Identification and Estimation of a Nonparametric Panel Data Model with Unobserved Heterogeneity

Kirill Evdokimov

Yale University

May, 2010

This paper is about:

- Unobserved Heterogeneity
- Nonparametric Panel Data Model
- Distribution of Heterogeneous Marginal Effects
- Nonparametric Identification and Estimation

Plan

- Model and Examples
- Related Literature
- Identification
- Estimation
- Empirical Illustration

Model

$$Y_{it} = m\left(X_{it}, \alpha_i\right) + U_{it}; \quad i = 1, \ldots, n;\ t = 1, \ldots, T$$

- i - individual, t - time period
- Observed: Y_{it} and X_{it}
- Unobserved: α_i - scalar heterogeneity, U_{it} - idiosyncratic shocks
- function $m\left(x, \alpha\right)$ is not known

- $\partial m\left(X_{it}, \alpha_i\right) / \partial x$ depends on α_i $\Rightarrow$ heterogeneous marginal effects
- will ID $m(x, \alpha)$ and $F_{\alpha_i}\left(\alpha | X_{it}\right)$
- distributions of the outcomes, counterfactuals
- "what percentage of the treated would be better/worse off?"

Model

$$Y_{it} = m\left(X_{it}, \alpha_i\right) + U_{it}; \quad i = 1, \ldots, n;\ t = 1, \ldots, T$$

- i - individual, t - time period
- Observed: Y_{it} and X_{it}
- Unobserved: α_i - scalar heterogeneity, U_{it} - idiosyncratic shocks
- function $m\left(x, \alpha\right)$ is not known

- $\partial m\left(X_{it}, \alpha_i\right) / \partial x$ depends on α_i $\Rightarrow$ heterogeneous marginal effects
- will ID $m(x, \alpha)$ and $F_{\alpha_i}\left(\alpha | X_{it}\right)$
- distributions of the outcomes, counterfactuals
- "what percentage of the treated would be better/worse off?"

Model

$$Y_{it} = m\left(X_{it}, \alpha_i\right) + U_{it}; \quad i = 1, \ldots, n;\ t = 1, \ldots, T$$

- i - individual, t - time period
- Observed: Y_{it} and X_{it}
- Unobserved: α_i - scalar heterogeneity, U_{it} - idiosyncratic shocks
- function $m\left(x, \alpha\right)$ is not known

- $\partial m\left(X_{it}, \alpha_i\right) / \partial x$ depends on α_i $\Rightarrow$ heterogeneous marginal effects
- will ID $m(x, \alpha)$ and $F_{\alpha_i}\left(\alpha | X_{it}\right)$
- distributions of the outcomes, counterfactuals
- "what percentage of the treated would be better/worse off?"

Model

$$Y_{it} = m(X_{it}, \alpha_i) + U_{it}; \quad i = 1, \ldots, n; \; t = 1, \ldots, T$$

- i - individual, t - time period
- Observed: Y_{it} and X_{it}
- Unobserved: α_i - scalar heterogeneity, U_{it} - idiosyncratic shocks
- function $m(x, \alpha)$ is not known

Some assumptions:

- $E[U_{it}|X_{it}] = 0$ and $U_{it} \perp \alpha_i | X_{it}$
- no restrictions on $F_{\alpha_i}(\alpha | X_{it} = x) \Rightarrow$ *fixed* effects
- no parametric assumptions; $T = 2$ is sufficient
- Y_{it} must be cont.; also, no lagged dep. variables

Motivation

Union Wage Premium

$$Y_{it} = m\left(X_{it}, \alpha_i\right) + U_{it}$$

- Y_{it} is (log)-wage,
- X_{it} is 1 if member of a union, 0 o/w,
- α_i is skill and U_{it} is luck

- $m(1,\alpha_i) - m(0,\alpha_i)$ is union wage premium
- When $m\left(X_{it}, \alpha_i\right) = X_{it}\beta + \alpha_i$, it becomes

$$m\left(1, \alpha_i\right) - m\left(0, \alpha_i\right) = \beta = const!$$

Motivation

Union Wage Premium

$$Y_{it} = m\left(X_{it}, \alpha_i\right) + U_{it}$$

- Y_{it} is (log)-wage,
- X_{it} is 1 if member of a union, 0 o/w,
- α_i is skill and U_{it} is luck

- $m(1,\alpha_i)-m(0,\alpha_i)$ is union wage premium may be NOT monotone in α_i
- When $m\left(X_{it}, \alpha_i\right) = X_{it}\beta + \alpha_i$, it becomes

$$m\left(1, \alpha_i\right) - m\left(0, \alpha_i\right) = \beta = const!$$

Motivation $(Y_{it} = m(X_{it}, \alpha_i) + U_{it})$

Life Cycle Models of Consumption and Labor Supply

- Heckman and MaCurdy (1980), MaCurdy (1981). Under the assumptions of these papers and $\rho = r$:

$$C_{it} = \underbrace{C(W_{it}, \lambda_i)}_{unknown\ fn.} + \underbrace{U_{it}}_{meas.err.}$$

  - C_{it} and W_{it} - (log) consumption and hourly wage
  - λ_i - Lagrange multiplier
- Parametric specification of utility $\Rightarrow$ additively separable λ_i
- Instead note that:
  - $C(w, \lambda)$ is unknown
  - $C(w, \lambda)$ is strictly increasing in λ
  - (unobserved) λ_i is likely to be correlated with W_{it}
  - can handle this nonparametrically

Motivation $(Y_{it} = m(X_{it}, \alpha_i) + U_{it})$

Life Cycle Models of Consumption and Labor Supply

- Heckman and MaCurdy (1980), MaCurdy (1981). Under the assumptions of these papers and $\rho = r$:

$$C_{it} = \underbrace{C(W_{it}, \lambda_i)}_{unknown\ \ fn.} + \underbrace{U_{it}}_{meas.err.}$$

  - C_{it} and W_{it} - (log) consumption and hourly wage
  - λ_i - Lagrange multiplier
- Parametric specification of utility $\Rightarrow$ additively separable λ_i
- Instead note that:
  - $C(w, \lambda)$ is unknown
  - $C(w, \lambda)$ is strictly increasing in λ
  - (unobserved) λ_i is likely to be correlated with W_{it}
  - can handle this nonparametrically

Motivation $(Y_{it} = m(X_{it}, \alpha_i) + U_{it})$

Life Cycle Models of Consumption and Labor Supply

- Heckman and MaCurdy (1980), MaCurdy (1981). Under the assumptions of these papers and $\rho = r$:

$$C_{it} = \underbrace{C(W_{it}, \lambda_i)}_{unknown\ fn.} + \underbrace{U_{it}}_{meas.err.}$$

  - C_{it} and W_{it} - (log) consumption and hourly wage
  - λ_i - Lagrange multiplier
- Parametric specification of utility $\Rightarrow$ additively separable λ_i
- Instead note that:
  - $C(w, \lambda)$ is unknown
  - $C(w, \lambda)$ is strictly increasing in λ
  - (unobserved) λ_i is likely to be correlated with W_{it}
  - can handle this nonparametrically

Motivation ($Y_{it} = m(X_{it}, \alpha_i) + U_{it}$)

Life Cycle Models of Consumption and Labor Supply

- Heckman and MaCurdy (1980), MaCurdy (1981). Under the assumptions of these papers and $\rho = r$:

$$C_{it} = \underbrace{C(W_{it}, \lambda_i)}_{unknown\ fn.} + \underbrace{U_{it}}_{meas.err.}$$

  - C_{it} and W_{it} - (log) consumption and hourly wage
  - λ_i - Lagrange multiplier
- Parametric specification of utility $\Rightarrow$ additively separable λ_i
- Instead note that:
  - $C(w, \lambda)$ is unknown
  - $C(w, \lambda)$ is strictly increasing in λ
  - (unobserved) λ_i is likely to be correlated with W_{it}
  - can handle this nonparametrically

Related Literature

Semiparametric Models:

- Arellano & Hahn (2006), Hahn & Newey (2004), Woutersen (2002) - bias reduction
- Honore & Tamer (2006), Chernozhukov et al. (2006) - set ID

Correlated Random Effects (exchangeability or index restrictions):

- Altonji & Matzkin (2005), Bester & Hansen (2009) - avg. derivative
- Altonji & Matzkin (2005) - struct. fn, but scalar unobs.

Related papers:

- Kitamura (2004) - finite mixture
- Horowitz & Markatou (1996) - linear panel, deconvolution

Related Literature

Semiparametric Models:

- Arellano & Hahn (2006), Hahn & Newey (2004), Woutersen (2002) - bias reduction
- Honore & Tamer (2006), Chernozhukov et al. (2006) - set ID

Correlated Random Effects (exchangeability or index restrictions):

- Altonji & Matzkin (2005), Bester & Hansen (2009) - avg. derivative
- Altonji & Matzkin (2005) - struct. fn, but scalar unobs.

Related papers:

- Kitamura (2004) - finite mixture
- Horowitz & Markatou (1996) - linear panel, deconvolution

Related Literature

Semiparametric Models:

- Arellano & Hahn (2006), Hahn & Newey (2004), Woutersen (2002) - bias reduction
- Honore & Tamer (2006), Chernozhukov et al. (2006) - set ID

Correlated Random Effects (exchangeability or index restrictions):

- Altonji & Matzkin (2005), Bester & Hansen (2009) - avg. derivative
- Altonji & Matzkin (2005) - struct. fn, but scalar unobs.

Related papers:

- Kitamura (2004) - finite mixture
- Horowitz & Markatou (1996) - linear panel, deconvolution

Related Literature

Semiparametric Models:

- Arellano & Hahn (2006), Hahn & Newey (2004), Woutersen (2002) - bias reduction
- Honore & Tamer (2006), Chernozhukov et al. (2006) - set ID

Correlated Random Effects (exchangeability or index restrictions):

- Altonji & Matzkin (2005), Bester & Hansen (2009) - avg. derivative
- Altonji & Matzkin (2005) - struct. fn, but scalar unobs.

Related papers:

- Kitamura (2004) - finite mixture
- Horowitz & Markatou (1996) - linear panel, deconvolution

Related Literature

Semiparametric Models:

- Arellano & Hahn (2006), Hahn & Newey (2004), Woutersen (2002) - bias reduction
- Honore & Tamer (2006), Chernozhukov et al. (2006) - set ID

Correlated Random Effects (exchangeability or index restrictions):

- Altonji & Matzkin (2005), Bester & Hansen (2009) - avg. derivative
- Altonji & Matzkin (2005) - struct. fn, but scalar unobs.

Related papers:

- Kitamura (2004) - finite mixture
- Horowitz & Markatou (1996) - linear panel, deconvolution

Related Literature

Semiparametric Models:

- Arellano & Hahn (2006), Hahn & Newey (2004), Woutersen (2002) - bias reduction
- Honore & Tamer (2006), Chernozhukov et al. (2006) - set ID

Correlated Random Effects (exchangeability or index restrictions):

- Altonji & Matzkin (2005), Bester & Hansen (2009) - avg. derivative
- Altonji & Matzkin (2005) - struct. fn, but scalar unobs.

Related papers:

- Kitamura (2004) - finite mixture
- Horowitz & Markatou (1996) - linear panel, deconvolution

Related Literature

Semiparametric Models:

- Arellano & Hahn (2006), Hahn & Newey (2004), Woutersen (2002) - bias reduction
- Honore & Tamer (2006), Chernozhukov et al. (2006) - set ID

Correlated Random Effects (exchangeability or index restrictions):

- Altonji & Matzkin (2005), Bester & Hansen (2009) - avg. derivative
- Altonji & Matzkin (2005) - struct. fn, but scalar unobs.

Related papers:

- Kitamura (2004) - finite mixture
- Horowitz & Markatou (1996) - linear panel, deconvolution

Related Literature - 2 (Very recent)

Linear:

- Arellano & Bonhomme (2008) - linear, distr. of coeffs
- Graham & Powell (2008) - linear, means of coeffs
- Graham & Hahn & Powell (2009) - linear, quantile effects

General Nonlinear:

- Chernozhukov & Fernandez-Val & Newey (2009) - set ID, effects
- Hoderlein & White (2009) - cont. covar., effects

Also:

- Evdokimov (2009) - time-varying struct. fn., full ID

Related Literature - 2 (Very recent)

Linear:

- Arellano & Bonhomme (2008) - linear, distr. of coeffs
- Graham & Powell (2008) - linear, means of coeffs
- Graham & Hahn & Powell (2009) - linear, quantile effects

General Nonlinear:

- Chernozhukov & Fernandez-Val & Newey (2009) - set ID, effects
- Hoderlein & White (2009) - cont. covar., effects

Also:

- Evdokimov (2009) - time-varying struct. fn., full ID

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

1. Subpopulation $X_{i1} = X_{i2} = x$: then
$m(X_{i1}, \alpha_i) = m(X_{i2}, \alpha_i) = m(x, \alpha_i)$
$$\begin{aligned} Y_{i1} &= m(x, \alpha_i) + U_{i1} \\ Y_{i2} &= m(x, \alpha_i) + U_{i2} \\ Y_{i1} - Y_{i2} &= U_{i1} - U_{i2} \implies \text{Obtain distr } \mathcal{L}(U_{it} | X_{it} = x) \end{aligned}$$
2. Deconvolve: $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$, $\alpha_i \perp U_{it}$ given $X_{it} = x$, thus
$$\underbrace{\mathcal{L}(Y_{it} | X_{it} = x)}_{\text{from data}} = \mathcal{L}(m(x, \alpha_i) | X_{it} = x) * \underbrace{\mathcal{L}(U_{it} | X_{it} = x)}_{\text{from step 1}},$$
where "$*$" means "convolution" and "$\mathcal{L}$" means "probability $\mathcal{L}$aw". Then obtain $\mathcal{L}(m(x, \alpha_i) | X_{it} = x)$ by deconvolution
3. Use $\mathcal{L}(m(x, \alpha_i) | X_{it} = x)$, eg. quantiles $Q_{m(x,\alpha_i)|X_{it}}(q|x)$ [for nonparametric between- or within- variation]

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

1. Subpopulation $X_{i1} = X_{i2} = x$: then
$m(X_{i1}, \alpha_i) = m(X_{i2}, \alpha_i) = m(x, \alpha_i)$
$$\begin{aligned} Y_{i1} &= m(x, \alpha_i) + U_{i1} \\ Y_{i2} &= m(x, \alpha_i) + U_{i2} \\ Y_{i1} - Y_{i2} &= U_{i1} - U_{i2} \implies \text{Obtain distr } \mathcal{L}(U_{it} | X_{it} = x) \end{aligned}$$
2. Deconvolve: $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$, $\alpha_i \perp U_{it}$ given $X_{it} = x$, thus
$$\underbrace{\mathcal{L}(Y_{it} | X_{it} = x)}_{\text{from data}} = \mathcal{L}(m(x, \alpha_i) | X_{it} = x) * \underbrace{\mathcal{L}(U_{it} | X_{it} = x)}_{\text{from step 1}},$$
where "$*$" means "convolution" and "$\mathcal{L}$" means "probability $\mathcal{L}$aw". Then obtain $\mathcal{L}(m(x, \alpha_i) | X_{it} = x)$ by deconvolution
3. Use $\mathcal{L}(m(x, \alpha_i) | X_{it} = x)$, eg. quantiles $Q_{m(x,\alpha_i)|X_{it}}(q|x)$ [for nonparametric between- or within- variation]

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

1. Subpopulation $X_{i1} = X_{i2} = x$: then
$m(X_{i1}, \alpha_i) = m(X_{i2}, \alpha_i) = m(x, \alpha_i)$
$$\begin{aligned} Y_{i1} &= m(x, \alpha_i) + U_{i1} \\ Y_{i2} &= m(x, \alpha_i) + U_{i2} \\ Y_{i1} - Y_{i2} &= U_{i1} - U_{i2} \implies \text{Obtain distr } \mathcal{L}(U_{it} | X_{it} = x) \end{aligned}$$
2. Deconvolve: $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$, $\alpha_i \perp U_{it}$ given $X_{it} = x$, thus
$$\underbrace{\mathcal{L}(Y_{it} | X_{it} = x)}_{\text{from data}} = \mathcal{L}(m(x, \alpha_i) | X_{it} = x) * \underbrace{\mathcal{L}(U_{it} | X_{it} = x)}_{\text{from step 1}},$$
where "$*$" means "convolution" and "$\mathcal{L}$" means "probability $\mathcal{L}$aw". Then obtain $\mathcal{L}(m(x, \alpha_i) | X_{it} = x)$ by deconvolution
3. Use $\mathcal{L}(m(x, \alpha_i) | X_{it} = x)$, eg. quantiles $Q_{m(x,\alpha_i)|X_{it}}(q|x)$ [for nonparametric between- or within- variation]

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

1. Subpopulation $X_{i1} = X_{i2} = x$: then
$m(X_{i1}, \alpha_i) = m(X_{i2}, \alpha_i) = m(x, \alpha_i)$
$$\begin{aligned} Y_{i1} &= m(x, \alpha_i) + U_{i1} \\ Y_{i2} &= m(x, \alpha_i) + U_{i2} \\ Y_{i1} - Y_{i2} &= U_{i1} - U_{i2} \implies \text{Obtain distr } \mathcal{L}(U_{it} | X_{it} = x) \end{aligned}$$
2. Deconvolve: $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$, $\alpha_i \perp U_{it}$ given $X_{it} = x$, thus
$$\underbrace{\mathcal{L}(Y_{it} | X_{it} = x)}_{\text{from data}} = \mathcal{L}(m(x, \alpha_i) | X_{it} = x) * \underbrace{\mathcal{L}(U_{it} | X_{it} = x)}_{\text{from step 1}},$$
where "$*$" means "convolution" and "$\mathcal{L}$" means "probability $\mathcal{L}$aw". Then obtain $\mathcal{L}(m(x, \alpha_i) | X_{it} = x)$ by deconvolution
3. Use $\mathcal{L}(m(x, \alpha_i) | X_{it} = x)$, eg. quantiles $Q_{m(x,\alpha_i)|X_{it}}(q|x)$ [for nonparametric between- or within- variation]

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

1. Subpopulation $X_{i1} = X_{i2} = x$: then
$m(X_{i1}, \alpha_i) = m(X_{i2}, \alpha_i) = m(x, \alpha_i)$
$$\begin{aligned} Y_{i1} &= m(x, \alpha_i) + U_{i1} \\ Y_{i2} &= m(x, \alpha_i) + U_{i2} \\ Y_{i1} - Y_{i2} &= U_{i1} - U_{i2} \implies \text{Obtain distr } \mathcal{L}(U_{it} | X_{it} = x) \end{aligned}$$
2. Deconvolve: $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$, $\alpha_i \perp U_{it}$ given $X_{it} = x$, thus
$$\underbrace{\mathcal{L}(Y_{it} | X_{it} = x)}_{\text{from data}} = \mathcal{L}(m(x, \alpha_i) | X_{it} = x) * \underbrace{\mathcal{L}(U_{it} | X_{it} = x)}_{\text{from step 1}},$$
where "$*$" means "convolution" and "$\mathcal{L}$" means "probability $\mathcal{L}$aw". Then obtain $\mathcal{L}(m(x, \alpha_i) | X_{it} = x)$ by deconvolution
3. Use $\mathcal{L}(m(x, \alpha_i) | X_{it} = x)$, eg. quantiles $Q_{m(x,\alpha_i)|X_{it}}(q|x)$ [for nonparametric between- or within- variation]

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

1. Subpopulation $X_{i1} = X_{i2} = x$: then
$m(X_{i1}, \alpha_i) = m(X_{i2}, \alpha_i) = m(x, \alpha_i)$
$$\begin{aligned} Y_{i1} &= m(x, \alpha_i) + U_{i1} \\ Y_{i2} &= m(x, \alpha_i) + U_{i2} \\ Y_{i1} - Y_{i2} &= U_{i1} - U_{i2} \implies \text{Obtain distr } \mathcal{L}(U_{it} | X_{it} = x) \end{aligned}$$
2. Deconvolve: $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$, $\alpha_i \perp U_{it}$ given $X_{it} = x$, thus
$$\underbrace{\mathcal{L}(Y_{it} | X_{it} = x)}_{\text{from data}} = \mathcal{L}(m(x, \alpha_i) | X_{it} = x) * \underbrace{\mathcal{L}(U_{it} | X_{it} = x)}_{\text{from step 1}},$$
where "$*$" means "convolution" and "$\mathcal{L}$" means "probability $\mathcal{L}$aw". Then obtain $\mathcal{L}(m(x, \alpha_i) | X_{it} = x)$ by deconvolution
3. Use $\mathcal{L}(m(x, \alpha_i) | X_{it} = x)$, eg. quantiles $Q_{m(x,\alpha_i)|X_{it}}(q|x)$ [for nonparametric between- or within- variation]

