A multivariate new promising approach for assessing balance of categorical covariates and measuring local effects in observational studies using the "potential outcome" frame

Furio Camillo – Università di Bologna – www.furiocamillo.it
(with Ida D'Attoma – Università di Bologna)

The evaluation problem

The evaluation process concerns the retrospective analysis of interventions to support business and management processes (private and public)

Decision makers need objective and automatic results not dependent from researcher subjectivity for what concerns research hypothesis, model specification and estimation method

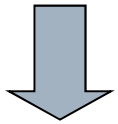
Potential outcome framework

Y ~ outcome

Z ~ treatment indicator

X ~ covariates (pre-intervention)

CAUSAL QUESTION



What would have happened to those who, in fact, received the treatment if they not have received treatment?

Potential outcome framework

A causal effect is the comparison of the outcome that would be observed with the interventions (treatment) and without intervention, both measured at the same point in time (D. B. Rubin, R.P. Waterman, 2006)

Each subject gets assigned one treatment and has a potential outcome

Outcome / **Treat**	Outcome for exposed Y(1)	Outcome for unexposed Y (0)
*Exposed to treatment :**Z=1***	**OBSERVED**	**MISSING**
*Unexposed to treatment : **Z=0***	**MISSING**	**OBSERVED**

First, potential outcomes and covariates are defined as scientific entities, no matter which design - experimental, observational or something else - researcher use. What Rubin calls *The Science*, in our approach is represented by the information matrix $\underline{X}$, and by observed potential outcomes.

Formally, the starting information of this multivariate approach is represented in table 4.1:

	1 ... j ... Q
1	
...	
i	x_{ij}
...	
n	

T
1
...
0
...
1

Y(0)	Y(1)
missing	observed
...	...
observed	missing
...	...
missing	observed

Table 4.1: Left: Information matrix; Center: assignment vector; Right: observed potential outcome

The essential role of assignment mechanism

If the assignment mechanism is not randomized, without a model for how treatments get assigned to units, formal causal inference, as least using probabilistic statement, is impossible. (Rubin, 1976, p.581)

Potential Outcomes (Illustration)

Constant treatment effect

subject i	control Y^0	treat. Y^1	ITE Y^1-Y^0
1	6	8	2
2	8	10	2
3	9	11	2
4	11	13	2
5	11	13	2
6	12	14	2
7	14	16	2
8	17	19	2
mean	11	13	2

- Potential outcomes: Y_i^0, Y_i^1
- *Constant* individual treatm. effect (ITE)
  $Y_i^1 - Y_i^0 = \tau_i = \tau = 2$
- Average Treatment Effect (ATE):
  $\bar{Y}^1 - \bar{Y}^0 =$
  $= 13 - 11 = 2$

Potential Outcomes (Illustration)

Heterogeneous treatment effect

subject	control	treat.	ITE
i	Y^0	Y^1	Y^1-Y^0
1	6	8	2
2	8	11	3
3	9	9	0
4	11	12	1
5	11	15	4
6	12	11	-1
7	14	17	3
8	17	21	4
mean	11	13	2

- Potential ouctomes Y_i^0, Y_i^1
- *Heterogeneous* indiv. causal effect (ITE) $Y_i^1 - Y_i^0 = \tau_i$
- Average Treatment Effect (ATE): $\bar{Y}^1 - \bar{Y}^0 = 13 - 11 = 2$

Fundamental Problem of Causal Inference (Holland, 1986)

- We only observe one of the two potential outcomes. Depending on treatment assignment Z it is either the treatment or control outcome:
  - For the *treated* we do not know Y_i^0, i.e., what their outcome would have been if they would have not been treated
  - For the *untreated* do not know Y_i^1, i.e., what their outcome would have been if they would have been treated

- We only observe Y_i:

$$Y_i = Y_i^0(1 - Z_i) + Y_i^1 Z_i$$

Fundamental Problem of Causal Inference (Holland, 1986)

- Without imposing assumptions, we can neither infer/estimate individual treatment effects nor average treatment effects:
  - *Individual causal effects* (ITE) are not directly accessible since only one of the two potential outcomes is observed
  - *Average treatment effect* (ATE) cannot be inferred since, in general, the *prima facie effect* (i.e., the mean difference between the treated and untreated subjects) is different to ATE

$$\tau = E(Y^1) - E(Y^0) \neq E(Y \mid Z=1) - E(Y \mid Z=0) =$$

$$E(Y^1 \mid Z=1) - E(Y^0 \mid Z=0)$$

ATE $\neq$ prima facie effect

Conditions Enabling Causal Inference

- Given the fundamental problem of causal inference, how can we infer causal effects?
- Here are the most important conditions (assumptions!) under which causal inference is possible:
  - Temporal stability and causal transience
  - Unit homogeneity
  - Independence (of treatment selection and potential outcomes)
- (Holland, 1986; West et al., 2000)

Random. Exeriment (potential outcomes)

subject	treat.	crtl.	treat.
	Z	Y^0	Y^1
1	1	6	8
2	0	8	10
3	1	9	11
4	0	11	13
5	0	11	13
6	1	12	14
7	0	14	16
8	1	17	19
mean C	0	11	13
mean T	1	11	13

- *Randomized experiment*
- Selection probabilities $P(Z_i = 1) = .5$
- Y^0, Y^1 are *independent* of treatment assign. Z
- Particularly, treatment and control means of the pot.outcomes *do not differ*

Randomized Exeriment (observed data)

subject	treat.	crtl.	treat.
	Z	Y^0	Y^1
1	1	■	8
2	0	8	■
3	1	■	11
4	0	11	■
5	0	11	■
6	1	■	14
7	0	14	■
8	1	■	19
mean		11	13

- *Randomized experiment*
- Selection probabilities $P(Z_i = 1) = .5$
- Average treatment effect (ATE):

$$\bar{Y}^1 - \bar{Y}^0 =$$

$$= 13 - 11 = 2$$

Observational study (potential outcomes)

subject	treat.	crtl.	treat.
	Z	Y^0	Y^1
1	0	6	8
2	1	8	10
3	0	9	11
4	0	11	13
5	1	11	13
6	0	12	14
7	1	14	16
8	1	17	19
mean C	0	9.5	11.5
mean T	1	12.5	14.5

