`cfbinout` & `xtdhazard`: Control-Function Estimation of Binary-Outcome Models and the Discrete-Time Hazard Model

Harald Tauchmann & Elena Yurkevich

Friedrich-Alexander-Universität Erlangen-Nürnberg

June 7, 2024

Outline

Motivation

- **Discrete-time hazard model** obvious approach to modeling transitions if covariates $\mathbf{x}_{it}$ **change** values **frequently**
- E.g. Panel data (for instance annual surveys)
  - New information about $\mathbf{x}_{it}$ in each panel wave $t = 1, \ldots, T$
  - Information about whether transition has occurred ($\tau_i = t$) or has not yet occurred ($\tau_i > t$) at same frequency
- Models prob. of unit i of transitioning in period t conditionally on not yet having transitioned (hazard λ_{it})

$$\lambda_{it} = \mathrm{P}\left(\tau_i = t | a_i, \mathbf{x}_{it}, \tau_i \geq t\right) = F\left(a_i + \mathbf{x}_{it}\boldsymbol{\beta}\right)$$

- **Focus**:
  - Transition into **absorbing state** (single-spell model)
  - Model with **unobserved heterogeneity**/frailty a_i

Estimation of Discrete-Time Hazard Model

- In terms of estimation, **coincide with binary outcome models** such as logit, probit, and cloglog (Jenkins, 1995; Tutz and Schmid, 2016)
  - Re-formulating model in terms of binary outcome y_{it}
  - $y_{it} = 1$ if $\tau_i = t$ and $y_{it} = 0$ if $\tau_i > t$
  - y_{it} with $t > \tau_i$ not considered (not informative about transition to absorbing state)
- Allows for modelling (unit-level**unobserved heterogeneity**/frailty, a_i) as **random effects** like in panel binary outcome models (`xtlogit`, `xtprobit`, `xtclolog`)

Modelling Unobserved Heterogeneity

- Random effects do not accommodate correlation of unobserved heterogeneity and covariates
- **Linear probability model** (LPM), which allows for **fixed effects** accommodating such correlation, as possible alternative
- LPM with unit-level **fixed effects** heavily **biased** in single-spell hazard model setting (Farbmacher and Tauchmann, 2023)
  - Not because of – possibly inappropriate – linear specification
  - Not because of failure to eliminate unobserved heterogeneity (biased even in its absence)
- Cond. expectation of error term $\varepsilon_{it}^{\mathrm{FE}}$ in fixed-effects model

$$\mathrm{E}\left(\varepsilon_{it}^{\mathrm{FE}} | a_i, \mathbf{x}_{i1}, \ldots, \mathbf{x}_{iT}, \tau_i \geq t\right) = a_i + \mathrm{E}(\bar{\mathbf{x}}_i)_t \boldsymbol{\beta}$$

Own-Differences IV Estimation

- Linear (internal) **instrumental variables** (IV) estimator to deal with unobserved heterogeneity (suggested by Farbmacher and Tauchmann, 2023)
  - All explanatory variables $\mathbf{x}_{it}$ instrumented by their own (**first**) **differences** $\Delta\mathbf{x}_{it}$
  - Unlike fixed-effects in standard setting, does not accommodate arbitrary correlation between a_i and $\mathbf{x}_{it}$
  - Rather assumption $\text{Cov}(\Delta\mathbf{x}_{it}, a_i) = \mathbf{0}$ required (while allowing for $\text{Cov}(\mathbf{x}_{it}, a_i) \neq \mathbf{0}$)
  - Still subject to – some sort of – survivor bias (rather small in many settings)
  - Does not suffer from bias caused by including unit-level fixed effects

Non-Linear Control Function Estimation

- Linear hazard specification (LPM) not too appealing
  - At least if one thinks of LPM in terms of modelling a data generating process
- Rationale for using (first) own-differences as an instrument may also apply to nonlinear models
  - Discussed in passing in Farbmacher and Tauchmann (2023)
  - **Control function** approach (cf. Wooldridge, 2015)
  - I.e. including first-stage residuals as additional second-stage regressors