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

1. Subpopulation $X_{i1} = X_{i2} = x$: then
$m(X_{i1}, \alpha_i) = m(X_{i2}, \alpha_i) = m(x, \alpha_i)$
$$\begin{aligned} Y_{i1} &= m(x, \alpha_i) + U_{i1} \\ Y_{i2} &= m(x, \alpha_i) + U_{i2} \\ Y_{i1} - Y_{i2} &= U_{i1} - U_{i2} \implies \text{Obtain distr } \mathcal{L}(U_{it} | X_{it} = x) \end{aligned}$$
2. Deconvolve: $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$, $\alpha_i \perp U_{it}$ given $X_{it} = x$, thus
$$\underbrace{\mathcal{L}(Y_{it} | X_{it} = x)}_{\text{from data}} = \mathcal{L}(m(x, \alpha_i) | X_{it} = x) * \underbrace{\mathcal{L}(U_{it} | X_{it} = x)}_{\text{from step 1}},$$
where "$*$" means "convolution" and "$\mathcal{L}$" means "probability $\mathcal{L}$aw".
Then obtain $\mathcal{L}(m(x, \alpha_i) | X_{it} = x)$ by deconvolution
3. Use $\mathcal{L}(m(x, \alpha_i) | X_{it} = x)$, eg. quantiles $Q_{m(x,\alpha_i)|X_{it}}(q|x)$ [for nonparametric between- or within- variation]

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

1. Subpopulation $X_{i1} = X_{i2} = x$: then
$m(X_{i1}, \alpha_i) = m(X_{i2}, \alpha_i) = m(x, \alpha_i)$
$$\begin{aligned} Y_{i1} &= m(x, \alpha_i) + U_{i1} \\ Y_{i2} &= m(x, \alpha_i) + U_{i2} \\ Y_{i1} - Y_{i2} &= U_{i1} - U_{i2} \implies \text{Obtain distr } \mathcal{L}(U_{it} | X_{it} = x) \end{aligned}$$
2. Deconvolve: $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$, $\alpha_i \perp U_{it}$ given $X_{it} = x$, thus
$$\underbrace{\mathcal{L}(Y_{it} | X_{it} = x)}_{\text{from data}} = \mathcal{L}(m(x, \alpha_i) | X_{it} = x) * \underbrace{\mathcal{L}(U_{it} | X_{it} = x)}_{\text{from step 1}},$$
where "$*$" means "convolution" and "$\mathcal{L}$" means "probability $\mathcal{L}$aw". Then obtain $\mathcal{L}(m(x, \alpha_i) | X_{it} = x)$ by deconvolution
3. Use $\mathcal{L}(m(x, \alpha_i) | X_{it} = x)$, eg. quantiles $Q_{m(x,\alpha_i)|X_{it}}(q|x)$ [for nonparametric between- or within- variation]

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

1. Subpopulation $X_{i1} = X_{i2} = x$: then
$m(X_{i1}, \alpha_i) = m(X_{i2}, \alpha_i) = m(x, \alpha_i)$
$$\begin{aligned} Y_{i1} &= m(x, \alpha_i) + U_{i1} \\ Y_{i2} &= m(x, \alpha_i) + U_{i2} \\ Y_{i1} - Y_{i2} &= U_{i1} - U_{i2} \implies \text{Obtain distr } \mathcal{L}(U_{it} | X_{it} = x) \end{aligned}$$
2. Deconvolve: $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$, $\alpha_i \perp U_{it}$ given $X_{it} = x$, thus
$$\underbrace{\mathcal{L}(Y_{it} | X_{it} = x)}_{\text{from data}} = \mathcal{L}(m(x, \alpha_i) | X_{it} = x) * \underbrace{\mathcal{L}(U_{it} | X_{it} = x)}_{\text{from step 1}},$$
where "$*$" means "convolution" and "$\mathcal{L}$" means "probability $\mathcal{L}$aw".
Then obtain $\mathcal{L}(m(x, \alpha_i) | X_{it} = x)$ by deconvolution
3. Use $\mathcal{L}(m(x, \alpha_i) | X_{it} = x)$, eg. quantiles $Q_{m(x,\alpha_i)|X_{it}}(q|x)$ [for nonparametric between- or within- variation]

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

1. Subpopulation $X_{i1} = X_{i2} = x$: then
$m(X_{i1}, \alpha_i) = m(X_{i2}, \alpha_i) = m(x, \alpha_i)$
$$\begin{aligned} Y_{i1} &= m(x, \alpha_i) + U_{i1} \\ Y_{i2} &= m(x, \alpha_i) + U_{i2} \\ Y_{i1} - Y_{i2} &= U_{i1} - U_{i2} \implies \text{Obtain distr } \mathcal{L}(U_{it} | X_{it} = x) \end{aligned}$$
2. Deconvolve: $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$, $\alpha_i \perp U_{it}$ given $X_{it} = x$, thus
$$\underbrace{\mathcal{L}(Y_{it} | X_{it} = x)}_{\text{from data}} = \mathcal{L}(m(x, \alpha_i) | X_{it} = x) * \underbrace{\mathcal{L}(U_{it} | X_{it} = x)}_{\text{from step 1}},$$
where "$*$" means "convolution" and "$\mathcal{L}$" means "probability $\mathcal{L}$aw". Then obtain $\mathcal{L}(m(x, \alpha_i) | X_{it} = x)$ by deconvolution
3. Use $\mathcal{L}(m(x, \alpha_i) | X_{it} = x)$, eg. quantiles $Q_{m(x,\alpha_i)|X_{it}}(q|x)$ [for nonparametric between- or within- variation]

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

1. Subpopulation $X_{i1} = X_{i2} = x$: then
$m(X_{i1}, \alpha_i) = m(X_{i2}, \alpha_i) = m(x, \alpha_i)$
$$\begin{aligned} Y_{i1} &= m(x, \alpha_i) + U_{i1} \\ Y_{i2} &= m(x, \alpha_i) + U_{i2} \\ Y_{i1} - Y_{i2} &= U_{i1} - U_{i2} \implies \text{Obtain distr } \mathcal{L}(U_{it} | X_{it} = x) \end{aligned}$$
2. Deconvolve: $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$, $\alpha_i \perp U_{it}$ given $X_{it} = x$, thus
$$\underbrace{\mathcal{L}(Y_{it} | X_{it} = x)}_{\text{from data}} = \mathcal{L}(m(x, \alpha_i) | X_{it} = x) * \underbrace{\mathcal{L}(U_{it} | X_{it} = x)}_{\text{from step 1}},$$
where "$*$" means "convolution" and "$\mathcal{L}$" means "probability $\mathcal{L}$aw". Then obtain $\mathcal{L}(m(x, \alpha_i) | X_{it} = x)$ by deconvolution
3. Use $\mathcal{L}(m(x, \alpha_i) | X_{it} = x)$, eg. quantiles $Q_{m(x,\alpha_i)|X_{it}}(q|x)$ [for nonparametric between- or within- variation]

Theorem (Identification)

Consider the model

$$Y_{it} = m(X_{it}, \alpha_i) + U_{it},$$

and suppose that

1. $f_{X_{i1},X_{i2}}(x,x) > 0$ *for all* x
2. $m(x,\alpha)$ *is strictly increasing in* α
3. $\mathcal{L}(U_{i1}|X_{i1},\alpha_i,X_{i2},U_{i2}) = \mathcal{L}(U_{i1}|X_{i1})$ *(and similarly for* U_{i2}*)*
4. $E[U_{it}|X_{it}] = 0$ *and some other assumptions...*

Then, model (1) *is identified, i.e.* $m(x,\alpha)$, $F_{\alpha_i}(\alpha|X_{it}=x)$, $F_{U_{it}}(u|X_{it}=x)$ *are identified.*

Theorem (Identification)

Consider the model

$$Y_{it} = m(X_{it}, \alpha_i) + U_{it},$$

and suppose that

1. $f_{X_{i1},X_{i2}}(x, x) > 0$ *for all* x
2. $m(x, \alpha)$ *is strictly increasing in* α
3. $\mathcal{L}(U_{i1}|X_{i1}, \alpha_i, X_{i2}, U_{i2}) = \mathcal{L}(U_{i1}|X_{i1})$ *(and similarly for* U_{i2}*)*
4. $E[U_{it}|X_{it}] = 0$ *and some other assumptions...*

Then, model (1) *is identified, i.e.* $m(x, \alpha)$, $F_{\alpha_i}(\alpha|X_{it} = x)$, $F_{U_{it}}(u|X_{it} = x)$ *are identified.*

Theorem (Identification)

Consider the model

$$Y_{it} = m(X_{it}, \alpha_i) + U_{it},$$

and suppose that

1. $f_{X_{i1},X_{i2}}(x,x) > 0$ *for all* x
2. $m(x,\alpha)$ *is strictly increasing in* α
3. $\mathcal{L}(U_{i1}|X_{i1}, \alpha_i, X_{i2}, U_{i2}) = \mathcal{L}(U_{i1}|X_{i1})$ *(and similarly for* U_{i2}*)*
4. $E[U_{it}|X_{it}] = 0$ *and some other assumptions...*

Then, model (1) *is identified, i.e.* $m(x,\alpha)$, $F_{\alpha_i}(\alpha|X_{it} = x)$, $F_{U_{it}}(u|X_{it} = x)$ *are identified.*

Theorem (Identification)

Consider the model

$$Y_{it} = m(X_{it}, \alpha_i) + U_{it},$$

and suppose that

1. $f_{X_{i1},X_{i2}}(x,x) > 0$ *for all* x
2. $m(x,\alpha)$ *is strictly increasing in* α
3. $\mathcal{L}(U_{i1}|X_{i1},\alpha_i,X_{i2},U_{i2}) = \mathcal{L}(U_{i1}|X_{i1})$ *(and similarly for* U_{i2}*)*
4. $E[U_{it}|X_{it}] = 0$ *and some other assumptions...*

Then, model (1) *is identified, i.e.* $m(x,\alpha)$, $F_{\alpha_i}(\alpha|X_{it}=x)$, $F_{U_{it}}(u|X_{it}=x)$ *are identified.*

Theorem (Identification)

Consider the model

$$Y_{it} = m(X_{it}, \alpha_i) + U_{it},$$

and suppose that

1. $f_{X_{i1},X_{i2}}(x,x) > 0$ *for all* x
2. $m(x,\alpha)$ *is strictly increasing in* α
3. $\mathcal{L}(U_{i1}|X_{i1}, \alpha_i, X_{i2}, U_{i2}) = \mathcal{L}(U_{i1}|X_{i1})$ *(and similarly for* U_{i2}*)*
4. $E[U_{it}|X_{it}] = 0$ *and some other assumptions...*

Then, model (1) *is identified, i.e.* $m(x,\alpha)$, $F_{\alpha_i}(\alpha|X_{it} = x)$, $F_{U_{it}}(u|X_{it} = x)$ *are identified.*

Theorem (Identification)

Consider the model

$$Y_{it} = m\left(X_{it}, \alpha_i\right) + U_{it},$$

and suppose that

1. $f_{X_{i1},X_{i2}}\left(x, x\right) > 0$ *for all* x
2. $m\left(x, \alpha\right)$ *is strictly increasing in* α
3. $\mathcal{L}\left(U_{i1}|X_{i1}, \alpha_i, X_{i2}, U_{i2}\right) = \mathcal{L}\left(U_{i1}|X_{i1}\right)$ *(and similarly for* U_{i2}*)*
4. $E\left[U_{it}|X_{it}\right] = 0$ *and some other assumptions...*

Then, model (1) *is identified, i.e.* $m\left(x, \alpha\right)$, $F_{\alpha_i}\left(\alpha|X_{it} = x\right)$, $F_{U_{it}}\left(u|X_{it} = x\right)$ *are identified.*

$$U_{it} = \sigma_t(X_{it}) \cdot \epsilon_{it}, \quad \epsilon_{it} \sim i.i.d(0,1)$$

Theorem (Identification)

Consider the model

$$Y_{it} = m(X_{it}, \alpha_i) + U_{it},$$

and suppose that

1. $f_{X_{i1},X_{i2}}(x,x) > 0$ *for all* x
2. $m(x,\alpha)$ *is strictly increasing in* α
3. $\mathcal{L}(U_{i1}|X_{i1},\alpha_i,X_{i2},U_{i2}) = \mathcal{L}(U_{i1}|X_{i1})$ *(and similarly for* U_{i2}*)*
4. $E[U_{it}|X_{it}] = 0$ *and some other assumptions...*

Then, model (1) *is identified, i.e.* $m(x,\alpha)$, $F_{\alpha_i}(\alpha|X_{it}=x)$, $F_{U_{it}}(u|X_{it}=x)$ *are identified.*

Theorem (Identification)

Consider the model

$$Y_{it} = m(X_{it}, \alpha_i) + U_{it},$$

and suppose that

1. $f_{X_{i1},X_{i2}}(x,x) > 0$ *for all* x
2. $m(x,\alpha)$ *is strictly increasing in* α
3. $\mathcal{L}(U_{i1}|X_{i1},\alpha_i,X_{i2},U_{i2}) = \mathcal{L}(U_{i1}|X_{i1})$ *(and similarly for* U_{i2}*)*
4. $E[U_{it}|X_{it}] = 0$ *and some other assumptions...*

Then, model (1) *is identified, i.e.* $m(x,\alpha)$, $F_{\alpha_i}(\alpha|X_{it}=x)$, $F_{U_{it}}(u|X_{it}=x)$ *are identified.*

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

Kotlarski's (1967) lemma (slightly improved): A, U_1, U_2 are not observed. Observe (Z_1, Z_2):

$$\begin{aligned} Z_1 &= A + U_1 \\ Z_2 &= A + U_2 \end{aligned}$$

- A, U_1, and U_2 are mutually independent
- characteristic functions $\phi_{U_t}(s)$ are nonvanishing
- $E[U_1] = 0$

$\Longrightarrow$ then the distributions of U_1, U_2, and A are ID'ed

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

- Kotlarski (1967): Observe (Z_1, Z_2), where

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} A + U_1 \\ A + U_2 \end{pmatrix}$$

- KEY: $m(X_{i1}, \alpha_i) = m(X_{i2}, \alpha_i)$ when $X_{i1} = X_{i2}$. Thus

$$\begin{pmatrix} Y_{i1} \\ Y_{i2} \end{pmatrix} \Bigg| \{X_{i1} = X_{i2} = x\} = \begin{pmatrix} m(x, \alpha_i) + U_{i1} \\ m(x, \alpha_i) + U_{i2} \end{pmatrix} \Bigg| \{X_{i1} = X_{i2} = x\}$$

- ID'ed $\mathcal{L}(U_{i1} | X_{i1} = X_{i2} = x)$
- $\mathcal{L}(U_{i1} | X_{i1} = X_{i2} = x) = \mathcal{L}(U_{i1} | X_{i1} = x)$, thus $\mathcal{L}(U_{i1} | X_{i1} = x)$ is ID'ed

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

- Kotlarski (1967): Observe (Z_1, Z_2), where

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} A + U_1 \\ A + U_2 \end{pmatrix}$$

- KEY: $m(X_{i1}, \alpha_i) = m(X_{i2}, \alpha_i)$ when $X_{i1} = X_{i2}$. Thus

$$\begin{pmatrix} Y_{i1} \\ Y_{i2} \end{pmatrix} \Bigg| \{X_{i1} = X_{i2} = x\} = \begin{pmatrix} m(x, \alpha_i) + U_{i1} \\ m(x, \alpha_i) + U_{i2} \end{pmatrix} \Bigg| \{X_{i1} = X_{i2} = x\}$$

- ID'ed $\mathcal{L}(U_{i1} | X_{i1} = X_{i2} = x)$
- $\mathcal{L}(U_{i1} | X_{i1} = X_{i2} = x) = \mathcal{L}(U_{i1} | X_{i1} = x)$, thus $\mathcal{L}(U_{i1} | X_{i1} = x)$ is ID'ed

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

- Kotlarski (1967): Observe (Z_1, Z_2), where

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} A + U_1 \\ A + U_2 \end{pmatrix}$$

- KEY: $m(X_{i1}, \alpha_i) = m(X_{i2}, \alpha_i)$ when $X_{i1} = X_{i2}$. Thus

$$\begin{pmatrix} Y_{i1} \\ Y_{i2} \end{pmatrix} \Bigg| \{X_{i1} = X_{i2} = x\} = \begin{pmatrix} m(x, \alpha_i) + U_{i1} \\ m(x, \alpha_i) + U_{i2} \end{pmatrix} \Bigg| \{X_{i1} = X_{i2} = x\}$$

- ID'ed $\mathcal{L}(U_{i1} | X_{i1} = X_{i2} = x)$
- $\mathcal{L}(U_{i1} | X_{i1} = X_{i2} = x) = \mathcal{L}(U_{i1} | X_{i1} = x)$, thus $\mathcal{L}(U_{i1} | X_{i1} = x)$ is ID'ed

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

- Kotlarski (1967): Observe (Z_1, Z_2), where

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} A + U_1 \\ A + U_2 \end{pmatrix}$$

- KEY: $m(X_{i1}, \alpha_i) = m(X_{i2}, \alpha_i)$ when $X_{i1} = X_{i2}$. Thus

$$\begin{pmatrix} Y_{i1} \\ Y_{i2} \end{pmatrix} \Bigg| \{X_{i1} = X_{i2} = x\} = \begin{pmatrix} m(x, \alpha_i) + U_{i1} \\ m(x, \alpha_i) + U_{i2} \end{pmatrix} \Bigg| \{X_{i1} = X_{i2} = x\}$$

- ID'ed $\mathcal{L}(U_{i1} | X_{i1} = X_{i2} = x)$
- $\mathcal{L}(U_{i1} | X_{i1} = X_{i2} = x) = \mathcal{L}(U_{i1} | X_{i1} = x)$, thus $\mathcal{L}(U_{i1} | X_{i1} = x)$ is ID'ed

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

- Kotlarski (1967): Observe (Z_1, Z_2), where

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} A + U_1 \\ A + U_2 \end{pmatrix}$$

- KEY: $m(X_{i1}, \alpha_i) = m(X_{i2}, \alpha_i)$ when $X_{i1} = X_{i2}$. Thus

$$\begin{pmatrix} Y_{i1} \\ Y_{i2} \end{pmatrix} \Bigg| \{X_{i1} = X_{i2} = x\} = \begin{pmatrix} m(x, \alpha_i) + U_{i1} \\ m(x, \alpha_i) + U_{i2} \end{pmatrix} \Bigg| \{X_{i1} = X_{i2} = x\}$$

- ID'ed $\mathcal{L}(U_{i1} | X_{i1} = X_{i2} = x)$
- $\mathcal{L}(U_{i1} | X_{i1} = X_{i2} = x) = \mathcal{L}(U_{i1} | X_{i1} = x)$, thus $\mathcal{L}(U_{i1} | X_{i1} = x)$ is ID'ed

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

- Kotlarski (1967): Observe (Z_1, Z_2), where

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} A + U_1 \\ A + U_2 \end{pmatrix}$$

- KEY: $m(X_{i1}, \alpha_i) = m(X_{i2}, \alpha_i)$ when $X_{i1} = X_{i2}$. Thus

$$\begin{pmatrix} Y_{i1} \\ Y_{i2} \end{pmatrix} \Bigg| \{X_{i1} = X_{i2} = x\} = \begin{pmatrix} m(x, \alpha_i) + U_{i1} \\ m(x, \alpha_i) + U_{i2} \end{pmatrix} \Bigg| \{X_{i1} = X_{i2} = x\}$$

- ID'ed $\mathcal{L}(U_{i1} | X_{i1} = X_{i2} = x)$
- $\mathcal{L}(U_{i1} | X_{i1} = X_{i2} = x) = \mathcal{L}(U_{i1} | X_{i1} = x)$, thus $\mathcal{L}(U_{i1} | X_{i1} = x)$ is ID'ed

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

- Kotlarski (1967): Observe (Z_1, Z_2), where

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} A + U_1 \\ A + U_2 \end{pmatrix}$$