- Y^0, Y^1 *depend* on treatment assign. Z
- Particularly, treatment and control means of the potential outcomes *differ*

Observational Study (observed data)

subject	treat.	crtl.	treat.
	Z	Y_0	Y_1
1	0	6	■
2	1	■	10
3	0	9	■
4	0	11	■
5	1	■	13
6	0	12	■
7	1	■	16
8	1	■	19
mean		9.5	14.5

- *Non-equivalent control group design* (self-/administrator selection)
- Selection probabilities $P(Z_i = 1)$ unknown
- Unadj. treatm. effect: $\overline{Y}^1 - \overline{Y}^0 = 14.5 - 9.5 = 5$
- Y^0, Y^1 *dependent* on treatment assign. Z

Theory of Strong Ignorability

- If all covariates $\mathbf{X} = (X_1, \ldots, X_p)'$ related to *both* treatment assignment and potential outcomes are observed, and
- If the selection probabilities, given $\mathbf{X}$, are strictly between zero and one $0 < P(Z = 1 \mid \mathbf{X}) < 1$ holds
- then, potential outcomes are independent of treatment assignment given observed covariates $\mathbf{X}$:

$$(Y^0, Y^1) \perp Z \mid \mathbf{X}$$

and treatment assignment is said to be strongly ignorable (*strong ignorability*; Rosenbaum & Rubin 1983).

Practice of Strong Ignorability

- Need to *measure all confounding covariates*! If *not* all covariates, that are simultaneously related to treatment selection and potential outcomes, are observed
  - the strong ignorability assumption is not met and
  - the average treatment effect will remain biased!
- Need to measure covariates *reliably* (with respect to the selection mechanism)
- Each subject must have a *positive probability (but less than one)* of being in the treatment group (overlap).

How to Estimate an Unbiased Treatment Effect?

- Assume that we observe all covariates **X** such that SI holds, selection bias can be removed with different approaches
- With original covariates **X**
  - Covariance adjustments (standard regression methods)
  - Case matching on observed covariates
  - Multivariate stratification
- With a composite of original covariates $b = f(\mathbf{X})$
  - *Propensity scores* (Rosenbaum & Rubin, 1983)
  - Other approaches we will not consider: first discriminant, prognosis scores

Potential outcome framework

Two fundamental aspects:

- Potential outcomes and covariates are defined as scientific entities

- A formal probabilistic model is defined to take into account the selection mechanism, the process that creates the missing and observed potential outcomes.

The propensity score model

Propensity score represents a formal model for the assignment mechanism, it explicitly defines the process that creates missing and observed potential outcomes.

The propensity score was first established in the seminal paper by Rosenbaum and Rubin (1983). Authors demonstrated that, given some pre-intervention characteristics of units X, it is possible to construct a de-conditioned indicator $b(X)$ that allows the comparison between treated and control units with respect to an outcome variable Y. They assume that conditioning on pre-intervention observable covariates, we can take the assignment to have been random.

$$y_0, y_1 \perp z \mid X \qquad y_0, y_1 \perp z \mid b(X)$$

The propensity score

Propensity score is the probability the units gets one treatment (vs. another) given covariates. (Rosenbaum and Rubin, 1983)

$$e(x) = pr(Z_i = 1 | X_i = x)$$

Propensity score: assumptions

- ***Stable-unit-treatment-value-assumption (SUTVA)***

  The response of unit i to the treat Z does not depend on the treatment given to unit j.

- ***Strongly ignorable treatment assumptions*** $(Y_1, Y_0) \perp Z|X$

  It means that the non-treated and treated outcomes are independent of the participation status, conditioning on the set of variable X

- ***Common support*** $0 < P(Z = 1|X) < 1$

  It means that all treated units have a counterpart on the population of the non treated and anyone is a possible participant

Propensity score: key results

- **Propensity score is a balancing score**

$$X \perp Z | e(X)$$

$$\Downarrow$$

$$P(Z=1|X,e(X)) = P(Z=1|e(X))$$

*treatment and control subgroups with the same scalar e (X) have the same distribution of all covariates entered in e (X), and thus the bias due to **X** has been controlled.*

The idea is that participants who have the same propensity score but who are in different conditions are comparable because the distributions of their covariates are balanced.

- **Average treatment effect at e (X)**

is the average difference between the observed responses in each treatment group at e (X)

$$E(Y_1, Y_0 | e(X)) = E(Y | e(X), Z=1) - E(Y | e(X), Z=0)$$

Propensity score matching

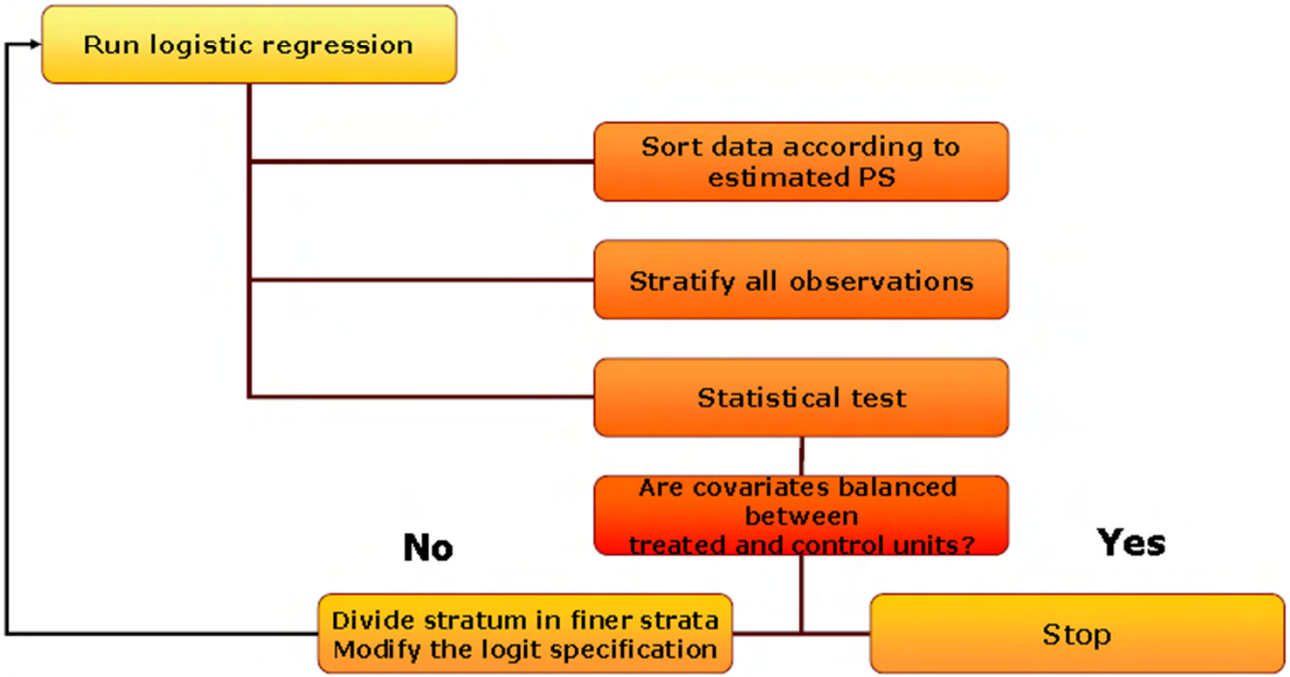
Given the propensity score for each unit:

$$e(\underline{X}_i) = P(Z_i = 1 | \underline{X}_i)$$

Propensity score matching algorithm then uses $e(\underline{X}_i)$, or a monotone function of it, to select control unit, i.e. choosing in turn for each of the treated subject the closest not yet chosen control subject.

Propensity score is not known
and
we have to estimate it!

A well-known algorithm for estimating the PS (Dehejia and Wahba,2002)



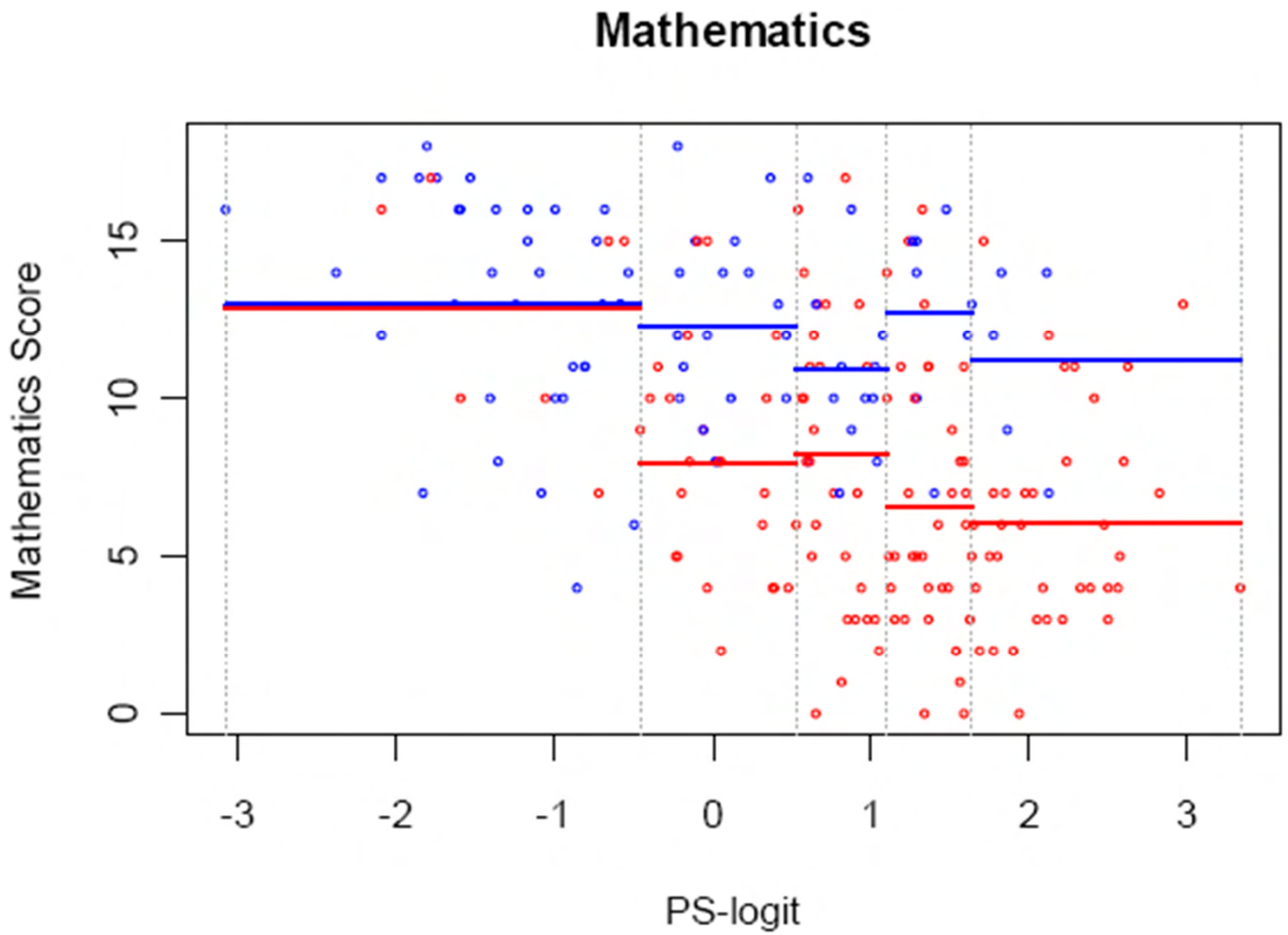
Review of Economics and Statistics
(Vol. 84, No.1, pp. 151-161, February 2002)

