2014

ntreatreg: A Stata module for estimation of treatment effects in the presence of neighborhood interactions

Giovanni Cerulli
Ceris-CNR

National Research Council
Institute for Economic Research on Firms and Growth
Unit of Rome, Italy
Email: g.cerulli@ceris.cnr.it

Science is when
we manage to pass from the
theoretically fanciful
to the empirically plausible

2014 UK Stata Users Group meeting
Cass Business School
London, UK

September 11–12, 2014

Outline of this presentation

- Statistical background and related studies
- The Rubin's *potential outcome model* with *neighborhood interactions*
- Model's estimation
- Stata implementation via `ntreatreg`
- Application to real data
- Conclusions

In standard **Econometrics of Program Evaluation** (aimed at estimating the **effect of a policy** on supported individuals) it is assumed the so-called **SUTVA** (Rubin 1978):

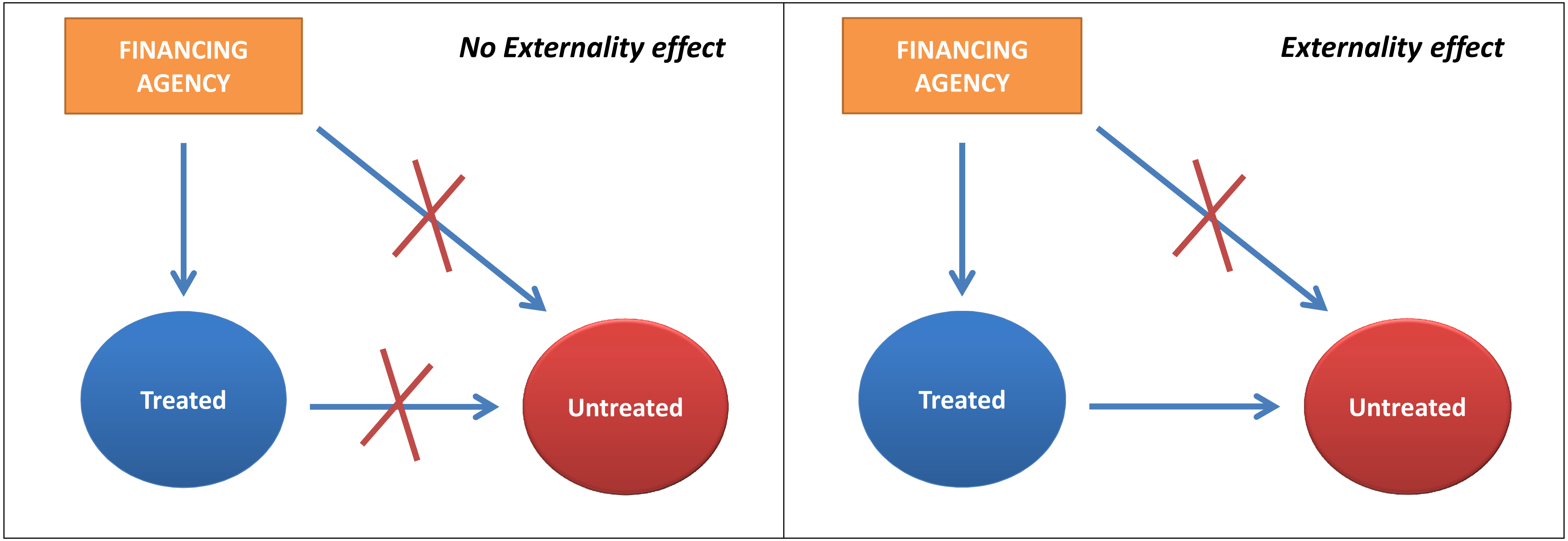
SUTVA: Stable-Unit-Treatment-Value-Assumption

"treatment received by one unit do not affect outcomes for another unit"

It means that: only the treatment applied to the specific individual is assumed to potentially affect the outcome for that particular individual.

===> We would like to *relax* this assumption and understand what happens to the estimation of the effect of a "treatment" in the presence of potential ***contagion* (or *neighborhood*, or *social*) effects**.

SUTVA and NO-SUTVA setting



- ✓ **Rubin** (1978): calls this important assumption as **Stable-Unit-Treatment-Value-Assumption (SUTVA)**
- ✓ **Manski** (2011): refers to **Individualistic-Treatment-Response (ITR)** to emphasize that this poses a restriction in the form of the treatment response function that the analyst considers.

- **AIM**: estimating the "Average Treatment Effects" (ATEs) of a policy program in a *non-experimental* setup in the presence of ***endogenous neighbourhood*** (or ***externality***) interactions (Manski, 1993), by assuming that ***Conditional Mean Independence*** (i.e., *selection-on-observables*) holds.

- **SETTING**: we consider a *binary* treatment variable w - taking value 1 for treated and 0 for untreated units - assumed to affect an *outcome* variable y that can take a variety of forms: binary, count, continuous, etc..

- **NOTATION:**

  – N = number of units involved in the (social) experiment

  – N_1 = number of treated units

  – N_0 = number of untreated units

  – w_i = treatment variable assuming value "1" if the unit is treated and "0" if untreated

  – y_{1i} = outcome of unit i when he is treated

  – y_{0i} = outcome of unit i when he is untreated

  – $\mathbf{x}_i = (x_{1i}, x_{2i}, x_{3i} \ldots x_{Mi})$ = row vector of M observable variables for unit i.

The notion of “endogenous” neighbourhood effects

Manski (1993) identifies three types of effects corresponding to three arguments of an individual outcome equation incorporating *social effects*:

1. ***Endogenous effects***: the outcome of an individual depends on the outcomes of other individuals belonging to his neighbourhood.

2. ***Exogenous (or contextual) effects***: the outcome of an individual is affected by the exogenous idiosyncratic characteristics of the individuals belonging to his neighbourhood.

