.13 |
| 222 | Mesa Police Dept | AZ | 0.20 | 289 | Macon Police Dept | GA | 0.12 |
| 223 | Tacoma Police Dept | WA | 0.20 | 290 | Miami Police Dept | FL | 0.12 |
| 224 | Hollywood Police Dept | FL | 0.20 | 291 | Baltimore City Police Dept | MD | 0.12 |
| 225 | Sacramento Police Dept | CA | 0.20 | 292 | Atlanta Police Dept | GA | 0.10 |
| 226 | Rockford Police Dept | IL | 0.20 | 293 | Chattanooga Police Dept | TN | 0.10 |
| 227 | Buffalo Police Dept | NY | 0.20 | 294 | St Joseph County Sheriff Department | IN | 0.10 |

## 4.7 References


1. Aigner, D., C.A.K. Lovell and P. Schmidt, Formulation and estimation of stochastic frontiers production function models. **Journal of Econometrics**, 6, 21-37, 1977.

2. Aragon, Y., A. Daouia, C. Thomas-Agnan, Nonparametric frontier estimation: a conditional quantile-based approach. **Econometric Theory**, 21, 358-389, 2005.

3. Buja, A., Hastie, T. and Tibshirani, R., Linear smoothers and additive models (with discussion). **Annals of Statistics**, 17, 453-555, 1989.

4. Daouia, A., Gardes, L. and S. Girard, 2009, Large Sample Approximation of the Distribution for Smoothed Monotone Frontier Estimators . Working paper.

5. Fan, J., Design adaptive nonparametric regression. **Journal of the American Statistical Association**, 87, 998-1004, 1992.

6. Fan, Y., Q. Li and A. Weersink, Semiparametric estimation of stochastic production frontier models. **Journal of Business and Economic Statistics**, 14, 460-468, 1996.

7. Hastie, T. J. and Tibshirani, R. J., **Generalized Additive Models.** London: Chapman and Hall, 1990.

8. Kumbhakar, S. C., B. U. Park, L. Simar and E. Tsionas, Nonparametric stochastic frontiers: a local maximum likelihood approach. **Journal of Econometrics**, 137, 1-27, 2007.

9. Kim, W., Linton, O. and Hengartner, N., A computationally efficient oracle estimator for additive non-parametric regression with bootstrap confidence intervals. **Journal of Computational and Graphical Statistics**, 8, 278-297, 1999.

10. Linton, O. B., Efficient estimation of additive nonparametric regression models. **Biometrika**, 84, 469-473, 1997.

11. Linton, O. B. and Nielsen, J. P., A kernel method of estimating structured nonparametric regression based on marginal integration. **Biometrika**, 82, 93-101, 1995.

12. Mammen, E., Linton, O. B. and Nielsen, J. P., The existence and asymptotic properties of a backfitting projection algorithm under weak conditions. **Annals of Statistics**, 27, 1443-1490, 1999.

13. Mammen, E., Marron, J. S. Turlach, B. and Wand, M. P., A general projection framework for constrained smoothing. **Statistical Science**, 16, 232-248, 2001.

14. Martins-Filho, C., Torrent, H., Ziegelmann, F., 2010, Nonparametric Frontier Estimation: Using Local Exponential Regression for Conditional Variance. Submitted paper.

15. Martins-Filho, C. and F. Yao, Nonparametric frontier estimation via local linear regression. Journal of Econometrics, 141, 283-319, 2007.

16. Martins-Filho, C. and F. Yao, A Smooth Nonparametric Conditional Quantile Frontier Estimator. **Journal of Econometrics**, 143, 317-333, 2008.

17. Martins-Filho, C. and F. Yao, 2010, Nonparametric stochastic frontier estimation via profile likelihood. Working paper, University of Colorado, Boulder.

18. Meeusen, W., van den Broeck, J., Efficiency estimation from Cobb-Douglas production functions with composed error. **International Economic Review**, 18, 435-444, 1977.

19. Nielsen, J. P. and Linton, O. B., An optimization interpretation of integration and back-fitting estimators for separable nonparametric models. **J. R. Statist. Soc. B**, 60, 217-222, 1998.

