xtfeis.ado: Linear Fixed Effects Models with Individual Slopes

Volker Ludwig[1]

[1]University of Mannheim, Mannheim Centre for European Social Research (MZES)
vludwig@mail.uni-mannheim.de

8th German Stata Users Group Meeting
June 25, 2010

Outline

The Problem. Consider growth curves

- Fixed Effect framework: the ultimate method for causal analysis with non-experimental data

The Problem. Consider growth curves

- Fixed Effect framework: the ultimate method for causal analysis with non-experimental data
- But many applications where conventional FE models fail because strict exogeneity is violated

The Problem. Consider growth curves

- Fixed Effect framework: the ultimate method for causal analysis with non-experimental data
- But many applications where conventional FE models fail because strict exogeneity is violated
  - here: time-constant unobserved factors correlate with observed factors
  - major example: unobserved effect changes over time

The Problem. Consider growth curves

- Fixed Effect framework: the ultimate method for causal analysis with non-experimental data
- But many applications where conventional FE models fail because strict exogeneity is violated
  - here: time-constant unobserved factors correlate with observed factors
  - major example: unobserved effect changes over time
- Problem recognized (Allison 1990, Heckman & Hotz 1989, Polachek & Kim 1994, Winship & Morgan 1999, Morgan & Winship 2007)

The Problem. Consider growth curves

- Fixed Effect framework: the ultimate method for causal analysis with non-experimental data
- But many applications where conventional FE models fail because strict exogeneity is violated
  - here: time-constant unobserved factors correlate with observed factors
  - major example: unobserved effect changes over time
- Problem recognized (Allison 1990, Heckman & Hotz 1989, Polachek & Kim 1994, Winship & Morgan 1999, Morgan & Winship 2007)
- but seldomly solved in practice

UNIVERSITÄT MANNHEIM

The Problem. Consider growth curves

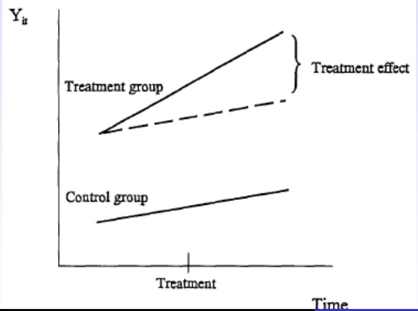
Allison (1990), Winship & Morgan (1999) consider 3 situations:

- 1. Absent treatment, potential outcomes of treatment and control group develop along the same path

The Problem. Consider growth curves

Allison (1990), Winship & Morgan (1999) consider 3 situations:

- 1. Absent treatment, potential outcomes of treatment and control group develop along the same path



The Problem. Consider growth curves

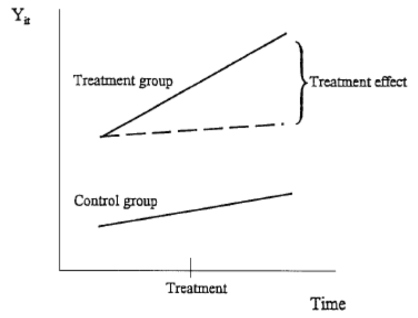
Allison (1990), Winship & Morgan (1999) consider 3 situations:

- 2. Absent treatment, potential outcomes converge

The Problem. Consider growth curves

Allison (1990), Winship & Morgan (1999) consider 3 situations:

- 2. Absent treatment, potential outcomes converge



The Problem. Consider growth curves

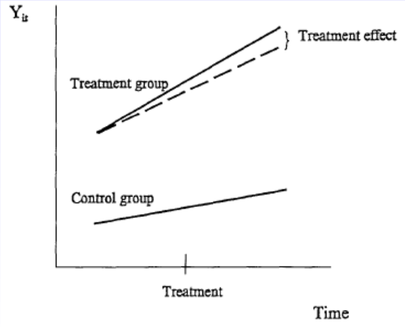
Allison (1990), Winship & Morgan (1999) consider 3 situations:

- 3. Absent treatment, potential outcomes diverge

The Problem. Consider growth curves

Allison (1990), Winship & Morgan (1999) consider 3 situations:

- 3. Absent treatment, potential outcomes diverge



The Problem. Consider growth curves

Allison (1990), Winship & Morgan (1999) consider 3 situations:

- 1. potential outcomes of treatment and control group develop along the same path

The Problem. Consider growth curves

Allison (1990), Winship & Morgan (1999) consider 3 situations:

- 1. potential outcomes of treatment and control group develop along the same path
  - ideal case, assumption of FE model (or "‘change score model"’) hold

The Problem. Consider growth curves

Allison (1990), Winship & Morgan (1999) consider 3 situations:

- 1. potential outcomes of treatment and control group develop along the same path
  - ideal case, assumption of FE model (or "‘change score model"’) hold
- 2. potential outcomes converge

UNIVERSITÄT MANNHEIM

The Problem. Consider growth curves

Allison (1990), Winship & Morgan (1999) consider 3 situations:

- 1. potential outcomes of treatment and control group develop along the same path
  - ideal case, assumption of FE model (or "'change score model"') hold
- 2. potential outcomes converge
  - assumption of FE model violated, use ANCOVA model instead

UNIVERSITÄT MANNHEIM

The Problem. Consider growth curves

Allison (1990), Winship & Morgan (1999) consider 3 situations:

- 1. potential outcomes of treatment and control group develop along the same path
  - ideal case, assumption of FE model (or "‘change score model"’) hold
- 2. potential outcomes converge
  - assumption of FE model violated, use ANCOVA model instead
- 3. potential outcomes diverge

UNIVERSITÄT MANNHEIM

The Problem. Consider growth curves

Allison (1990), Winship & Morgan (1999) consider 3 situations:

- 1. potential outcomes of treatment and control group develop along the same path
  - ideal case, assumption of FE model (or "‘change score model"’) hold
- 2. potential outcomes converge
  - assumption of FE model violated, use ANCOVA model instead
- 3. potential outcomes diverge
  - assumption of FE model and ANCOVA violated

The Problem. Consider growth curves

Allison (1990), Winship & Morgan (1999) consider 3 situations:

- 1. potential outcomes of treatment and control group develop along the same path
  - ideal case, assumption of FE model (or "‘change score model"’) hold
- 2. potential outcomes converge
  - assumption of FE model violated, use ANCOVA model instead
- 3. potential outcomes diverge
  - assumption of FE model and ANCOVA violated
  - solution ???

UNIVERSITÄT MANNHEIM

The Problem. Consider growth curves

Allison (1990), Winship & Morgan (1999) consider 3 situations:

- 1. potential outcomes of treatment and control group develop along the same path
  - ideal case, assumption of FE model (or "‘change score model"’) hold
- 2. potential outcomes converge
  - assumption of FE model violated, use ANCOVA model instead
- 3. potential outcomes diverge
  - assumption of FE model and ANCOVA violated
  - solution ???
- 1., 2., 3.? Theories seldomly strong enough to decide

FE with individual Slopes, notation first

- Let Y_{it} : Outcome Y of individual i at time point t

UNIVERSITÄT MANNHEIM

FE with individual Slopes, notation first

- Let Y_{it} : Outcome Y of individual i at time point t
- D_{it} : Binary treatment indicator (time-varying),
  D_i : treatment indicator (time-constant)

FE with individual Slopes, notation first

- Let Y_{it} : Outcome Y of individual i at time point t
- D_{it} : Binary treatment indicator (time-varying),
  D_i : treatment indicator (time-constant)

FE with individual Slopes, notation first

- Let Y_{it} : Outcome Y of individual i at time point t
- D_{it} : Binary treatment indicator (time-varying),
  D_i : treatment indicator (time-constant)
- α_{1i} : individual constants, i.e. time-constant individual heterogeneity, usually correlated with D_{it}

UNIVERSITÄT MANNHEIM

FE with individual Slopes, notation first

- Let Y_{it} : Outcome Y of individual i at time point t
- D_{it} : Binary treatment indicator (time-varying),
  D_i : treatment indicator (time-constant)
- α_{1i} : individual constants, i.e. time-constant individual heterogeneity, usually correlated with D_{it}
- Z_{it} : further independent variable(s), e.g. time (T)

FE with individual Slopes, notation first

- Let Y_{it} : Outcome Y of individual i at time point t
- D_{it} : Binary treatment indicator (time-varying), D_i : treatment indicator (time-constant)
- α_{1i} : individual constants, i.e. time-constant individual heterogeneity, usually correlated with D_{it}
- Z_{it} : further independent variable(s), e.g. time (T)
- α_{2i} : individual slopes, time-constant individual heterogeneity, correlated with D_{it} and interacts with Z_{it} to produce the outcome

FE with individual Slopes, notation first

- Let Y_{it} : Outcome Y of individual i at time point t
- D_{it} : Binary treatment indicator (time-varying), D_i : treatment indicator (time-constant)
- α_{1i} : individual constants, i.e. time-constant individual heterogeneity, usually correlated with D_{it}
- Z_{it} : further independent variable(s), e.g. time (T)
- α_{2i} : individual slopes, time-constant individual heterogeneity, correlated with D_{it} and interacts with Z_{it} to produce the outcome
- ϵ_{it} : idiosyncratic time-varying disturbance term

Extending the FE framework

- Conventional FE

Extending the FE framework

- Conventional FE
  - within transform aka time-demeaning eliminates α_{1i}

Extending the FE framework

- Conventional FE
  - within transform aka time-demeaning eliminates α_{1i}
  - estimate
    $(Y_{it} - \overline{Y_i}) = \beta(D_{it} - \overline{D_i}) + \gamma(Z_{it} - \overline{Z_i}) + (\alpha_{1i} - \alpha_{1i}) + (\epsilon_{it} - \overline{\epsilon_{it}})$