3. ***Correlated effects***: due to belonging to a specific group and thus sharing some institutional/normative condition (that one can loosely define as “environment”).

Contextual and correlated effects are to be assumed as exogenous, as they clearly depend on pre-determined characteristics of the individuals in the neighbourhood (case 2) or of the neighbourhood itself (case 3).

Endogenous effects are of broader interest: they depend on the behaviour (measured as "outcome") of other individuals involved in the same neighbourhood.

Endogenous effects both comprise ***direct*** and ***indirect*** effects linked to a given external intervention on individuals.

The model presented here incorporates the presence of ***endogenous neighbourhood effects*** as defined by Manski within a traditional ***binary counterfactual model*** and provides both an identification and an estimation procedure of the **Average Treatment Effects (ATEs)** in a simple parametric case.

Some related literature

Rosenbaum (2007) discusses methods for testing *null hypotheses* on the presence of interference in trials where *random assignment occurs* within groups and interference *does not cross* group boundaries.

Hudgens and Halloran (2008) extend the previous work in the setting of a *two-stage randomized trial* in which some groups are randomly assigned to host treatments, and then treatments are assigned at random within the selected groups. Interference is presumed to operate only *within groups*.

Tchetgen-Tchetgen and VanderWeele (2010) extend Hudgens and Halloran's results, providing conservative variance estimators, a framework for finite sample inference and extensions to observational studies. Hierarchical treatment assignment and interference limited to groups greatly simplifies the estimation problem, as inference can proceed assuming *independence across groups*.

===> Sobel (2006) analyzes the potential for *bias* when no-interference is mistakenly assumed, and then defines a number of direct and indirect effects that may be identifiable.

He characterizes the usual estimators of treatment effects developing their form when *interference* is allowed.

> *"When interference is present, the difference between a* ***treatment group*** *mean and a* ***control group*** *mean (unadjusted or adjusted for covariates) estimates not an average treatment effect, but rather the difference between two effects defined on two distinct subpopulations. This result is of great importance, for a researcher who fails to recognize this could easily infer that a treatment is beneficial when in fact it is universally harmful"* (p. 1398).

=== > Application: **social experiment** (with *randomization*)

=== > MTO program ("Move To Opportunity")

=== > LATE estimator (*à la* Angrist)

Position of this paper within previous literature

Previous literature assumes:

1. Randomized assignment
2. Multiple treatment
3. Non-parametric form for the *potential outcome* and *interaction*

This paper assumes:

1. Non-randomized assignment
2. Binary treatment
3. Parametric form for the *potential outcome* and *interaction*

=== > **Therefore**: this paper suggests a *simpler* and *less general* way to relax SUTVA, but one that is easy to implement in many contexts of application.

DEFINITION OF (AVERAGE) TREATMENT EFFECTS (ATEs)

Unit i **Treatment Effect**:

$$TE_i = y_{1i} - y_{0i}$$

we observe just *one* of the two quantities (y_{1i} ; y_{0i}), but *never* both: missing observation problem (Holland, 1986).

What is *observable* to the analyst is the single status of unit i, that is:

$$y_i = y_{0i} + w_i (y_{1i} - y_{0i})$$

called the **Potential Outcome Model**, and it links *unobservable* with *observable* outcomes.

Since recovering the entire distributions of y_{1i} and y_{0i} is too demanding, we focus on the population **Average Treatment Effects** (hereinafter **ATEs**) and on ATEs *conditional on* **x** (i.e., ATE(**x**)) of a policy intervention, defined as:

$$\text{ATE} = \text{E}(y_{i1} - y_{i0})$$
$$\text{ATE}(\mathbf{x}_i) = \text{E}(y_{i1} - y_{i0} \mid \mathbf{x}_i)$$

$$\text{ATET} = \text{E}(y_{i1} - y_{i0} \mid w_i = 1)$$
$$\text{ATET}(\mathbf{x}_i) = \text{E}(y_{i1} - y_{i0} \mid \mathbf{x}_i, w_i = 1)$$

$$\text{ATENT} = \text{E}(y_{i1} - y_{i0} \mid w_i = 0)$$
$$\text{ATENT}(\mathbf{x}_i) = \text{E}(y_{i1} - y_{i0} \mid \mathbf{x}_i, w_i = 0)$$

where E(·) is the mean operator. These parameters are equal to the difference between the average of the target variable when the individual is treated (y_1), and the average of the target variable when the same individual is untreated (y_0). Observe that by LIE: ATE $=\text{E}_{\mathbf{x}}\{\text{ATE}(\mathbf{x})\}$, ATET $=\text{E}_{\mathbf{x}}\{\text{ATET}(\mathbf{x})\}$, ATENT $=\text{E}_{\mathbf{x}}\{\text{ATENT}(\mathbf{x})\}$.

A NEIGHBORHOOD-EFFECT TREATMENT MODEL

y_{0i} and y_{1i} need to have a representation including the **neighborhood effect** from *treated* to *untreated* units. We start by this **parametric model system**:

$y_{1i} = \mu_1 + \mathbf{x}_i \boldsymbol{\beta}_1 + e_{1i}$	Outcome equation for the treated status
$y_{0i} = \mu_0 + \mathbf{x}_i \boldsymbol{\beta}_0 + \gamma s_i + e_{0i}$	Outcome equation for the non-treated status with *neighbourhood effect* "*s*"
$s_i = \begin{cases} \sum_{j=1}^{N_1} \omega_{ij} y_{1j} & \text{if } i \in \{w=0\} \\ 0 & \text{if } i \in \{w=1\} \end{cases}$	Form of the *neighbourhood effect* of treated *j*s on unit *i* (*weighted mean*)
$y_i = y_{0i} + w_i (y_{1i} - y_{0i})$	Potential Outcome Equation (POM)
$\sum_{j=1}^{N_1} \omega_{ij} = 1$	Weights add to one
$i = 1,\ldots,N \quad \text{and} \quad j = 1,\ldots,N_1$	*i*: index for all units; *j*: index for treated units

and **Conditional Mean Independence (CMI)** holds:

$$\mathrm{E}(y_{ig} \mid w_i, \mathbf{x}_i) = \mathrm{E}(y_{ig} \mid \mathbf{x}_i) \quad \text{with} \quad g = \{0,1\}$$