- KEY: $m(X_{i1}, \alpha_i) = m(X_{i2}, \alpha_i)$ when $X_{i1} = X_{i2}$. Thus

$$\begin{pmatrix} Y_{i1} \\ Y_{i2} \end{pmatrix} \Bigg| \{X_{i1} = X_{i2} = x\} = \begin{pmatrix} m(x, \alpha_i) + U_{i1} \\ m(x, \alpha_i) + U_{i2} \end{pmatrix} \Bigg| \{X_{i1} = X_{i2} = x\}$$

- ID'ed $\mathcal{L}(U_{i1} | X_{i1} = X_{i2} = x)$
- $\mathcal{L}(U_{i1} | X_{i1} = X_{i2} = x) = \mathcal{L}(U_{i1} | X_{i1} = x)$, thus $\mathcal{L}(U_{i1} | X_{i1} = x)$ is ID'ed

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

- Kotlarski (1967): Observe (Z_1, Z_2), where

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} A + U_1 \\ A + U_2 \end{pmatrix}$$

- KEY: $m(X_{i1}, \alpha_i) = m(X_{i2}, \alpha_i)$ when $X_{i1} = X_{i2}$. Thus

$$\begin{pmatrix} Y_{i1} \\ Y_{i2} \end{pmatrix} \Bigg| \{X_{i1} = X_{i2} = x\} = \begin{pmatrix} m(x, \alpha_i) + U_{i1} \\ m(x, \alpha_i) + U_{i2} \end{pmatrix} \Bigg| \{X_{i1} = X_{i2} = x\}$$

- ID'ed $\mathcal{L}(U_{i1} | X_{i1} = X_{i2} = x)$
- $\mathcal{L}(U_{i1} | X_{i1} = X_{i2} = x) = \mathcal{L}(U_{i1} | X_{i1} = x)$, thus $\mathcal{L}(U_{i1} | X_{i1} = x)$ is ID'ed

Identification, Random Effects

- Random Effects: α_i is independent of X_{it}
- *Normalization*: $\alpha_i \sim U[0,1]$
- For any x consider:

$$\underbrace{\mathcal{L}\left(Y_{i1}|X_{i1}=x\right)=}_{\text{from data}}\mathcal{L}\left(m\left(x,\alpha_i\right)|X_{i1}=x\right)*\underbrace{\mathcal{L}\left(U_{i1}|X_{i1}=x\right)}_{\text{from step 1}}$$

$$\downarrow$$

$$Q_{m(x,\alpha_i)}\left(q|X_{i1}=x\right)$$

- Then, for any $q \in (0,1)$:

$$Q_{m(x,\alpha_i)}\left(q|X_{i1}=x\right)=m\left(x,Q_{\alpha_i}\left(q|X_{i1}=x\right)\right)=m\left(x,q\right)$$

Identification, Random Effects

- Random Effects: α_i is independent of X_{it}
- *Normalization*: $\alpha_i \sim U[0,1]$
- For any x consider:

$$\underbrace{\mathcal{L}\left(Y_{i1}|X_{i1}=x\right)=}_{\text{from data}}\mathcal{L}\left(m\left(x,\alpha_i\right)|X_{i1}=x\right)*\underbrace{\mathcal{L}\left(U_{i1}|X_{i1}=x\right)}_{\text{from step 1}}$$

$$\downarrow$$

$$Q_{m(x,\alpha_i)}\left(q|X_{i1}=x\right)$$

- Then, for any $q \in (0,1)$:

$$Q_{m(x,\alpha_i)}\left(q|X_{i1}=x\right)=m\left(x,Q_{\alpha_i}\left(q|X_{i1}=x\right)\right)=m\left(x,q\right)$$

Identification, Random Effects

- Random Effects: α_i is independent of X_{it}
- *Normalization*: $\alpha_i \sim U[0,1]$
- For any x consider:

$$\underbrace{\mathcal{L}\left(Y_{i1}|X_{i1}=x\right)=}_{\text{from data}} \mathcal{L}\left(m\left(x,\alpha_i\right)|X_{i1}=x\right) * \underbrace{\mathcal{L}\left(U_{i1}|X_{i1}=x\right)}_{\text{from step 1}}$$

$$\downarrow$$

$$Q_{m(x,\alpha_i)}\left(q|X_{i1}=x\right)$$

- Then, for any $q \in (0,1)$:

$$Q_{m(x,\alpha_i)}\left(q|X_{i1}=x\right)=m\left(x,Q_{\alpha_i}\left(q|X_{i1}=x\right)\right)=m\left(x,q\right)$$

Identification, Random Effects

- Random Effects: α_i is independent of X_{it}
- *Normalization*: $\alpha_i \sim U[0,1]$
- For any x consider:

$$\underbrace{\mathcal{L}\left(Y_{i1}|X_{i1}=x\right)=}_{\text{from data}}\mathcal{L}\left(m\left(x,\alpha_i\right)|X_{i1}=x\right)*\underbrace{\mathcal{L}\left(U_{i1}|X_{i1}=x\right)}_{\text{from step 1}}$$

$$\downarrow$$

$$Q_{m(x,\alpha_i)}\left(q|X_{i1}=x\right)$$

- Then, for any $q \in (0,1)$:

$$Q_{m(x,\alpha_i)}\left(q|X_{i1}=x\right)=m\left(x,Q_{\alpha_i}\left(q|X_{i1}=x\right)\right)=m\left(x,q\right)$$

Identification, Random Effects

- Random Effects: α_i is independent of X_{it}
- *Normalization*: $\alpha_i \sim U[0,1]$
- For any x consider:

$$\underbrace{\mathcal{L}\left(Y_{i1}|X_{i1}=x\right)=}_{\text{from data}} \mathcal{L}\left(m\left(x,\alpha_i\right)|X_{i1}=x\right) * \underbrace{\mathcal{L}\left(U_{i1}|X_{i1}=x\right)}_{\text{from step 1}}$$

$$\downarrow$$

$$Q_{m(x,\alpha_i)}\left(q|X_{i1}=x\right)$$

- Then, for any $q \in (0,1)$:

$$Q_{m(x,\alpha_i)}\left(q|X_{i1}=x\right)=m\left(x,Q_{\alpha_i}\left(q|X_{i1}=x\right)\right)=m\left(x,q\right)$$

Identification, Random Effects

- Random Effects: α_i is independent of X_{it}
- *Normalization*: $\alpha_i \sim U[0,1]$
- For any x consider:

$$\underbrace{\mathcal{L}\left(Y_{i1}|X_{i1}=x\right)=}_{\text{from data}} \mathcal{L}\left(m\left(x,\alpha_i\right)|X_{i1}=x\right) * \underbrace{\mathcal{L}\left(U_{i1}|X_{i1}=x\right)}_{\text{from step 1}}$$

$$\downarrow$$

$$Q_{m(x,\alpha_i)}\left(q|X_{i1}=x\right)$$

- Then, for any $q \in (0,1)$:

$$Q_{m(x,\alpha_i)}\left(q|X_{i1}=x\right) = m\left(x, Q_{\alpha_i}\left(q|X_{i1}=x\right)\right) = m\left(x,q\right)$$

Identification, Random Effects

- Random Effects: α_i is independent of X_{it}
- *Normalization*: $\alpha_i \sim U[0,1]$
- For any x consider:

$$\underbrace{\mathcal{L}\left(Y_{i1}|X_{i1}=x\right)=}_{\text{from data}} \mathcal{L}\left(m\left(x,\alpha_i\right)|X_{i1}=x\right) * \underbrace{\mathcal{L}\left(U_{i1}|X_{i1}=x\right)}_{\text{from step 1}}$$

$$\downarrow$$

$$Q_{m(x,\alpha_i)}\left(q|X_{i1}=x\right)$$

- Then, for any $q \in (0,1)$:

$$Q_{m(x,\alpha_i)}\left(q|X_{i1}=x\right) = m\left(x, Q_{\alpha_i}\left(q|X_{i1}=x\right)\right) = m\left(x,q\right)$$

Identification, Random Effects

- Random Effects: α_i is independent of X_{it}
- *Normalization*: $\alpha_i \sim U[0,1]$
- For any x consider:

$$\underbrace{\mathcal{L}\left(Y_{i1}|X_{i1}=x\right)=}_{\text{from data}} \mathcal{L}\left(m\left(x,\alpha_i\right)|X_{i1}=x\right) * \underbrace{\mathcal{L}\left(U_{i1}|X_{i1}=x\right)}_{\text{from step 1}}$$

$$\downarrow$$

$$Q_{m(x,\alpha_i)}\left(q|X_{i1}=x\right)$$

- Then, for any $q \in (0,1)$:

$$Q_{m(x,\alpha_i)}\left(q|X_{i1}=x\right) = m\left(x, Q_{\alpha_i}\left(q|X_{i1}=x\right)\right) = m\left(x,q\right)$$

Identification, Random Effects

- Random Effects: α_i is independent of X_{it}
- *Normalization*: $\alpha_i \sim U[0,1]$
- For any x consider:

$$\underbrace{\mathcal{L}\left(Y_{i1}|X_{i1}=x\right)=}_{\text{from data}} \mathcal{L}\left(m\left(x,\alpha_i\right)|X_{i1}=x\right) * \underbrace{\mathcal{L}\left(U_{i1}|X_{i1}=x\right)}_{\text{from step 1}}$$

$$\downarrow$$

$$Q_{m(x,\alpha_i)}\left(q|X_{i1}=x\right)$$

- Then, for any $q \in (0,1)$:

$$Q_{m(x,\alpha_i)}\left(q|X_{i1}=x\right) = m\left(x, Q_{\alpha_i}\left(q|X_{i1}=x\right)\right) = m\left(x,q\right)$$

Identification, Fixed Effects

Normalization: for some $\overline{x}$, $m(\overline{x}, \alpha) = \alpha$ for all α

For any x denote $\mathcal{G}_x = \{X_{i1} = x, X_{i2} = \overline{x}\}$:

$$\begin{aligned} \mathcal{L}(Y_{i1}|\mathcal{G}_x) &= \mathcal{L}(m(x, \alpha_i)|\mathcal{G}_x) * \mathcal{L}(U_{i1}|\mathcal{G}_x) \\ \underbrace{\mathcal{L}(Y_{i2}|\mathcal{G}_x)}_{\text{from data}} &= \mathcal{L}(m(\overline{x}, \alpha_i)|\mathcal{G}_x) * \underbrace{\mathcal{L}(U_{i2}|\mathcal{G}_x)}_{\text{from step 1}} \end{aligned}$$

$$\begin{aligned} Y_{i1}: \quad Q_{m(x,\alpha_i)}(q|x, \overline{x}) &= m(x, Q_{\alpha_i}(q|x, \overline{x})) \\ & \updownarrow \\ Y_{i2}: \quad Q_{m(\overline{x},\alpha_i)}(q|x, \overline{x}) &= Q_{\alpha_i}(q|x, \overline{x}) \end{aligned}$$

Then $m(x, a)$ is ID'ed by

$$m(x, a) = Q_{m(x,\alpha_i)}\left(F_{m(\overline{x},\alpha_i)}(a|x, \overline{x})|x, \overline{x}\right)$$

Identification, Fixed Effects

Normalization: for some $\overline{x}$, $m(\overline{x}, \alpha) = \alpha$ for all α

For any x denote $\mathcal{G}_x = \{X_{i1} = x, X_{i2} = \overline{x}\}$:

$$\begin{aligned} \mathcal{L}(Y_{i1}|\mathcal{G}_x) &= \mathcal{L}(m(x, \alpha_i)|\mathcal{G}_x) * \mathcal{L}(U_{i1}|\mathcal{G}_x) \\ \underbrace{\mathcal{L}(Y_{i2}|\mathcal{G}_x)}_{\text{from data}} &= \mathcal{L}(m(\overline{x}, \alpha_i)|\mathcal{G}_x) * \underbrace{\mathcal{L}(U_{i2}|\mathcal{G}_x)}_{\text{from step 1}} \end{aligned}$$

$$\begin{aligned} Y_{i1}: \quad Q_{m(x,\alpha_i)}(q|x, \overline{x}) &= m(x, Q_{\alpha_i}(q|x, \overline{x})) \\ & \updownarrow \\ Y_{i2}: \quad Q_{m(\overline{x},\alpha_i)}(q|x, \overline{x}) &= Q_{\alpha_i}(q|x, \overline{x}) \end{aligned}$$

Then $m(x, a)$ is ID'ed by

$$m(x, a) = Q_{m(x,\alpha_i)}\left(F_{m(\overline{x},\alpha_i)}(a|x, \overline{x})|x, \overline{x}\right)$$

Identification, Fixed Effects

Normalization: for some $\overline{x}$, $m(\overline{x}, \alpha) = \alpha$ for all α

For any x denote $\mathcal{G}_x = \{X_{i1} = x, X_{i2} = \overline{x}\}$:

$$\underbrace{\mathcal{L}(Y_{i1}|\mathcal{G}_x)}_{\text{from data}} = \mathcal{L}(m(x, \alpha_i)|\mathcal{G}_x) * \mathcal{L}(U_{i1}|\mathcal{G}_x)$$
$$\underbrace{\mathcal{L}(Y_{i2}|\mathcal{G}_x)}_{\text{from data}} = \mathcal{L}(m(\overline{x}, \alpha_i)|\mathcal{G}_x) * \underbrace{\mathcal{L}(U_{i2}|\mathcal{G}_x)}_{\text{from step 1}}$$

$$Y_{i1}: \quad Q_{m(x,\alpha_i)}(q|x, \overline{x}) = m(x, Q_{\alpha_i}(q|x, \overline{x}))$$
$$\updownarrow$$
$$Y_{i2}: \quad Q_{m(\overline{x},\alpha_i)}(q|x, \overline{x}) = Q_{\alpha_i}(q|x, \overline{x})$$

Then $m(x, a)$ is ID'ed by

$$m(x, a) = Q_{m(x,\alpha_i)}\left(F_{m(\overline{x},\alpha_i)}(a|x, \overline{x})|x, \overline{x}\right)$$

Identification, Fixed Effects

Normalization: for some $\overline{x}$, $m(\overline{x}, \alpha) = \alpha$ for all α

For any x denote $\mathcal{G}_x = \{X_{i1} = x, X_{i2} = \overline{x}\}$:

$$\underbrace{\begin{aligned}\mathcal{L}(Y_{i1}|\mathcal{G}_x) &= \\ \mathcal{L}(Y_{i2}|\mathcal{G}_x) &=\end{aligned}}_{\text{from data}} \begin{aligned}\mathcal{L}(m(x,\alpha_i)|\mathcal{G}_x) &* \mathcal{L}(U_{i1}|\mathcal{G}_x) \\ \mathcal{L}(m(\overline{x},\alpha_i)|\mathcal{G}_x) &* \underbrace{\mathcal{L}(U_{i2}|\mathcal{G}_x)}_{\text{from step 1}}\end{aligned}$$

$$\begin{aligned} Y_{i1}: \quad & Q_{m(x,\alpha_i)}(q|x,\overline{x}) = m(x, Q_{\alpha_i}(q|x,\overline{x})) \\ & \qquad\qquad\qquad\qquad \updownarrow \\ Y_{i2}: \quad & Q_{m(\overline{x},\alpha_i)}(q|x,\overline{x}) = Q_{\alpha_i}(q|x,\overline{x}) \end{aligned}$$

Then $m(x,a)$ is ID'ed by

$$m(x,a) = Q_{m(x,\alpha_i)}\left(F_{m(\overline{x},\alpha_i)}(a|x,\overline{x})\,|x,\overline{x}\right)$$

Identification, Fixed Effects

Normalization: for some $\overline{x}$, $m(\overline{x}, \alpha) = \alpha$ for all α

For any x denote $\mathcal{G}_x = \{X_{i1} = x, X_{i2} = \overline{x}\}$:

$$\underbrace{\begin{array}{l} \mathcal{L}(Y_{i1}|\mathcal{G}_x) \\ \mathcal{L}(Y_{i2}|\mathcal{G}_x) \end{array}}_{\text{from data}} \begin{array}{l} = \mathcal{L}(m(x, \alpha_i)|\mathcal{G}_x) * \mathcal{L}(U_{i1}|\mathcal{G}_x) \\ = \mathcal{L}(m(\overline{x}, \alpha_i)|\mathcal{G}_x) * \underbrace{\mathcal{L}(U_{i2}|\mathcal{G}_x)}_{\text{from step 1}} \end{array} \begin{array}{l} \longmapsto Q_{m(x,\alpha_i)}(q|x, \overline{x}) \\ \longmapsto Q_{m(\overline{x},\alpha_i)}(q|x, \overline{x}) \end{array}$$

$$\begin{array}{lll} Y_{i1}: & Q_{m(x,\alpha_i)}(q|x, \overline{x}) & = \quad m(x, Q_{\alpha_i}(q|x, \overline{x})) \\ & & \quad\quad \updownarrow \\ Y_{i2}: & Q_{m(\overline{x},\alpha_i)}(q|x, \overline{x}) & = \quad Q_{\alpha_i}(q|x, \overline{x}) \end{array}$$

Then $m(x, a)$ is ID'ed by

$$m(x, a) = Q_{m(x,\alpha_i)}\left(F_{m(\overline{x},\alpha_i)}(a|x, \overline{x})\,|x, \overline{x}\right)$$

Identification, Fixed Effects

Normalization: for some $\overline{x}$, $m(\overline{x}, \alpha) = \alpha$ for all α

For any x denote $\mathcal{G}_x = \{X_{i1} = x, X_{i2} = \overline{x}\}$:

$$\underbrace{\begin{array}{l} \mathcal{L}(Y_{i1}|\mathcal{G}_x) \\ \mathcal{L}(Y_{i2}|\mathcal{G}_x) \end{array}}_{\text{from data}} \begin{array}{l} = \mathcal{L}(m(x, \alpha_i)|\mathcal{G}_x) * \mathcal{L}(U_{i1}|\mathcal{G}_x) \\ = \mathcal{L}(m(\overline{x}, \alpha_i)|\mathcal{G}_x) * \underbrace{\mathcal{L}(U_{i2}|\mathcal{G}_x)}_{\text{from step 1}} \end{array} \begin{array}{l} \longmapsto Q_{m(x,\alpha_i)}(q|x, \overline{x}) \\ \longmapsto Q_{m(\overline{x},\alpha_i)}(q|x, \overline{x}) \end{array}$$

$$\begin{array}{lll} Y_{i1}: & Q_{m(x,\alpha_i)}(q|x, \overline{x}) & = \quad m(x, Q_{\alpha_i}(q|x, \overline{x})) \\ & & \quad\quad \updownarrow \\ Y_{i2}: & Q_{m(\overline{x},\alpha_i)}(q|x, \overline{x}) & = \quad Q_{\alpha_i}(q|x, \overline{x}) \end{array}$$

Then $m(x, a)$ is ID'ed by

$$m(x, a) = Q_{m(x,\alpha_i)}\left(F_{m(\overline{x},\alpha_i)}(a|x, \overline{x})\,|x, \overline{x}\right)$$

Identification, Fixed Effects

Normalization: for some $\overline{x}$, $m(\overline{x}, \alpha) = \alpha$ for all α

For any x denote $\mathcal{G}_x = \{X_{i1} = x, X_{i2} = \overline{x}\}$:

$$\underbrace{\begin{array}{l} \mathcal{L}(Y_{i1}|\mathcal{G}_x) \\ \mathcal{L}(Y_{i2}|\mathcal{G}_x) \end{array}}_{\text{from data}} \begin{array}{l} = \mathcal{L}(m(x, \alpha_i)|\mathcal{G}_x) * \mathcal{L}(U_{i1}|\mathcal{G}_x) \\ = \mathcal{L}(m(\overline{x}, \alpha_i)|\mathcal{G}_x) * \underbrace{\mathcal{L}(U_{i2}|\mathcal{G}_x)}_{\text{from step 1}} \end{array} \begin{array}{l} \longmapsto Q_{m(x,\alpha_i)}(q|x, \overline{x}) \\ \longmapsto Q_{m(\overline{x},\alpha_i)}(q|x, \overline{x}) \end{array}$$

$$\begin{array}{lll} Y_{i1}: & Q_{m(x,\alpha_i)}(q|x, \overline{x}) & = \quad m(x, Q_{\alpha_i}(q|x, \overline{x})) \\ & & \quad\quad \updownarrow \\ Y_{i2}: & Q_{m(\overline{x},\alpha_i)}(q|x, \overline{x}) & = \quad Q_{\alpha_i}(q|x, \overline{x}) \end{array}$$

Then $m(x, a)$ is ID'ed by

$$m(x, a) = Q_{m(x,\alpha_i)}\left(F_{m(\overline{x},\alpha_i)}(a|x, \overline{x})\,|x, \overline{x}\right)$$