Existing Stata Commands

- Instrumental variables and control function estimators (naturally) **implemented** in **Stata**
  1. `ivregress 2sls`
  2. `ivprobit, twostep`
  3. `ivreg210` (contributed by C.F. Baum, M.E. Schaffer, Steven Stillman)
  4. `ivcloglog` (contributed by W. Liu)

  ... probably more
- Generating numerous internal instruments **cumbersome**
  - In particular if factor variables syntax used
- Existing commands **not specific** to single spell hazard model setting
  - User needs to check actively if data warrants using first-differences IV estimator
- No command for control function logit

Two New Stata Commands

1. **`xtdhazard`**
   - Checks data for being consistent with transition to absorbing state
   - Temporarily generates internal instruments
   - Calls either `ivregress 2sls` or `cfbinout`
2. **`cfbinout`**
   - Runs logit, probit, or cloglog control function estimator (following Wooldridge, 2015)
     - Stata implementation follows Terza (2017)
   - If called by `xtdhazard` uses internal instruments (first- or higher-oder differences)
   - Can also be used as stand-alone command using user-specified instruments

The `xtdhazard` Command

- Just a wrapper for more convenient IV estimation of discrete-time hazard model
- Straight-forward with estimator (two-part command name) `2sls` (calls `ivregress 2sls`)
  - In terms of theory, in detail discussed in Farbmacher and Tauchmann (2023)
  - Just `ivregress 2sls` in terms of Stata
- Less straight-forward with estimators `logit`, `probit`, or `cloglog` (calls `cfbinout`)
  - Only some simulation-based discussion in online appendix to Farbmacher and Tauchmann (2023)
  - Some issues that do not apply to linear model (binary rhs-variables, quasi-complete separation)

The `cfbinout` Command

- Implements control function estimators, i.e. first stage residuals included as additional regressors
- Follows and draws on Terza (2017) "Two-stage residual inclusion estimation: A practitioners guide to Stata implementation [st0505]"
  - Considering ML estimation in second stage
- Allows for logit, probit, or cloglog link (two part command name `cfbinout link`)
  - For probit link (largely) equivalent to `ivprobit, twostep`
- Estimates a first stage for all regressors if called by `xtdhazard`

Specific issues with `cfbinout`

- Discrete endogeneous variables
  - Not accommodated by standard control function approach (e.g. not allowed with `ivprobit, twostep`)
  - According to Wooldridge (2015) ‘average structural form’ can – under certain assumptions – still be estimated by including **generalized residuals** from **binary outcome first stage**
  - `cfbinout` allows specifying the link-function for first stage (probit, logit, linear) and uses generalized residuals for the former
- Quasi complete separation
  - First stage may be subject to quasi complete separation (prone to if `cfbinout` is called by `xtdhazard`)
  - `cfbinout` optionally allowed to switch to linear first stage

Syntax of `xtdhazard`

`xtdhazard` *`estimator depvar indepvars`* `[`*`if`*`] [`*`in`*`] [`*`weight`*`] [,` *`options`*`]`

- Requires data to be `xtset` (declared panel data)
  - » *`panelvar`* and *`timevar`* required
- Estimators
  - (i) `2sls` ($\rightarrow$ linear 2SLS, calls `ivregress 2sls`)
  - (ii) `logit` ($\rightarrow$ control function logit, calls `cfbinout`)
  - (iii) `probit` ($\rightarrow$ control function probit, calls `cfbinout`)
  - (iv) `cloglog` ($\rightarrow$ control function complementary log-log, calls `cfbinout`)

Selected Options for `xtdhazard`

- General Options
  - `difference(numlist)`: Sets order of differencing
    - `difference(1)`, i.e. (only) using first-differences as instruments, is the default
    - Yet, for instance, also `difference(1 3)` is possible (yet makes probably little sense)
  - `instruments(varlist)`: Specifies additional, non-internal instruments
  - `noabsorbing`: Forces estimation if `depvar` does not indicate absorbing state
- Options for estimator 2sls
  - `interactinst`: Use squares and interactions of instruments as additional instruments
  - `nofirststage`: Do not save first-stage coefficients in `e(G)` and do not perform checks regarding first-stage
  - `underid(string)`: Calls `underid` from within `xtdhazard`

Selected Options for `xtdhazard` (cont.)