UNIVERSITÄT MANNHEIM

Extending the FE framework

- Conventional FE
  - within transform aka time-demeaning eliminates α_{1i}
  - estimate
    $(Y_{it} - \overline{Y_i}) = \beta(D_{it} - \overline{D_i}) + \gamma(Z_{it} - \overline{Z_i}) + (\alpha_{1i} - \alpha_{1i}) + (\epsilon_{it} - \overline{\epsilon_{it}})$
- Extended FE, including individual slopes

Extending the FE framework

- Conventional FE
  - within transform aka time-demeaning eliminates α_{1i}
  - estimate
    $(Y_{it} - \overline{Y_i}) = \beta(D_{it} - \overline{D_i}) + \gamma(Z_{it} - \overline{Z_i}) + (\alpha_{1i} - \alpha_{1i}) + (\epsilon_{it} - \overline{\epsilon_{it}})$
- Extended FE, including individual slopes
  - substract not the mean, but time-varying estimate to eliminate α_{1i} and α_{2i}

UNIVERSITÄT MANNHEIM

Extending the FE framework

- Conventional FE
  - within transform aka time-demeaning eliminates α_{1i}
  - estimate
    $(Y_{it} - \overline{Y_i}) = \beta(D_{it} - \overline{D_i}) + \gamma(Z_{it} - \overline{Z_i}) + (\alpha_{1i} - \alpha_{1i}) + (\epsilon_{it} - \overline{\epsilon_{it}})$
- Extended FE, including individual slopes
  - substract not the mean, but time-varying estimate to eliminate α_{1i} and α_{2i}
  - estimate
    $\widetilde{Y_{it}} = \beta \widetilde{D_{it}} + \widetilde{\epsilon_{it}}$, where $\widetilde{Y_{it}}$ is residual from individual time-series regression (OLS) of Y_{it} on Z_{it},
    and analog for indep. variables

Extending the FE framework

- Conventional FE
  - within transform aka time-demeaning eliminates α_{1i}
  - estimate
    $(Y_{it} - \overline{Y_i}) = \beta(D_{it} - \overline{D_i}) + \gamma(Z_{it} - \overline{Z_i}) + (\alpha_{1i} - \alpha_{1i}) + (\epsilon_{it} - \overline{\epsilon_{it}})$
- Extended FE, including individual slopes
  - substract not the mean, but time-varying estimate to eliminate α_{1i} and α_{2i}
  - estimate
    $\widetilde{Y_{it}} = \beta\widetilde{D_{it}} + \widetilde{\epsilon_{it}}$, where $\widetilde{Y_{it}}$ is residual from individual time-series regression (OLS) of Y_{it} on Z_{it},
    and analog for indep. variables
  - transform "'by hand"' possible, but very very slow

Extending the FE framework

- General approach to within transform (Wooldridge 2002):

Extending the FE framework

- General approach to within transform (Wooldridge 2002):
  - Premultiply all variables by matrix $\Omega_i = I_T - Z_i'(Z_i'Z_i)^{-1}Z_i'$

Extending the FE framework

- General approach to within transform (Wooldridge 2002):
  - Premultiply all variables by matrix $\Omega_i = I_T - Z_i'(Z_i'Z_i)^{-1}Z_i'$
  - Since $\Omega_i Z_i = 0$, α_{1i} and α_{2i} are eliminated

UNIVERSITÄT MANNHEIM

Extending the FE framework

- General approach to within transform (Wooldridge 2002):
  - Premultiply all variables by matrix $\Omega_i = I_T - Z_i'(Z_i'Z_i)^{-1}Z_i'$
  - Since $\Omega_i Z_i = 0$, α_{1i} and α_{2i} are eliminated
- Conventional FE is a special case where $Z_i = (1)$ is $(N \times 1)$ vector of constants

UNIVERSITÄT MANNHEIM

Extending the FE framework

- General approach to within transform (Wooldridge 2002):
  - Premultiply all variables by matrix $\Omega_i = I_T - Z_i'(Z_i'Z_i)^{-1}Z_i'$
  - Since $\Omega_i Z_i = 0$, α_{1i} and α_{2i} are eliminated
- Conventional FE is a special case where $Z_i = (1)$ is $(N \times 1)$ vector of constants
- Random growth model is another special case where $Z_i = (1, t)$ is $(N \times 2)$ matrix of constants and time variable

Stata ado `xtfeis.ado`

- based on mata

UNIVERSITÄT MANNHEIM

Stata ado xtfeis.ado

- based on mata
- appropriate when $T_i > J$, where J is number of Z variables (including possibly individual constants)

Stata ado xtfeis.ado

- based on mata
- appropriate when $T_i > J$, where J is number of Z variables (including possibly individual constants)
- automatically selects estimation sample

Stata ado xtfeis.ado

- based on mata
- appropriate when $T_i > J$, where J is number of Z variables (including possibly individual constants)
- automatically selects estimation sample
- panels may be unbalanced (missing data)

Stata ado `xtfeis.ado`

- based on mata
- appropriate when $T_i > J$, where J is number of Z variables (including possibly individual constants)
- automatically selects estimation sample
- panels may be unbalanced (missing data)
- collapses to conventional FE model w/o specifying Z variables