20. Nielsen, J. P. and Sperlich, S., Smooth backfitting in practice. **J. R. Statist. Soc. B**, 67, 43-61, 2005.

21. Opsomer, J. D., Asymptotic properties of backitting estimators. **J. Multiv. Anal.**, 73, 166-179, 2000.

22. Opsomer, J. D. and Ruppert, D., Fitting a bivariate additive model by local polynomial regression. **Annals of Statistics**, 25, 186–211, 1997.

23. Simar, L. and P. Wilson, Statistical inference in nonparametric frontier models: recent developments and perspectives, in: H. Fried, C.A.K. Lovell, and S.S. Schmidt, (Eds.), **The Measurement of Productive Efficiency**, 2nd edition. Oxford University Press, Oxford, 2007.

24. Tjøstheim, D. and Auestad, B., Nonparametric identification of nonlinear time series projections. **Journal of the American Statistical Association**, 89, 1398–1409, 1994.

25. Torrent, H. S. and Ziegelmann, F., 2010, Nonparametric Frontier Estimation in Two Steps. Working paper.

26. Ziegelmann, F., Nonparametric estimation of volatility functions: the local exponential approach. **Econometric Theory**, 18, 985-991, 2002.

# 5 Considerações Finais

Conforme estabelecido no Capítulo 1, o principal objetivo desta tese é propor e analisar estimadores não-paramétricos e semi-paramétricos para fronteiras de produção e consequentemente para eficiência de unidades produtivas. Considerou-se a abordagem determinística para este problema, o que permite a estimação dos modelos de fronteira de produção sob hipóteses bastante razoáveis sobre a distribuição de probabilidade conjunta das variáveis produto e insumos. O modelo estatístico e o estimador inicialmente analisados foram aqueles propostos por Martins-Filho e Yao (2007). Esse estimador, chamado NP, caracteriza-se por ser de fácil implementação e por possuir vantagens importantes em relação aos estimadores não-paramétricos de fronteira de produção tradicionais. Nesta tese são apresentados três artigos.

No primeiro artigo, Capítulo 2 desta tese, o estimador exponencial local, proposto por Ziegelmann (2002), foi utilizado para melhorar o estimador NP. A estratégia de estimação que caracteriza o estimador NP não elimina a possibilidade de que valores negativos sejam estimados para a variância condicional. O estimador exponencial local proposto no Capítulo 2, chamado NPE, elimina essa possibilidade, mantendo, contudo, as propriedades desejáveis do estimador NP. Além disso, mostrou-se que o estimador proposto é assintoticamente normal e consistente sob hipóteses razoáveis no contexto de estimação não-paramétrica. Um estudo de Monte Carlo foi apresentado, em que há evidências de ganhos em amostra fininta ao se utilizar o estimador NPE, vis-à-vis o estimador NP. Esses ganhos parecem ser particularmente importantes na estimação do parâmetro locacional, $\sigma_R$, presente no modelo de fronteira de produção considerado.

No segundo artigo, Capítulo 3 desta tese, é proposta uma abordagem original para a estimação do modelo de fronteira de produção de Martins-Filho e Yao. A estratégia utilizada por esses autores, bem como a estratégia proposta no Capítulo 2 desta tese, consiste em estimar a fronteira em três estágios. Os dois primeiros estágios são responsáveis por estimar o formato da fronteira, enquanto o terceiro estágio é responsável por posicioná-la. Contudo, no Capítulo 3, foi proposta uma modificação desse procedimento, eliminando a necessidade do segundo estágio. Ou seja, o procedimento proposto consiste em dois estágios, sendo o primeiro responsável por capturar o formato da fronteira e o segundo por posicioná-la. Como resultado, obteve-se um procedimento de estimação mais simples, chamado NP2S, que mantém todas as vantagens inerentes aos estimadores NPE e NP. Além disso, mostrou-se que o estimador proposto é assintoticamente normal e consistente sob hipóteses razoáveis no contexto de estimação não-paramétrica. Um estudo de Monte Carlo foi apresentado comparando os três estimadores: NP2S, NPE e NP. Os resultados mostraram que o estimador proposto, NP2S, obteve melhor performance vis-à-vis seus concorrentes quando o processo gerador dos dados apresenta eficiência média superior a 50%, i.e., $\mu_R = 0,60$ e $\mu_R = 0,75$ nas simulações consideradas.