We need to solve the previous **SYSTEM** to recover an estimation of ATEs. By substitutions within the **previous system**, we eventually get that:

$$y_{0i} = \mu_0 + \mathbf{x}_i \boldsymbol{\beta}_0 + \gamma \sum_{j=1}^{N_1} \omega_{ij} y_{1j} + e_{0i}$$

Hence, ATE is equal to:

$$\text{ATE} = \text{E}(y_{1i} - y_{0i}) = \text{E}\left[\left(\mu_1 + \mathbf{x}_i \boldsymbol{\beta}_1 + e_{1i}\right) - \left(\mu_0 + \mathbf{x}_i \boldsymbol{\beta}_0 + \gamma \sum_{j=1}^{N_1} \omega_{ij} y_{1j} + e_{0i} \right) \right]$$

After some manipulations, we get that:

$$\boxed{\text{ATE} = \mu + \overline{\mathbf{x}}_i \boldsymbol{\delta} - \left(\sum_{j=1}^{N_1} \omega_{ij} \overline{\mathbf{x}}_j \right) \gamma \boldsymbol{\beta}_1}$$

where: $\overline{\mathbf{x}}_i = \text{E}(\mathbf{x}_i)$.

We are also interested in estimating ATE(**x**). Using the previous results, we finally get that:

$$\mathrm{ATE}(\mathbf{x}_i) = \mathrm{ATE} + (\mathbf{x}_i - \bar{\mathbf{x}})\boldsymbol{\delta} + \sum_{j=1}^{N_1} \omega_{ij}(\bar{\mathbf{x}} - \mathbf{x}_j)\gamma\boldsymbol{\beta}_1$$

where it is clear that ATE(**x**) depends on **x**.

Once the formulas for ATE and ATE($\mathbf{x}$) are available, it is also possible to recover the Average Treatment Effect on Treated (ATET) and on non-Treated (ATENT), that is:

$$\text{ATET} = \text{ATE} + \frac{1}{\sum_{i=1}^{N} w_i} \sum_{i=1}^{N} w_i \left[(\mathbf{x}_i - \overline{\mathbf{x}})\boldsymbol{\delta} + \sum_{j=1}^{N_1} \omega_{ij} (\overline{\mathbf{x}} - \mathbf{x}_j) \gamma \boldsymbol{\beta}_1 \right]$$

and:

$$\text{ATENT} = \text{ATE} + \frac{1}{\sum_{i=1}^{N} (1 - w_i)} \sum_{i=1}^{N} (1 - w_i) \left[(\mathbf{x}_i - \overline{\mathbf{x}})\boldsymbol{\delta} + \sum_{j=1}^{N_1} \omega_{ij} (\overline{\mathbf{x}} - \mathbf{x}_j) \gamma \boldsymbol{\beta}_1 \right]$$

These quantities are functions of observable components and parameters to be firstly consistently estimated. Once these estimates are available, standard errors for ATET and ATENT can be obtained via bootstrapping (Wooldridge, 2010, Ch. 21).

How to get consistent estimation of ATEs ?

Using an *i.i.d.* sample of observed variables for each individual *i*:

$$\{y_i, w_i, \mathbf{x}_i\} \text{ with } i = 1, \ldots, N$$

and by substitution into the POM, we get this ***Switching Random Coefficient*** Model:

$$y_i = \left(\mu_0 + \mathbf{x}_i \boldsymbol{\beta}_0 + \gamma \sum_{j=1}^{N_1} \omega_{ij} y_{1j} + e_{0i} \right) + w \left[\left(\mu_1 + \mathbf{x}_i \boldsymbol{\beta}_1 + e_{1i} \right) - \left(\mu_0 + \mathbf{x}_i \boldsymbol{\beta}_0 + \gamma \sum_{j=1}^{N_1} \omega_{ij} y_{1j} + e_{0i} \right) \right]$$

After sorting out previous formula, we finally get that:

$$y_i = \eta + w_i \cdot \text{ATE} + \mathbf{x}_i \boldsymbol{\beta}_0 + w_i(\mathbf{x}_i - \overline{\mathbf{x}})\boldsymbol{\delta} + w_i \sum_{j=1}^{N} \omega_{ij} w_j (\overline{\mathbf{x}} - \mathbf{x}_j)\gamma \boldsymbol{\beta}_1 + e_i$$

with:

$$\mu = \mu_1 - \mu_0 - \gamma\mu_1; \;\; \eta = \mu_0 + \gamma\mu_1$$

$$e_i = \gamma \sum_{j=1}^{N_1} \omega_{ij} e_{1j} + e_{0i} + w_i(e_{1i} - e_{0i}) - w_i \gamma \sum_{j=1}^{N_1} \omega_{ij} e_{1j}$$

This is a usual regression model whose parameters – **under CMI** – can be estimated *consistently* by **Ordinary Least Squares (OLS)**. With an estimation of the parameters at hand we can estimate ATE (directly from the regression) and ATEs by plugging parameters into their formulas. Observe, however, that a **matrix of distance weights Ω**=[ω_{ij}] needs beforehand to be provided by the analyst.