- Options for estimators logit, probit, and cloglog
  - `order(#)`: Specify order of control-function polynomial
    - `order(1)`, i.e. just including (generalized) residuals, is the default
    - `order(2)`, `order(3)`, ... means also including higher powers ($\to$ `ivcloglog`)
  - `noresgenerate`: First-stage residuals are only temporarily generated; coefficients of first-stage residuals are not reported; first-stage coefficients are not saved
  - `resname(stub)`: First-stage residual permanently saved as *stub_varname2*; the default for *stub* is *res*
  - `replace`: Variable *stub_varname2* replaced if already it exists

Syntax of `cfbinout`

`cfbinout` *link depvar* [*varlist1*] (*varlist2* = *varlist_iv*) [*if*] [*in*] [*weight*], [*options*]

- Link functions
  - (i) `probit`: normal CDF, $\Phi(\boldsymbol{\beta x})$
  - (ii) `logit`: logistic CDF, $\frac{1}{1+\exp(-\boldsymbol{\beta x})}$
  - (iii) `cloglog`: Gumbel CDF, $1-\exp(-\exp(\boldsymbol{\beta x}))$

Selected Options for `cfbinout`

- Largely the same as for *`xtdhazard logit`* (and *`xtdhazard probit/cloglog`*)
- Options not available with *`xtdhazard logit`* (*`xtdhazard probit/cloglog`*)
  - `fslink(`*name*`)`: Specifies first-stage link function (`logit` [default] `probit`, `linear`)
  - `fsswitch`: Switch equation-wise to `fslink(linear)` in case of quasi-complete separation in first-stage

Data Generating Process for Outcome y_{it}

$$\begin{aligned}\lambda_{it} &= \mathrm{P}\left(y_{it} = 1 \middle| a_i, \mathbf{x}_{it}, \tau_i \geq t\right) \\ &= \mathrm{F}\left(a_i + \beta_{con} x_{con,it} + \beta_{bin} x_{bin,it} + \beta_{tiv} x_{tiv,i}\right)\end{aligned}$$

- $y_{it} =$ "*missing*" if $\tau_i < t$ (i.e. $y_{it-1} \neq 0$)
- a_i: unobserved heterogeneity (beta distr.)
- x_{con}: **continuous** (normal distr.)
- x_{bin}: **binary** (Bernoulli distr.)
- x_{tiv}: **time-invariant** and contin. (wt. sum of beta distr.)
- $\mathrm{F}(\cdot) = \Phi(\cdot)$ (alternatively logistic CDF, Gumbel CDF)
- β_{con}, β_{bin}, β_{tiv} take value of 1 (rescaled for logit and cloglog) or 0 ($\rightarrow$ various **exclusion restrictions**)

Correlation of Unobserved Heterogeneity and **x** Vars.

```
. correlate a x_tiv x_bin x_con d.x_bin d.x_con
(obs=16,000)

             |                                            D.       D.
             |        a    x_tiv    x_bin    x_con    x_bin    x_con
-------------+------------------------------------------------------
           a |   1.0000
       x_tiv |  -0.7517   1.0000
       x_bin |  -0.5843   0.4349   1.0000
       x_con |  -0.8131   0.6103   0.4727   1.0000
       x_bin |
         D1. |   0.0034  -0.0000   0.5698  -0.0053   1.0000
       x_con |
         D1. |  -0.0028   0.0018  -0.0054   0.4134  -0.0103   1.0000
```

- **Unobserved heterogeneity** *a* (negatively) **correlated** with all **x** variables, yet **uncorrelated** with Δ**x**
- **x** variables **stationary** (time series properties matter Farbmacher and Tauchmann, 2023)