Identification, Fixed Effects

Normalization: for some $\overline{x}$, $m(\overline{x}, \alpha) = \alpha$ for all α

For any x denote $\mathcal{G}_x = \{X_{i1} = x, X_{i2} = \overline{x}\}$:

$$\begin{aligned}
\underbrace{\mathcal{L}(Y_{i1}|\mathcal{G}_x)}_{\text{from data}} &= \mathcal{L}(m(x, \alpha_i)|\mathcal{G}_x) * \mathcal{L}(U_{i1}|\mathcal{G}_x) &\longmapsto\ & Q_{m(x,\alpha_i)}(q|x, \overline{x}) \\
\underbrace{\mathcal{L}(Y_{i2}|\mathcal{G}_x)}_{\text{from data}} &= \mathcal{L}(m(\overline{x}, \alpha_i)|\mathcal{G}_x) * \underbrace{\mathcal{L}(U_{i2}|\mathcal{G}_x)}_{\text{from step 1}} &\longmapsto\ & Q_{m(\overline{x},\alpha_i)}(q|x, \overline{x})
\end{aligned}$$

$$\begin{aligned}
Y_{i1}: \quad Q_{m(x,\alpha_i)}(q|x, \overline{x}) &= m(x, Q_{\alpha_i}(q|x, \overline{x})) \\
& \updownarrow \\
Y_{i2}: \quad Q_{m(\overline{x},\alpha_i)}(q|x, \overline{x}) &= Q_{\alpha_i}(q|x, \overline{x})
\end{aligned}$$

Then $m(x, a)$ is ID'ed by

$$m(x, a) = Q_{m(x,\alpha_i)}\left(F_{m(\overline{x},\alpha_i)}(a|x, \overline{x})\,|x, \overline{x}\right)$$

Identification, Fixed Effects

Normalization: for some $\overline{x}$, $m(\overline{x}, \alpha) = \alpha$ for all α

For any x denote $\mathcal{G}_x = \{X_{i1} = x, X_{i2} = \overline{x}\}$:

$$\underbrace{\begin{array}{l} \mathcal{L}(Y_{i1}|\mathcal{G}_x) \\ \mathcal{L}(Y_{i2}|\mathcal{G}_x) \end{array}}_{\text{from data}} \begin{array}{l} = \mathcal{L}(m(x, \alpha_i)|\mathcal{G}_x) * \mathcal{L}(U_{i1}|\mathcal{G}_x) \\ = \mathcal{L}(m(\overline{x}, \alpha_i)|\mathcal{G}_x) * \underbrace{\mathcal{L}(U_{i2}|\mathcal{G}_x)}_{\text{from step 1}} \end{array} \begin{array}{l} \longmapsto Q_{m(x,\alpha_i)}(q|x, \overline{x}) \\ \longmapsto Q_{m(\overline{x},\alpha_i)}(q|x, \overline{x}) \end{array}$$

$$\begin{array}{lcl} Y_{i1}: \quad Q_{m(x,\alpha_i)}(q|x, \overline{x}) & = & m(x, Q_{\alpha_i}(q|x, \overline{x})) \\ & & \updownarrow \\ Y_{i2}: \quad Q_{m(\overline{x},\alpha_i)}(q|x, \overline{x}) & = & Q_{\alpha_i}(q|x, \overline{x}) \end{array}$$

Then $m(x, a)$ is ID'ed by

$$m(x, a) = Q_{m(x,\alpha_i)}\left(F_{m(\overline{x},\alpha_i)}(a|x, \overline{x})\,|x, \overline{x}\right)$$

Identification, Fixed Effects

Normalization: for some $\overline{x}$, $m(\overline{x}, \alpha) = \alpha$ for all α

For any x denote $\mathcal{G}_x = \{X_{i1} = x, X_{i2} = \overline{x}\}$:

$$\begin{aligned}
\underbrace{\mathcal{L}(Y_{i1}|\mathcal{G}_x)}_{\text{from data}} &= \mathcal{L}(m(x, \alpha_i)|\mathcal{G}_x) * \mathcal{L}(U_{i1}|\mathcal{G}_x) &\longmapsto\ & Q_{m(x,\alpha_i)}(q|x, \overline{x}) \\
\underbrace{\mathcal{L}(Y_{i2}|\mathcal{G}_x)}_{\text{from data}} &= \mathcal{L}(m(\overline{x}, \alpha_i)|\mathcal{G}_x) * \underbrace{\mathcal{L}(U_{i2}|\mathcal{G}_x)}_{\text{from step 1}} &\longmapsto\ & Q_{m(\overline{x},\alpha_i)}(q|x, \overline{x})
\end{aligned}$$

$$\begin{aligned}
Y_{i1}: \quad Q_{m(x,\alpha_i)}(q|x, \overline{x}) &= m(x, Q_{\alpha_i}(q|x, \overline{x})) \\
& \qquad\qquad \updownarrow \\
Y_{i2}: \quad Q_{m(\overline{x},\alpha_i)}(q|x, \overline{x}) &= Q_{\alpha_i}(q|x, \overline{x})
\end{aligned}$$

Then $m(x, a)$ is ID'ed by

$$m(x, a) = Q_{m(x,\alpha_i)}\left(F_{m(\overline{x},\alpha_i)}(a|x, \overline{x})\,|x, \overline{x}\right)$$

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

The structural function $m(x, a)$ is ID'ed by

$$m(x, a) = Q_{m(x,\alpha_i)}\left(F_{m(\overline{x},\alpha_i)}(a|x, \overline{x})\,|x, \overline{x}\right)$$

- ID'ed $m(x, a)$ for all x and all a from $Supp(\alpha_i|X_{i1} = x, X_{i2} = \overline{x})$
- Want ID for all α from $Supp(\alpha_i)$
- *May assume*: $Supp(\alpha_i) = Supp(\alpha_i|X_{i1} = x, X_{i2} = \overline{x})$ (cf. Altonji and Matzkin, 2005)
- $h(x, q) \equiv m(x, Q_{\alpha_i}(q))$

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

The structural function $m(x, a)$ is ID'ed by

$$m(x, a) = Q_{m(x,\alpha_i)}\left(F_{m(\overline{x},\alpha_i)}(a|x, \overline{x})\,|x, \overline{x}\right)$$

- ID'ed $m(x, a)$ for all x and all a from $Supp(\alpha_i|X_{i1} = x, X_{i2} = \overline{x})$
- Want ID for all α from $Supp(\alpha_i)$
- *May assume*: $Supp(\alpha_i) = Supp(\alpha_i|X_{i1} = x, X_{i2} = \overline{x})$ (cf. Altonji and Matzkin, 2005)
- $h(x, q) \equiv m(x, Q_{\alpha_i}(q))$

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

The structural function $m(x, a)$ is ID'ed by

$$m(x, a) = Q_{m(x,\alpha_i)}\left(F_{m(\overline{x},\alpha_i)}(a|x, \overline{x})\,|x, \overline{x}\right)$$

- ID'ed $m(x, a)$ for all x and all a from $Supp(\alpha_i|X_{i1} = x, X_{i2} = \overline{x})$
- Want ID for all α from $Supp(\alpha_i)$
- *May assume*: $Supp(\alpha_i) = Supp(\alpha_i|X_{i1} = x, X_{i2} = \overline{x})$ (cf. Altonji and Matzkin, 2005)
- $h(x, q) \equiv m(x, Q_{\alpha_i}(q))$

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

The structural function $m(x, a)$ is ID'ed by

$$m(x, a) = Q_{m(x,\alpha_i)}\left(F_{m(\overline{x},\alpha_i)}(a|x, \overline{x})\,|x, \overline{x}\right)$$

- ID'ed $m(x, a)$ for all x and all a from $Supp(\alpha_i|X_{i1} = x, X_{i2} = \overline{x})$
- Want ID for all α from $Supp(\alpha_i)$
- *May assume*: $Supp(\alpha_i) = Supp(\alpha_i|X_{i1} = x, X_{i2} = \overline{x})$ (cf. Altonji and Matzkin, 2005)
- $h(x, q) \equiv m(x, Q_{\alpha_i}(q))$

Identification, $Y_{it} = m(X_{it}, \alpha_i) + U_{it}$

The structural function $m(x, a)$ is ID'ed by

$$m(x, a) = Q_{m(x,\alpha_i)}\left(F_{m(\overline{x},\alpha_i)}(a|x, \overline{x})\,|x, \overline{x}\right)$$

- ID'ed $m(x, a)$ for all x and all a from $Supp(\alpha_i|X_{i1} = x, X_{i2} = \overline{x})$
- Want ID for all α from $Supp(\alpha_i)$
- *May assume*: $Supp(\alpha_i) = Supp(\alpha_i|X_{i1} = x, X_{i2} = \overline{x})$ (cf. Altonji and Matzkin, 2005)
- $h(x, q) \equiv m(x, Q_{\alpha_i}(q))$

Identification, Comparison RE vs FE

- Random Effects (α_i and X_{it} are independent)

$$\mathcal{L}\left(Y_{i1}|X_{i1}=x\right)=\mathcal{L}\left(m\left(x,\alpha_i\right)|X_{i1}=x\right)*\mathcal{L}\left(U_{i1}|X_{i1}=x\right)$$

  - Uses between-variation
  - The effect of time-invariant covariates is identified
  - $m(x,\alpha)$ can be weakly increasing in α

- Fixed Effects (α_i and X_{it} are not independent)

$$\mathcal{L}\left(Y_{i1}|x,\overline{x}\right)=\mathcal{L}\left(m\left(x,\alpha_i\right)|x,\overline{x}\right)*\mathcal{L}\left(U_{i1}|x,\overline{x}\right)$$
$$\mathcal{L}\left(Y_{i2}|x,\overline{x}\right)=\mathcal{L}\left(m\left(\overline{x},\alpha_i\right)|x,\overline{x}\right)*\mathcal{L}\left(U_{i2}|x,\overline{x}\right)$$

  - Uses only within-variation
  - The effect of time-invariant covariates is *not* identified

Identification, Comparison RE vs FE

- Random Effects (α_i and X_{it} are independent)

$$\mathcal{L}\left(Y_{i1}|X_{i1}=x\right)=\mathcal{L}\left(m\left(x,\alpha_i\right)|X_{i1}=x\right)*\mathcal{L}\left(U_{i1}|X_{i1}=x\right)$$

  - Uses between-variation
  - The effect of time-invariant covariates is identified
  - $m(x,\alpha)$ can be weakly increasing in α

- Fixed Effects (α_i and X_{it} are not independent)

$$\mathcal{L}\left(Y_{i1}|x,\overline{x}\right)=\mathcal{L}\left(m\left(x,\alpha_i\right)|x,\overline{x}\right)*\mathcal{L}\left(U_{i1}|x,\overline{x}\right)$$
$$\mathcal{L}\left(Y_{i2}|x,\overline{x}\right)=\mathcal{L}\left(m\left(\overline{x},\alpha_i\right)|x,\overline{x}\right)*\mathcal{L}\left(U_{i2}|x,\overline{x}\right)$$

  - Uses only within-variation
  - The effect of time-invariant covariates is *not* identified

Extensions

- Time Effects

$$Y_{it} = m(X_{it}, \alpha_i) + \eta_t(X_{it}) + U_{it}, \quad \eta_1(x) \equiv 0$$

Remember that $m(X_{it}, \alpha_i) = m(X_{i1}, \alpha_i)$ when $X_{it} = X_{i1} = x$,

$$\begin{aligned} & E[Y_{it} - Y_{i1} | X_{it} = X_{i1} = x] \\ = \quad & E[m(X_{it}, \alpha_i) + \eta_t(X_{it}) + U_{it} \\ & - (m(X_{i1}, \alpha_i) + U_{i1}) | X_{it} = X_{i1} = x] \\ = \quad & E[\eta_t(X_{it}) + U_{it} - U_{i1} | X_{it} = X_{i1} = x] \quad = \quad \eta_t(x) \end{aligned}$$

- Similarly, can identify β in the model

$$Y_{it} = m(X_{it}, \alpha_i) + W_{it}'\beta + U_{it}$$

- ID when $U_{it} = \rho U_{it-1} + \varepsilon_{it}$ and $T \geq 3$

Extensions

- Time Effects

$$Y_{it} = m(X_{it}, \alpha_i) + \eta_t(X_{it}) + U_{it}, \quad \eta_1(x) \equiv 0$$

Remember that $m(X_{it}, \alpha_i) = m(X_{i1}, \alpha_i)$ when $X_{it} = X_{i1} = x$,

$$\begin{aligned} & E[Y_{it} - Y_{i1} | X_{it} = X_{i1} = x] \\ = \quad & E[m(X_{it}, \alpha_i) + \eta_t(X_{it}) + U_{it} \\ & - (m(X_{i1}, \alpha_i) + U_{i1}) | X_{it} = X_{i1} = x] \\ = \quad & E[\eta_t(X_{it}) + U_{it} - U_{i1} | X_{it} = X_{i1} = x] \quad = \quad \eta_t(x) \end{aligned}$$

- Similarly, can identify β in the model

$$Y_{it} = m(X_{it}, \alpha_i) + W_{it}'\beta + U_{it}$$

- ID when $U_{it} = \rho U_{it-1} + \varepsilon_{it}$ and $T \geq 3$

Extensions

- Time Effects

$$Y_{it} = m(X_{it}, \alpha_i) + \eta_t(X_{it}) + U_{it}, \quad \eta_1(x) \equiv 0$$

Remember that $m(X_{it}, \alpha_i) = m(X_{i1}, \alpha_i)$ when $X_{it} = X_{i1} = x$,

$$\begin{aligned} & E[Y_{it} - Y_{i1} | X_{it} = X_{i1} = x] \\ = \quad & E[m(X_{it}, \alpha_i) + \eta_t(X_{it}) + U_{it} \\ & - (m(X_{i1}, \alpha_i) + U_{i1}) | X_{it} = X_{i1} = x] \\ = \quad & E[\eta_t(X_{it}) + U_{it} - U_{i1} | X_{it} = X_{i1} = x] \quad = \quad \eta_t(x) \end{aligned}$$

- Similarly, can identify β in the model

$$Y_{it} = m(X_{it}, \alpha_i) + W_{it}'\beta + U_{it}$$

- ID when $U_{it} = \rho U_{it-1} + \varepsilon_{it}$ and $T \geq 3$

Extensions

- Time Effects

$$Y_{it} = m(X_{it}, \alpha_i) + \eta_t(X_{it}) + U_{it}, \quad \eta_1(x) \equiv 0$$

Remember that $m(X_{it}, \alpha_i) = m(X_{i1}, \alpha_i)$ when $X_{it} = X_{i1} = x$,

$$\begin{aligned} & E[Y_{it} - Y_{i1} | X_{it} = X_{i1} = x] \\ = \quad & E[m(X_{it}, \alpha_i) + \eta_t(X_{it}) + U_{it} \\ & - (m(X_{i1}, \alpha_i) + U_{i1}) | X_{it} = X_{i1} = x] \\ = \quad & E[\eta_t(X_{it}) + U_{it} - U_{i1} | X_{it} = X_{i1} = x] \quad = \quad \eta_t(x) \end{aligned}$$

- Similarly, can identify β in the model

$$Y_{it} = m(X_{it}, \alpha_i) + W_{it}'\beta + U_{it}$$

- ID when $U_{it} = \rho U_{it-1} + \varepsilon_{it}$ and $T \geq 3$

Extensions

- Time Effects

$$Y_{it} = m(X_{it}, \alpha_i) + \eta_t(X_{it}) + U_{it}, \quad \eta_1(x) \equiv 0$$

Remember that $m(X_{it}, \alpha_i) = m(X_{i1}, \alpha_i)$ when $X_{it} = X_{i1} = x$,

$$\begin{aligned} & E[Y_{it} - Y_{i1} | X_{it} = X_{i1} = x] \\ = \quad & E[m(X_{it}, \alpha_i) + \eta_t(X_{it}) + U_{it} \\ & - (m(X_{i1}, \alpha_i) + U_{i1}) | X_{it} = X_{i1} = x] \\ = \quad & E[\eta_t(X_{it}) + U_{it} - U_{i1} | X_{it} = X_{i1} = x] \quad = \quad \eta_t(x) \end{aligned}$$

- Similarly, can identify β in the model

$$Y_{it} = m(X_{it}, \alpha_i) + W_{it}'\beta + U_{it}$$

- ID when $U_{it} = \rho U_{it-1} + \varepsilon_{it}$ and $T \geq 3$

Extensions

- Time Effects

$$Y_{it} = m(X_{it}, \alpha_i) + \eta_t(X_{it}) + U_{it}, \quad \eta_1(x) \equiv 0$$

Remember that $m(X_{it}, \alpha_i) = m(X_{i1}, \alpha_i)$ when $X_{it} = X_{i1} = x$,

$$\begin{aligned} & E[Y_{it} - Y_{i1} | X_{it} = X_{i1} = x] \\ = \; & E[m(X_{it}, \alpha_i) + \eta_t(X_{it}) + U_{it} \\ & - (m(X_{i1}, \alpha_i) + U_{i1}) | X_{it} = X_{i1} = x] \\ = \; & E[\eta_t(X_{it}) + U_{it} - U_{i1} | X_{it} = X_{i1} = x] \quad = \quad \eta_t(x) \end{aligned}$$

- Similarly, can identify β in the model

$$Y_{it} = m(X_{it}, \alpha_i) + W_{it}'\beta + U_{it}$$

- ID when $U_{it} = \rho U_{it-1} + \varepsilon_{it}$ and $T \geq 3$

Extensions

- Time Effects

$$Y_{it} = m(X_{it}, \alpha_i) + \eta_t(X_{it}) + U_{it}, \quad \eta_1(x) \equiv 0$$

Remember that $m(X_{it}, \alpha_i) = m(X_{i1}, \alpha_i)$ when $X_{it} = X_{i1} = x$,

$$\begin{aligned} & E[Y_{it} - Y_{i1} | X_{it} = X_{i1} = x] \\ = \quad & E[m(X_{it}, \alpha_i) + \eta_t(X_{it}) + U_{it} \\ & - (m(X_{i1}, \alpha_i) + U_{i1}) | X_{it} = X_{i1} = x] \\ = \quad & E[\eta_t(X_{it}) + U_{it} - U_{i1} | X_{it} = X_{i1} = x] \quad = \quad \eta_t(x) \end{aligned}$$

- Similarly, can identify β in the model

$$Y_{it} = m(X_{it}, \alpha_i) + W_{it}'\beta + U_{it}$$

- ID when $U_{it} = \rho U_{it-1} + \varepsilon_{it}$ and $T \geq 3$

Misclassification

- Suppose $X_{it}^* \in \{0,1\}$ and

$$X_{it} = \begin{cases} X_{it}^*, & \text{with probability } p, \\ 1 - X_{it}^*, & \text{with probability } 1-p. \end{cases}$$

- Suppose p is *known* (eg. from a validation study). Write

$$\begin{aligned} \phi_{Y_{i1},Y_{i2}|X_{i1},X_{i2}}(s_1,s_2|0,0) &= p^2 \phi_{Y_{i1},Y_{i2}|X_{i1}^*,X_{i2}^*}(s_1,s_2|0,0) \\ &+ p(1-p)\phi_{Y_{i1},Y_{i2}|X_{i1}^*,X_{i2}^*}(s_1,s_2|0,1) \\ &+ p(1-p)\phi_{Y_{i1},Y_{i2}|X_{i1}^*,X_{i2}^*}(s_1,s_2|1,0) \\ &+ (1-p)^2 \phi_{Y_{i1},Y_{i2}|X_{i1}^*,X_{i2}^*}(s_1,s_2|1,1), \end{aligned}$$

  and similarly for $\phi_{Y_{i1},Y_{i2}|X_{i1},X_{i2}}(s_1,s_2|0,1)$, ...

- Then can solve for $\phi_{Y_{i1},Y_{i2}|X_{i1}^*,X_{i2}^*}(s_1,s_2|0,0)$, ...

