Regressão Quantílica

Seminário Metodológico

Profa. Maria Eduarda Tannuri-Pianto

EPRG/UnB – Abril 2018

“Desde o século 19, as ciências socias tratam o “indivíduo médio” como um ícone, aquele indivíduo hipotético sem quantis que está confortável com a cabeça no forno e os pés em um balde de gelo.

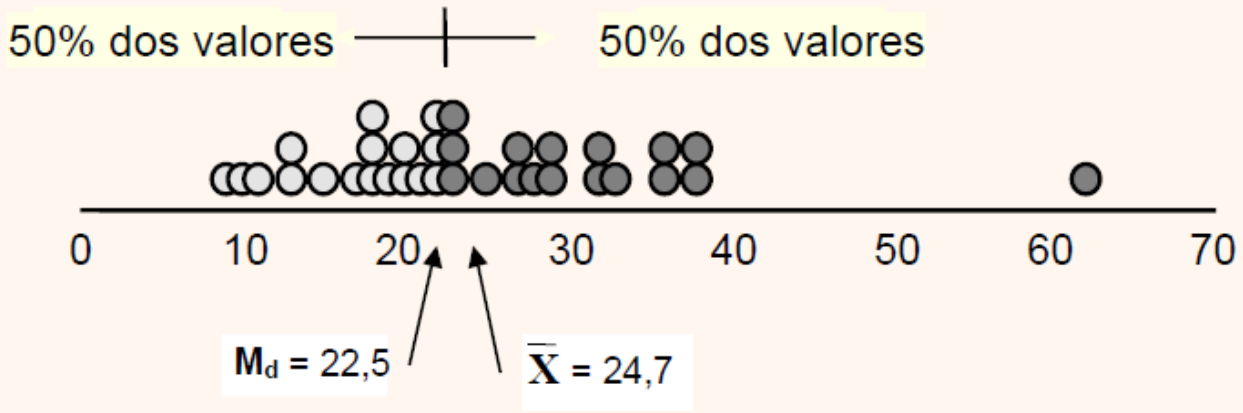
Métodos estatísticos convencionais procuraram estimar os efeitos dos tratamentos de política nesse “indivíduo médio”.

Mas esses efeitos são frequentemente heterogeneos: tratamentos médicos podem melhorar a expectativa de vida , mas também impõem sérios riscos no curto prazo; reduzir o tamanho das salas de aula pode melhorar o desempenho dos bons alunos , mas não ajudar os alunos mais fracos, ou vice-versa.

Métodos de Regressão Quantílica podem ajudar a explorar esses efeitos heterogêneos. “

Roger Koenker

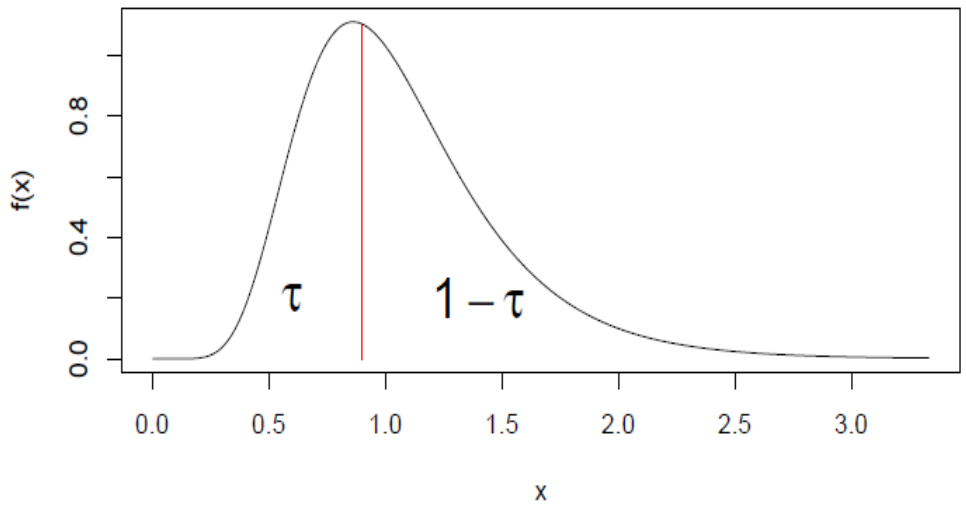
Média e Mediana



PEDRO A. BARBETTA – Estatística Aplicada às Ciências Sociais 6ed. Editora da UFSC, 2006.

Quantis Univariados

Viewed from the perspective of densities, the τth quantile splits the area under the density into two parts: one with area τ below the τth quantile and the other with area $1 - \tau$ above it:



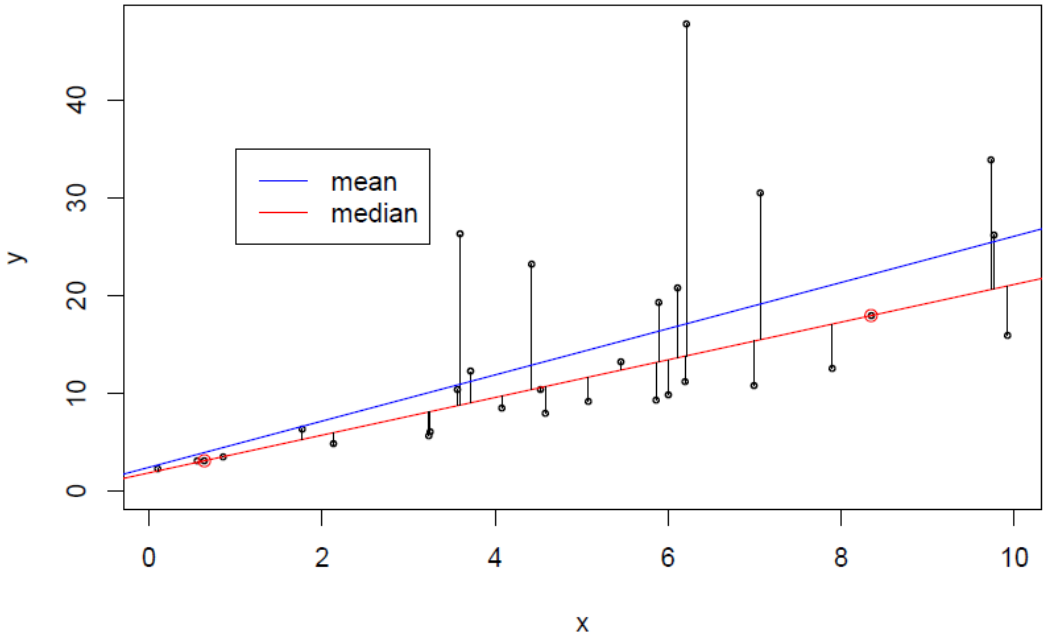
Regressão – MQO

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + u$$

$$y = X\beta + u$$

$$E[y|X] = X\beta$$

Médias x Medianas condicionais



MQO x RQ

The unconditional mean solves

$$\mu = \min_m E(Y - m)^2$$

The conditional mean $\mu(x) = E(Y|X = x)$ solves

$$\mu(x) = \min_m E_{Y|X=x}(Y - m(X))^2.$$

Similarly, the unconditional τth quantile solves

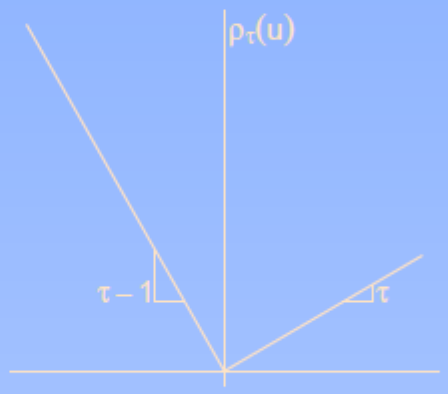
$$\alpha_\tau = \min_a E\rho_\tau(Y - a)$$

and the conditional τth quantile solves

$$\alpha_\tau(x) = \min_a E_{Y|X=x}\rho_\tau(Y - a(X))$$

Função Perda

$\rho_\tau(u) = (\tau - I(u < 0))u$ as illustrated below.



Regressão Quantílica

* Métodos de regressão quantílica complementam métodos de regressão na média (MQO);
* Ao focar em locais específicos da distribuição condicional, representa uma ampliação dos modelos de média condicional ;
* Permite um papel mais flexível do efeito das covariadas, que podem influenciar a localização, a escala e o formato da distribuição resposta;
* Em aplicações de RQ, uma variedade de heterogeneidade não observada se faz observar