PROPENSITY SCORE MATCHING METHODS FOR NON-EXPERIMENTAL CAUSAL STUDIES

Rajeev H. Dehejia and Sadek Wahba*

A Simple Algorithm for Estimating the Propensity Score

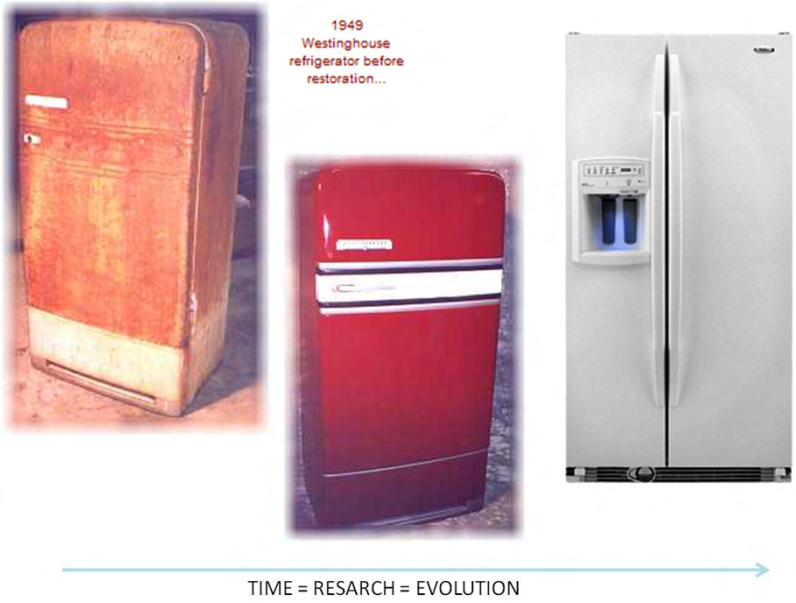
- Start with a parsimonious logit specification to estimate the score.
- Sort data according to estimated propensity score (ranking from lowest to highest).
- Stratify all observations such that estimated propensity scores within a stratum for treated and comparison units are close (no significant difference); e.g., start by dividing observations into strata of equal score range (0-0.2, ..., 0.8-1).
- Statistical test: for all covariates, differences in means across treated and comparison units within each stratum are not significantly different from zero.
  1. If covariates are balanced between treated and comparison observations for all strata, stop.
  2. If covariates are not balanced for some stratum, divide the stratum into finer strata and re-evaluate.
  3. If a covariate is not balanced for many strata, modify the logit by adding interaction terms and/or higher-order terms of the covariate and re-evaluate.


Mathematics
Mathematics Score
0
5
10
15
-3
-2
-1
0
1
2
3
PS-logit

The Data Mining approach

- Researchers and analysts don't need any a priori hypothesis about variables distribution
- We can analyze high dimensional data in a easy way
- DM algorithms aim to minimize the complexity, the time and costs of elaborations
- It generates results easy to understand

The data miner produces a "black-box", that is like an automatic tool, that aims to meet decision makers daily requirements, but in a flexible way (U. Fayaad, 2001)



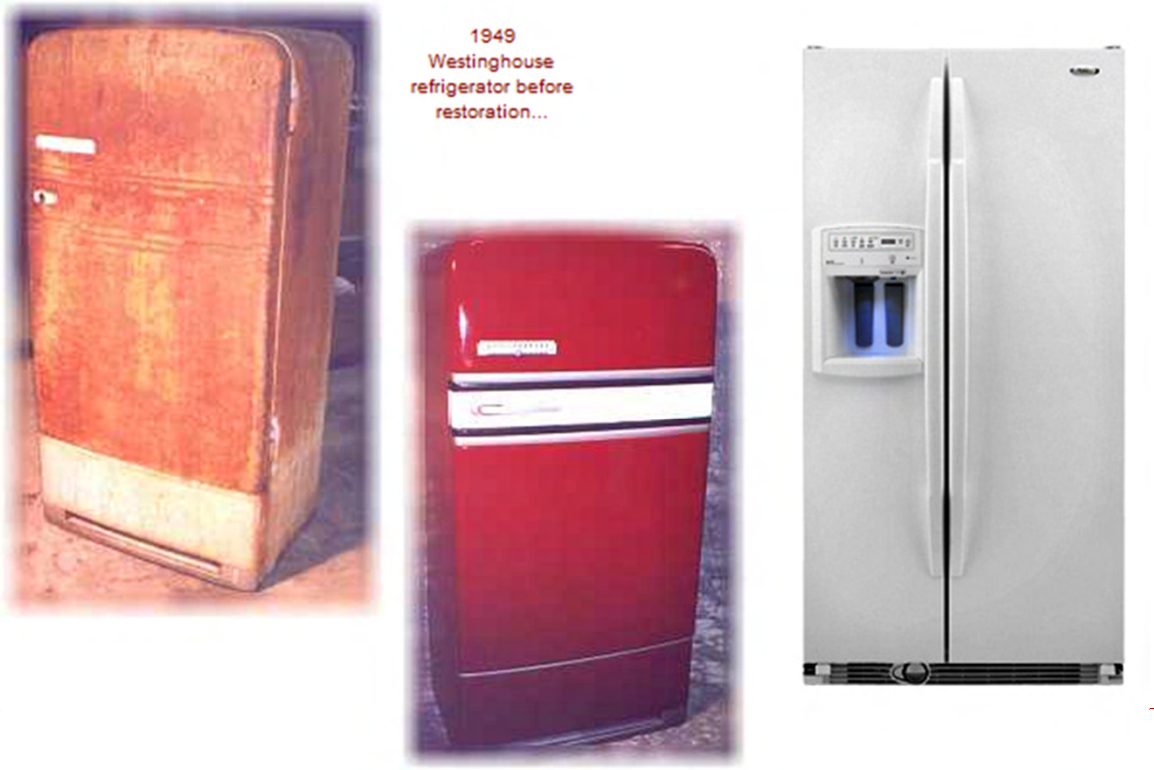
boston

3.3 Limitation of propensity score estimation

Propensity score methods differ from economic model in the sense that they do not require any model for outcome. But both PS matching methods and economic selection models are model dependent: economists use a model for both the selection process and outcomes; whereas PS methods use a model for the assignment mechanism.
Many authors agree in considering that small variations in choice during the estimation stage could yield to different results for what concerns bias reduction and size of treatment effect estimation:these choices concern control variables, functional forms, model assumptions, (see for example, W.Shadish, M.H.Clark, P.M.Steiner, 2008). When researchers use parametric methods, they do not know the true parametric model, and many different specification could be plausible.
According to Ho et al. (2007) we consider the PS as a *tautology*. In fact, in order to use non parametric matching to avoid parametric modeling researchers must know the parametric functional form of the propensity score equation. PS is a tautology also in the sense that to be a balancing score, analysts must know a consistent estimate of the true PS; but researchers know to have a consistent estimate of the PS when matching on the PS balances the covariates. Obviously, a wrong or not unique PS estimate will affect all sub-sequent analysis based on the estimated PS. Sekhon and Grieve (2008) noted that if the PS model is wrong then PS matching makes covariate balance worse, and