A PROTOCOL FOR ESTIMATING PARAMETRICALLY ATEs UNDER "NEIGHBORHOOD INTERACTIONS"

1. Provide a **matrix of distance weights $\boldsymbol{\Omega}$**=[ω_{ij}] between the generic unit i (untreated) and unit j (treated).

2. Estimate the regression model by an OLS of:

$$y_i \quad \text{on} \quad \left\{ 1,\ w_i,\ \mathbf{x}_i,\ w_i(\mathbf{x}_i - \bar{\mathbf{x}}),\ w_i \sum_{j=1}^{N} \omega_{ij} w_j (\bar{\mathbf{x}} - \mathbf{x}_j) \right\}$$

3. Obtain $\left\{\hat{\boldsymbol{\beta}}_0,\ \hat{\boldsymbol{\delta}},\ \hat{\gamma},\ \hat{\boldsymbol{\beta}}_1\right\}$ and put them into the formulas for ATEs.

INTERPRETATION OF THE "NEIGHBOURHOOD BIAS"

By comparing the formula of ATE *with* ($\gamma \neq 0$) and *without* ($\gamma = 0$) neighbourhood effect, we get the so-called **Neighbourhood Bias** (Sobel, 2006):

$$\text{Bias} = \left| \text{ATE}_{\text{no-neigh}} - \text{ATE}_{\text{with-neigh}} = \left(\sum_{j=1}^{N_1} \omega_{ij} \bar{\mathbf{x}}_j \right) \gamma \boldsymbol{\beta}_1 \right|$$

This can also be seen as the ***externality effect*** produced by the policy: it depends on:

1. weights
2. mean of **x**
3. magnitude and sign of coefficients γ and $\boldsymbol{\beta}_1$.

Observe that it can be ***positive*** as well as ***negative***.

Observe that the **NEIGHBOURHOOD BIAS** can also be interpreted as a **SPECIFICATION ERROR** in the outcome equation arising when potential outcomes are modelled without taking into account externality effects.

Finally, by defining:

$$\gamma \boldsymbol{\beta}_1 = \boldsymbol{\lambda}$$

one can (parametrically) test whether this *bias* is or is not statistically significant by testing this null:

$$\mathrm{H}_0: \quad \lambda_1 = \lambda_2 = \ldots = \lambda_M = 0$$

STATA implementation: the "ntreatreg" command

The syntax of **ntreatreg** is a very common one for a STATA command:

```
ntreatreg outcome treatment varlist , hetero(varlist_h)
spill(matrix) graphic
```

where:

outcome: *y*

treatment: *w*

varlist: x

varlist_h: subset of x

matrix: distance matrix Ω

Stata help-file for **ntreatreg**

help ntreatreg

Title

ntreatreg — Stata module for estimation treatment effects in the presence of neighbourhood Interactions

Syntax

ntreatreg *outcome treatment* [*varlist*] [*if*] [*in*] [*weight*], [spill(*matrix*) hetero(*varlist_h*) conf(*number*) graphic vce(robust) const(noconstant) head(noheader)]

fweights, iweights, and pweights are allowed; see weight.

Description

ntreatreg estimates Average Treatment Effects (ATEs) under Conditional Mean Independence (CMI) when neighbourhood interactions may be present. It incorporates such externalities within the traditional Rubin's potential outcome model. As such, it provides an attempt to relax the Stable Unit Treatment Value Assumption (SUTVA) generally used in observational studies.

Options

spill(*matrix*) specifies the adjacent (weighted) matrix used to define presence and strength of units' relationship. It could be a distance matrix, with distance loosely defined either as vector or spatial.

hetero(*varlist_h*) specifies the variables over which to calculate the idyosincratic Average Treatment Effect ATE(x), ATET(x) and ATENT(x), where x=*varlist_h*. It is optional. When this option is not specified, the command estimates the specified model without heterogeneous average effect. Observe that *varlist_h* should be the same set or a subset of the variables specified in *varlist*.

graphic allows for a graphical representation of the density distributions of ATE(x), ATET(x) and ATENT(x). It is optional for all models and gives an outcome only if variables into hetero() are specified.

vce(robust) allows for robust regression standard errors. It is optional for all models.

beta reports standardized beta coefficients. It is optional for all models.

const(noconstant) suppresses regression constant term. It is optional for all models.

conf(*number*) sets the confidence level equal to the specified *number*. The default is *number*=95.

ntreatreg creates a number of variables:

_ws_*varname_h* are the additional regressors used in model's regression when hetero(*varlist_h*) is specified.

z_ws_*varname_h* are the spillover additional regressors used in model's regression when hetero(*varlist_h*) is specified.

ATE(x) is an estimate of the idiosyncratic Average Treatment Effect.

ATET(x) is an estimate of the idiosyncratic Average Treatment Effect on treated.

ATENT(x) is an estimate of the idiosyncratic Average Treatment Effect on Non-Treated.

ntreatreg returns the following scalars:

r(N_tot) is the total number of (used) observations.

r(N_treat) is the number of (used) treated units.

r(N_untreat) is the number of (used) untreated units.

r(ate) is the value of the Average Treatment Effect.

r(atet) is the value of the Average Treatment Effect on Treated.

r(atent) is the value of the Average Treatment Effect on Non-treated.

Remarks

The treatment has to be a 0/1 binary variable (1 = treated, 0 = untreated).

When option hetero is not specified, ATE(x), ATET(x) and ATENT(x) are one singleton number equal to ATE=ATET=ATENT.

Please remember to use the update query command before running this program to make sure you have an up-to-date version of Stata installed.

Example

```
. ssc install ntreatreg
. use "FERTIL2_200.DTA"
. matrix dissimilarity dist = age agesq urban electric tv , corr
. matewmf dist dist_abs, f(abs)
. ntreatreg children educ7 age agesq evermarr electric tv , ///
  hetero(age agesq evermarr) spill(dist_abs) graphic
. test  z_ws_age1 = z_ws_agesq1 = z_ws_evermarr1 = 0
```

References

Cerulli, G. 2014. Identification and Estimation of Treatment Effects in the Presence of Neighbourhood Interactions, *Working Paper Cnr-Ceris*, N° 04/2014.

Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data, 2nd Edition*. Chapter 21. The MIT Press, Cambridge.

Example 1: effect of *location* on *crime*

Dataset. “SPATIAL_COLUMBUS.DTA” provided by Anselin (1988) containing information (22 variables) on property crimes in 49 neighbourhoods in Columubus, Ohio, in 1980.

Objective. Evaluating the impact of *housing location* on *crimes*, i.e. the causal effect of the variable “cp” - taking value 1 if the neighbourhood is located in the “core” of the city and 0 if located in the “periphery” - on the number of residential burglaries and vehicle thefts per thousand households (i.e., the variable “crime”).

Confounding observables. Only two main factors: the *household income* in $1,000 (“inc”) and the *housing value* in $1,000 (“hoval”).

===> We are interested in detecting the effect of housing location on the number of crimes in such a setting, by taking into account possible ***interactions among neighbourhoods***.

STEP 0. INPUT DATA FOR THE REGRESSION MODEL

y: `crime`

w: `cp`

x: `inc hoval`

Matrix **Ω**: `W`

STEP 1. LOAD THE STATA ROUTINE "NTREATREG" AND THE DATASET

```
. ssc install ntreatreg
. ssc install spatwmat // see package: sg162 from
http://www.stata.com/stb/stb60
. use "SPATIAL_COLUMBUS.DTA "
```

STEP 2. PROVIDE THE MATRIX "OMEGA" (HERE WE CALL IT "W")

```
. spatwmat, name(W) xcoord($xcoord) ycoord($ycoord) band(0 $band) ///
standardize eigenval(E)   // this generates the inverse distance matrix W
```

```
The following matrices have been created:
1. Inverse distance weights matrix W (row-standardized)
   Dimension: 49x49
   Distance band: 0 < d <= 10
   Friction parameter: 1
   Minimum distance: 0.7
   1st quartile distance: 6.0
   Median distance: 9.5
   3rd quartile distance: 13.6
   Maximum distance: 27.0
   Largest minimum distance: 3.37
   Smallest maximum distance: 14.51
2. Eigenvalues matrix E
   Dimension: 49x1
```

STEP 3. ESTIMATE THE MODEL USING "NTREATREG" TO GET THE “ATE” WITH NEIGHBORHOOD-INTERACTIONS

```
. set more off

. xi: ntreatreg  crime cp inc hoval , hetero(inc hoval) spill(W) graphic
```

```
      Source |       SS       df       MS              Number of obs =      49
-------------+------------------------------           F(  7,     41) =   15.74
       Model |  9793.37437     7  1399.05348           Prob > F      =  0.0000
    Residual |  3644.84518    41  88.8986629           R-squared     =  0.7288
-------------+------------------------------           Adj R-squared =  0.6825
       Total |  13438.2195    48  279.962907           Root MSE      =  9.4286
------------------------------------------------------------------------------
       crime |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          cp |   9.492458   4.816401     1.97   0.056    -.2344611    19.21938
         inc |  -.4968051   .3653732    -1.36   0.181    -1.234691     .241081
       hoval |  -.2133293    .101395    -2.10   0.042    -.4181006    -.008558
     _ws_inc |   -1.19053   .9911119    -1.20   0.237    -3.192121    .8110612
   _ws_hoval |   .1440651   .2268815     0.63   0.529    -.3141313    .6022616
   z_ws_inc1 |  -5.719737   2.934276    -1.95   0.058    -11.64563    .2061538
 z_ws_hoval1 |   .3889889   .9016162     0.43   0.668    -1.431862     2.20984
       _cons |   34.78312   8.655264     4.02   0.000     17.30346    52.26279
```

ATE

```
. scalar ate_neigh = _b[cp]      // put ATE into a scalar
. rename ATE_x _ATE_x_spill      // rename ATE_x as _ATE_x_spill
. rename ATET_x _ATET_x_spill
. rename ATENT_x _ATENT_x_spill
```

STEP 4. DO A TEST TO SEE IF THE COEFFICIENTS OF THE NEIGHBOURHOOD-EFFECT ARE JOINTLY ZERO

4.1. if one accepts the null Ho: $\gamma\boldsymbol{\beta}_0 = 0$ => the neighbourhood-effect is negligible;

4.2. if one does not accept the null => the neighbourhood-effect effect is relevant.

```
. test  z_ws_inc1 = z_ws_hoval1 = 0

 ( 1)  z_ws_inc1 - z_ws_hoval1 = 0
 ( 2)  z_ws_inc1 = 0

       F(  2,    41) =    2.35
            Prob > F =    0.1078  //  externality effect seems not significant
```

STEP 5. ESTIMATE THE MODEL USING "IVTREATREG" (TO GET ATE "WITHOUT" NEIGHBOURHOOD-INTERACTIONS)

```
. xi: ivtreatreg crime cp inc hoval , hetero(inc hoval) model(cf-ols) graphic

      Source |       SS       df       MS              Number of obs =      49
-------------+------------------------------           F(  5,    43) =   19.84
       Model |  9375.05895     5  1875.01179           Prob > F      =  0.0000
    Residual |  4063.1606     43  94.4921069           R-squared     =  0.6976
-------------+------------------------------           Adj R-squared =  0.6625
       Total |  13438.2195    48  279.962907           Root MSE      =  9.7207
------------------------------------------------------------------------------
       crime |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          cp |   13.59008   4.119155     3.30   0.002     5.283016    21.89715
         inc |  -.8335211   .3384488    -2.46   0.018    -1.516068   -.1509741
       hoval |  -.1885477   .1036879    -1.82   0.076    -.3976543    .0205588
     _ws_inc |   -1.26008   1.004873    -1.25   0.217    -3.286599    .7664396
   _ws_hoval |   .2021829   .2300834     0.88   0.384    -.2618246    .6661904
       _cons |   46.52524   6.948544     6.70   0.000     32.51217    60.53832
------------------------------------------------------------------------------
. scalar ate_no_neigh = _b[educ7] // put ATE into a scalar
. di ate_no_neigh
```