Stata ado xtfeis.ado

- based on mata
- appropriate when $T_i > J$, where J is number of Z variables (including possibly individual constants)
- automatically selects estimation sample
- panels may be unbalanced (missing data)
- collapses to conventional FE model w/o specifying Z variables
- fully robust s.e. on request

Stata ado `xtfeis.ado`

- based on mata
- appropriate when $T_i > J$, where J is number of Z variables (including possibly individual constants)
- automatically selects estimation sample
- panels may be unbalanced (missing data)
- collapses to conventional FE model w/o specifying Z variables
- fully robust s.e. on request
- Syntax: `xtfeis` *`varlist`*`, [slope(`*`varlist`*`)] [noconstant] [cluster(`*`clustvar`*`)]`

UNIVERSITÄT MANNHEIM

Simulation setup

- set up panel data set with 3000 cases ($N = 1000,\ T = 3$)

UNIVERSITÄT MANNHEIM

Simulation setup

- set up panel data set with 3000 cases ($N = 1000,\ T = 3$)
- 500 receive treatment between $t = 2$ and $t = 3$

UNIVERSITÄT MANNHEIM

Simulation setup

- set up panel data set with 3000 cases ($N = 1000,\ T = 3$)
- 500 receive treatment between $t = 2$ and $t = 3$
- choose true parameters:

Simulation setup

- set up panel data set with 3000 cases ($N = 1000,\ T = 3$)
- 500 receive treatment between $t = 2$ and $t = 3$
- choose true parameters:
  - treatment effect β

UNIVERSITÄT MANNHEIM

Simulation setup

- set up panel data set with 3000 cases ($N = 1000,\ T = 3$)
- 500 receive treatment between $t = 2$ and $t = 3$
- choose true parameters:
  - treatment effect β
  - normally distributed individual constants α_{1i}

UNIVERSITÄT MANNHEIM

Simulation setup

- set up panel data set with 3000 cases ($N = 1000, \ T = 3$)
- 500 receive treatment between $t = 2$ and $t = 3$
- choose true parameters:
  - treatment effect β
  - normally distributed individual constants α_{1i}
  - normally distributed individual slopes α_{2i}

UNIVERSITÄT MANNHEIM

Simulation setup

- set up panel data set with 3000 cases ($N = 1000,\ T = 3$)
- 500 receive treatment between $t = 2$ and $t = 3$
- choose true parameters:
  - treatment effect β
  - normally distributed individual constants α_{1i}
  - normally distributed individual slopes α_{2i}
- Generate outcome: $Y_{it} = D_{it} + \alpha_{1i} + \alpha_{2i}T + \epsilon_{it}$

Simulation setup

- set up panel data set with 3000 cases ($N = 1000,\ T = 3$)
- 500 receive treatment between $t = 2$ and $t = 3$
- choose true parameters:
  - treatment effect β
  - normally distributed individual constants α_{1i}
  - normally distributed individual slopes α_{2i}
- Generate outcome: $Y_{it} = D_{it} + \alpha_{1i} + \alpha_{2i} T + \epsilon_{it}$
- Estimate treatment effect $\hat{\beta}$ using 5 different models

Simulation setup

- set up panel data set with 3000 cases ($N = 1000,\ T = 3$)
- 500 receive treatment between $t = 2$ and $t = 3$
- choose true parameters:
  - treatment effect β
  - normally distributed individual constants α_{1i}
  - normally distributed individual slopes α_{2i}
- Generate outcome: $Y_{it} = D_{it} + \alpha_{1i} + \alpha_{2i}T + \epsilon_{it}$
- Estimate treatment effect $\hat{\beta}$ using 5 different models
- repeat 1000 times

UNIVERSITÄT MANNHEIM

Simulation setup

- set up panel data set with 3000 cases ($N = 1000,\ T = 3$)
- 500 receive treatment between $t = 2$ and $t = 3$
- choose true parameters:
  - treatment effect β
  - normally distributed individual constants α_{1i}
  - normally distributed individual slopes α_{2i}
- Generate outcome: $Y_{it} = D_{it} + \alpha_{1i} + \alpha_{2i}T + \epsilon_{it}$
- Estimate treatment effect $\hat{\beta}$ using 5 different models
- repeat 1000 times
- get mean of $\hat{\beta}$, s.e., % coefs. diff. from true β

UNIVERSITÄT MANNHEIM

Simulation: 5 different models

Simulation: 5 different models

Model	Equation

UNIVERSITÄT MANNHEIM

Simulation: 5 different models

Model	Equation
Pooled OLS	$Y_{it} = \beta D_{it} + \gamma T + \epsilon_{it}$

UNIVERSITÄT MANNHEIM

Simulation: 5 different models

Model	Equation
Pooled OLS	$Y_{it} = \beta D_{it} + \gamma T + \epsilon_{it}$
ANCOVA	$Y_{it+1} = \delta Y_{it} + \beta D_i + \gamma T + \epsilon_{it}$