Data mining: key result

Data miners produce a “black-box”, that is like an automatic tool, that aims to meet decision makers daily requirements, but in a flexible way



TIME = RESARCH = EVOLUTION

Our proposal

- Our point of view: the Rubin Science Matrix is a multivariate data system
- A probabilistic tool: the "Dependence detachée"
- It is possible to transform data using eigenvalues-eigenvectors transformation
- Brigitte Escofier: conditional MCA (1988)
- Inertia decomposition: conditioned – no conditioned
- Tomas Aluja and others (2005) computes Between / Within in a CondMCA starting from the Burt table

First, potential outcomes and covariates are defined as scientific entities, no matter which design - experimental, observational or something else - researcher use. What Rubin calls *The Science*, in our approach is represented by the information matrix $\underline{X}$, and by observed potential outcomes.

Formally, the starting information of this multivariate approach is represented in table 4.1:

	1	...	j	...	Q
1					
...					
i			x_{ij}		
...					
n					

T
1
...
0
...
1

Y(0)	Y(1)
missing	observed
...	...
observed	missing
...	...
missing	observed

Table 4.1: Left: Information matrix; Center: assignment vector; Right: observed potential outcome

JEAN-JACQUES DAUDIN

Analyse factorielle des dépendances partielles

Revue de statistique appliquée, tome 29, n° 2 (1981), p. 15-29.

<http://www.numdam.org/item?id=RSA_1981__29_2_15_0>

I. DEPENDANCE PARTIELLE ATTACHEE ET DEPENDANCE PARTIELLE DETACHEE

A. Présentation des définitions de J.N. DARROCH

Soient X, Y, Z trois variables aléatoires discrètes et appelons

$$P_{ijk} = P(X = i,\ Y = j,\ Z = k)\ ,\ (i = 1..I, j = 1..J, k = 1..K)$$

$$P_{ij.} = \sum_{k} P_{ijk}\ , P_{.jk} = \sum_{i} P_{ijk}\ ; P_{i.k}\ ;\ P_{i.k} = \sum_{j} P_{ijk}$$

$$P_{..k} = \sum_{ij} P_{ijk}$$

X et Y sont indépendants conditionnellement à Z si (1) est vérifié :

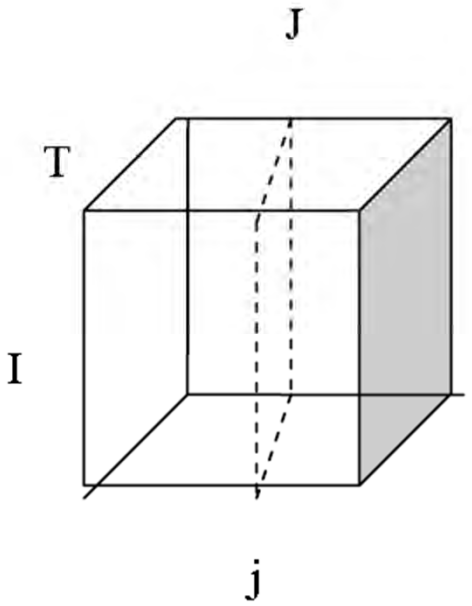
$$\forall ijk\ ,\ P_{ijk} = P_{i.k}\, P_{.jk} / P_{..k} \qquad (1)$$

JEAN-JACQUES DAUDIN

Analyse factorielle des dépendances partielles

Revue de statistique appliquée, tome 29, n° 2 (1981), p. 15-29.

<http://www.numdam.org/item?id=RSA_1981__29_2_15_0>



III. ANALYSE FACTORIELLE DE LA DEPENDANCE DETACHEE

2. Analyse factorielle de la dépendance détachée

Une solution plus simple consiste à faire l'analyse factorielle des correspondances de la table de contingence suivante :

$$N^{*}_{ij} = N_{i..}\, N_{.j.}/N + (N_{ij.} - M_{ij}) \qquad \text{Ou:} \quad M_{ij} = \sum_{k} N_{i.k}\, N_{.jk}/N_{..k}$$