Coefficient Values and Sample

- $N = 4000$ (# of units), $T = 5$ (# of periods)
- Patters of included $\mathbf{x}$ vars.
  (i) Only continuous x_{con}
  (ii) Continuous x_{con} and time-invariant x_{tiv}
  (iii) Continuous x_{con} and binary x_{bin}
  (iv) All three x_{con}, x_{tiv}, x_{bin}
- Exclusion restrictions taken into account in estimation
- Average hazard $\bar{\lambda} = 0.25$ (alternatively $\bar{\lambda} = 0.05$)

distribution of λ

Estimators and Replications

- Three estimating procedures
  - (i) 'Naive' `probit` (alternatively `logit` and `cloglog`)
  - (ii) Linear 2SLS ($\rightarrow$ `xtdhazard 2sls`)
  - (iii) Control function probit ($\rightarrow$ `xtdhazard probit`, alternatively `xtdhazard logit` and `xtdhazard cloglog`)
- Focus on (sample) average partial effects of x_{con} and x_{bin}
  - » β_{tiv} not identified in 2SLS and control function estimation
- Monte Carlo simulations using 2000 replications

Simulation Results: Probit, $\overline{\lambda} = 0.25$

Table: Average Partial Effects (probit, $\overline{\lambda} = 0.25$)

Av. partial effect of	Estimator	Scenarios: rhs vars. inclusion			
		x_{con}	x_{con}, x_{tiv}	x_{con}, x_{bin}	$x_{con}, x_{bin}, x_{tiv}$
x_{con}	Naive probit	0.106	0.153	0.113	0.126
	CF probit	0.306	0.282	0.249	0.214
	2SLS	0.306	0.282	0.249	0.216
	True value	0.305	0.282	0.260	0.219
x_{bin}	Naive probit			0.219	0.205
	CF probit			0.282	0.248
	2SLS			0.280	0.250
	True value			0.284	0.248

Note: The table presents simulation results comparing estimated and true mean-marginal effects in a discrete time hazard framework. The estimations were based on 2000 simulation runs. "True value" refers to the actual mean-marginal effects used in the simulation. The constants chosen for different scenarios were selected to achieve a mean probability of 0.25.

Simulation Results: Alternative Simulations

Pattern of results generally the same:

(i) `logit` estimation, $\overline{\lambda} = 0.25$ ▸ logit, $\overline{\lambda} = 0.25$

(ii) `cloglog` estimation, $\overline{\lambda} = 0.25$ ▸ cloglog, $\overline{\lambda} = 0.25$

(iii) `probit` estimation, $\overline{\lambda} = 0.05$ ▸ probit, $\overline{\lambda} = 0.05$

Real Data Application: Replication of Cantoni (2012)

Cantoni (2012, EJ): "Adopting a New Religion: The Case of Protestantism in 16th Century Germany"

- Research Question: Which factors explain territories adopting protestantism?
- Uses historical panel data (74 territories in 16th century Germany; 5 years from 1532 to 1600)
- Key results: (i) distance to Wittenberg and (ii) neighbours' religious choices matter for adoption of protestantism

- Focus on fixed-effects specification Cantoni (2012, p. 522, Table 6, Column 3)
  - Subject to minor change regarding clustering

Original Result by Cantoni (2012)

```
. xtreg refuntil lagrefneighbors, cluster(kreis) nonest fe

Fixed-effects (within) regression               Number of obs     =        370
Group variable: tcode                           Number of groups  =         74

R-squared:                                      Obs per group:
     Within  = 0.2348                                         min =          5
     Between = 0.0751                                         avg =        5.0
     Overall = 0.0974                                         max =          5

                                                F(1, 9)           =      37.76
corr(u_i, Xb) = 0.0471                          Prob > F          =     0.0002

                                  (Std. err. adjusted for 10 clusters in kreis)
--------------------------------------------------------------------------------
              |               Robust
     refuntil |      Coef.    Std. Err.      t    P>|t|     [95% Conf. Interval]
--------------+-----------------------------------------------------------------
lagrefneig~s  |   .6792987    .1105446     6.15   0.000     .4292295    .9293678
        _cons |   .1832339    .033076      5.54   0.000     .1084108    .2580569
--------------------------------------------------------------------------------
      sigma_u |   .4024456
      sigma_e |  .26168501
          rho |  .70283546   (fraction of variance due to u_i)
--------------------------------------------------------------------------------
```