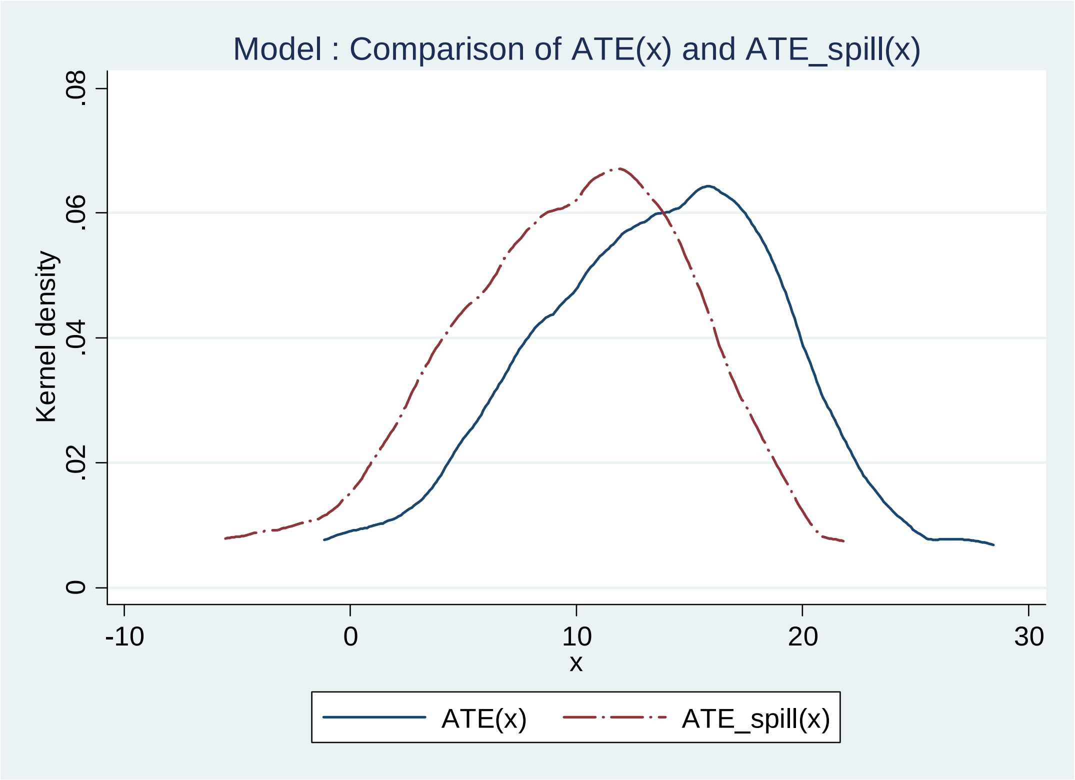
ATE

STEP 6. SEE THE MAGNITUDE OF THE NEIGHBORHOOD-INTERACTIONS BIAS

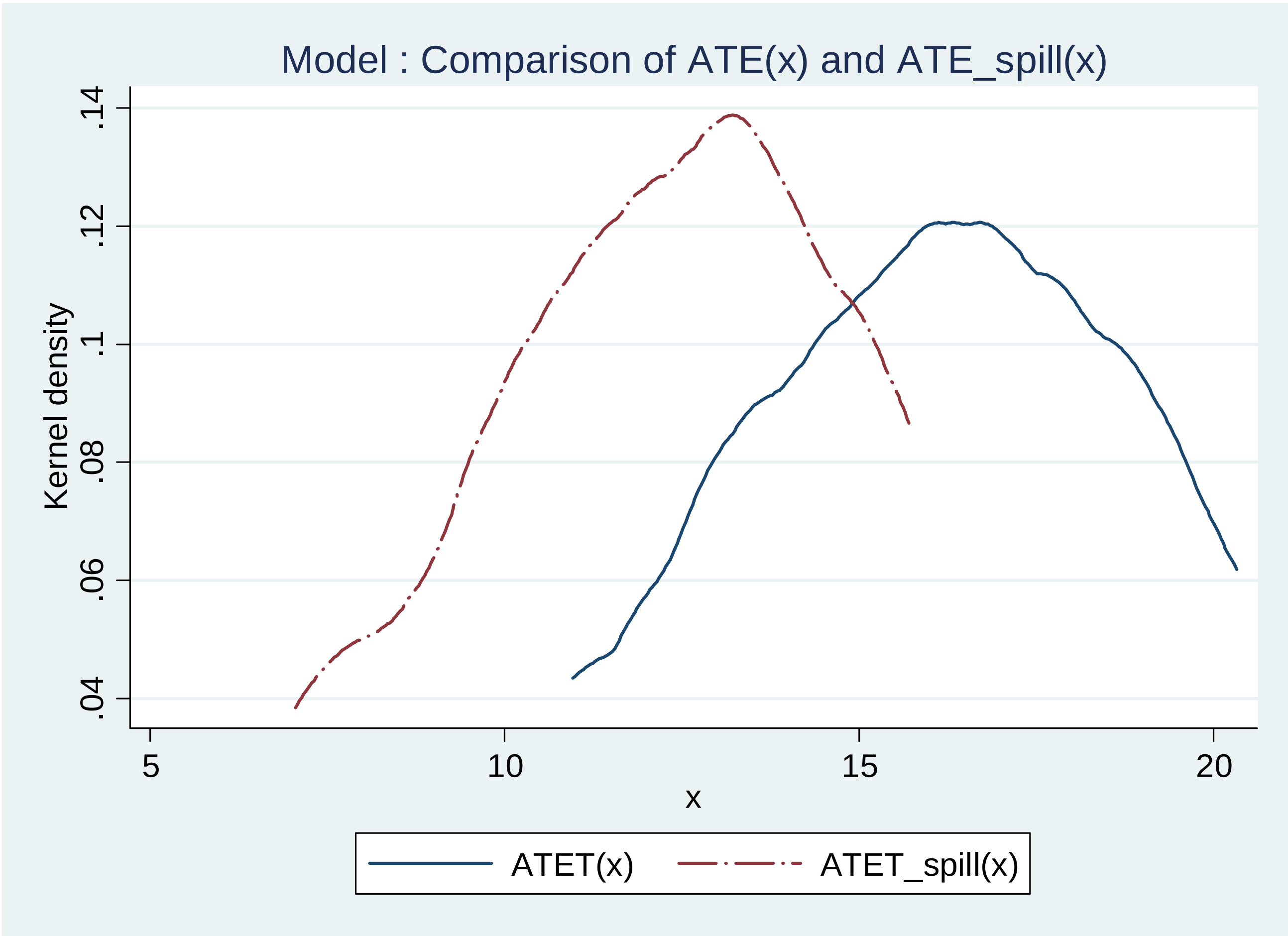
```
. scalar bias= ate_no_neigh - ate_neigh  // in level
. di bias
4.09  // the difference in level is around four crimes
. scalar bias_perc=(bias/ate_no_neigh)*100  // in percentage
. di bias_perc
30.15 // there is a 30% of bias due to neighbourhood interaction
```