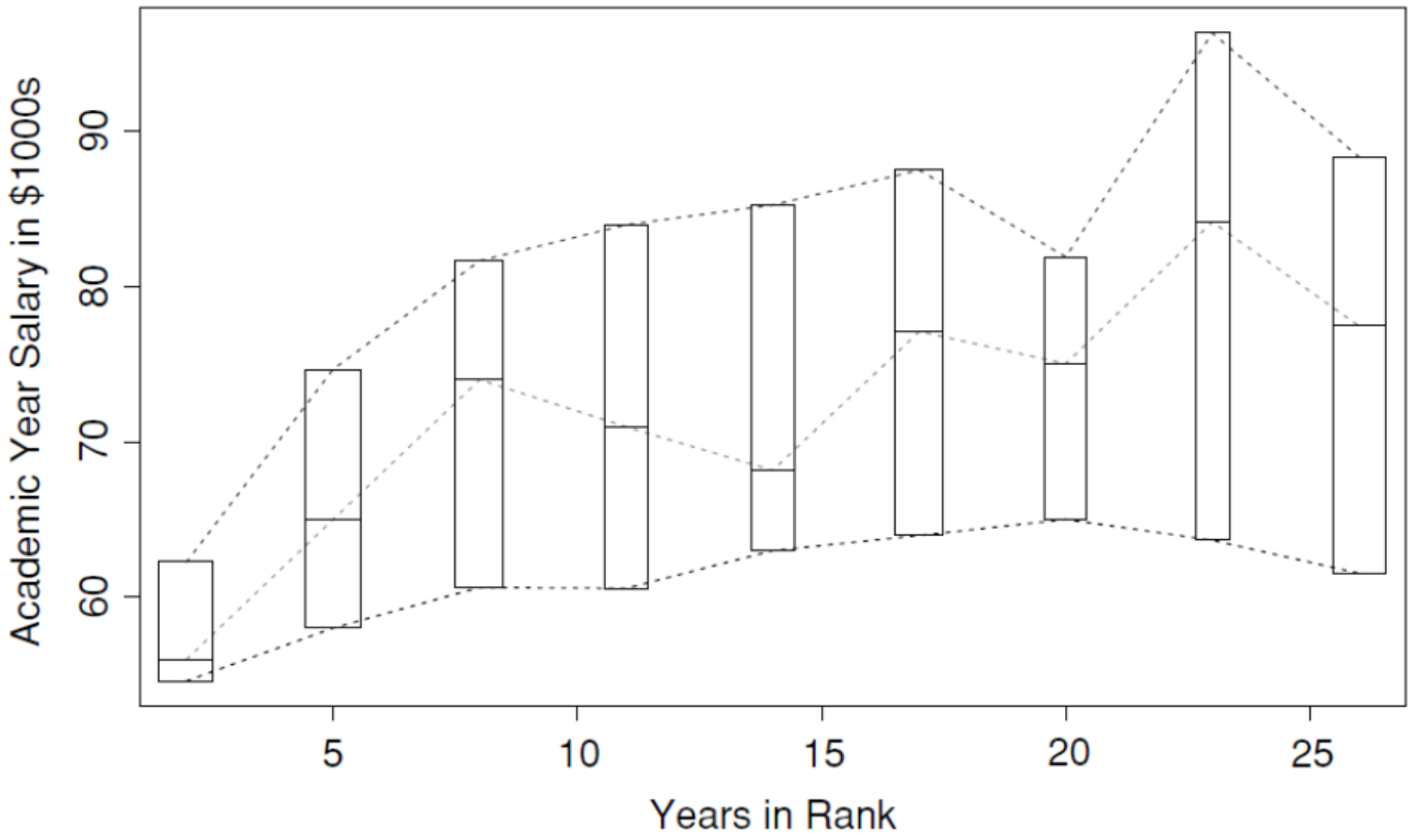


Figure 1.8. Boxplots of 1995 ASA academic salary survey for full professors of statistics in U.S. colleges and universities.

Aplicações

* Modelos Quantílicos Condicionais
  * Efeito tratamento binário
  * Múltiplos tratamentos, covariadas e interações
  * RQ não linear (nos parâmetros)
* RQ não paramétrica
* Modelos Séries Temporais
  * Quantile Auto-Regressive (QAR)
  * Testes Raiz Unitária
  * Value at Risk (VaR)

Aplicações

* Modelos Dados Longitudinais
* Modelos de Duração ou Sobrevivência
* Modelos Causais (QTE – quantile treatment effect)
* Modelos Variáveis Instrumentais
* Erros em Variáveis, Dados Faltantes e Seleção Amostral

Engel's Food Expenditure Data

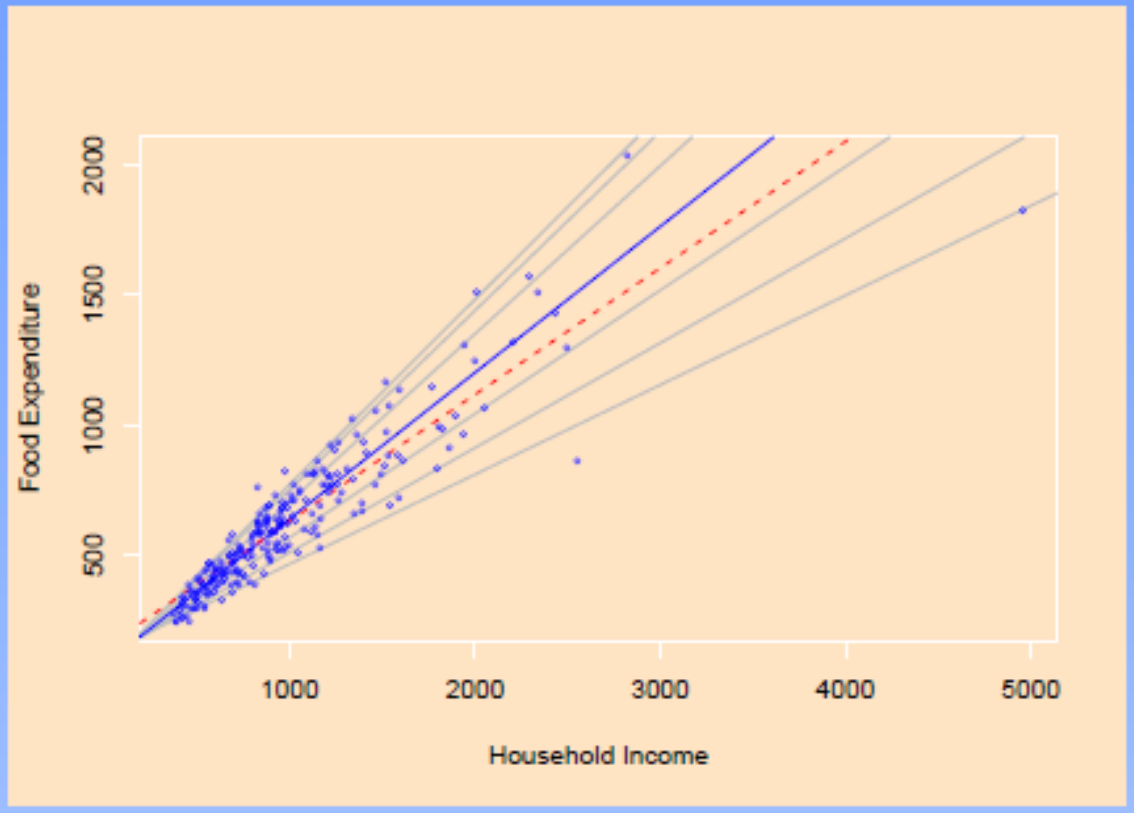


Figure 1: Engel Curves for Food: This figure plots data taken from Engel's (1857) study of the dependence of households' food expenditure on household income. Seven estimated quantile regression lines for $\tau \in \{.05, .1, .25, .5, .75, .9, .95\}$ are superimposed on the scatterplot. The median $\tau = .5$ fit is indicated by the darker solid line; the least squares estimate of the conditional mean function is indicated by the dashed line.