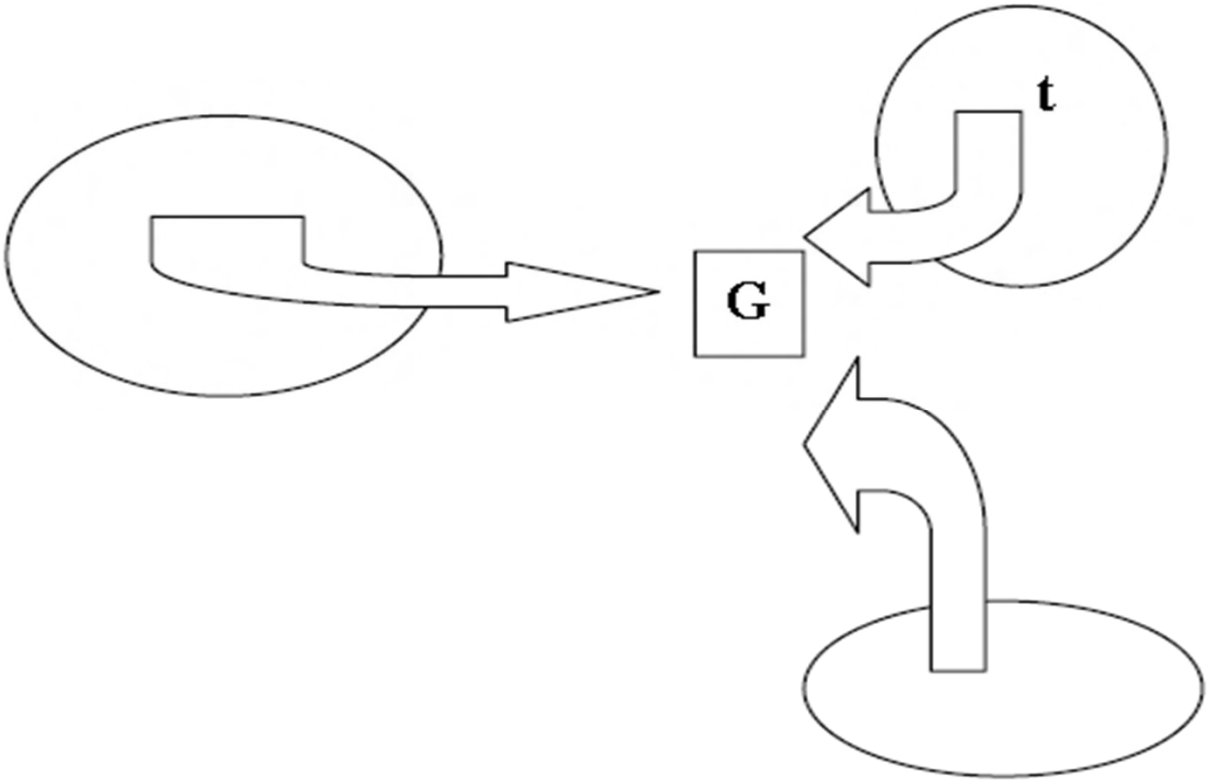


Figure 4.2: The unit space in the CORCO model

$$\frac{1}{k_t} \sum_{i \in I_{n(t)}} \frac{k_{ij}}{Q} = \frac{1}{QK_{.t}} \sum_{i \in I_{n(t)}} k_{ij} = \frac{b_{tj}}{Qk_{.t}}$$

The inertia decomposition (1)

$$
\begin{aligned}
I_{total} &= \sum_t D_p^{(T)} \|g_t - g\|^2 + \sum_t \sum_{i \in I_{n(t)}} m_i^t \|x_i^t - g_t\|^2 \\
&= \sum_t \sum_{i \in I_{n(t)}} m_i^t (x_i^t - g)' D_p^{(k)^{-1}} (x_i^t - g) \\
&= \sum_t D_p^{(T)} (g_t - g)' D_p^{-1} (g_t - g) + \sum_t m_i \sum_{i \in I_{n(t)}} m_i^t (x_i^t - g_t)' D_p^{-1} (x_i^t - g_t) \\
&= I_{between} + I_{within}
\end{aligned}
\tag{4.13}
$$

The inertia decomposition (2)

$$I_{between} = \sum_{t=1}^{T} D_p^{(T)} (g_t - g)' D_p^{(k)^{-1}} (g_t - g) \qquad (4.16)$$

with the metric $D_p^{(k)^{-1}}$ and weights $D_p^{(T)}$, is thus:

$$\begin{aligned} I_{between} &= \sum_t \frac{k_{.t}}{n} \sum_j \frac{nQ}{k_{.j}} \left(\frac{b_{tj}}{Qk_t} - \frac{k_{.j}}{nQ} \right)^2 \\ &= \frac{1}{Q} \sum_t \sum_j \frac{b_{tj}^2}{k_{.t} k_{.j}} - 1 \end{aligned} \qquad (4.17)$$

Therefore, the inertia within group is:

$$\begin{aligned} I_{within} &= I_{total} - I_{between} \\ &= \frac{J}{Q} - 1 - \frac{1}{Q} \sum_t \sum_j \frac{b_{tj}^2}{k_{.t} k_{.j}} - 1 \end{aligned}$$

The inertia decomposition (3)

4.8 Some properties of the de-conditioned space

As Escofier (1988) has shown, the CORCO model has the same properties as the MCA:

- Constructing and projecting two spaces (R^n and R^p) on their main principal axes.
- Duality and transition formula from units space to variable space and vice versa (the conventional barycentric formula hold)
- Equivalence with the analysis of a table like a Burt table where the contingency tables are conditioned to T.

The bias elimination coefficient (BEC) $$BEC = \frac{I_{within}}{I_{total}}$$

The bias elimination coefficient (BEC)

$$BEC = \frac{I_{within}}{I_{total}} \tag{4.36}$$

Then to determine how important is the inertia between with respect to the hypothetical case of a random partition ($I_{between}$ equals zero), we need to conduct an hypothesis test. We specify the null hypothesis we would like to test as follows:

$$H_0 : I_{within} = I_{total} \Longrightarrow \text{ no dependence between covariates and selection into treatment} \tag{4.37}$$

Computational Statistics (2005) 20:449-463

Distribution of the inter and intra inertia in conditional MCA

Josep Daunis-i-Estadella[1], Tomàs Aluja-Banet[2] and Santiago Thió-Henestrosa[1]

$$I_{between} \sim \frac{\chi^2}{nQ} \tag{4.41}$$

If the variables of study are independent, the degrees of freedom of the obtained χ^2 distribution are $(T-1)(J-Q)$. When there is unknown association, they assume that:

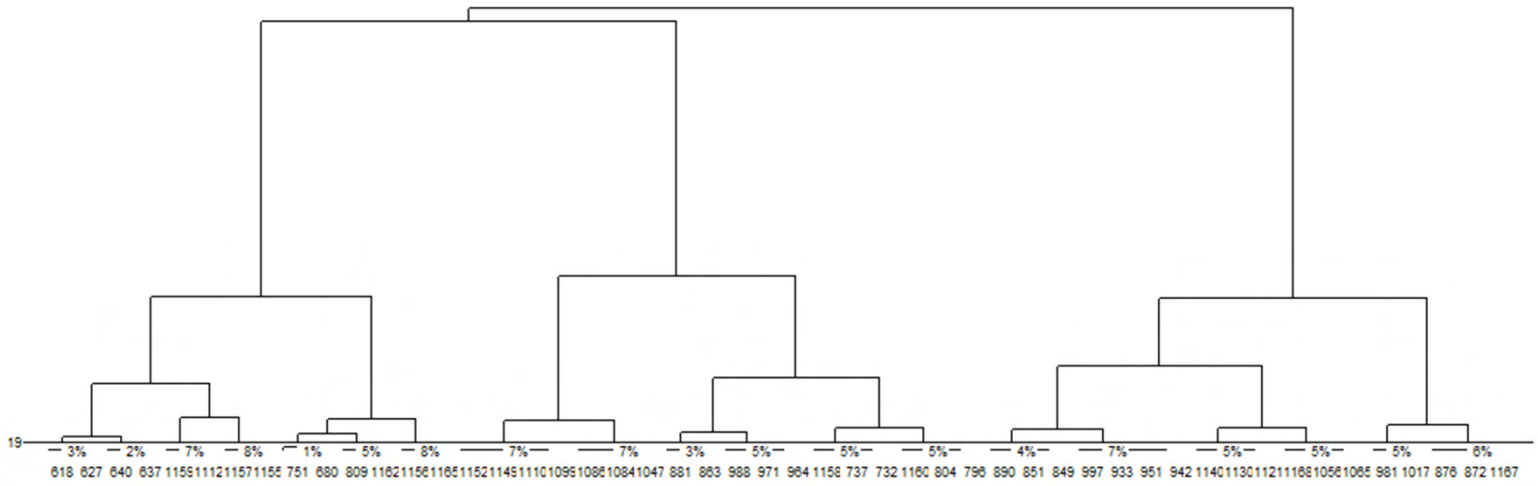
$$I_{between} \sim \frac{\chi^2_{(T-1)(J-1)}}{n-Q} \tag{4.42}$$

which provides a larger confidence interval and hence a more conservative test.