No terceiro artigo, Capítulo 4 desta tese, é proposta uma extensão original do modelo de fronteira de produção e do estimador NP2S apresentados no Capítulo 3. Essa extensão permite a estimação de fronteiras de produção e de eficiência de unidades produtivas no contexto de múltiplos insumos sem, contudo, incorrer no fenômeno de *curse of dimensionality*, i.e., perda de performance de estimadores totalmente não-paramétricos no contexto multivariado. A fim de superar esse problema, o modelo do Capítulo 3 foi reescrito em uma estrutura conveniente para sua estimação via modelos aditivos. Para isso foram feitas hipóteses sobre como os insumos se combinam no processo produtivo. Em particular, forma considerados os casos de insumos aditivos e multiplicativos. É importante ressaltar que o estimador NP2S possui uma estrutura conveniente para a estimação via modelos aditivos, visto que para se reescrever esse estimador no contexto de modelos aditivos basta supor que a fronteira é aditiva. Tal extensão não ocorre para os estimadores NP e NPE. Além disso, é estabelecido no capítulo 4, a extensão do estimador NP2S para o caso de múltiplos insumos sob a hipótese de fronteira de produção multiplicativa no contexto de estimação de modelos aditivos. Essa extensão é particularmente importante, visto que a teoria econômica considera amplamente tecnologias em que os insumos se combinam de forma multiplicativa. Um estudo de Monte Carlo foi realizado comparando-se a performance dos estimadores *Classical Backfitting* (CBE) e *Smooth Backfitting* (SBE) como estimadores do modelo aditivo em questão. Em termos gerais, o estimador CBE apresentou melhor performance vis-à-vis o estimador SBE nas simulações consideradas.

# 6 Referências


1. Aigner, D., C.A.K. Lovell and P. Schmidt, Formulation and estimation of stochastic frontiers production function models. **Journal of Econometrics**, 6, 21-37, 1977.

2. Aragon, Y., A. Daouia, C. Thomas-Agnan, Nonparametric frontier estimation: a conditional quantile-based approach. **Econometric Theory**, 21, 358-389, 2005.

3. Buja, A., Hastie, T. and Tibshirani, R., Linear smoothers and additive models (with discussion). **Annals of Statistics**, 17, 453-555, 1989.

4. Daouia, A., Gardes, L. and S. Girard, 2009, Large Sample Approximation of the Distribution for Smoothed Monotone Frontier Estimators . Working paper.

5. Fan, J., Design adaptive nonparametric regression. **Journal of the American Statistical Association**, 87, 998-1004, 1992.

6. Fan, J. and I. Gijbels, Data driven bandwidth selection in local polynomial fitting: variable bandwidth and spatial adaptation. **Journal of the Royal Statistical Society B**, 57, 371-394, 1995.

7. Fan, Y., Q. Li and A. Weersink, Semiparametric estimation of stochastic production frontier models. **Journal of Business and Economic Statistics**, 14, 460-468, 1996.

8. Fan, J., and Q. Yao, Efficient estimation of conditional variance functions in stochastic regression. **Biometrika**, 85, 645-660, 1998.

9. Farrell, M., The measurement of productive efficiency. **Journal of the Royal Statistical Society A**, 120, 253-290, 1957.

10. Gorman, M.F., Ruggiero, J., Evaluating US state police performance using data envelopment analysis. **Int. J. Production Economics**, 113, 1031-1037, 2008.

11. Hall, P., Wolff, R. and Q. Yao, Methods for estimating a conditional distribution function. **Journal of the American Statistical Association**, 94, 154-163, 1999.