Engel's Food Expenditure Data

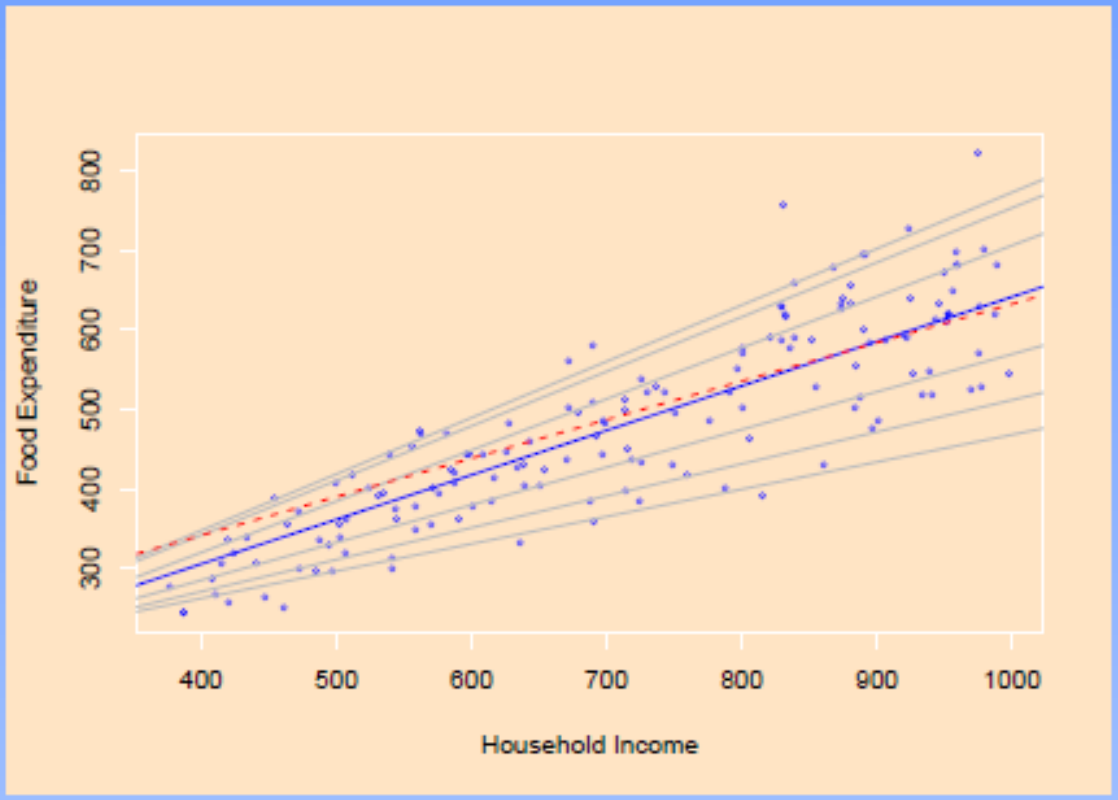
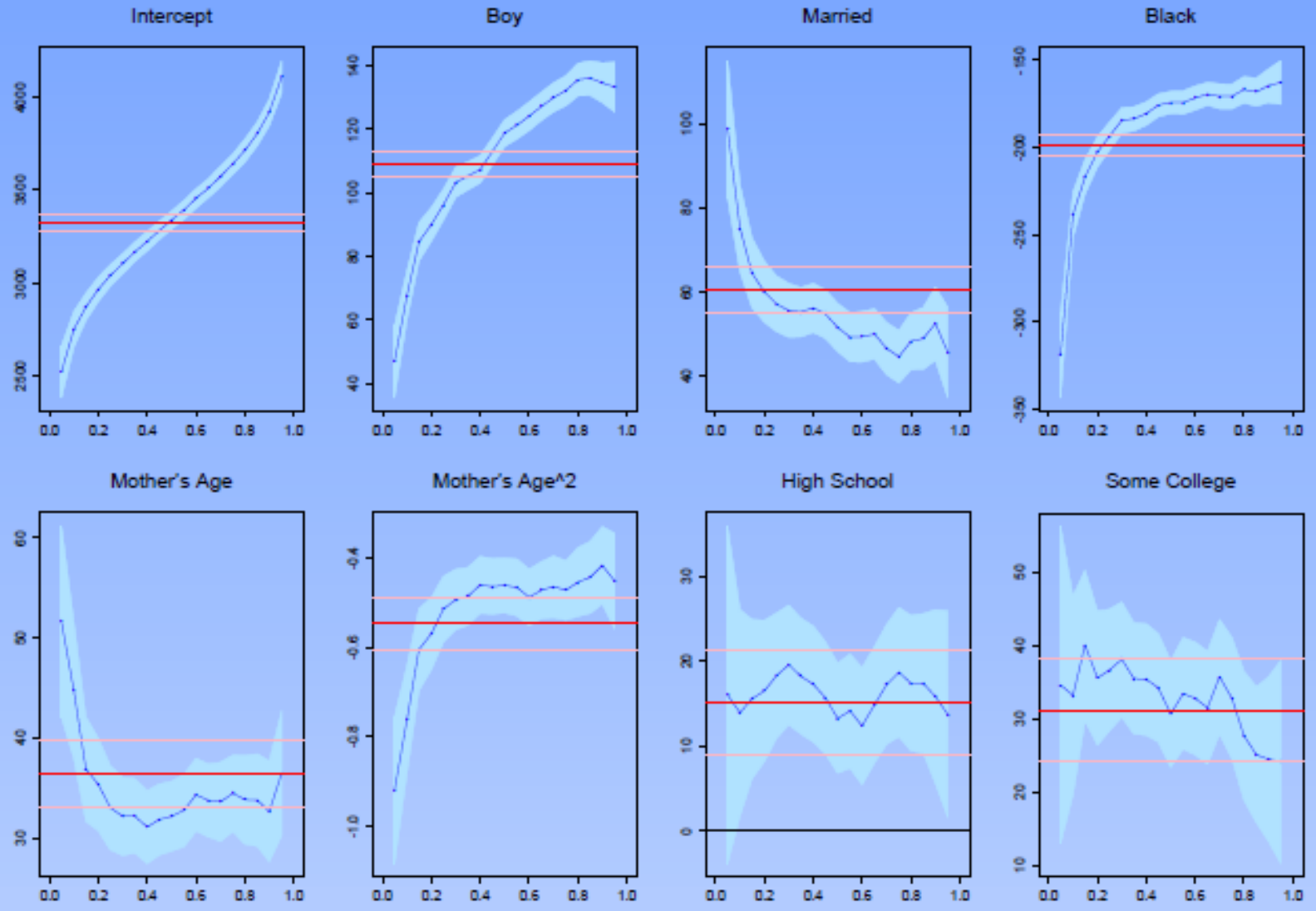


Figure 2: Engel Curves for Food: This figure plots data taken from Engel's (1857) study of the dependence of households' food expenditure on household income. Seven estimated quantile regression lines for $\tau \in \{.05, .1, .25, .5, .75, .9, .95\}$ are superimposed on the scatterplot. The median $\tau = .5$ fit is indicated by the darker solid line; the least squares estimate of the conditional mean function is indicated by the dashed line.