- Neighbouring territories' confession (`lagrefneighbors`) matters

How is Absorbing State Dealt With?

```
 . gen l_refuntil = l.refuntil
(74 missing values generated)

. tab refuntil l_refuntil if e(sample)

           |      l_refuntil
  refuntil |         0          1 |     Total
-----------+----------------------+----------
         0 |       166          0 |       166
         1 |        28        102 |       130
-----------+----------------------+----------
     Total |       194        102 |       296
```

- “Selection into Protestantism was effectively an absorbing state” (Cantoni, 2012, p. 523)
- Not taken into account in Fixed Effects Specification

Territories at Risk Only

```
 . xtreg refuntil lagrefneighbors if l_refuntil != 1, cluster(kreis) fe

Fixed-effects (within) regression               Number of obs     =        268
Group variable: tcode                           Number of groups  =         74

R-squared:                                      Obs per group:
     Within  = 0.1442                                         min =          1
     Between = 0.8669                                         avg =        3.6
     Overall = 0.0004                                         max =          5

                                                F(1, 9)           =      10.27
corr(u_i, Xb) = -0.4007                         Prob > F          =     0.0107

                                 (Std. err. adjusted for 10 clusters in kreis)
------------------------------------------------------------------------------
              |               Robust
    refuntil  |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
--------------+---------------------------------------------------------------
lagrefneig_s  |   .5722313    .1785385     3.21   0.011     .1683491    .9761135
       _cons  |   .0104711    .044465      0.24   0.819    -.0901157    .1110578
--------------+---------------------------------------------------------------
     sigma_u  |  .43093158
     sigma_e  |  .28265007
         rho  |   .6991975   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

- Only considering territories at risk makes little difference

Linear First-Differences IV Estimation

```
 . xtdhazard 2sls refuntil lagrefneighbors if l_refuntil != 1, cluster(kreis)
> underid(underid)


Linear discrete-time hazard model            Number of obs        =       194
first-differences IV estimation              Number of groups     =        61
                                             Wald chi2(1)         =      0.25
                                             Prob > chi2          =     0.619
                                             R-sq                 =         .

                                 (Std. err. adjusted for 10 clusters in kreis)
------------------------------------------------------------------------------
             |              Clustered
   refuntil  |     Coef.    Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lagrefneig_s |  -.7880119    1.582831    -0.50   0.619    -3.890303     2.31428
       _cons |   .4154438    .5556795     0.75   0.455    -.6736681    1.504556
------------------------------------------------------------------------------
Underidentification test: j =      6.35; Chi-sq(   1); p-value = 0.0117
```

- `xtdhazard 2sls` used for estimation
- Does not confirm that neighbours' religious choices matter
- Underidentification (`d.lagrefneig_s` weak instr.) rejected

Control Function Logit Estimation

```
. xtdhazard logit refuntil lagrefneighbors if l_refuntil != 1, cluster(kreis)

Logit discrete-time hazard model              Number of obs        =       194
first-differences CF estimation               Number of groups     =        61
                                              Wald chi2(1)         =      0.10
                                              Prob > chi2          =     0.749
Log pseudolikelihood =   -79.136

                                 (Std. err. adjusted for 10 clusters in kreis)
--------------------------------------------------------------------------------
              |               Clustered
   refuntil   |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
--------------+-----------------------------------------------------------------
lagrefneig_s  |  -5.736172   17.90578    -0.32   0.749    -40.83086    29.35851
res_lagref_s  |   6.672166   15.71915     0.42   0.671     -24.1368    37.48113
       _cons  |   .1694307   6.219791     0.03   0.978    -12.02113       12.36
--------------------------------------------------------------------------------
```

- `xtdhazard logit` used for estimation
- Results qualitatively the same as counterparts from 2SLS

Control Function Logit Estimation (cont.)