UNIVERSITÄT MANNHEIM

Simulation: 5 different models

Model	Equation
Pooled OLS	$Y_{it} = \beta D_{it} + \gamma T + \epsilon_{it}$
ANCOVA	$Y_{it+1} = \delta Y_{it} + \beta D_i + \gamma T + \epsilon_{it}$
Change score	$Y_{it+1} - Y_{it} = \beta D_i + \gamma T + \epsilon_{it}$

Simulation: 5 different models

Model	Equation
Pooled OLS	$Y_{it} = \beta D_{it} + \gamma T + \epsilon_{it}$
ANCOVA	$Y_{it+1} = \delta Y_{it} + \beta D_i + \gamma T + \epsilon_{it}$
Change score	$Y_{it+1} - Y_{it} = \beta D_i + \gamma T + \epsilon_{it}$
Fixed Effects	$\ddot{Y}_{it} = \beta \ddot{D}_{it} + \gamma \ddot{T} + \ddot{\epsilon}_{it}$

UNIVERSITÄT MANNHEIM

Simulation: 5 different models

Model	Equation
Pooled OLS	$Y_{it} = \beta D_{it} + \gamma T + \epsilon_{it}$
ANCOVA	$Y_{it+1} = \delta Y_{it} + \beta D_i + \gamma T + \epsilon_{it}$
Change score	$Y_{it+1} - Y_{it} = \beta D_i + \gamma T + \epsilon_{it}$
Fixed Effects	$\ddot{Y}_{it} = \beta \ddot{D}_{it} + \gamma \ddot{T} + \ddot{\epsilon}_{it}$
Fixed Effects IS	$\widetilde{Y_{it}} = \beta \widetilde{D_{it}} + \widetilde{\epsilon_{it}}$

Simulation results

- Scenario 1: Absent treatment, potential outcomes follow same path

UNIVERSITÄT MANNHEIM

Simulation results

- Scenario 1: Absent treatment, potential outcomes follow same path
  - $\beta = 0$, $\alpha_{1i} = \mathrm{N}(.25D_i, .1)$, $\alpha_{2i} = 0$, $\epsilon_{it} = \mathrm{N}(0, .1)$

Simulation results

- Scenario 1: Absent treatment, potential outcomes follow same path
  - $\beta = 0$, $\alpha_{1i} = \mathrm{N}(.25D_i, .1)$, $\alpha_{2i} = 0$, $\epsilon_{it} = \mathrm{N}(0, .1)$

Model	$\hat{\beta}$	$s.e.$	**%** $p > .05$

Simulation results

- Scenario 1: Absent treatment, potential outcomes follow same path
  - $\beta = 0$, $\alpha_{1i} = \mathrm{N}(.25D_i, .1)$, $\alpha_{2i} = 0$, $\epsilon_{it} = \mathrm{N}(0, .1)$

Model	$\hat{\beta}$	$s.e.$	**%** $p > .05$
Pooled OLS	.214	.010	100

Simulation results

- Scenario 1: Absent treatment, potential outcomes follow same path
  - $\beta = 0$, $\alpha_{1i} = \mathrm{N}(.25D_i, .1)$, $\alpha_{2i} = 0$, $\epsilon_{it} = \mathrm{N}(0, .1)$

Model	$\hat{\beta}$	$s.e.$	**%** $p > .05$
Pooled OLS	.214	.010	100
ANCOVA	.125	.007	100

UNIVERSITÄT MANNHEIM

Simulation results

- Scenario 1: Absent treatment, potential outcomes follow same path
  - $\beta = 0$, $\alpha_{1i} = \mathrm{N}(.25D_i, .1)$, $\alpha_{2i} = 0$, $\epsilon_{it} = \mathrm{N}(0, .1)$

Model	$\hat{\beta}$	$s.e.$	**%** $p > .05$
Pooled OLS	.214	.010	100
ANCOVA	.125	.007	100
Change score	$<.001$	.006	1.1

UNIVERSITÄT MANNHEIM

Simulation results

- Scenario 1: Absent treatment, potential outcomes follow same path
  - $\beta = 0$, $\alpha_{1i} = \mathrm{N}(.25D_i, .1)$, $\alpha_{2i} = 0$, $\epsilon_{it} = \mathrm{N}(0, .1)$

Model	$\hat{\beta}$	$s.e.$	**%** $p > .05$
Pooled OLS	.214	.010	100
ANCOVA	.125	.007	100
Change score	<.001	.006	1.1
Fixed Effects	<.001	.007	4.7

UNIVERSITÄT MANNHEIM

Simulation results

- Scenario 1: Absent treatment, potential outcomes follow same path
  - $\beta = 0$, $\alpha_{1i} = \mathrm{N}(.25D_i, .1)$, $\alpha_{2i} = 0$, $\epsilon_{it} = \mathrm{N}(0, .1)$

Model	$\hat{\beta}$	*s.e.*	**%** $p > .05$
Pooled OLS	.214	.010	100
ANCOVA	.125	.007	100
Change score	<.001	.006	1.1
Fixed Effects	<.001	.007	4.7
Fixed Effects IS	<.001	.011	5.1