12. Hastie, T. J. and Tibshirani, R. J., **Generalized Additive Models.** London: Chapman and Hall, 1990.

13. Kumbhakar, S. C., B. U. Park, L. Simar and E. Tsionas, Nonparametric stochastic frontiers: a local maximum likelihood approach. **Journal of Econometrics**, 137, 1-27, 2007.

14. Kim, W., Linton, O. and Hengartner, N., A computationally efficient oracle estimator for additive non-parametric regression with bootstrap confidence intervals. **Journal of Computational and Graphical Statistics**, 8, 278-297, 1999.

15. Linton, O. B., Efficient estimation of additive nonparametric regression models. **Biometrika**, 84, 469-473, 1997.

16. Linton, O. B. and Nielsen, J. P., A kernel method of estimating structured nonparametric regression based on marginal integration. **Biometrika**, 82, 93-101, 1995.

17. Mammen, E., Linton, O. B. and Nielsen, J. P., The existence and asymptotic properties of a backfitting projection algorithm under weak conditions. **Annals of Statistics**, 27, 1443-1490, 1999.

18. Mammen, E., Marron, J. S. Turlach, B. and Wand, M. P., A general projection framework for constrained smoothing. **Statistical Science**, 16, 232-248, 2001.

19. Martins-Filho, C., Torrent, H., Ziegelmann, F., 2010, Nonparametric Frontier Estimation: Using Local Exponential Regression for Conditional Variance. Submitted paper.

20. Martins-Filho, C. and F. Yao, Nonparametric frontier estimation via local linear regression. Journal of Econometrics, 141, 283-319, 2007.

21. Martins-Filho, C. and F. Yao, A Smooth Nonparametric Conditional Quantile Frontier Estimator. **Journal of Econometrics**, 143, 317-333, 2008.

22. Martins-Filho, C. and F. Yao, 2010, Nonparametric stochastic frontier estimation via profile likelihood. Working paper, University of Colorado, Boulder.

23. Meeusen, W., van den Broeck, J., Efficiency estimation from Cobb-Douglas production functions with composed error. **International Economic Review**, 18, 435-444, 1977.

24. Nielsen, J. P. and Linton, O. B., An optimization interpretation of integration and back-fitting estimators for separable nonparametric models. **J. R. Statist. Soc. B**, 60, 217-222, 1998.

25. Nielsen, J. P. and Sperlich, S., Smooth backfitting in practice. **J. R. Statist. Soc. B**, 67, 43-61, 2005.

26. Opsomer, J. D., Asymptotic properties of backitting estimators. **J. Multiv. Anal.**, 73, 166-179, 2000.

27. Opsomer, J. D. and Ruppert, D., Fitting a bivariate additive model by local polynomial regression. **Annals of Statistics**, 25, 186–211, 1997.

28. Ruppert D., S. Sheather, M. Wand, An effective bandwidth selector for least squares regression. **Journal of the American Statistical Association**, 90, 1257-1270, 1995.

29. Silverman, B.W., **Density estimation for statistics and data analysis**. Chapman and Hall, London, 1986.

30. Simar, L. and P. Wilson, Statistical inference in nonparametric frontier models: recent developments and perspectives, in: H. Fried, C.A.K. Lovell, and S.S. Schmidt, (Eds.), **The Measurement of Productive Efficiency**, 2nd edition. Oxford University Press, Oxford, 2007.

31. Stone, C. J., Consistent nonparametric regression. **Annals of Statistics** 5, 595-620, 1977.

32. Tjøstheim, D. and Auestad, B., Nonparametric identification of nonlinear time series projections. **Journal of the American Statistical Association**, 89, 1398–1409, 1994.

33. Torrent, H. S. and Ziegelmann, F., 2010, Nonparametric Frontier Estimation in Two Steps. Working paper.

34. Torrent, H. S. and Ziegelmann, F., 2010b, Nonparametric Frontier Estimation: Using Additive Models in a Multivariate input case. Working paper.

35. Ziegelmann, F., Nonparametric estimation of volatility functions: the local exponential approach. **Econometric Theory**, 18, 985-991, 2002.