Average causal effect by clusters

- Used in subgroup analysis to detect treatment group heterogeneity (L.R.Peck,2005)
- According to a very fine clustering process on the de-conditioned coordinates we can compare, for each cluster, each non-treated individual with the cluster-benchmark treated individual

19-clusters partition



parti19=12

cond	N	Lower CL Mean	Mean	Upper CL Mean	Lower CL Std Dev	Std Dev	Upper CL Std Dev	Std Err
1	16	0.2034	0.35	0.4966	0.2032	0.2751	0.4258	0.0688
2	14	0.4245	0.6513	0.8781	0.2848	0.3929	0.6329	0.105
Diff (1-2)		-0.552	-0.301	-0.05	0.2658	0.335	0.453	0.1226

Average causal effect

T-Tests					
Variable	Method	Variances	DF	Valore t	Pr > |t|
outcome	Pooled	Equal	28	-2.46	0.0204
outcome	Satterthwaite	Unequal	22.9	-2.40	0.0249

Equality of Variances					
Variable	Method	Num DF	Den DF	Valore F	Pr > F
outcome	Folded F	13	15	2.04	0.1880

The brain reaction to spot with testimonial is 0.301 higher than reaction to spot without testimonial

LaLonde data

Rajeev H. Dehejia and Sadek Wahba
"Causal Effects in Non-Experimental Studies: Reevaluating the Evaluation of Training Programs,"
Journal of the American Statistical Association, Vol. 94, No. 448 (December 1999), pp. 1053-1062.

The data is drawn from a paper by **Robert Lalonde,** "Evaluating the Econometric Evaluations of Training Programs," *American Economic Review*, Vol. 76, pp. 604-620. We are grateful to him for allowing us to use this data, assistance in reading his original data tapes, and permission to publish it here.

Data Files

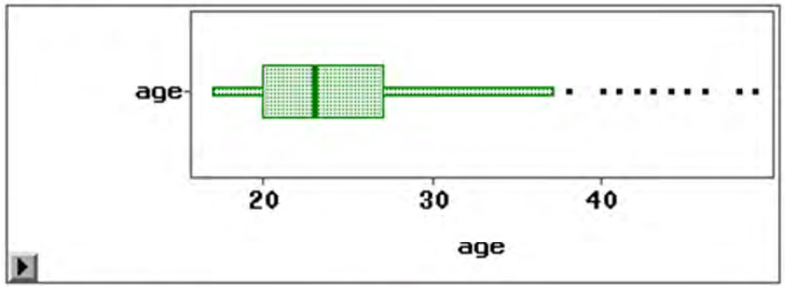
NSW_TREATED.TXT (297 observations)

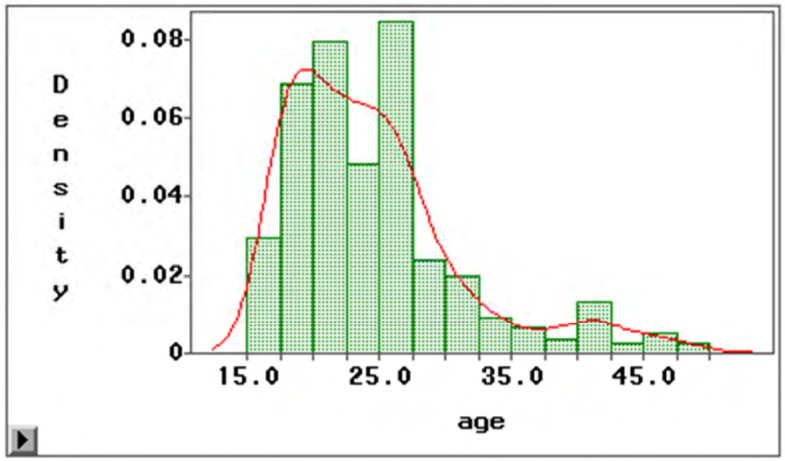
NSW_CONTROL.TXT (425 observations)

These files contain the treated and control units from the male sub-sample from the National Supported Work Demonstration as used by Lalonde in his paper. These are text files. The order of the variables from left to right is: treatment indicator (1 if treated, 0 if not treated), age, education, Black (1 if black, 0 otherwise), Hispanic (1 if Hispanic, 0 otherwise), married (1 if married, 0 otherwise), nodegree (1 if no degree, 0 otherwise), RE75 (earnings in 1975), and RE78 (earnings in 1978). The last variable is the outcome; other variables are pre-treatment.