UNIVERSITÄT MANNHEIM

Simulation results

- Scenario 2: Absent treatment, potential outcomes converge

Simulation results

- Scenario 2: Absent treatment, potential outcomes converge
  - $\beta = 0$, $\alpha_{1i} = \mathrm{N}(.25D_i, .1)$, $\alpha_{2i} = \mathrm{N}(-.05D_i, .1)$, $\epsilon_{it} = \mathrm{N}(0, .1)$

UNIVERSITÄT MANNHEIM

Simulation results

- Scenario 2: Absent treatment, potential outcomes converge
  - $\beta = 0$, $\alpha_{1i} = \mathrm{N}(.25D_i, .1)$, $\alpha_{2i} = \mathrm{N}(-.05D_i, .1)$, $\epsilon_{it} = \mathrm{N}(0, .1)$

Model	$\hat{\beta}$	$s.e.$	**%** $p > .05$

Simulation results

- Scenario 2: Absent treatment, potential outcomes converge
  - $\beta = 0$, $\alpha_{1i} = \mathrm{N}(.25D_i, .1)$, $\alpha_{2i} = \mathrm{N}(-.05D_i, .1)$, $\epsilon_{it} = \mathrm{N}(0, .1)$

Model	$\hat{\beta}$	*s.e.*	**%** $p > .05$
Pooled OLS	.128	.012	100

Simulation results

- Scenario 2: Absent treatment, potential outcomes converge
  - $\beta = 0$, $\alpha_{1i} = \mathrm{N}(.25D_i, .1)$, $\alpha_{2i} = \mathrm{N}(-.05D_i, .1)$, $\epsilon_{it} = \mathrm{N}(0, .1)$

Model	$\hat{\beta}$	$s.e.$	**%** $p > .05$
Pooled OLS	.128	.012	100
ANCOVA	-.005	.009	10.8

Simulation results

- Scenario 2: Absent treatment, potential outcomes converge
  - $\beta = 0$, $\alpha_{1i} = \mathrm{N}(.25D_i, .1)$, $\alpha_{2i} = \mathrm{N}(-.05D_i, .1)$, $\epsilon_{it} = \mathrm{N}(0, .1)$

Model	$\hat{\beta}$	*s.e.*	**%** $p > .05$
Pooled OLS	.128	.012	100
ANCOVA	-.005	.009	10.8
Change score	-.050	.008	100

Simulation results

- Scenario 2: Absent treatment, potential outcomes converge
  - $\beta = 0$, $\alpha_{1i} = \mathrm{N}(.25D_i, .1)$, $\alpha_{2i} = \mathrm{N}(-.05D_i, .1)$, $\epsilon_{it} = \mathrm{N}(0, .1)$

Model	$\hat{\beta}$	*s.e.*	**%** $p > .05$
Pooled OLS	.128	.012	100
ANCOVA	-.005	.009	10.8
Change score	-.050	.008	100
Fixed Effects	-.060	.010	100

Simulation results

- Scenario 2: Absent treatment, potential outcomes converge
  - $\beta = 0$, $\alpha_{1i} = \mathrm{N}(.25D_i, .1)$, $\alpha_{2i} = \mathrm{N}(-.05D_i, .1)$, $\epsilon_{it} = \mathrm{N}(0, .1)$

Model	$\hat{\beta}$	*s.e.*	**%** $p > .05$
Pooled OLS	.128	.012	100
ANCOVA	-.005	.009	10.8
Change score	-.050	.008	100
Fixed Effects	-.060	.010	100
Fixed Effects IS	$<$.001	.011	4.6

UNIVERSITÄT MANNHEIM

Simulation results

- Scenario 2: Absent treatment, potential outcomes converge
  - $\beta = 0$, $\alpha_{1i} = \mathrm{N}(.25D_i, .1)$, $\alpha_{2i} = \mathrm{N}(-.05D_i, .1)$, $\epsilon_{it} = \mathrm{N}(0, .1)$

Model	$\hat{\beta}$	*s.e.*	**%** $p > .05$
Pooled OLS	.128	.012	100
ANCOVA	-.005	.009	10.8
Change score	-.050	.008	100
Fixed Effects	-.060	.010	100
Fixed Effects IS	$<$.001	.011	4.6

Simulation results

- Scenario 3: Absent treatment, potential outcomes diverge

Simulation results

- Scenario 3: Absent treatment, potential outcomes diverge
  - $\beta = 0$, $\alpha_{1i} = \mathrm{N}(.25D_i, .1)$, $\alpha_{2i} = \mathrm{N}(.05D_i, .1)$, $\epsilon_{it} = \mathrm{N}(0, .1)$

Simulation results

- Scenario 3: Absent treatment, potential outcomes diverge
  - $\beta = 0$, $\alpha_{1i} = \mathrm{N}(.25D_i, .1)$, $\alpha_{2i} = \mathrm{N}(.05D_i, .1)$, $\epsilon_{it} = \mathrm{N}(0, .1)$