```
 . margins, dydx(lagrefneighbors)

Average marginal effects                                   Number of obs = 194
Model VCE: Clustered

Expression: Pr(refuntil), predict()
dy/dx wrt:  lagrefneighbors

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lagrefneig_s |  -.7007748   2.234062    -0.31   0.754    -5.079456    3.677906
------------------------------------------------------------------------------
```

- Result in terms of average marginal effects

Control Function Cloglog Estimation

```
 . xtdhazard cloglog refuntil lagrefneighbors if l_refuntil != 1, cluster(kreis)
> difference(2) replace

Cloglog discrete-time hazard model          Number of obs        =        133
2nd-differences CF estimation               Number of groups     =         50
                                            Wald chi2(1)         =       0.53
                                            Prob > chi2          =      0.467
Log pseudolikelihood =   -50.153

                                 (Std. err. adjusted for 10 clusters in kreis)
------------------------------------------------------------------------------
             |               Clustered
    refuntil |      Coef.    Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lagrefneig_s |  -4.825455    6.626972    -0.73   0.467    -17.81408    8.163171
res_lagref_s |    7.00504    9.303846     0.75   0.451    -11.23016    25.24024
       _cons |  -.0397897    2.861651    -0.01   0.989    -5.648523    5.568944
------------------------------------------------------------------------------
```

- `xtdhazard cloglog` used for estimation
- Difference-in-Differences used as instrument

Control Function Cloglog Estimation (cont.)

```
 . margins, dydx(lagrefneighbors)

Average marginal effects                                   Number of obs = 133
Model VCE: Clustered