Estimation, Conditional Deconvolution

- $\phi_{Y_{i1}}(s|X_{i1}=x) = E\left[\exp\left\{\mathrm{i}sY_{i1}\right\}|X_{it}=x\right], \mathrm{i}=\sqrt{-1},$

$$\hat{\phi}_{Y_{it}}(s|x) = \frac{\sum_{i=1}^{n}\exp\left(\mathrm{i}sY_{it}\right)K\left(\left(X_{it}-x\right)/h_Y\right)}{\sum_{i=1}^{n}K\left(\left(X_{it}-x\right)/h_Y\right)},$$

where $h_Y \to 0$ is a bandwidth

Estimator:

$$\hat{F}_{m(x,\alpha_i)}(\omega|x) = \frac{1}{2} + \int_{-\infty}^{\infty}\frac{e^{-\mathrm{i}s\omega}}{2\pi\mathrm{i}s}\phi_w\left(h_w s\right)\frac{\hat{\phi}_{Y_{it}}(s|x)}{\hat{\phi}_{U_{it}}(s|x)}ds$$

- Conditional deconvolution (cf. Fan, 1991; Diggle and Hall, 1993)
- Derive the rates of convergence

Estimation, Conditional Deconvolution

- $\phi_{Y_{i1}}(s|X_{i1}=x) = E\left[\exp\left\{\mathrm{i}sY_{i1}\right\}|X_{it}=x\right], \mathrm{i}=\sqrt{-1},$

$$\hat{\phi}_{Y_{it}}(s|x) = \frac{\sum_{i=1}^{n}\exp\left(\mathrm{i}sY_{it}\right)K\left(\left(X_{it}-x\right)/h_Y\right)}{\sum_{i=1}^{n}K\left(\left(X_{it}-x\right)/h_Y\right)},$$

where $h_Y \to 0$ is a bandwidth

Estimator:

$$\hat{F}_{m(x,\alpha_i)}(\omega|x) = \frac{1}{2} + \int_{-\infty}^{\infty}\frac{e^{-\mathrm{i}s\omega}}{2\pi\mathrm{i}s}\phi_w\left(h_w s\right)\frac{\hat{\phi}_{Y_{it}}(s|x)}{\hat{\phi}_{U_{it}}(s|x)}ds$$

- Conditional deconvolution (cf. Fan, 1991; Diggle and Hall, 1993)
- Derive the rates of convergence

Estimation, Conditional Deconvolution

- $\phi_{Y_{i1}}(s|X_{i1}=x) = E\left[\exp\left\{\mathrm{i}sY_{i1}\right\}|X_{it}=x\right], \mathrm{i}=\sqrt{-1},$

$$\hat{\phi}_{Y_{it}}(s|x) = \frac{\sum_{i=1}^{n}\exp\left(\mathrm{i}sY_{it}\right)K\left(\left(X_{it}-x\right)/h_Y\right)}{\sum_{i=1}^{n}K\left(\left(X_{it}-x\right)/h_Y\right)},$$

where $h_Y \to 0$ is a bandwidth

Estimator:

$$\hat{F}_{m(x,\alpha_i)}(\omega|x) = \frac{1}{2} + \int_{-\infty}^{\infty}\frac{e^{-\mathrm{i}s\omega}}{2\pi\mathrm{i}s}\phi_w\left(h_w s\right)\frac{\hat{\phi}_{Y_{it}}(s|x)}{\hat{\phi}_{U_{it}}(s|x)}ds$$

- Conditional deconvolution (cf. Fan, 1991; Diggle and Hall, 1993)
- Derive the rates of convergence

Estimation, Conditional Deconvolution

- $\phi_{Y_{i1}}(s|X_{i1}=x) = E\left[\exp\{\mathrm{i}sY_{i1}\}|X_{it}=x\right], \mathrm{i}=\sqrt{-1},$

$$\hat{\phi}_{Y_{it}}(s|x) = \frac{\sum_{i=1}^{n}\exp(\mathrm{i}sY_{it})K((X_{it}-x)/h_Y)}{\sum_{i=1}^{n}K((X_{it}-x)/h_Y)},$$

where $h_Y \to 0$ is a bandwidth

Estimator:

$$\hat{F}_{m(x,\alpha_i)}(\omega|x) = \frac{1}{2} + \int_{-\infty}^{\infty}\frac{e^{-\mathrm{i}s\omega}}{2\pi\mathrm{i}s}\phi_w(h_w s)\frac{\hat{\phi}_{Y_{it}}(s|x)}{\hat{\phi}_{U_{it}}(s|x)}ds$$

- Conditional deconvolution (cf. Fan, 1991; Diggle and Hall, 1993)
- Derive the rates of convergence

Estimation, Conditional Deconvolution

Estimator:

$$\hat{F}_{m(x,\alpha_i)}(\omega|x) = \frac{1}{2} + \int_{-\infty}^{\infty} \frac{e^{-is\omega}}{2\pi i s} \phi_w(h_w s) \frac{\hat{\phi}_{Y_{it}}(s|x)}{\hat{\phi}_{U_{it}}(s|x)} ds$$

- Conditional deconvolution (cf. Fan, 1991; Diggle and Hall, 1993)
- Derive the rates of convergence
- Complications: zeros in the denominators
- $\phi_{U_{it}}(s|x) \to 0$ as $s \to \pm\infty$. Two cases:
  - Ordinary-smooth: $\inf_x |\phi_{U_{it}}(s|x)| \sim C|s|^{-\lambda_U}$,
    - eg: Laplace, χ^2_k, Exp, Gamma
  - Super-smooth: $\inf_x |\phi_{U_{it}}(s|x)| \sim C\exp(-C_2|s|^{\lambda_U})$,
    - eg: Normal, Student-t_k, Cauchy

Estimation, Conditional Deconvolution

Estimator:

$$\hat{F}_{m(x,\alpha_i)}(\omega|x) = \frac{1}{2} + \int_{-\infty}^{\infty} \frac{e^{-is\omega}}{2\pi i s} \phi_w(h_w s) \frac{\hat{\phi}_{Y_{it}}(s|x)}{\hat{\phi}_{U_{it}}(s|x)} ds$$

- Conditional deconvolution (cf. Fan, 1991; Diggle and Hall, 1993)
- Derive the rates of convergence
- Complications: zeros in the denominators
- $\phi_{U_{it}}(s|x) \to 0$ as $s \to \pm\infty$. Two cases:
  - Ordinary-smooth: $\inf_x |\phi_{U_{it}}(s|x)| \sim C|s|^{-\lambda_U}$,
    - eg: Laplace, χ^2_k, Exp, Gamma
  - Super-smooth: $\inf_x |\phi_{U_{it}}(s|x)| \sim C\exp(-C_2|s|^{\lambda_U})$,
    - eg: Normal, Student-t_k, Cauchy

Estimation, Conditional Deconvolution

Estimator:

$$\hat{F}_{m(x,\alpha_i)}(\omega|x) = \frac{1}{2} + \int_{-\infty}^{\infty} \frac{e^{-is\omega}}{2\pi i s} \phi_w(h_w s) \frac{\hat{\phi}_{Y_{it}}(s|x)}{\hat{\phi}_{U_{it}}(s|x)} ds$$

- Conditional deconvolution (cf. Fan, 1991; Diggle and Hall, 1993)
- Derive the rates of convergence
- Complications: zeros in the denominators
- $\phi_{U_{it}}(s|x) \to 0$ as $s \to \pm\infty$. Two cases:
  - Ordinary-smooth: $\inf_x |\phi_{U_{it}}(s|x)| \sim C|s|^{-\lambda_U}$,
    - eg: Laplace, χ^2_k, Exp, Gamma
  - Super-smooth: $\inf_x |\phi_{U_{it}}(s|x)| \sim C\exp(-C_2|s|^{\lambda_U})$,
    - eg: Normal, Student-t_k, Cauchy

Estimation, Conditional Deconvolution

Estimator:

$$\hat{F}_{m(x,\alpha_i)}(\omega|x) = \frac{1}{2} + \int_{-\infty}^{\infty} \frac{e^{-is\omega}}{2\pi i s} \phi_w(h_w s) \frac{\hat{\phi}_{Y_{it}}(s|x)}{\hat{\phi}_{U_{it}}(s|x)} ds$$

- Conditional deconvolution (cf. Fan, 1991; Diggle and Hall, 1993)
- Derive the rates of convergence
- Complications: zeros in the denominators
- $\phi_{U_{it}}(s|x) \to 0$ as $s \to \pm\infty$. Two cases:
  - Ordinary-smooth: $\inf_x |\phi_{U_{it}}(s|x)| \sim C|s|^{-\lambda_U}$,
    - eg: Laplace, χ^2_k, Exp, Gamma
  - Super-smooth: $\inf_x |\phi_{U_{it}}(s|x)| \sim C\exp(-C_2|s|^{\lambda_U})$,
    - eg: Normal, Student-t_k, Cauchy

Estimation, Conditional Deconvolution

Estimator:

$$\hat{F}_{m(x,\alpha_i)}(\omega|x) = \frac{1}{2} + \int_{-\infty}^{\infty} \frac{e^{-is\omega}}{2\pi i s} \phi_w(h_w s) \frac{\hat{\phi}_{Y_{it}}(s|x)}{\hat{\phi}_{U_{it}}(s|x)} ds$$

- Conditional deconvolution (cf. Fan, 1991; Diggle and Hall, 1993)
- Derive the rates of convergence
- Complications: zeros in the denominators
- $\phi_{U_{it}}(s|x) \to 0$ as $s \to \pm\infty$. Two cases:
  - Ordinary-smooth: $\inf_x |\phi_{U_{it}}(s|x)| \sim C|s|^{-\lambda_U}$,
    - eg: Laplace, χ^2_k, Exp, Gamma
  - Super-smooth: $\inf_x |\phi_{U_{it}}(s|x)| \sim C\exp(-C_2|s|^{\lambda_U})$,
    - eg: Normal, Student-t_k, Cauchy

Decomposition of $\hat{F}_{m(x,\alpha_i)}(w|x) - F_{m(x,\alpha_i)}(w|x_t)$

$$
\begin{aligned}
& \frac{1}{2} + \int_{-\infty}^{\infty} \frac{e^{-is\omega}}{2\pi i s} \phi_w(h_w s) \frac{\hat{\phi}_{Y_{it}}(s|x)}{\hat{\phi}_{U_{it}}(s|x)} ds - F_{m(x,\alpha_i)}(\omega|x) \\
= \quad & F_{m(x,\alpha_i)*K_w(\cdot/h_w)}(\omega|x) - F_{m(x,\alpha_i)}(\omega|x) \\
& + \int_{-\infty}^{\infty} \frac{e^{-is\omega}}{2\pi i s} \phi_w(h_w s) \left[\frac{\hat{\phi}_{Y_{it}}(s|x) - \phi_{Y_{it}}(s|x)}{\phi_{U_{it}}(s|x)} \right] ds \\
& + \int_{-\infty}^{\infty} \frac{e^{-is\omega}}{2\pi i s} \phi_w(h_w s) \phi_{m(x,\alpha_i)} \left[\frac{\hat{\phi}_{U_{it}}(s|x) - \phi_{U_{it}}(s|x)}{\phi_{U_{it}}(s|x)} \right] ds \\
= \quad & \text{deterministic error} \\
& + \text{stochastic error} \\
& + \text{error from est}'\text{n of } \phi_{U_{it}}(s|x)
\end{aligned}
$$

Decomposition of $\hat{F}_{m(x,\alpha_i)}(w|x) - F_{m(x,\alpha_i)}(w|x_t)$

$$
\begin{aligned}
& \frac{1}{2} + \int_{-\infty}^{\infty} \frac{e^{-is\omega}}{2\pi i s} \phi_w(h_w s) \frac{\hat{\phi}_{Y_{it}}(s|x)}{\hat{\phi}_{U_{it}}(s|x)} ds - F_{m(x,\alpha_i)}(\omega|x) \\
= \quad & F_{m(x,\alpha_i)*K_w(\cdot/h_w)}(\omega|x) - F_{m(x,\alpha_i)}(\omega|x) \\
& + \int_{-\infty}^{\infty} \frac{e^{-is\omega}}{2\pi i s} \phi_w(h_w s) \left[\frac{\hat{\phi}_{Y_{it}}(s|x) - \phi_{Y_{it}}(s|x)}{\phi_{U_{it}}(s|x)} \right] ds \\
& + \int_{-\infty}^{\infty} \frac{e^{-is\omega}}{2\pi i s} \phi_w(h_w s) \phi_{m(x,\alpha_i)} \left[\frac{\hat{\phi}_{U_{it}}(s|x) - \phi_{U_{it}}(s|x)}{\phi_{U_{it}}(s|x)} \right] ds \\
= \quad & \text{deterministic error} \\
& + \text{stochastic error} \\
& + \text{error from est}'\text{n of } \phi_{U_{it}}(s|x)
\end{aligned}
$$

Decomposition of $\hat{F}_{m(x,\alpha_i)}(w|x) - F_{m(x,\alpha_i)}(w|x_t)$

$$
\begin{aligned}
& \frac{1}{2} + \int_{-\infty}^{\infty} \frac{e^{-is\omega}}{2\pi i s} \phi_w(h_w s) \frac{\hat{\phi}_{Y_{it}}(s|x)}{\hat{\phi}_{U_{it}}(s|x)} ds - F_{m(x,\alpha_i)}(\omega|x) \\
= \quad & F_{m(x,\alpha_i) * K_w(\cdot/h_w)}(\omega|x) - F_{m(x,\alpha_i)}(\omega|x) \\
& + \int_{-\infty}^{\infty} \frac{e^{-is\omega}}{2\pi i s} \phi_w(h_w s) \left[\frac{\hat{\phi}_{Y_{it}}(s|x) - \phi_{Y_{it}}(s|x)}{\phi_{U_{it}}(s|x)} \right] ds \\
& + \int_{-\infty}^{\infty} \frac{e^{-is\omega}}{2\pi i s} \phi_w(h_w s) \phi_{m(x,\alpha_i)} \left[\frac{\hat{\phi}_{U_{it}}(s|x) - \phi_{U_{it}}(s|x)}{\phi_{U_{it}}(s|x)} \right] ds \\
= \quad & \text{deterministic error} \\
& + \text{stochastic error} \\
& + \text{error from est}'\text{n of } \phi_{U_{it}}(s|x)
\end{aligned}
$$

Decomposition of $\hat{F}_{m(x,\alpha_i)}(w|x) - F_{m(x,\alpha_i)}(w|x_t)$

$$
\begin{aligned}
& \frac{1}{2} + \int_{-\infty}^{\infty} \frac{e^{-is\omega}}{2\pi i s} \phi_w(h_w s) \frac{\hat{\phi}_{Y_{it}}(s|x)}{\hat{\phi}_{U_{it}}(s|x)} ds - F_{m(x,\alpha_i)}(\omega|x) \\
= \quad & F_{m(x,\alpha_i) * K_w(\cdot/h_w)}(\omega|x) - F_{m(x,\alpha_i)}(\omega|x) \\
& + \int_{-\infty}^{\infty} \frac{e^{-is\omega}}{2\pi i s} \phi_w(h_w s) \left[\frac{\hat{\phi}_{Y_{it}}(s|x) - \phi_{Y_{it}}(s|x)}{\phi_{U_{it}}(s|x)} \right] ds \\
& + \int_{-\infty}^{\infty} \frac{e^{-is\omega}}{2\pi i s} \phi_w(h_w s) \phi_{m(x,\alpha_i)} \left[\frac{\hat{\phi}_{U_{it}}(s|x) - \phi_{U_{it}}(s|x)}{\phi_{U_{it}}(s|x)} \right] ds \\
= \quad & \text{deterministic error} \\
& + \text{stochastic error} \\
& + \text{error from est}'\text{n of } \phi_{U_{it}}(s|x)
\end{aligned}
$$

Assume $x \in [0,1]^d$. Want to obtain $\varkappa_n$

$$\sup_{(x,w)\in[\delta,1-\delta]^d\times\mathbb{R}} \left|\hat{F}_{m(x,\alpha_i)}(w|x) - F_{m(x,\alpha_i)}(w|x)\right| = O_p(\varkappa_n)$$

which depends on

k_w - smooth. of $m(x,\alpha)$ and $F_{\alpha_i}(\alpha|x)$ in α
k_x - smooth. of $m(x,\alpha)$ and $F_{\alpha_i}(\alpha|x)$ in x
λ_U - $\left|\phi_{U_{it}}(s|x)\right| \sim C|s|^{-\lambda_U}$ as $|s| \to \infty$
d - dimension of x

Theorem

Under some regularity conditions $\varkappa_n = (\log(n)/n)^{\rho_{OS}}$

$$\rho_{OS} = \frac{1}{2}\frac{k_w}{k_w + 1 + \lambda_U + k_x d/(2k_x + d)}\frac{2k_x}{2k_x + d}$$

In the super-smooth case the rate is $\varkappa_n = (\log n)^{-k_w/\lambda_U}$

Assume $x \in [0,1]^d$. Want to obtain $\varkappa_n$

$$\sup_{(x,w)\in[\delta,1-\delta]^d\times\mathbb{R}} \left|\hat{F}_{m(x,\alpha_i)}(w|x) - F_{m(x,\alpha_i)}(w|x)\right| = O_p(\varkappa_n)$$

which depends on

k_w	- smooth. of $m(x,\alpha)$ and $F_{\alpha_i}(\alpha\|x)$ in α
k_x	- smooth. of $m(x,\alpha)$ and $F_{\alpha_i}(\alpha\|x)$ in x
λ_U	- $\left\|\phi_{U_{it}}(s\|x)\right\| \sim C\|s\|^{-\lambda_U}$ as $\|s\| \to \infty$
d	- dimension of x

Theorem

Under some regularity conditions $\varkappa_n = (\log(n)/n)^{\rho_{OS}}$

$$\rho_{OS} = \frac{1}{2}\frac{k_w}{k_w + 1 + \lambda_U + k_x d/(2k_x + d)}\frac{2k_x}{2k_x + d}$$

In the super-smooth case the rate is $\varkappa_n = (\log n)^{-k_w/\lambda_U}$

Assume $x \in [0,1]^d$. Want to obtain $\varkappa_n$

$$\sup_{(x,w)\in[\delta,1-\delta]^d\times\mathbb{R}} \left|\hat{F}_{m(x,\alpha_i)}(w|x) - F_{m(x,\alpha_i)}(w|x)\right| = O_p(\varkappa_n)$$

which depends on

k_w - smooth. of $m(x,\alpha)$ and $F_{\alpha_i}(\alpha|x)$ in α
k_x - smooth. of $m(x,\alpha)$ and $F_{\alpha_i}(\alpha|x)$ in x
λ_U - $\left|\phi_{U_{it}}(s|x)\right| \sim C|s|^{-\lambda_U}$ as $|s| \to \infty$
d - dimension of x

Theorem

Under some regularity conditions $\varkappa_n = (\log(n)/n)^{\rho_{OS}}$

$$\rho_{OS} = \frac{1}{2}\frac{k_w}{k_w + 1 + \lambda_U + k_x d/(2k_x + d)}\frac{2k_x}{2k_x + d}$$

In the super-smooth case the rate is $\varkappa_n = (\log n)^{-k_w/\lambda_U}$

Assume $x \in [0,1]^d$. Want to obtain $\varkappa_n$

$$\sup_{(x,w)\in[\delta,1-\delta]^d\times\mathbb{R}} \left|\hat{F}_{m(x,\alpha_i)}(w|x) - F_{m(x,\alpha_i)}(w|x)\right| = O_p(\varkappa_n)$$

which depends on

k_w	- smooth. of $m(x,\alpha)$ and $F_{\alpha_i}(\alpha\|x)$ in α
k_x	- smooth. of $m(x,\alpha)$ and $F_{\alpha_i}(\alpha\|x)$ in x
λ_U	- $\left\|\phi_{U_{it}}(s\|x)\right\| \sim C\|s\|^{-\lambda_U}$ as $\|s\| \to \infty$
d	- dimension of x

Theorem

Under some regularity conditions $\varkappa_n = (\log(n)/n)^{\rho_{OS}}$

$$\rho_{OS} = \frac{1}{2}\frac{k_w}{k_w + 1 + \lambda_U + k_x d/(2k_x + d)}\frac{2k_x}{2k_x + d}$$

In the super-smooth case the rate is $\varkappa_n = (\log n)^{-k_w/\lambda_U}$

Assume $x \in [0,1]^d$. Want to obtain $\varkappa_n$

$$\sup_{(x,w)\in[\delta,1-\delta]^d\times\mathbb{R}} \left|\hat{F}_{m(x,\alpha_i)}(w|x) - F_{m(x,\alpha_i)}(w|x)\right| = O_p(\varkappa_n)$$

which depends on

k_w	- smooth. of $m(x,\alpha)$ and $F_{\alpha_i}(\alpha\|x)$ in α
k_x	- smooth. of $m(x,\alpha)$ and $F_{\alpha_i}(\alpha\|x)$ in x
λ_U	- $\left\|\phi_{U_{it}}(s\|x)\right\| \sim C\|s\|^{-\lambda_U}$ as $\|s\| \to \infty$
d	- dimension of x