Model	$\hat{\beta}$	*s.e.*	**%** $p > .05$

UNIVERSITÄT MANNHEIM

Simulation results

- Scenario 3: Absent treatment, potential outcomes diverge
  - $\beta = 0$, $\alpha_{1i} = \mathrm{N}(.25D_i, .1)$, $\alpha_{2i} = \mathrm{N}(.05D_i, .1)$, $\epsilon_{it} = \mathrm{N}(0, .1)$

Model	$\hat{\beta}$	$s.e.$	**%** $p > .05$
Pooled OLS	.300	.013	100

UNIVERSITÄT MANNHEIM

Simulation results

- Scenario 3: Absent treatment, potential outcomes diverge
  - $\beta = 0$, $\alpha_{1i} = \mathrm{N}(.25D_i, .1)$, $\alpha_{2i} = \mathrm{N}(.05D_i, .1)$, $\epsilon_{it} = \mathrm{N}(0, .1)$

Model	$\hat{\beta}$	$s.e.$	**%** $p > .05$
Pooled OLS	.300	.013	100
ANCOVA	.104	.010	100

UNIVERSITÄT MANNHEIM

Simulation results

- Scenario 3: Absent treatment, potential outcomes diverge
  - $\beta = 0$, $\alpha_{1i} = \mathrm{N}(.25D_i, .1)$, $\alpha_{2i} = \mathrm{N}(.05D_i, .1)$, $\epsilon_{it} = \mathrm{N}(0, .1)$

Model	$\hat{\beta}$	*s.e.*	**%** $p > .05$
Pooled OLS	.300	.013	100
ANCOVA	.104	.010	100
Change score	.050	.008	100

UNIVERSITÄT MANNHEIM

Simulation results

- Scenario 3: Absent treatment, potential outcomes diverge
  - $\beta = 0$, $\alpha_{1i} = \mathrm{N}(.25D_i, .1)$, $\alpha_{2i} = \mathrm{N}(.05D_i, .1)$, $\epsilon_{it} = \mathrm{N}(0, .1)$

Model	$\hat{\beta}$	*s.e.*	**%** $p > .05$
Pooled OLS	.300	.013	100
ANCOVA	.104	.010	100
Change score	.050	.008	100
Fixed Effects	.060	.010	100

UNIVERSITÄT MANNHEIM

Simulation results

- Scenario 3: Absent treatment, potential outcomes diverge
  - $\beta = 0$, $\alpha_{1i} = \mathrm{N}(.25D_i, .1)$, $\alpha_{2i} = \mathrm{N}(.05D_i, .1)$, $\epsilon_{it} = \mathrm{N}(0, .1)$

Model	$\hat{\beta}$	*s.e.*	**%** $p > .05$
Pooled OLS	.300	.013	100
ANCOVA	.104	.010	100
Change score	.050	.008	100
Fixed Effects	.060	.010	100
Fixed Effects IS	$<.001$	.011	4.6

Simulation results

- Scenario 3: Absent treatment, potential outcomes diverge
  - $\beta = 0$, $\alpha_{1i} = \mathrm{N}(.25D_i, .1)$, $\alpha_{2i} = \mathrm{N}(.05D_i, .1)$, $\epsilon_{it} = \mathrm{N}(0, .1)$

Model	$\hat{\beta}$	*s.e.*	**%** $p > .05$
Pooled OLS	.300	.013	100
ANCOVA	.104	.010	100
Change score	.050	.008	100
Fixed Effects	.060	.010	100
Fixed Effects IS	$<.001$	.011	4.6

UNIVERSITÄT MANNHEIM

Simulation Summary

- FE model: fine when $\alpha_{1i} = 0$, but fails when potential outcomes w/o treatment do not follow same path ($\alpha_{2i} = 0$ required)

Simulation Summary

- FE model: fine when $\alpha_{1i} = 0$, but fails when potential outcomes w/o treatment do not follow same path ($\alpha_{2i} = 0$ required)
- Changes Score model: alternative to FE, but also requires $\alpha_{2i} = 0$

UNIVERSITÄT MANNHEIM

Simulation Summary

- FE model: fine when $\alpha_{1i} = 0$, but fails when potential outcomes w/o treatment do not follow same path ($\alpha_{2i} = 0$ required)
- Changes Score model: alternative to FE, but also requires $\alpha_{2i} = 0$
- ANCOVA model: appropriate when potential outcomes converge, but fails when they diverge or when $\alpha_{2i} = 0$

UNIVERSITÄT MANNHEIM

Simulation Summary

- FE model: fine when $\alpha_{1i} = 0$, but fails when potential outcomes w/o treatment do not follow same path ($\alpha_{2i} = 0$ required)
- Changes Score model: alternative to FE, but also requires $\alpha_{2i} = 0$
- ANCOVA model: appropriate when potential outcomes converge, but fails when they diverge or when $\alpha_{2i} = 0$
- FE-IS is the only model which performs nicely regardles of α_{1i} and α_{2i}

UNIVERSITÄT MANNHEIM

The story

- Married men earn higher (hourly) wages than never-married, cohabiting, divorced

The story

- Married men earn higher (hourly) wages than never-married, cohabiting, divorced
- "'... one of the most well documented phenomena in social science"' (Waite & Gallagher 2000)