STEP 7. COMPARE GRAPHICALLY THE DISTRIBUTION OF ATE(**x**), ATET(**x**) and ATENT(**x**) WITH AND WITHOUT NEIGHBOURHOOD-INTERACTION

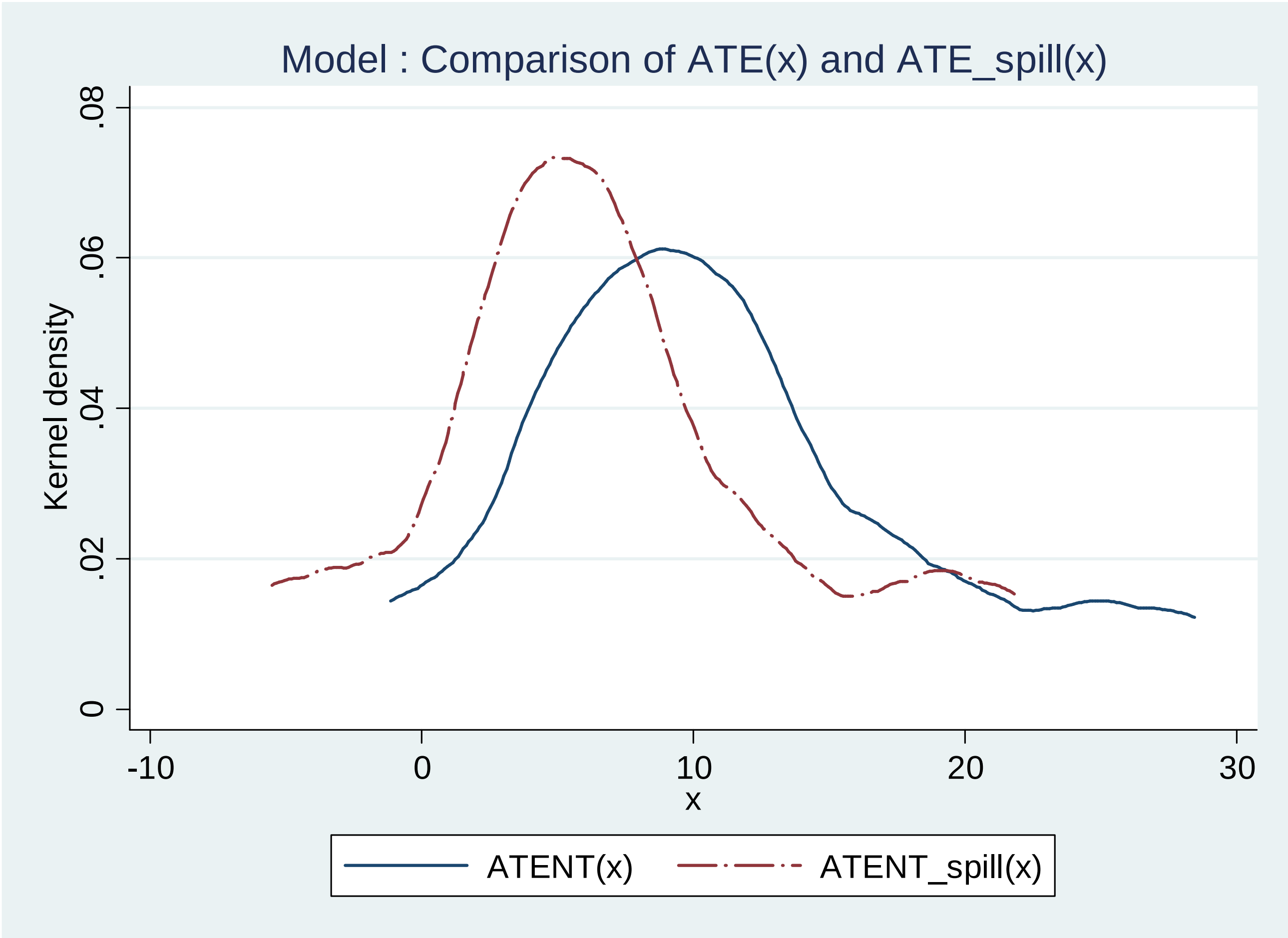
```
* ATE
twoway kdensity ATE_x , ///
|| ///
kdensity _ATE_x_spill ,lpattern(longdash_dot) xtitle() ///
ytitle(Kernel density) legend(order(1 "ATE(x)" 2 "ATE_spill(x)")) ///
title("Model `model': Comparison of ATE(x) and ATE_spill(x)", size(medlarge))
```



```
* ATET
twoway kdensity ATET_x , ///
|| ///
kdensity _ATET_x_spill ,lpattern(longdash_dot) xtitle() ///
ytitle(Kernel density) legend(order(1 "ATET(x)" 2 "ATET_spill(x)")) ///
title("Model `model': Comparison of ATE(x) and ATE_spill(x)", size(medlarge))
```



```
* ATENT
twoway kdensity ATENT_x , ///
|| ///
kdensity _ATENT_x_spill ,lpattern(longdash_dot) xtitle() ///
ytitle(Kernel density) legend(order(1 "ATENT(x)" 2 "ATENT_spill(x)")) ///
title("Model `model': Comparison of ATE(x) and ATE_spill(x)", size(medlarge))
```