Expression: Pr(refuntil), predict()
dy/dx wrt:  lagrefneighbors

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lagrefneig_s |  -.5720407    .8870068    -0.64   0.519    -2.310542    1.166461
------------------------------------------------------------------------------
```

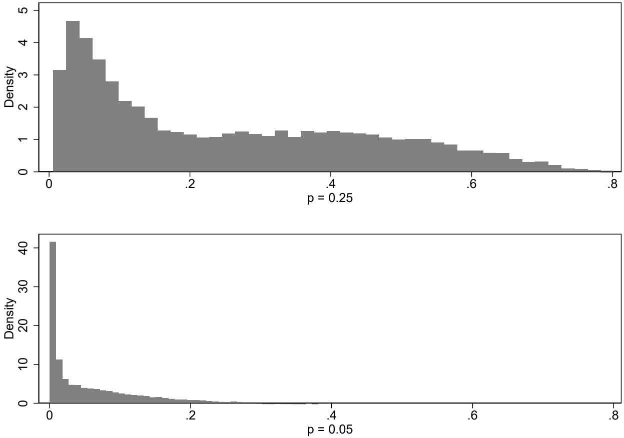
- Result in terms of average marginal effect
- Point estimate changes only marginally

Conclusions

- Using 2SLS or non-linear control function estimation probably better suited for dealing with unobserved heterogeneity in a single-spell hazard setting (than including unit fixed effects)
- These well-known estimators already implemented in Stata
- Using numerous internal instruments renders implementation through existing commands cumbersome
- `xtdhazard` eases using theses estimators for Stata users
- `cfbinout` can be used as stand-alone estimator, which complements `ivprobit` and `ivcloglog`

Distribution of Hazard

- Comparison of simulations with mean $\overline{\lambda} = 0.25$ and $\overline{\lambda} = 0.05$



Simulation Results: Logit, $\overline{\lambda} = 0.25$

Table: Average Partial Effects (logit, $\overline{\lambda} = 0.25$)

Av. partial effect of	Estimator	Scenarios: rhs vars. inclusion			
		x_{con}	$x_{\text{con}}, x_{\text{tiv}}$	$x_{\text{con}}, x_{\text{bin}}$	$x_{\text{con}}, x_{\text{bin}}, x_{\text{tiv}}$
x_{con}	Naive logit	0.035	0.055	0.044	0.058
	CF logit	0.102	0.101	0.099	0.097
	2SLS	0.102	0.101	0.099	0.097
	True value	0.103	0.102	0.101	0.098
x_{bin}	Naive logit			0.079	0.083
	CF logit			0.102	0.100
	2SLS			0.102	0.100
	True value			0.102	0.100

Note: The table presents simulation results comparing estimated and true mean-marginal effects in a discrete time hazard framework. The estimations were based on 2000 simulation runs. "True value" refers to the actual mean-marginal effects used in the simulation. The constants chosen for different scenarios were selected to achieve a mean probability of 0.25.

Simulation Results: Cloglog, $\overline{\lambda} = 0.25$

Table: Average Partial Effects (cloglog, $\overline{\lambda} = 0.25$)

Av. partial effect of	Estimator	Scenarios: rhs vars. inclusion			
		x_{con}	$x_{\text{con}}, x_{\text{tiv}}$	$x_{\text{con}}, x_{\text{bin}}$	$x_{\text{con}}, x_{\text{bin}}, x_{\text{tiv}}$
x_{con}	Naive cloglog	0.057	0.088	0.070	0.088
	CF cloglog	0.166	0.161	0.156	0.148
	2SLS	0.166	0.161	0.156	0.148
	True value	0.167	0.163	0.160	0.150
x_{bin}	Naive cloglog			0.125	0.127
	CF cloglog			0.162	0.153
	2SLS			0.162	0.153
	True value			0.162	0.153

Note: The table presents simulation results comparing estimated and true mean-marginal effects in a discrete time hazard framework. The estimations were based on 2000 simulation runs. "True value" refers to the actual mean-marginal effects used in the simulation. The constants chosen for different scenarios were selected to achieve a mean probability of 0.25.

Simulation Results: Probit, $\overline{\lambda} = 0.05$

Table: Average Partial Effects (probit, $\overline{\lambda} = 0.05$)

Av. partial effect of	Estimator	Scenarios: rhs vars. inclusion			
		x_{con}	$x_{\text{con}}, x_{\text{tiv}}$	$x_{\text{con}}, x_{\text{bin}}$	$x_{\text{con}}, x_{\text{bin}}, x_{\text{tiv}}$
x_{con}	Naive probit	0.035	0.052	0.039	0.050
	CF probit	0.102	0.095	0.087	0.083
	2SLS	0.101	0.094	0.086	0.083
	True value	0.104	0.096	0.089	0.084
x_{bin}	Naive probit			0.060	0.057
	CF probit			0.080	0.069
	2SLS			0.080	0.069
	True value			0.080	0.069

Note: The table presents simulation results comparing estimated and true mean-marginal effects in a discrete time hazard framework. The estimations were based on 2000 simulation runs. "True value" refers to the actual mean-marginal effects used in the simulation. The constants chosen for different scenarios were selected to achieve a mean probability of 0.05.

References

Cantoni, D. (2012). Adopting a New Religion: The Case of Protestantism in 16th Century Germany, *The Economic Journal* **122**(560): 502–531.

Farbmacher, H. and Tauchmann, H. (2023). Linear fixed-effects estimation with nonrepeated outcomes, *Econometric Reviews* **42**(8): 635–654.

Jenkins, S. P. (1995). Easy Estimation Methods for Discrete-Time Duration Models, *Oxford Bulletin of Economics and Statistics* **57**(1): 129–136.

Terza, J. (2017). Two-stage residual inclusion estimation: A practitioners guide to Stata implementation, *Stata Journal* **17**(4): 916–938.

Tutz, G. and Schmid, M. (2016). *Modeling Discrete Time-to-Event Data*, Springer Series in Statistics, Springer International Publishing, Cham.

Wooldridge, J. M. (2015). Control Function Methods in Applied Econometrics, *The Journal of Human Resources* **50**(2): 420–445.