The story

- Married men earn higher (hourly) wages than never-married, cohabiting, divorced
- "'... one of the most well documented phenomena in social science"' (Waite & Gallagher 2000)
- Studies show (using conventional FE):

UNIVERSITÄT MANNHEIM

The story

- Married men earn higher (hourly) wages than never-married, cohabiting, divorced
- "'... one of the most well documented phenomena in social science"' (Waite & Gallagher 2000)
- Studies show (using conventional FE):
  - selection of high earners into marriage explains at least half of the premium

UNIVERSITÄT MANNHEIM

The story

- Married men earn higher (hourly) wages than never-married, cohabiting, divorced
- "'... one of the most well documented phenomena in social science"' (Waite & Gallagher 2000)
- Studies show (using conventional FE):
  - selection of high earners into marriage explains at least half of the premium
  - a significant positive effect still remains (2-10% after controlling for α_{1i})

The story

- Married men earn higher (hourly) wages than never-married, cohabiting, divorced
- "'... one of the most well documented phenomena in social science"' (Waite & Gallagher 2000)
- Studies show (using conventional FE):
  - selection of high earners into marriage explains at least half of the premium
  - a significant positive effect still remains (2-10% after controlling for α_{1i})
  - not clear whether the effect is causal

UNIVERSITÄT MANNHEIM

The story

- Married men earn higher (hourly) wages than never-married, cohabiting, divorced
- "'... one of the most well documented phenomena in social science"' (Waite & Gallagher 2000)
- Studies show (using conventional FE):
  - selection of high earners into marriage explains at least half of the premium
  - a significant positive effect still remains (2-10% after controlling for α_{1i})
  - not clear whether the effect is causal
  - possible bias because of α_{2i}

UNIVERSITÄT MANNHEIM

Real data

- Two country studies: West Germany and US

Real data

- Two country studies: West Germany and US
- GSOEP 1984-2006 and NLSY 1979-2004 data

Real data

- Two country studies: West Germany and US
- GSOEP 1984-2006 and NLSY 1979-2004 data
- Use working males, never-married when first observed, observe (log) hourly wages before and after marriage

Real data

- Two country studies: West Germany and US
- GSOEP 1984-2006 and NLSY 1979-2004 data
- Use working males, never-married when first observed, observe (log) hourly wages before and after marriage
- Z includes slope variables work experience, and experience squared, and individual constants

Real data

- Two country studies: West Germany and US
- GSOEP 1984-2006 and NLSY 1979-2004 data
- Use working males, never-married when first observed, observe (log) hourly wages before and after marriage
- Z includes slope variables work experience, and experience squared, and individual constants
- Therefore, at least 4 years per person

UNIVERSITÄT MANNHEIM

Real data

- Two country studies: West Germany and US
- GSOEP 1984-2006 and NLSY 1979-2004 data
- Use working males, never-married when first observed, observe (log) hourly wages before and after marriage
- Z includes slope variables work experience, and experience squared, and individual constants
- Therefore, at least 4 years per person
- controls: divorce, remarriage, number of children, yrs. education, tenure, year dummies

Results: Effect of marriage on male wages

GSOEP			**NLSY**		
Model	$\hat{\beta}$	robust s.e.	Model	$\hat{\beta}$	robust s.e.
POLS	.078**	(.014)	POLS	.146**	(.010)
FE	.036**	(.013)	FE	.082**	(.008)
FE-IS	.015	(.010)	FE-IS	.021*	(.008)

UNIVERSITÄT MANNHEIM

Conclusions

- FE-IS valid under more general conditions than alternative models
  - besides linearity, no further assumptions needed regarding heterogeneous slopes

UNIVERSITÄT MANNHEIM

Conclusions

- FE-IS valid under more general conditions than alternative models
  - besides linearity, no further assumptions needed regarding heterogeneous slopes
  - outcomes may develop differently over time in treatment and control group (may diverge or converge)

UNIVERSITÄT MANNHEIM

Conclusions

- FE-IS valid under more general conditions than alternative models
  - besides linearity, no further assumptions needed regarding heterogeneous slopes
  - outcomes may develop differently over time in treatment and control group (may diverge or converge)
  - supported by simulation results

Conclusions

- FE-IS valid under more general conditions than alternative models
  - besides linearity, no further assumptions needed regarding heterogeneous slopes
  - outcomes may develop differently over time in treatment and control group (may diverge or converge)
  - supported by simulation results
- possible extension: implement a test for heterogeneous slopes

Conclusions

- FE-IS valid under more general conditions than alternative models
  - besides linearity, no further assumptions needed regarding heterogeneous slopes
  - outcomes may develop differently over time in treatment and control group (may diverge or converge)
  - supported by simulation results
- possible extension: implement a test for heterogeneous slopes

Conclusions

- FE-IS valid under more general conditions than alternative models
  - besides linearity, no further assumptions needed regarding heterogeneous slopes
  - outcomes may develop differently over time in treatment and control group (may diverge or converge)
  - supported by simulation results
- possible extension: implement a test for heterogeneous slopes