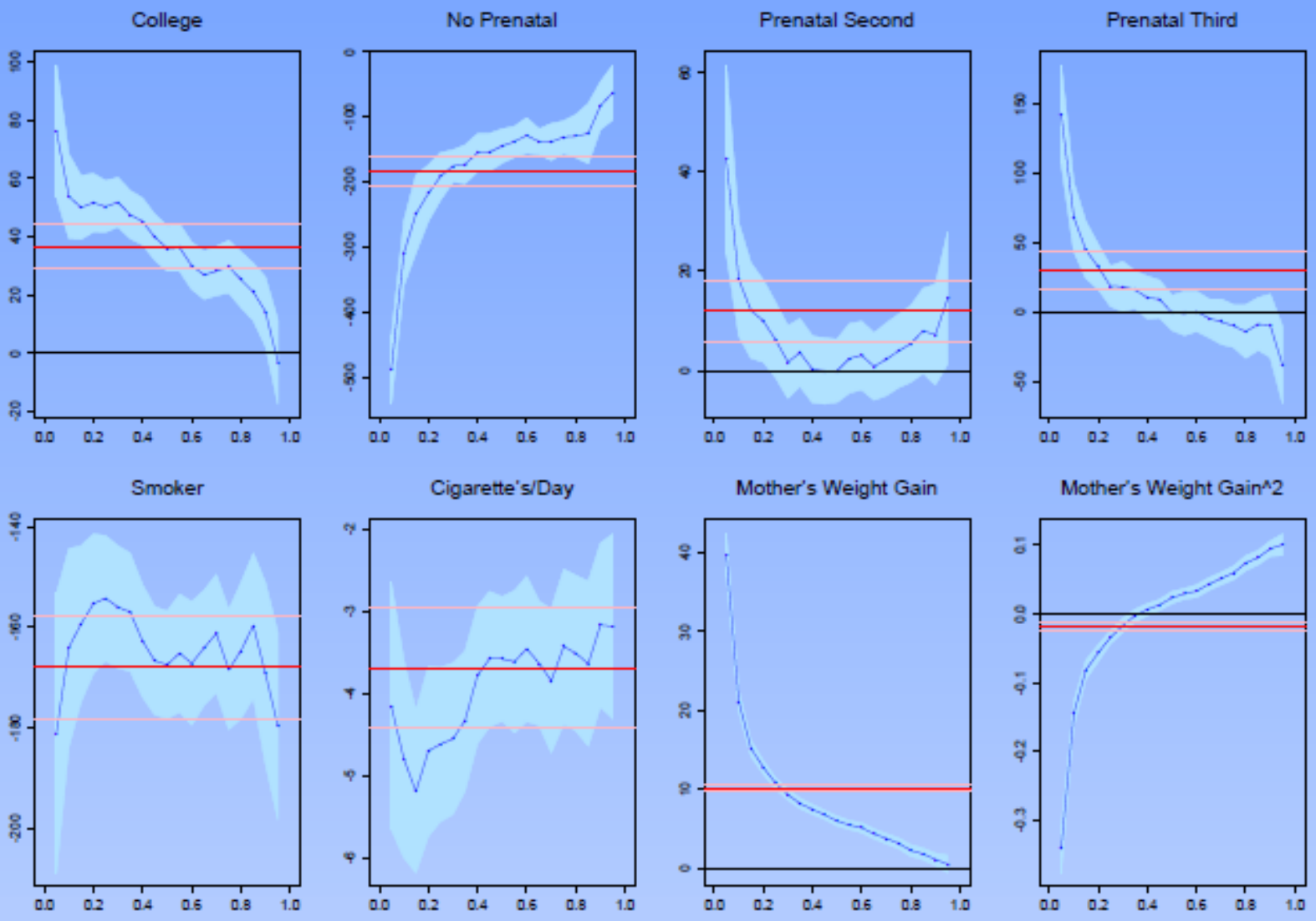
A Model of Infant Birthweight

- Reference: Abreveya (2001), Koenker and Hallock (2001)
- Data: June, 1997, Detailed Natality Data of the US. Live, singleton births, with mothers recorded as either black or white, between 18-45, and residing in the U.S. Sample size: 198,377.
- Response: Infant Birthweight (in grams)
- Covariates:
  - Mother's Education
  - Mother's Prenatal Care
  - Mother's Smoking
  - Mother's Age
  - Mother's Weight Gain