Theorem

Under some regularity conditions $\varkappa_n = (\log(n)/n)^{\rho_{OS}}$

$$\rho_{OS} = \frac{1}{2}\frac{k_w}{k_w + 1 + \lambda_U + k_x d/(2k_x + d)}\frac{2k_x}{2k_x + d}$$

In the super-smooth case the rate is $\varkappa_n = (\log n)^{-k_w/\lambda_U}$

Assume $x \in [0,1]^d$. Want to obtain $\varkappa_n$

$$\sup_{(x,w)\in[\delta,1-\delta]^d\times\mathbb{R}} \left|\hat{F}_{m(x,\alpha_i)}(w|x) - F_{m(x,\alpha_i)}(w|x)\right| = O_p(\varkappa_n)$$

which depends on

$$\begin{array}{ll} k_w & \text{- smooth. of } m(x,\alpha) \text{ and } F_{\alpha_i}(\alpha|x) \text{ in } \alpha \\ k_x & \text{- smooth. of } m(x,\alpha) \text{ and } F_{\alpha_i}(\alpha|x) \text{ in } x \\ \lambda_U & \text{- } \left|\phi_{U_{it}}(s|x)\right| \sim C|s|^{-\lambda_U} \text{ as } |s| \to \infty \\ d & \text{- dimension of } x \end{array}$$

Theorem

Under some regularity conditions $\varkappa_n = (\log(n)/n)^{\rho_{OS}}$

$$\rho_{OS} = \frac{1}{2}\frac{k_w}{k_w + 1 + \lambda_U + k_x d/(2k_x + d)}\frac{2k_x}{2k_x + d}$$

In the super-smooth case the rate is $\varkappa_n = (\log n)^{-k_w/\lambda_U}$

Assume $x \in [0,1]^d$. Want to obtain $\varkappa_n$

$$\sup_{(x,w)\in[\delta,1-\delta]^d\times\mathbb{R}} \left|\hat{F}_{m(x,\alpha_i)}(w|x) - F_{m(x,\alpha_i)}(w|x)\right| = O_p(\varkappa_n)$$

which depends on

k_w - smooth. of $m(x,\alpha)$ and $F_{\alpha_i}(\alpha|x)$ in α
k_x - smooth. of $m(x,\alpha)$ and $F_{\alpha_i}(\alpha|x)$ in x
λ_U - $\left|\phi_{U_{it}}(s|x)\right| \sim C|s|^{-\lambda_U}$ as $|s| \to \infty$
d - dimension of x

Theorem

Under some regularity conditions $\varkappa_n = (\log(n)/n)^{\rho_{OS}}$

$$\rho_{OS} = \frac{1}{2}\underbrace{\frac{k_w}{k_w + 1 + \lambda_U + k_x d/(2k_x + d)}}_{Fan(1991)\ (when\ d=0)}\underbrace{\frac{2k_x}{2k_x + d}}_{Stone(1982)}$$

In the super-smooth case the rate is $\varkappa_n = (\log n)^{-k_w/\lambda_U}$

Assume $x \in [0,1]^d$. Want to obtain $\varkappa_n$

$$\sup_{(x,w)\in[\delta,1-\delta]^d\times\mathbb{R}} \left|\hat{F}_{m(x,\alpha_i)}(w|x) - F_{m(x,\alpha_i)}(w|x)\right| = O_p(\varkappa_n)$$

which depends on

k_w - smooth. of $m(x,\alpha)$ and $F_{\alpha_i}(\alpha|x)$ in α
k_x - smooth. of $m(x,\alpha)$ and $F_{\alpha_i}(\alpha|x)$ in x
λ_U - $\left|\phi_{U_{it}}(s|x)\right| \sim C|s|^{-\lambda_U}$ as $|s| \to \infty$
d - dimension of x

Theorem

Under some regularity conditions $\varkappa_n = (\log(n)/n)^{\rho_{OS}}$

$$\rho_{OS} = \frac{1}{2}\underbrace{\frac{k_w}{k_w + 1 + \lambda_U + k_x d/(2k_x + d)}}_{\text{Fan(1991) (when } d=0)}\underbrace{\frac{2k_x}{2k_x + d}}_{\text{Stone(1982)}}$$

In the super-smooth case the rate is $\varkappa_n = (\log n)^{-k_w/\lambda_U}$

Assume $x \in [0,1]^d$. Want to obtain $\varkappa_n$

$$\sup_{(x,w)\in[\delta,1-\delta]^d\times\mathbb{R}} \left|\hat{F}_{m(x,\alpha_i)}(w|x) - F_{m(x,\alpha_i)}(w|x)\right| = O_p(\varkappa_n)$$

which depends on

k_w - smooth. of $m(x,\alpha)$ and $F_{\alpha_i}(\alpha|x)$ in α
k_x - smooth. of $m(x,\alpha)$ and $F_{\alpha_i}(\alpha|x)$ in x
λ_U - $\left|\phi_{U_{it}}(s|x)\right| \sim C|s|^{-\lambda_U}$ as $|s| \to \infty$
d - dimension of x

Theorem

Under some regularity conditions $\varkappa_n = (\log(n)/n)^{\rho_{OS}}$

$$\rho_{OS} = \frac{1}{2}\underbrace{\frac{k_w}{k_w + 1 + \lambda_U + k_x d/(2k_x + d)}}_{Fan(1991)\ (when\ d=0)}\underbrace{\frac{2k_x}{2k_x + d}}_{Stone(1982)}$$

In the super-smooth case the rate is $\varkappa_n = (\log n)^{-k_w/\lambda_U}$

Assume $x \in [0,1]^d$. Want to obtain $\varkappa_n$

$$\sup_{(x,w)\in[\delta,1-\delta]^d\times\mathbb{R}} \left|\hat{F}_{m(x,\alpha_i)}(w|x) - F_{m(x,\alpha_i)}(w|x)\right| = O_p(\varkappa_n)$$

which depends on

k_w - smooth. of $m(x,\alpha)$ and $F_{\alpha_i}(\alpha|x)$ in α
k_x - smooth. of $m(x,\alpha)$ and $F_{\alpha_i}(\alpha|x)$ in x
λ_U - $\left|\phi_{U_{it}}(s|x)\right| \sim C|s|^{-\lambda_U}$ as $|s| \to \infty$
d - dimension of x

Theorem

Under some regularity conditions $\varkappa_n = (\log(n)/n)^{\rho_{OS}}$

$$\rho_{OS} = \frac{1}{2}\underbrace{\frac{k_w}{k_w + 1 + \lambda_U + k_x d/(2k_x + d)}}_{Fan(1991)\ (when\ d=0)}\underbrace{\frac{2k_x}{2k_x + d}}_{Stone(1982)}$$

In the super-smooth case the rate is $\varkappa_n = (\log n)^{-k_w/\lambda_U}$

Monte Carlo

- $Y_{it} = m(X_{it}, \alpha_i) + U_{it}, \quad n = 2500,\ T = 2$
- $X_{it} \sim i.i.d. Uniform\,[0,1]$
- $\alpha_i \sim \frac{\rho}{\sqrt{T}} \sum_{t=1}^{T} \sqrt{12}\,(X_{it} - 0.5) + \sqrt{1-\rho^2}\psi_i,\ \rho = 0.5$
- $m(x, \alpha) = 2\alpha + (2+\alpha)(2x-1)^3$

- where $U_{it} = \sigma_0^2 (1 + X_{it})^2 \varepsilon_{it}$,
- $\varepsilon_{it} \sim i.i.d.\ Laplace\ /\ N(0,1)$
- $\psi_i \sim i.i.d.\ N(0,1)$ (skewed and bimodal distr'ns are in the paper)

Monte Carlo

- $Y_{it} = m(X_{it}, \alpha_i) + U_{it}, \quad n = 2500,\ T = 2$
- $X_{it} \sim i.i.d. Uniform\,[0,1]$
- $\alpha_i \sim \frac{\rho}{\sqrt{T}} \sum_{t=1}^{T} \sqrt{12}\,(X_{it} - 0.5) + \sqrt{1-\rho^2}\psi_i$, $\rho = 0.5$
- $m(x, \alpha) = 2\alpha + (2+\alpha)(2x-1)^3$

- where $U_{it} = \sigma_0^2 (1 + X_{it})^2 \varepsilon_{it}$,
- $\varepsilon_{it} \sim i.i.d.$ *Laplace* / $N(0,1)$
- $\psi_i \sim i.i.d.\ N(0,1)$ (skewed and bimodal distr'ns are in the paper)

Monte Carlo

- $Y_{it} = m(X_{it}, \alpha_i) + U_{it}, \quad n = 2500,\ T = 2$
- $X_{it} \sim i.i.d. Uniform\,[0,1]$
- $\alpha_i \sim \frac{\rho}{\sqrt{T}} \sum_{t=1}^{T} \sqrt{12}\,(X_{it} - 0.5) + \sqrt{1-\rho^2}\psi_i,\ \rho = 0.5$
- $m(x, \alpha) = 2\alpha + (2+\alpha)(2x-1)^3$

- where $U_{it} = \sigma_0^2 (1 + X_{it})^2 \varepsilon_{it}$,
- $\varepsilon_{it} \sim i.i.d.\ Laplace\ /\ N(0,1)$
- $\psi_i \sim i.i.d.\ N(0,1)$ (skewed and bimodal distr'ns are in the paper)

Monte Carlo

- $Y_{it} = m(X_{it}, \alpha_i) + U_{it}, \quad n = 2500,\ T = 2$
- $X_{it} \sim i.i.d. Uniform\,[0,1]$
- $\alpha_i \sim \frac{\rho}{\sqrt{T}} \sum_{t=1}^{T} \sqrt{12}\,(X_{it} - 0.5) + \sqrt{1-\rho^2}\psi_i,\ \rho = 0.5$
- $m(x, \alpha) = 2\alpha + (2+\alpha)(2x-1)^3$

- where $U_{it} = \sigma_0^2 (1 + X_{it})^2 \varepsilon_{it}$,
- $\varepsilon_{it} \sim i.i.d.\ Laplace\ /\ N(0,1)$
- $\psi_i \sim i.i.d.\ N(0,1)$ (skewed and bimodal distr'ns are in the paper)

Monte Carlo

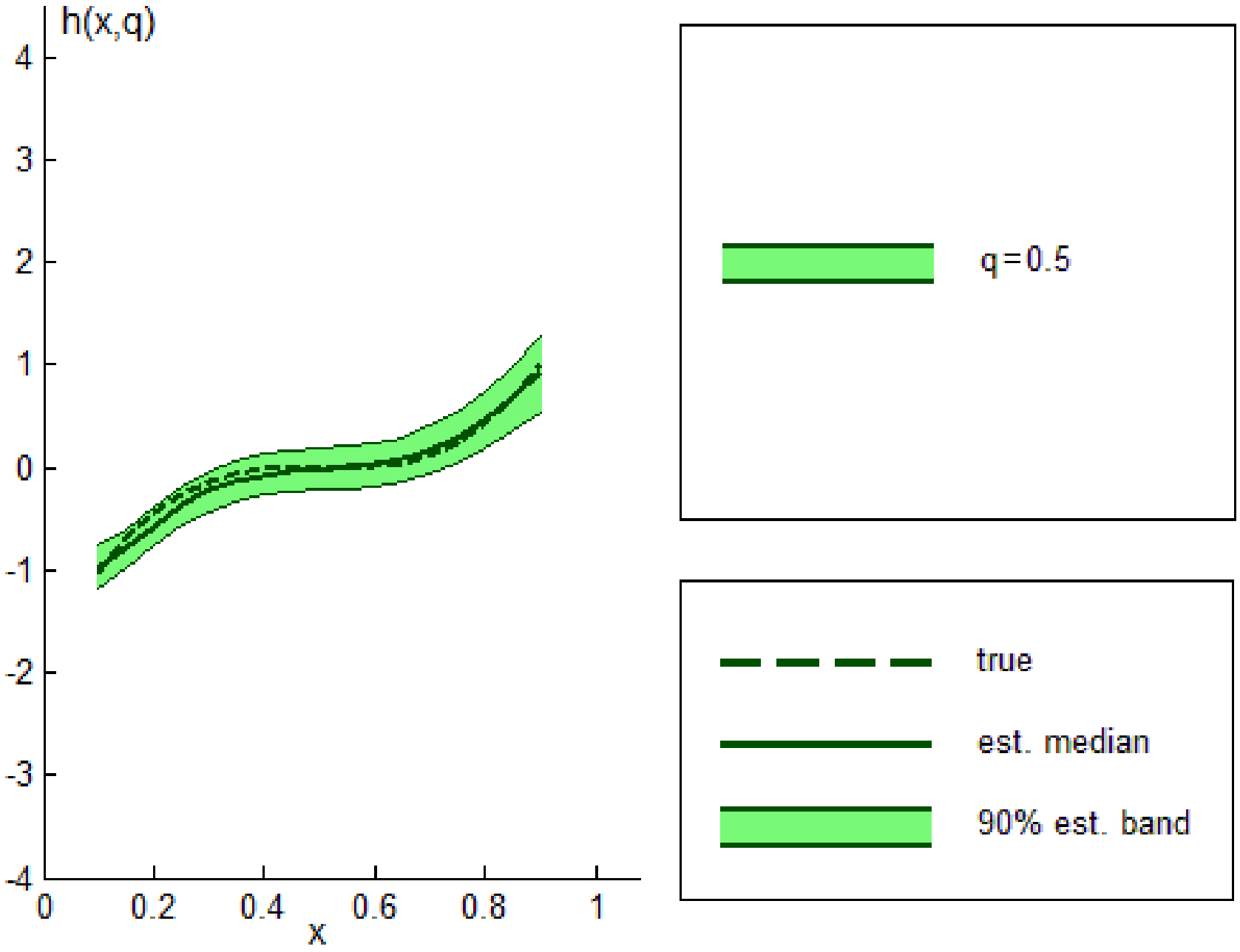
- $Y_{it} = m(X_{it}, \alpha_i) + U_{it}, \quad n = 2500,\ T = 2$
- $X_{it} \sim i.i.d. Uniform\,[0,1]$
- $\alpha_i \sim \frac{\rho}{\sqrt{T}} \sum_{t=1}^{T} \sqrt{12}\,(X_{it} - 0.5) + \sqrt{1-\rho^2}\psi_i,\ \rho = 0.5$
- $m(x, \alpha) = 2\alpha + (2+\alpha)(2x-1)^3$

- where $U_{it} = \sigma_0^2 (1 + X_{it})^2 \varepsilon_{it}$,
- $\varepsilon_{it} \sim i.i.d.\ Laplace\ /\ N(0,1)$
- $\psi_i \sim i.i.d.\ N(0,1)$ (skewed and bimodal distr'ns are in the paper)

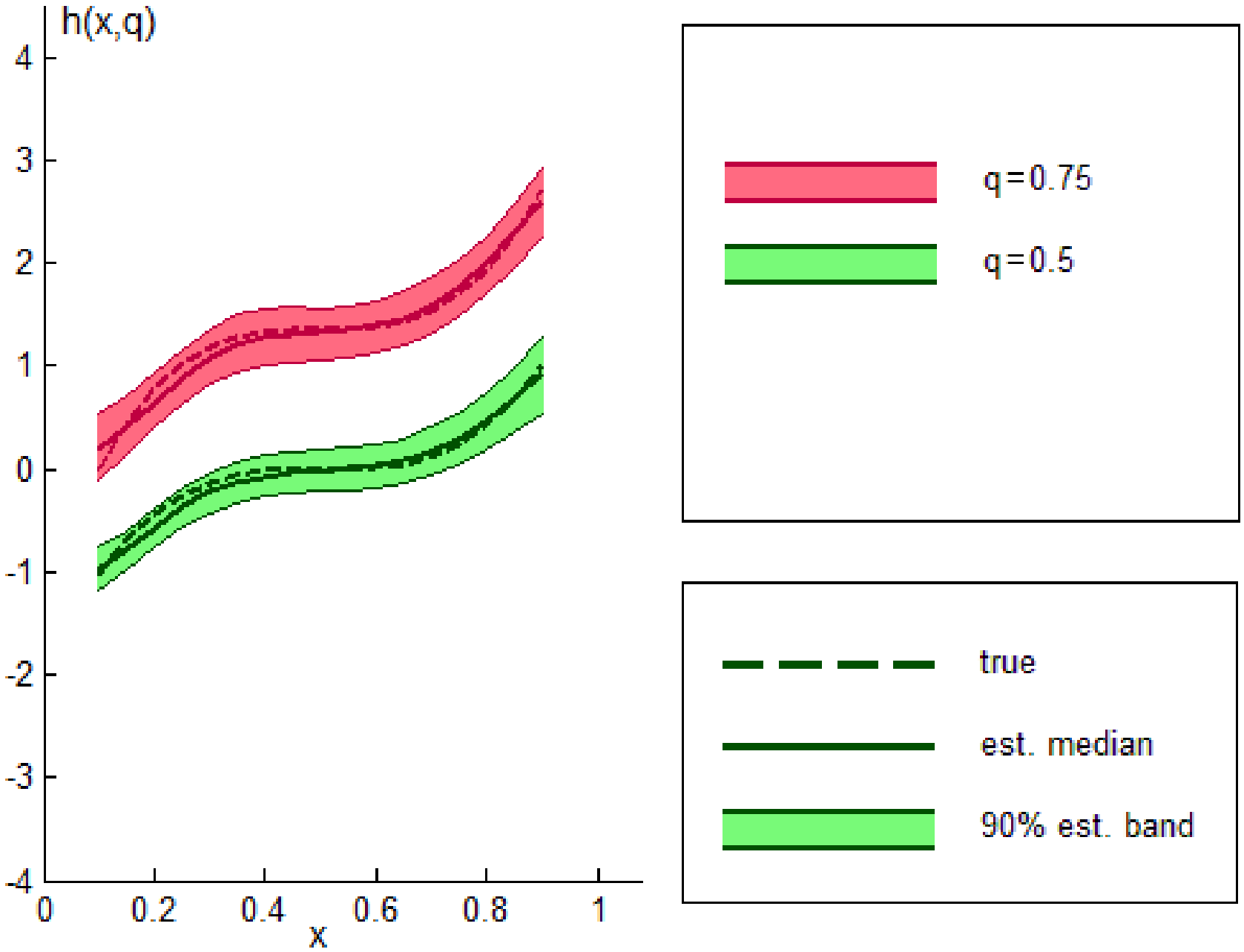
Monte Carlo

- $Y_{it} = m(X_{it}, \alpha_i) + U_{it}, \quad n = 2500,\ T = 2$
- $X_{it} \sim i.i.d. Uniform\,[0,1]$
- $\alpha_i \sim \frac{\rho}{\sqrt{T}} \sum_{t=1}^{T} \sqrt{12}\,(X_{it} - 0.5) + \sqrt{1-\rho^2}\psi_i,\ \rho = 0.5$
- $m(x, \alpha) = 2\alpha + (2+\alpha)(2x-1)^3$