STEP 8. COMPARING UNCONSTRAINED (i.e., WITH SPILLOVER) VS. UNCONSTRAINED (i.e., WITHOUT SPILLOVER) PREDICTIONS

We write a program, "`_marg`", returning the difference between the constrained and the unconstrained prediction, when `cp`=1:

```
cap prog drop _marg
program _marg , rclass
qui ntreatreg crime cp inc hoval , hetero(inc hoval) spill(W)
* uncontrained prediction
margins , at(cp= 1)
mat A=r(table)
mat B=A["b","_cons"]
return scalar _marg1=B[1,1]
* contrained prediction
margins , at(cp= 1 z_ws_inc1=0 z_ws_hoval1=0)
mat A=r(table)
mat B=A["b","_cons"]
return scalar _marg2=B[1,1]
end
```

We test:

$$H_0: \ E(y_1| \text{with spillover}) - E(y_1| \text{without spillover}) = 0$$

We can use "`_marg`" to test whether predictions are different by bootstrap:

```
bootstrap t=(r(_marg2)-r(_marg1)), rep(10): _marg
```

```
. bootstrap t=(r(_marg2)-r(_marg1)), rep(10): _marg
(running _marg on estimation sample)


Bootstrap results                               Number of obs      =        49
                                                Replications       =        10

      command:  _marg
            t:  r(_marg2)-r(_marg1)

------------------------------------------------------------------------------
             |   Observed   Bootstrap                         Normal-based
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           t |  -9.715185   2.203703    -4.41   0.000    -14.03436   -5.396007
------------------------------------------------------------------------------
```

The average difference in prediction is around -10 and it is significant. This entails that, in terms of prediction, the neighbourhood effect accounts for **10 fewer burglaries**.

===> **Conclusion**: not considering "neighbourhood effects" leads to "over-estimate" the actual effect of housing location on crime of around a 30%. Although, the Wald-test seems to show that the neighbourhood effect is not significant, if we accept the model with spillovers as the actual one, the average difference in prediction without and with spillovers is around -10 and it is also significant.

Limits and further developments

- Extending the model to "multiple" or "continuous" treatment (i.e., *w* no more binary, but **multi-valued** or **continuous**), by still holding CMI.

- Identifying the model when *w* is **endogenous** (i.e., CMI does not hold), by implementing some GMM-IV estimation.

- So far we have assumed the weighting matrix $\mathbf{\Omega}$ to be "exogenous". But: what happens if individuals **strategically** modify their "distance weights" to better profit of others' treatment? In this case weights become ***endogenous***. It poses severe identification problems.

- Providing **Monte Carlo** studies to see how the model is **robust** under different specification-errors in the weighting matrix $\mathbf{\Omega}$ provided.

- Going towards a **semi-parametric** approach

References

Anselin, L. (1988), *Spatial Econometrics: Methods and Models*. Boston: Kluwer Academic Publishers.

Holland, P.W. (1986), Statistics and causal inference, *Journal of the American Statistical Association*, 81, 396, 945–960.

Hudgens, M. G. and Halloran, M.E. (2008), Toward causal inference with interference, *Journal of the American Statistical Association*, 103, 482, 832–842.

Manski, C.F.(1993), Identification of endogenous social effects: The reflection problem, *The Review of Economic Studies*, 60, 3, 531–542.

Manski, C.F. (2013), Identification of treatment response with social interactions, *The Econometrics Journal*, 16, 1, S1–S23.

Robins, J. M., Hernan, M.A. and Brumback, B. (2000), Marginal structural models and causal inference in epidemiology, *Epidemiology*, 11, 5, 550–560.

Rosenbaum, P.R. (2007). Interference between units in randomized experiments. *Journal of the American Statistical Association*, 102, 477, 191–200.

Rubin, D.B. (1974), Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies, *Journal of Educational Psychology*, 66, 5, 688–701.

Rubin, D.B. (1977), Assignment to treatment group on the basis of a covariate, *Journal of Educational Statistics*, 2, 1, 1–26.

Rubin, D.B. (1978), Bayesian inference for causal effects: The role of randomization, *Annals of Statistics*, 6, 1, 34–58.

Sobel, M.E. (2006), What do randomized studies of housing mobility demonstrate?: Causal inference in the face of interference, *Journal of the American Statistical Association*, 101, 476, 1398–1407.

Tchetgen-Tchetgen, E. J. and VanderWeele, T.J. (2010), On causal inference in the presence of interference, *Statistical Methods in Medical Research*, 1–21.