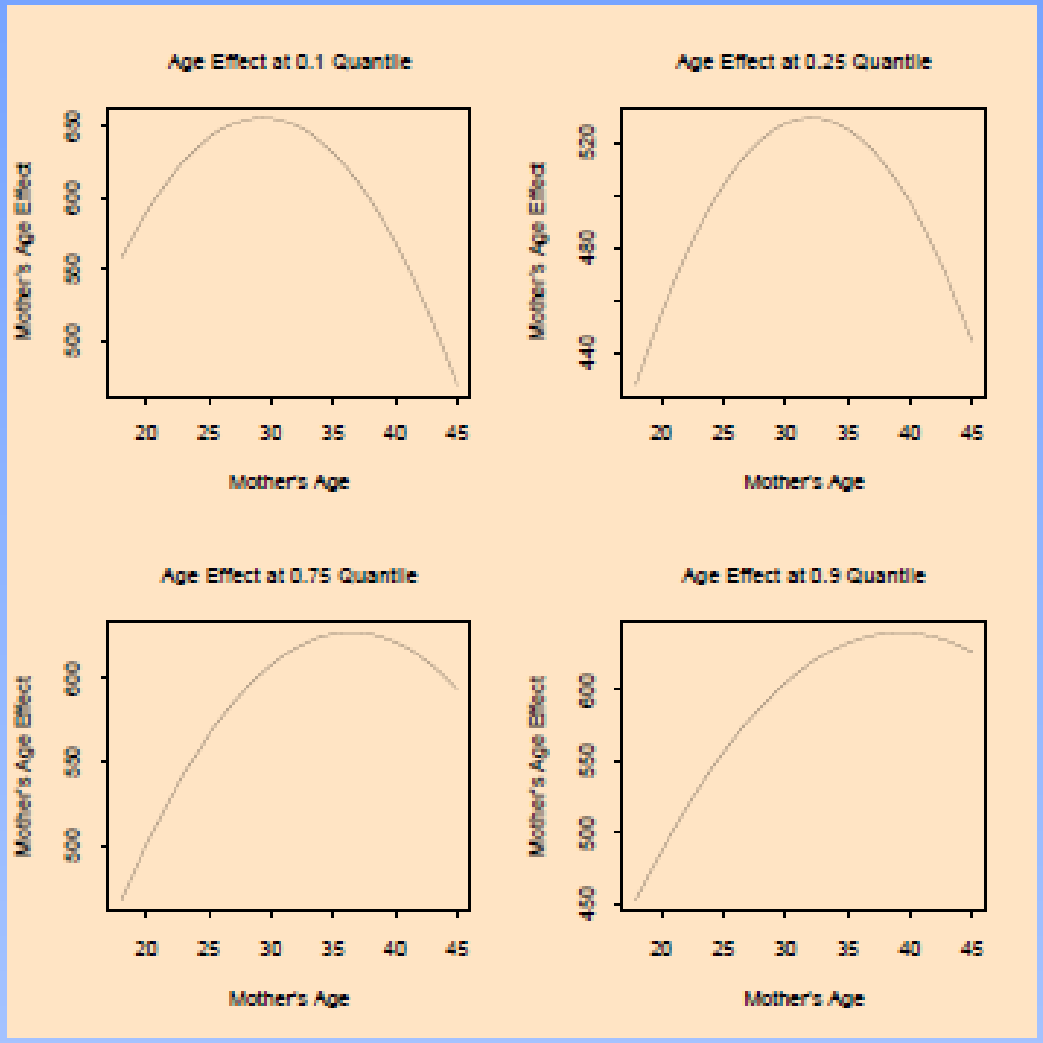
Quantile Regression Birthweight Model I



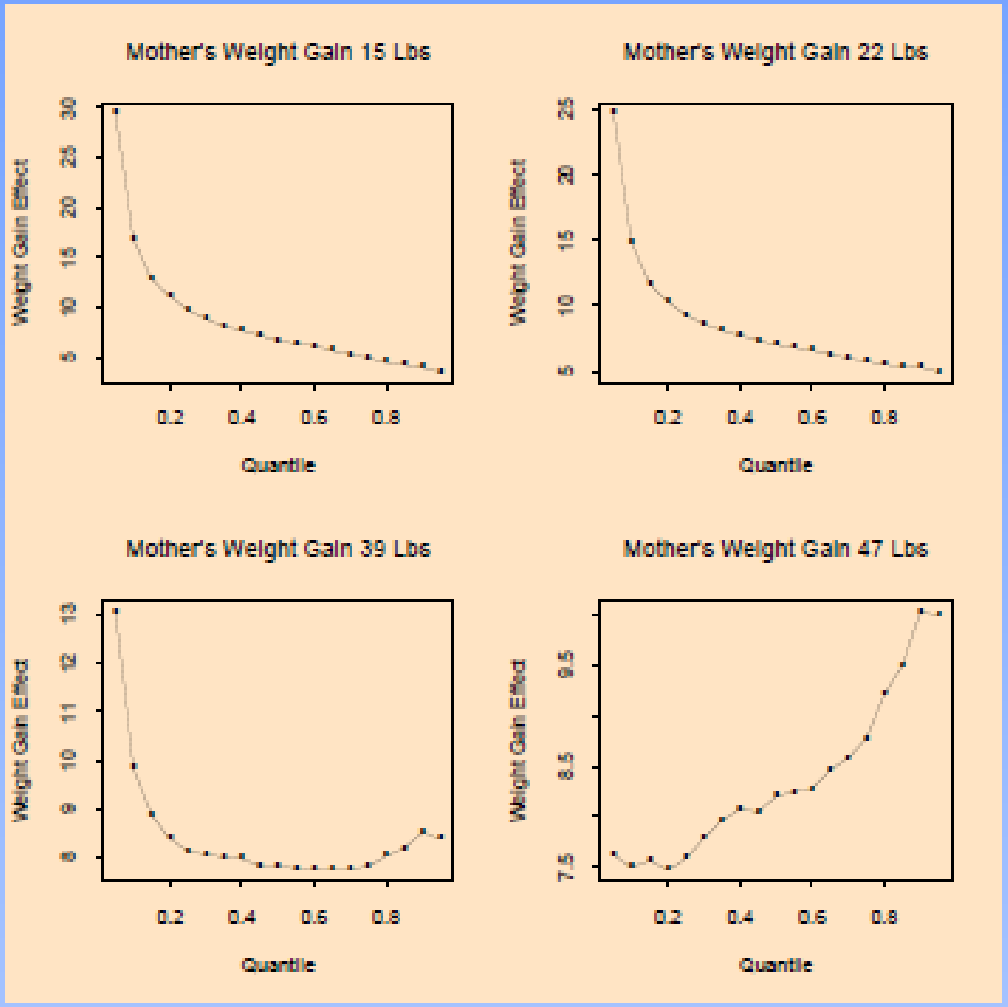
Quantile Regression Birthweight Model II