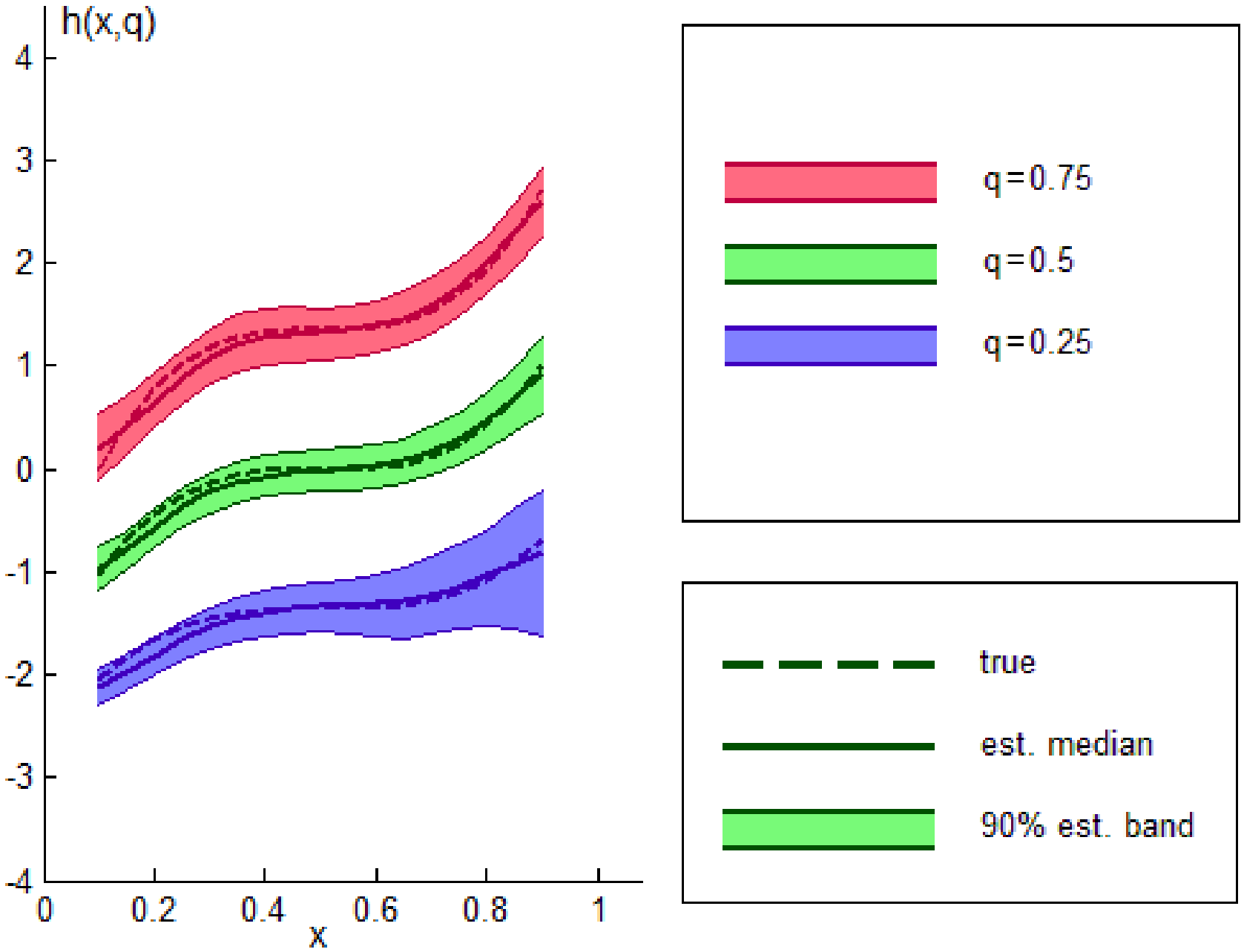
- where $U_{it} = \sigma_0^2 (1 + X_{it})^2 \varepsilon_{it}$,
- $\varepsilon_{it} \sim i.i.d.\ Laplace\ /\ N(0,1)$
- $\psi_i \sim i.i.d.\ N(0,1)$ (skewed and bimodal distr'ns are in the paper)



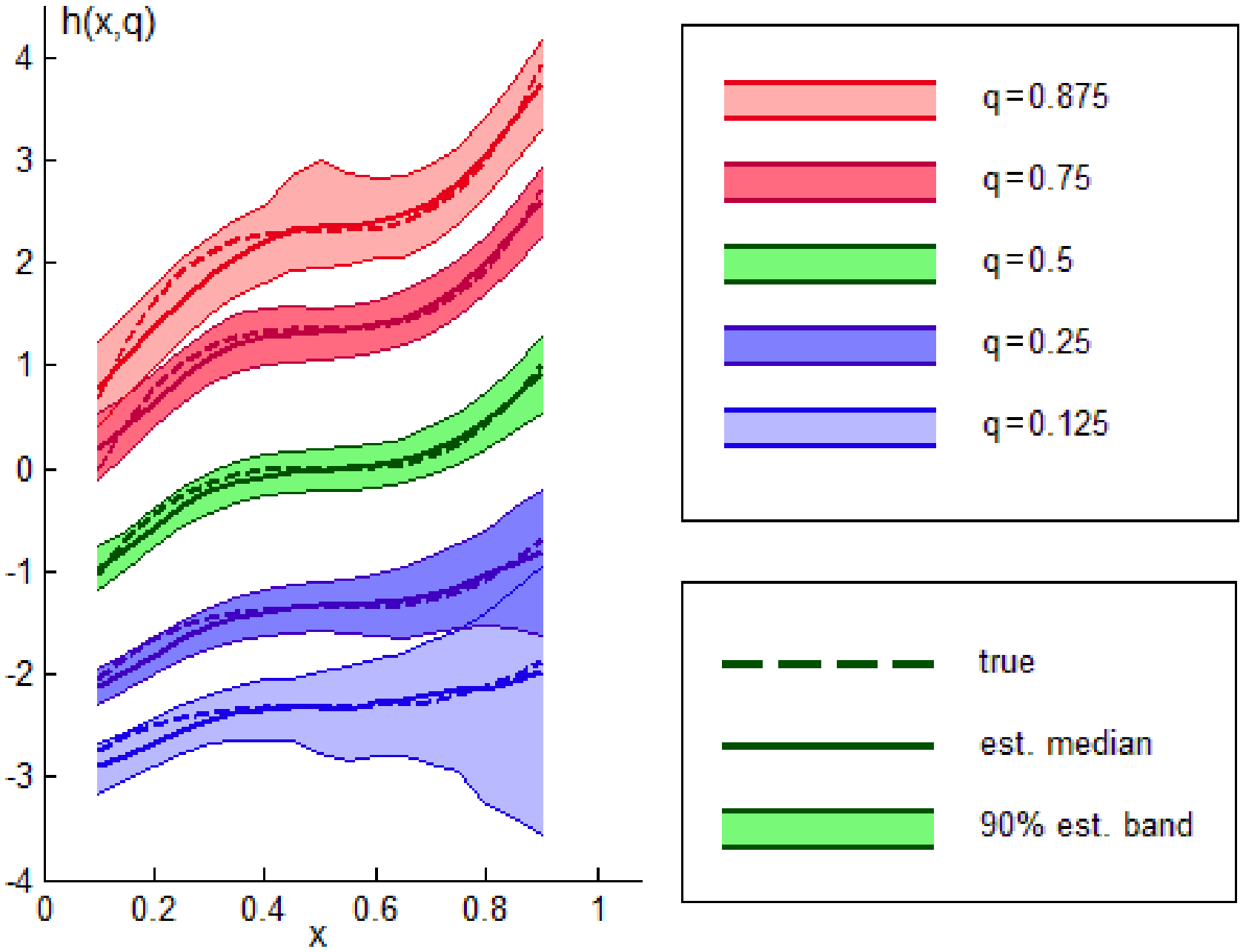
Laplace, N=2500, T=2



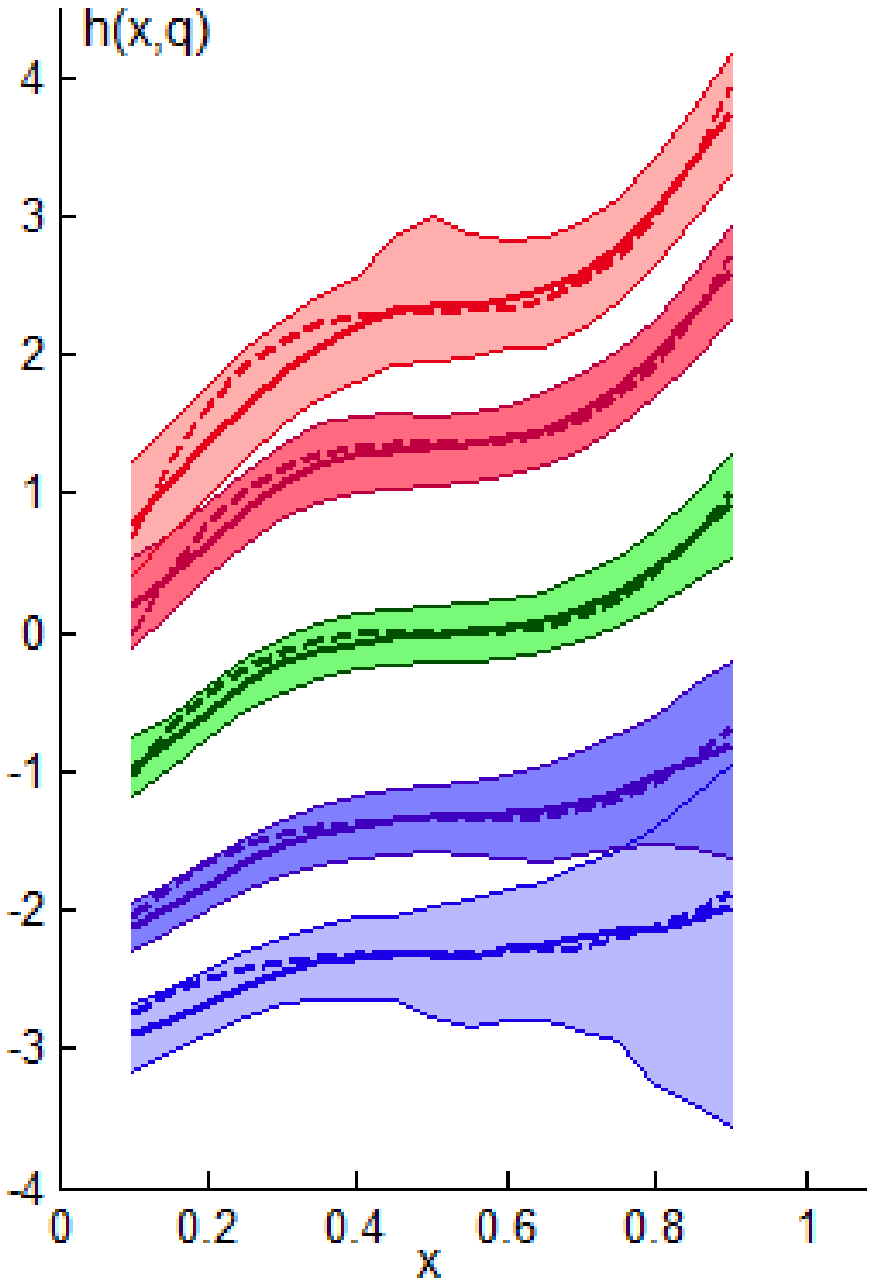
Laplace, N=2500, T=2



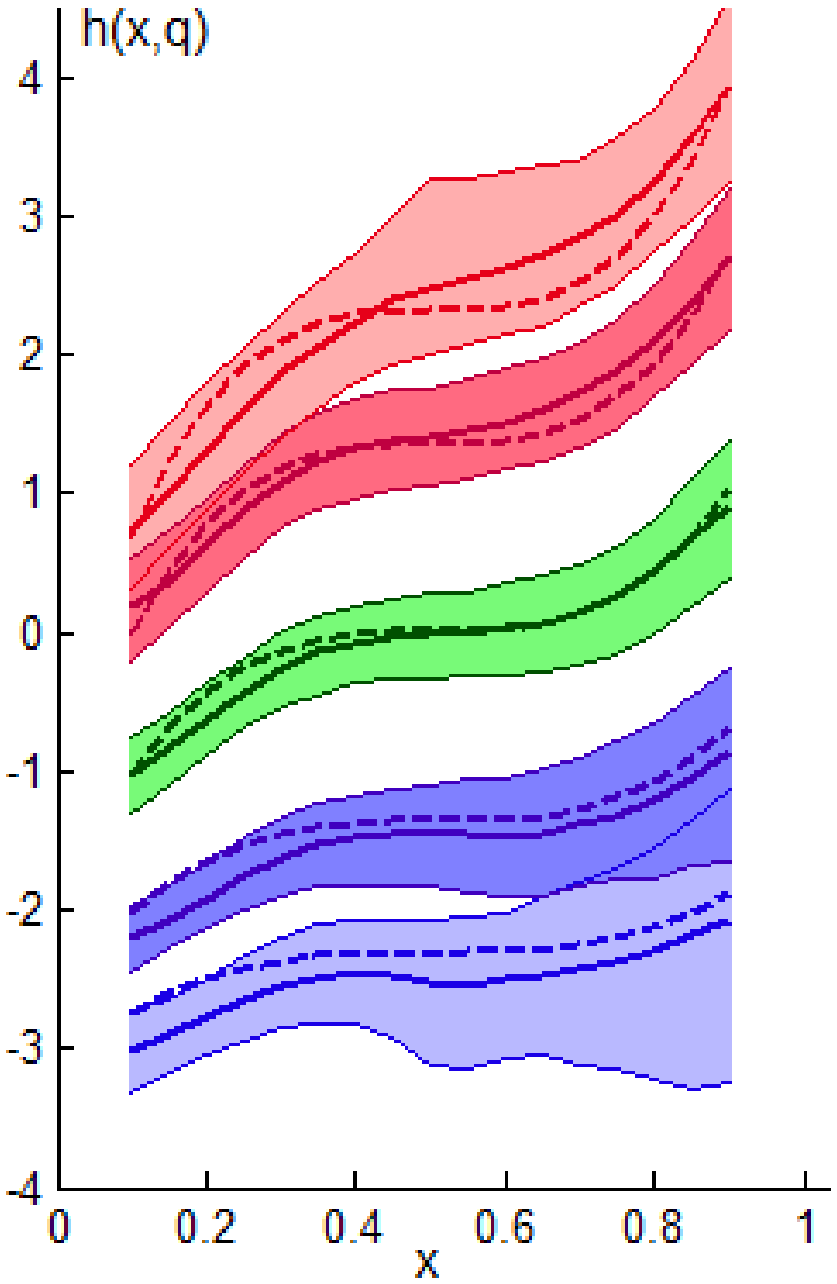
Laplace, N=2500, T=2



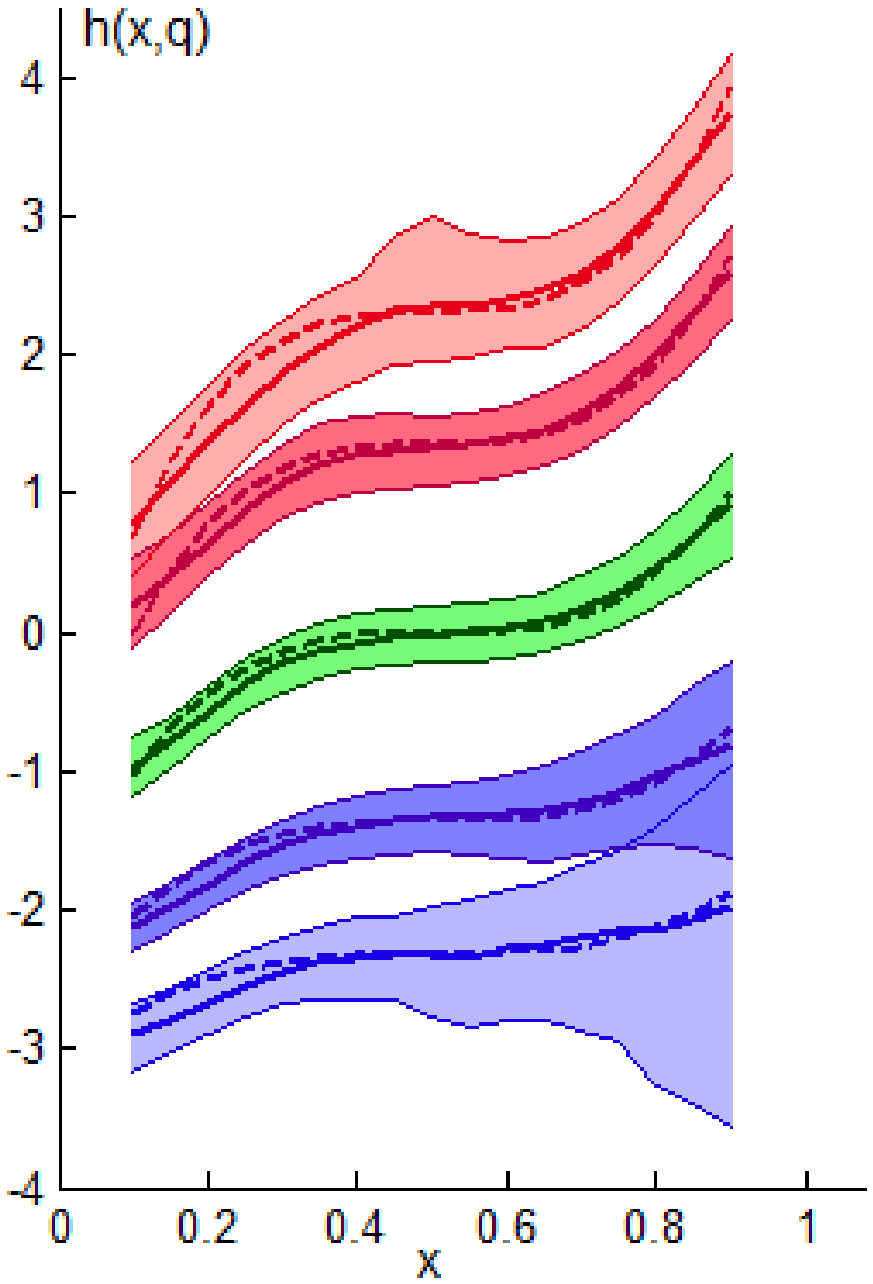
Laplace, N=2500, T=2



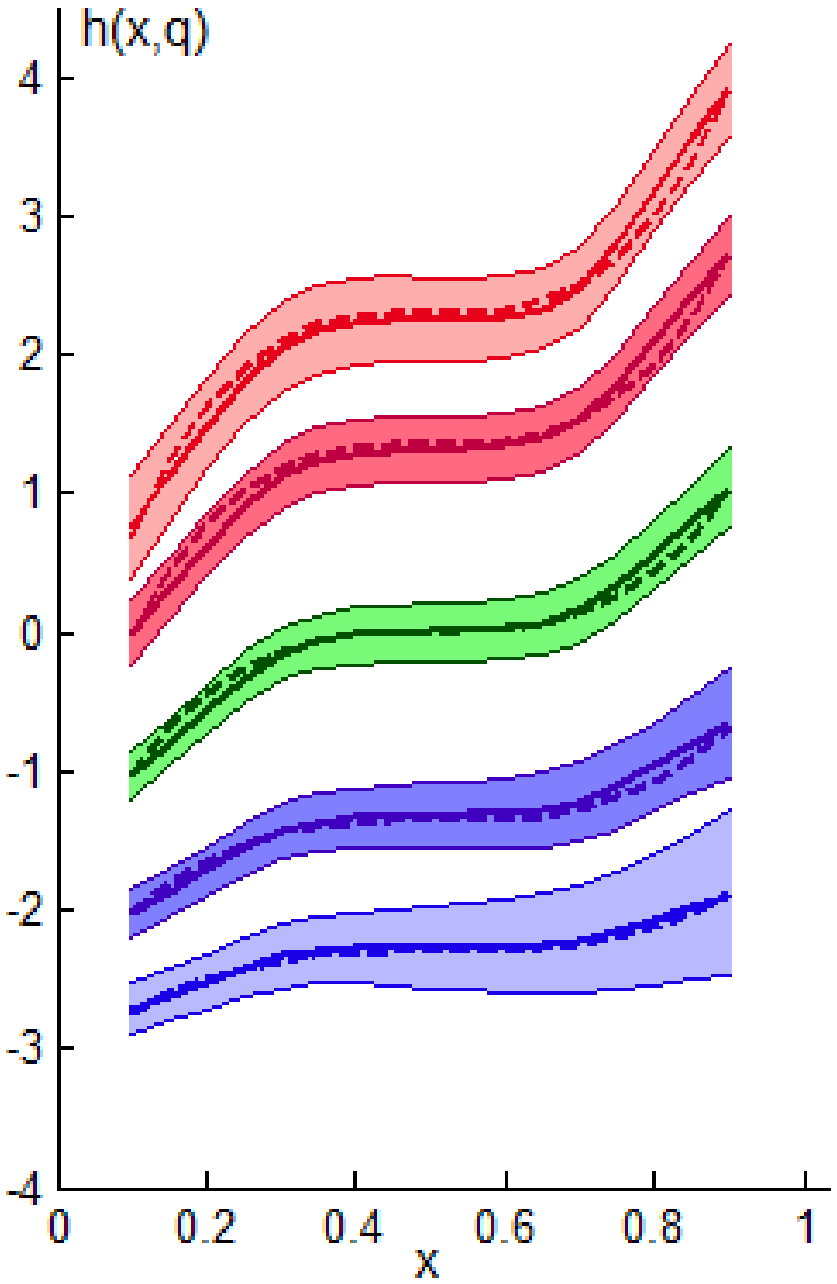
Laplace, N=2500, T=2



Normal, N=2500, T=2



Laplace, N=2500, T=2



Infeasible, N=2500, T=2

Empirical Illustration - Union Wage Premium

- Data: Current Population Survey (CPS), $T = 2$
  - matching
  - migration
  - measurement error
  - limited set of variables
  - 1987/1988, $n = 17,756$ (replicating Card, 1996)
- Most related papers: Card (1996) and Lemieux (1998)

Union Wage Premium

$$Y_{it} = m(X_{it}, \alpha_i) + U_{it}$$

- Y_{it} is (log)-wage,
- X_{it} is 1 if member of a union, 0 o/w,
- α_i is skill and U_{it} is luck.