Marginal Effect of Mother's Age



Marginal Effect of Mother's Weight Gain



AR(1) Model of Melbourne Daily Temperature

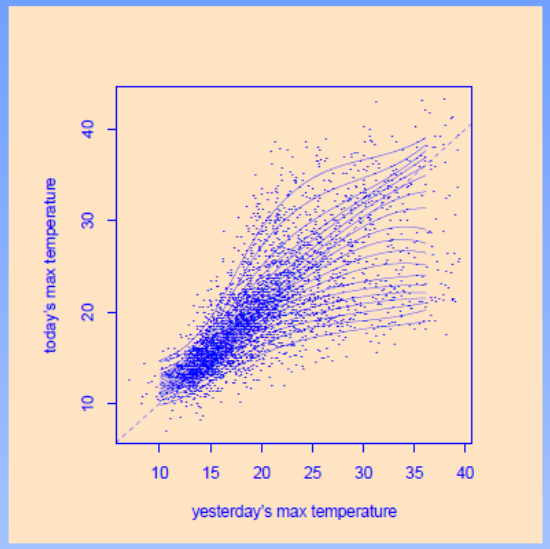
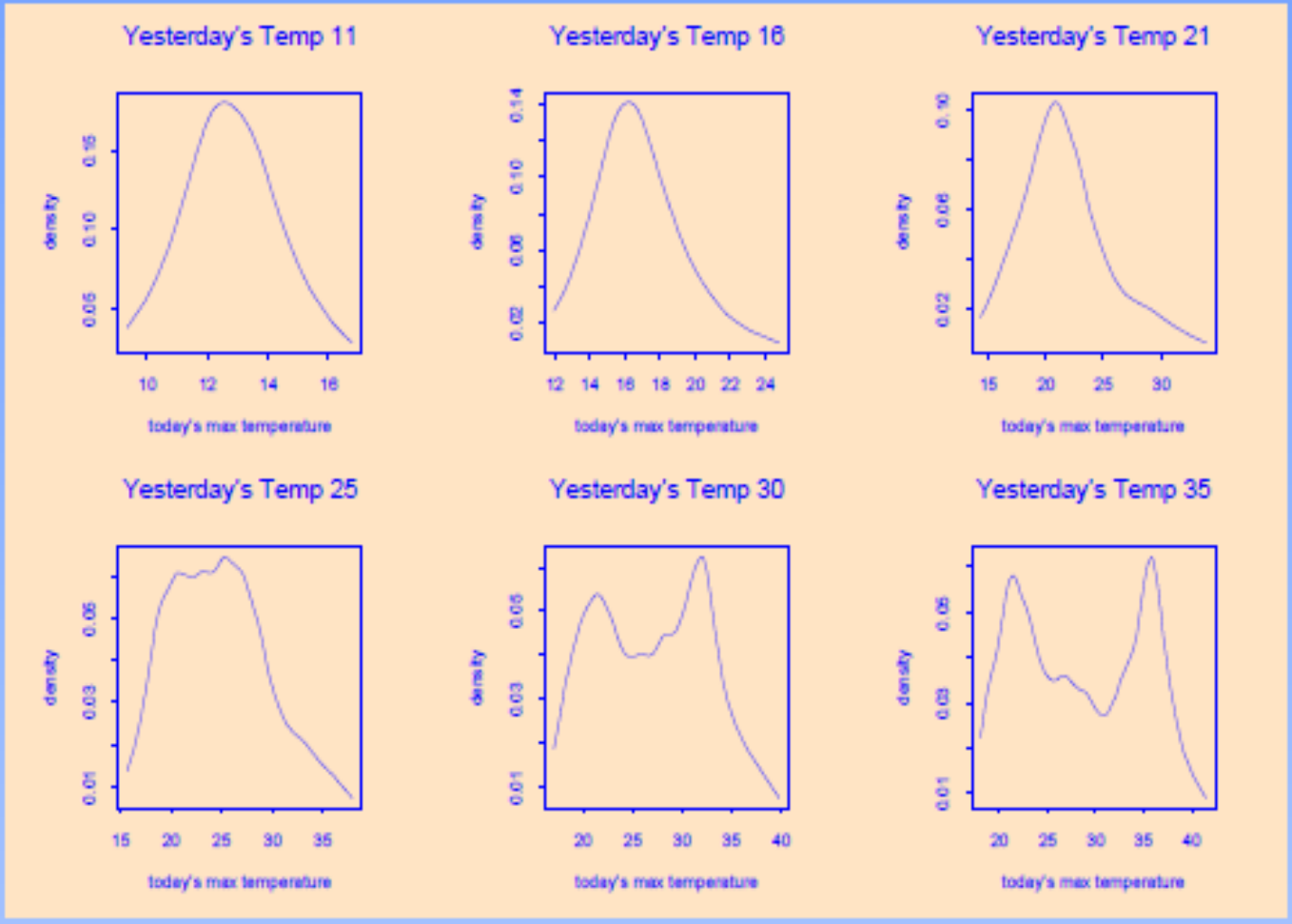


Figure 3: The plot illustrates 10 years of daily maximum temperature data for Melbourne, Australia as an AR(1) scatterplot. Superimposed are estimated conditional quantile functions for $\tau \in \{.05, .10, ..., .95\}$.

Conditional Densities of Melbourne Daily Temperature



Efeito Tratamento Quantílico

G-treat
F-control

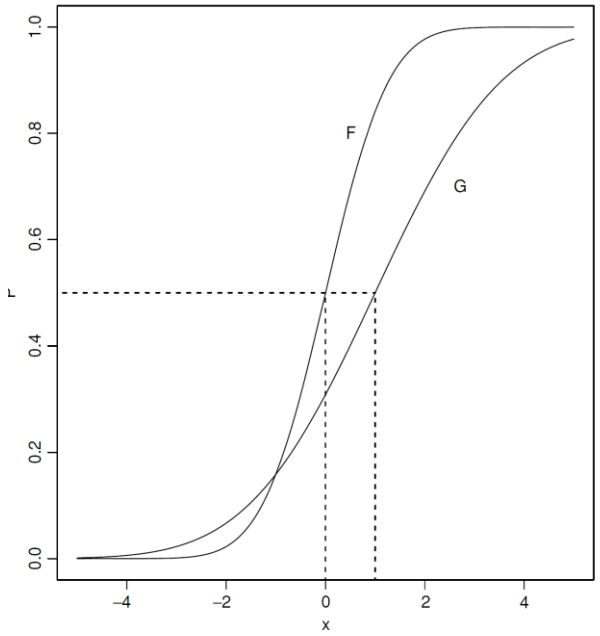
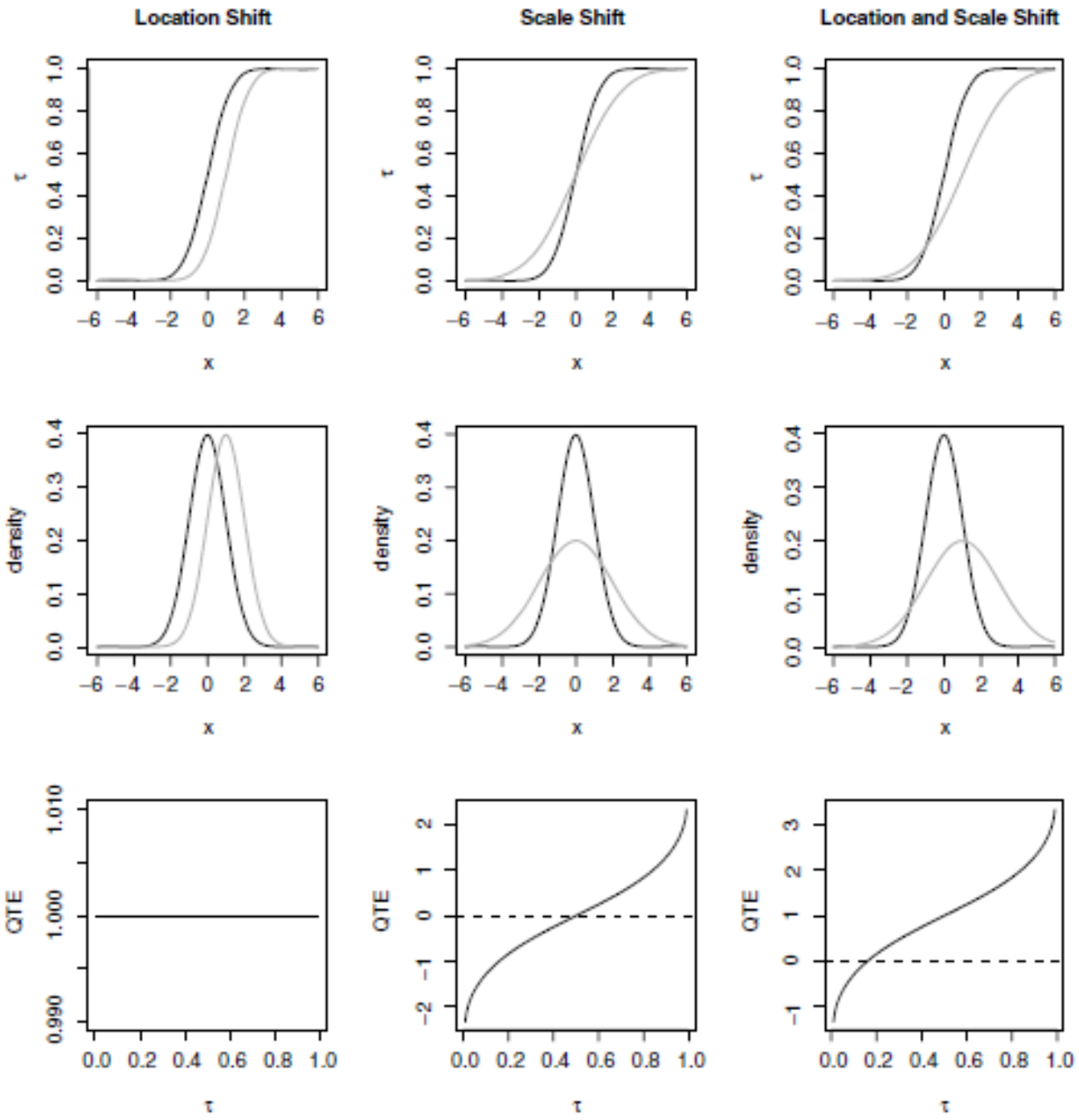


Figure 2.1. Lehmann quantile treatment effect. Horizontal distance between the treatment and control (marginal) distribution functions.



treat (cinza)

Control (preto)

Bibliografia

* Handbook of Quantile Regression (Chapman & Hall/CRC Handbooks of Modern Statistical Methods) 1st Edition, by Roger Koenker (Editor), Victor Chernozhukov (Editor), Xuming He (Editor), Limin Peng (Editor), 2018.

* Quantile Regression (Econometric Society Monographs) , by Roger Koenker (Author), 2005

* http://www.econ.uiuc.edu/~roger/courses

Obrigada!

maria.tannuri@gmail.com