- $\Delta(\alpha) = m(1, \alpha) - m(0, \alpha)$ is the union wage premium.

Empirical Illustration - Union Wage Premium

- Data: Current Population Survey (CPS), $T = 2$
  - matching
  - migration
  - measurement error
  - limited set of variables
  - 1987/1988, $n = 17,756$ (replicating Card, 1996)
- Most related papers: Card (1996) and Lemieux (1998)

Union Wage Premium

$$Y_{it} = m(X_{it}, \alpha_i) + U_{it}$$

- Y_{it} is (log)-wage,
- X_{it} is 1 if member of a union, 0 o/w,
- α_i is skill and U_{it} is luck.

- $\Delta(\alpha) = m(1, \alpha) - m(0, \alpha)$ is the union wage premium.

Empirical Illustration - Union Wage Premium

- Data: Current Population Survey (CPS), $T = 2$
  - matching
  - migration
  - measurement error
  - limited set of variables
  - 1987/1988, $n = 17,756$ (replicating Card, 1996)
- Most related papers: Card (1996) and Lemieux (1998)

Union Wage Premium

$$Y_{it} = m(X_{it}, \alpha_i) + U_{it}$$

- Y_{it} is (log)-wage,
- X_{it} is 1 if member of a union, 0 o/w,
- α_i is skill and U_{it} is luck.

- $\Delta(\alpha) = m(1, \alpha) - m(0, \alpha)$ is the union wage premium.

Empirical Illustration - Union Wage Premium

- Data: Current Population Survey (CPS), $T = 2$
  - matching
  - migration
  - measurement error
  - limited set of variables
  - 1987/1988, $n = 17,756$ (replicating Card, 1996)
- Most related papers: Card (1996) and Lemieux (1998)

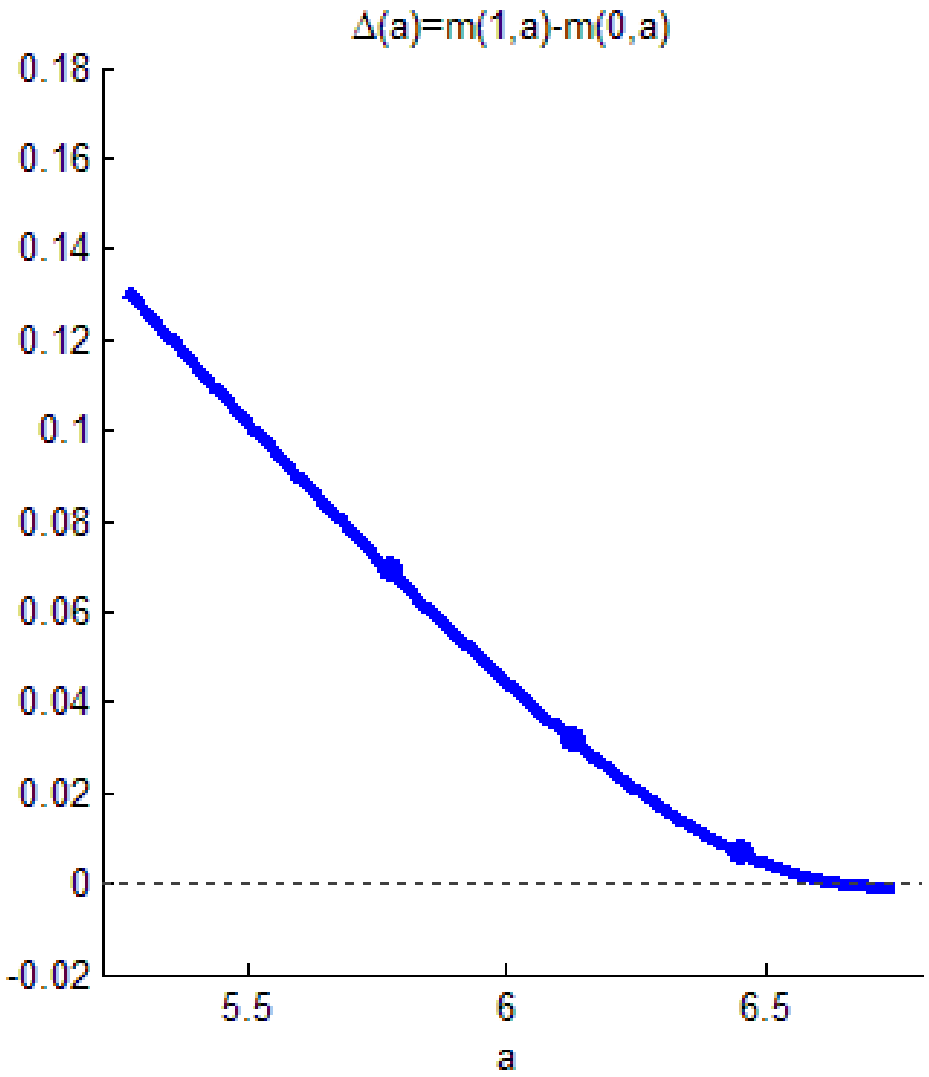
Union Wage Premium

$$Y_{it} = m(X_{it}, \alpha_i) + U_{it}$$

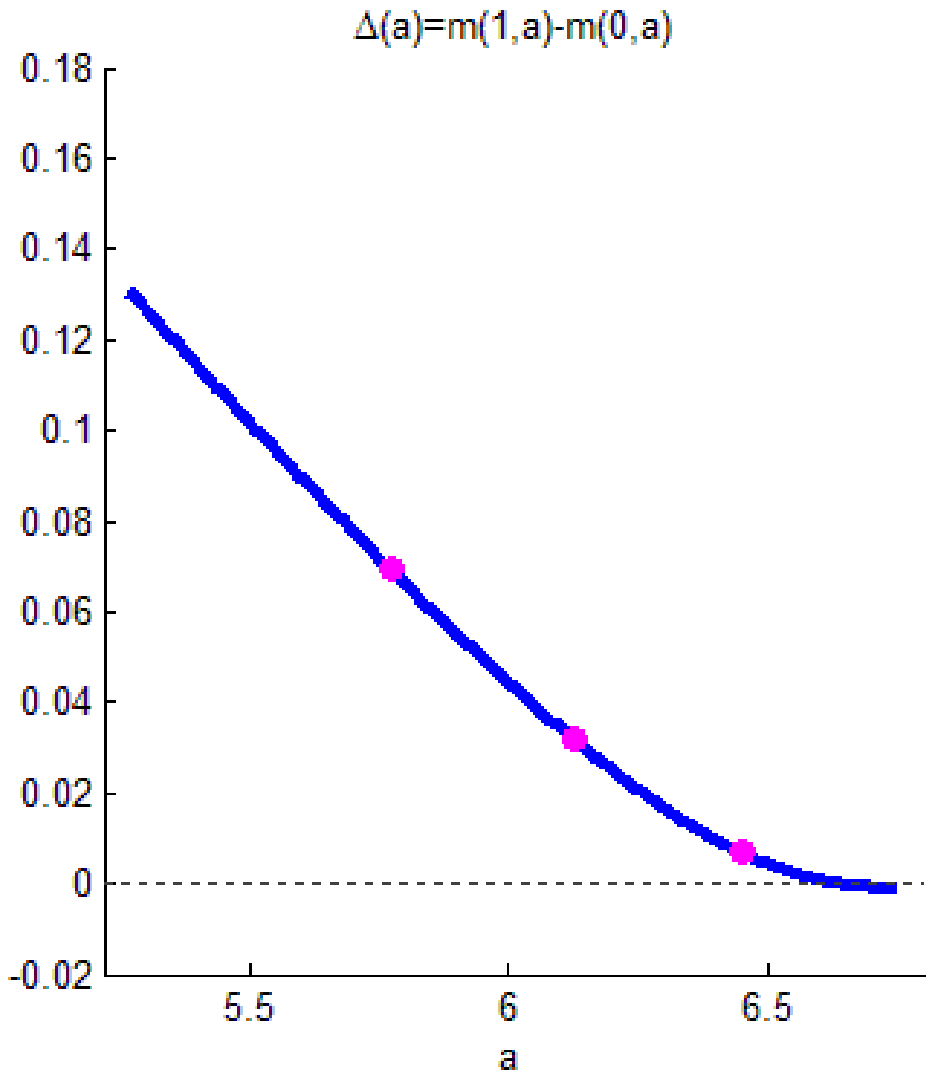
- Y_{it} is (log)-wage,
- X_{it} is 1 if member of a union, 0 o/w,
- α_i is skill and U_{it} is luck.

- $\Delta(\alpha) = m(1, \alpha) - m(0, \alpha)$ is the union wage premium.

Empirical Illustration - Union Wage Premium



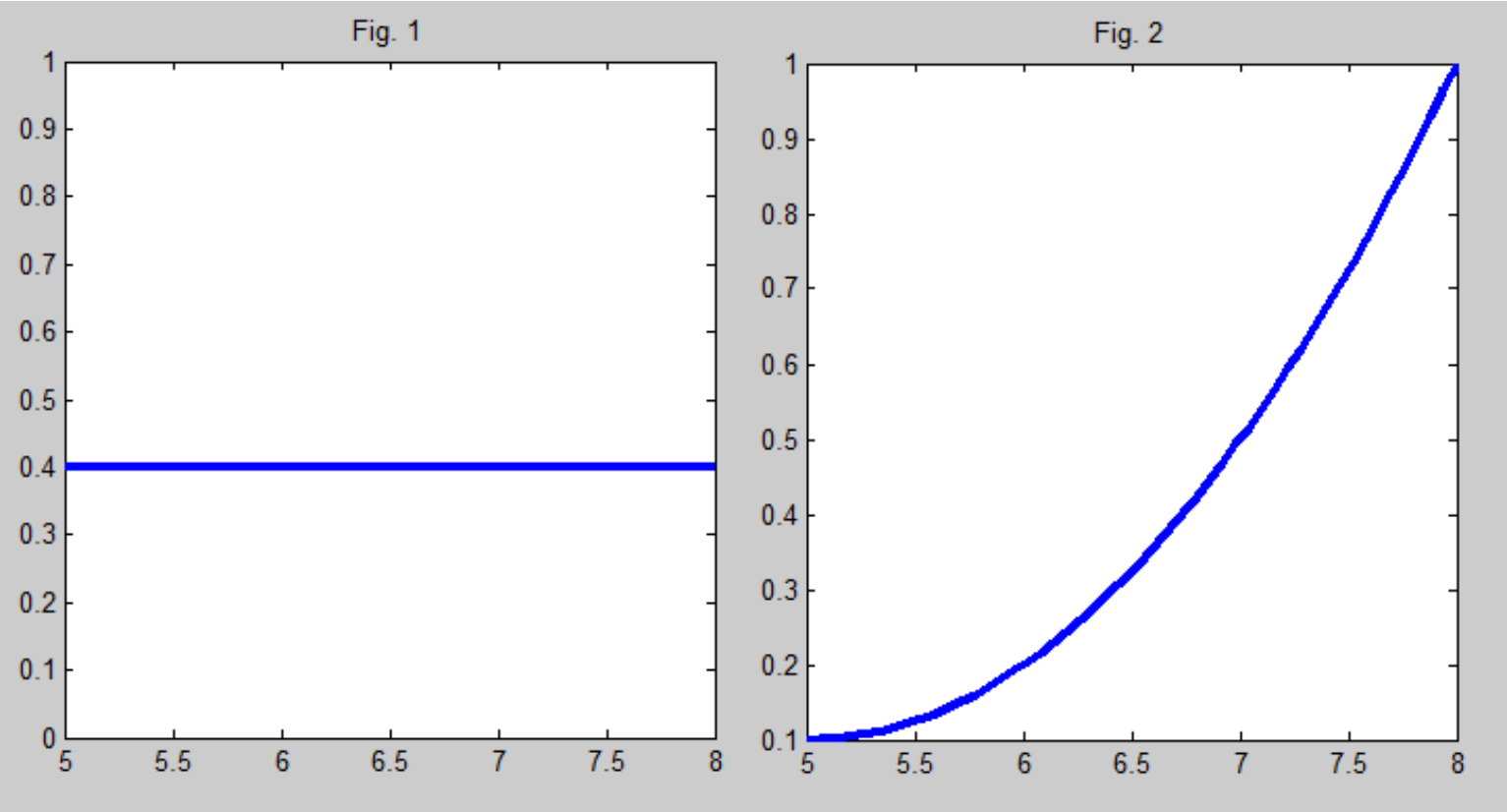
Empirical Illustration - Union Wage Premium



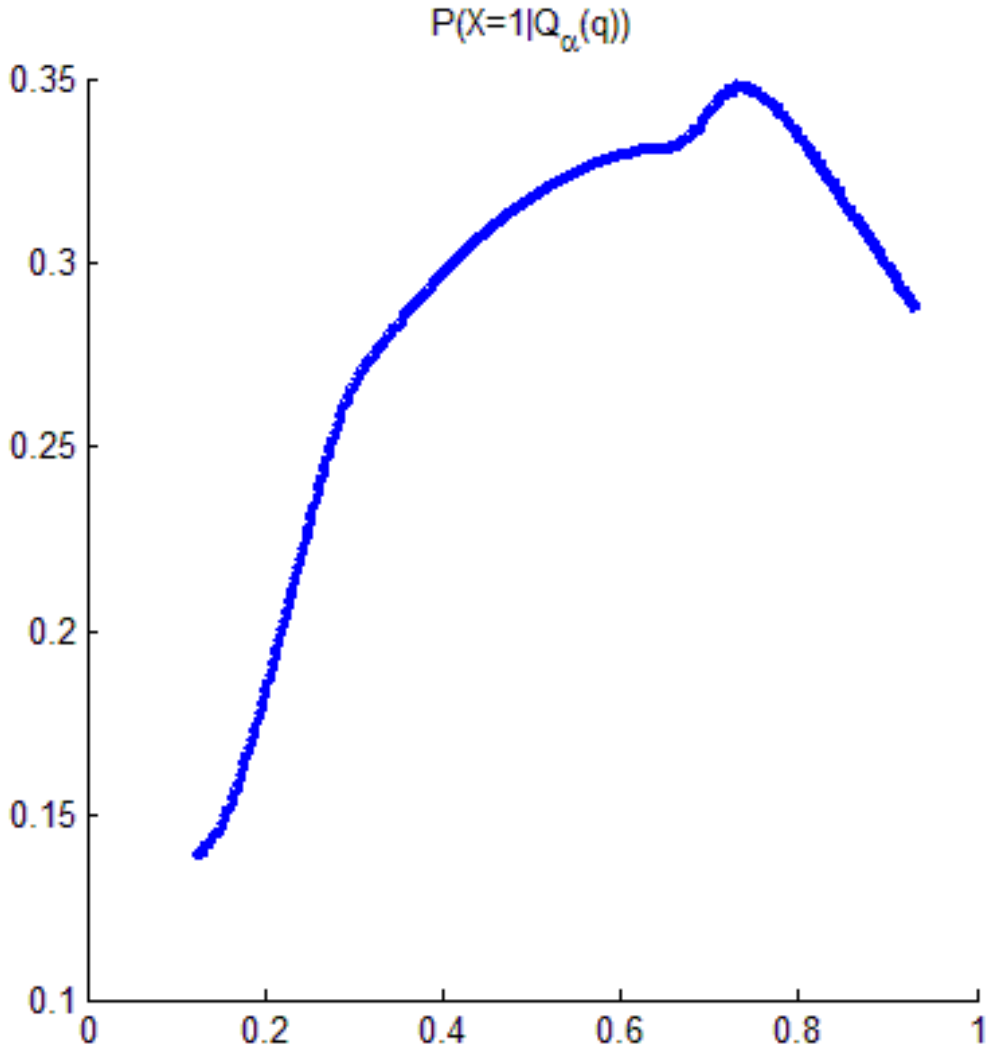
$$\Delta\left(Q_{\alpha_i}(0.25)\right) = 7.0\%,\ \Delta\left(Q_{\alpha_i}(0.5)\right) = 3.2\%,\ \Delta\left(Q_{\alpha_i}(0.75)\right) = 0.7\%$$

Empirical Illustration - Propensity Score

Also interested in $P\left(X_{it}=1|\alpha_i=a\right)$. Examples:



Empirical Illustration - Propensity Score



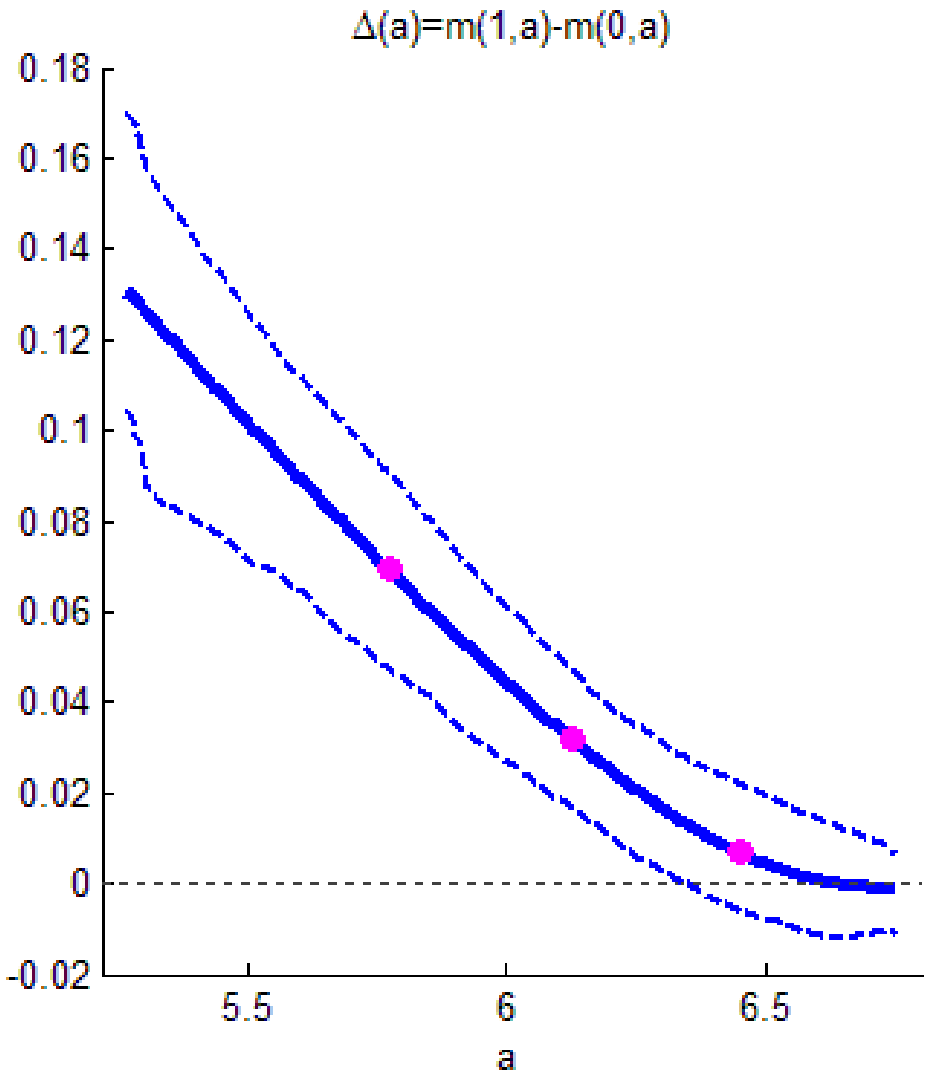
Summary

- Identification of a nonparametric panel with nonseparable heterogeneity

$$Y_{it} = m\left(X_{it}, \alpha_i\right) + U_{it} \qquad (1)$$

- Allows identification of distributional, policy, and counterfactual effects
- Practical estimation procedure
- Rates of convergence for the problem of conditional deconvolution

Empirical Illustration - Union Wage Premium



$$\Delta\left(Q_{\alpha_i}(0.25)\right) = 7.0\%,\ \Delta\left(Q_{\alpha_i}(0.5)\right) = 3.2\%,\ \Delta\left(Q_{\alpha_i}(0.75)\right) = 0.7\%$$