t = 1

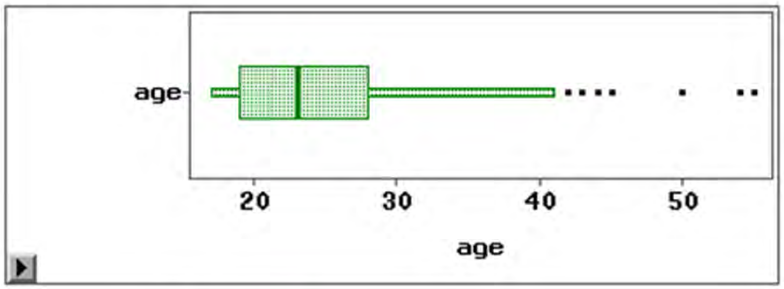
age

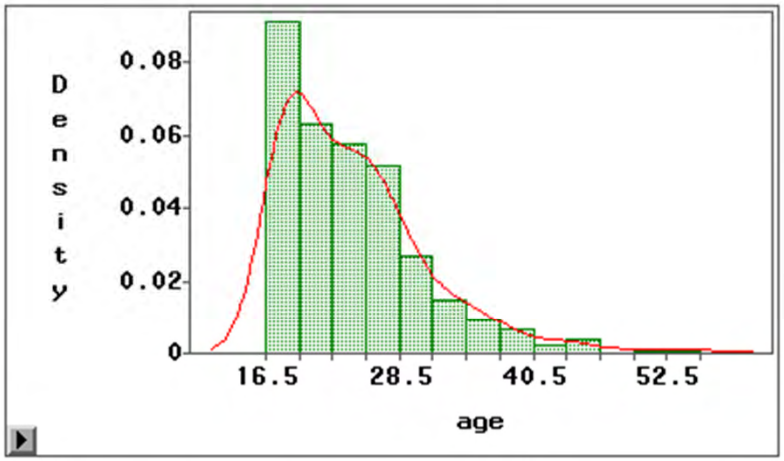




t = 2

age



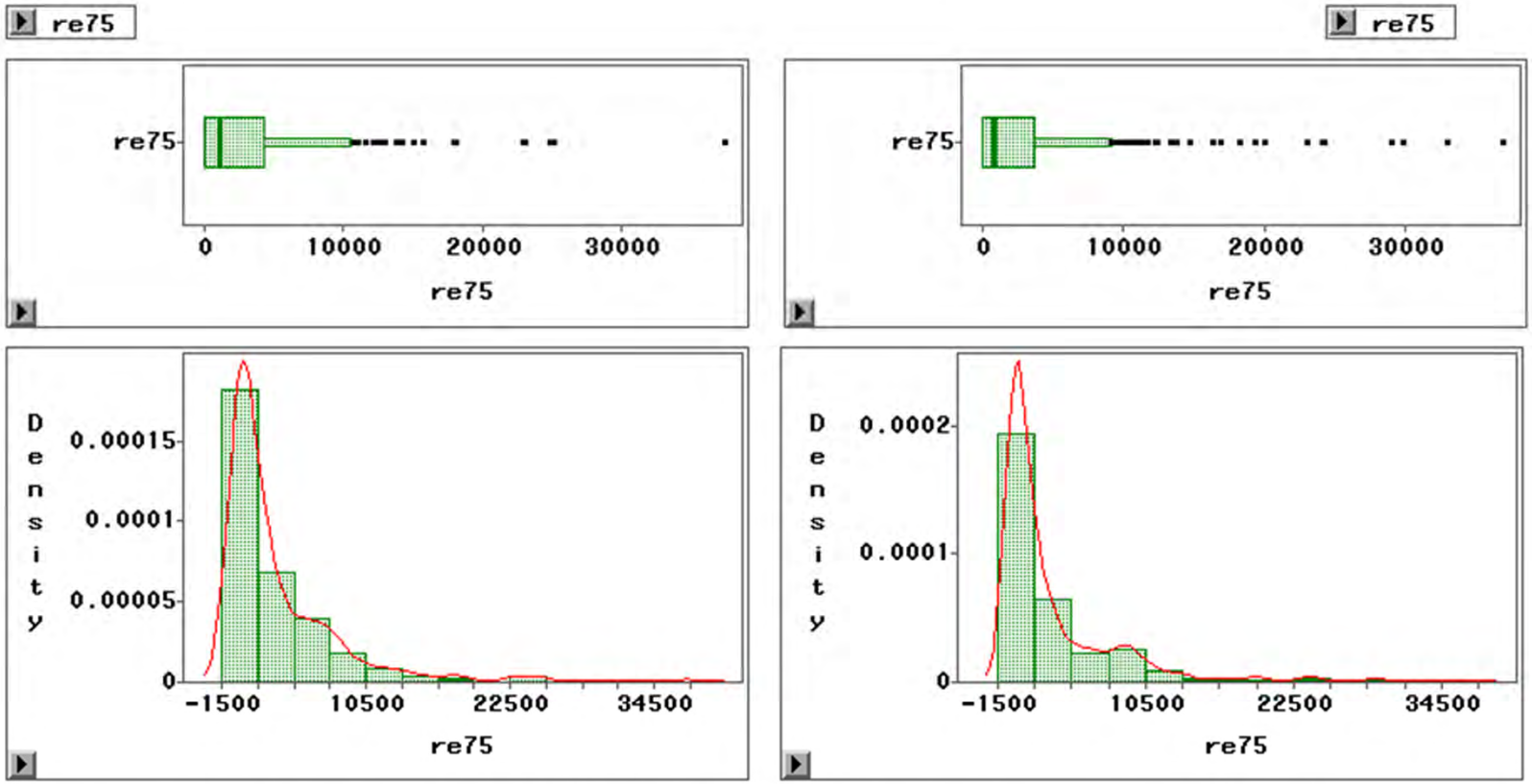


Kernel Density Estimation

Curve	Weight	Method	C Value	Bandwidth	Mode	AMISE (Normal)
———	Normal	AMISE	0.7852	1.7601	19.5000	0.0007

Kernel Density Estimation

Curve	Weight	Method	C Value	Bandwidth	Mode	AMISE (Normal)
———	Normal	AMISE	0.7852	2.1064	19.3750	0.0004



Kernel Density Estimation

Curve	Weight	Method	C Value	Bandwidth	Mode	AMISE (Normal)
	Normal	AMISE	0.7852	1083.8123	292.4348	1.095E-06

Kernel Density Estimation

Curve	Weight	Method	C Value	Bandwidth	Mode	AMISE (Normal)
	Normal	AMISE	0.7852	854.2202	288.6037	9.713E-07

Frequenza Prevista Percentuale Pct riga Pct col	Tabella di nodegree per t		
	t(t)		
nodegree(nodegree)	1	2	Totale
1	217 231.59 30.06 38.54 73.06	346 331.41 47.92 61.46 81.41	563 77.98
2	80 65.406 11.08 50.31 26.94	79 93.594 10.94 49.69 18.59	159 22.02
Totale	297 41.14	425 58.86	722 100.00

Statistica	DF	Valore	Prob
Chi-quadrato	1	7.0945	0.0077

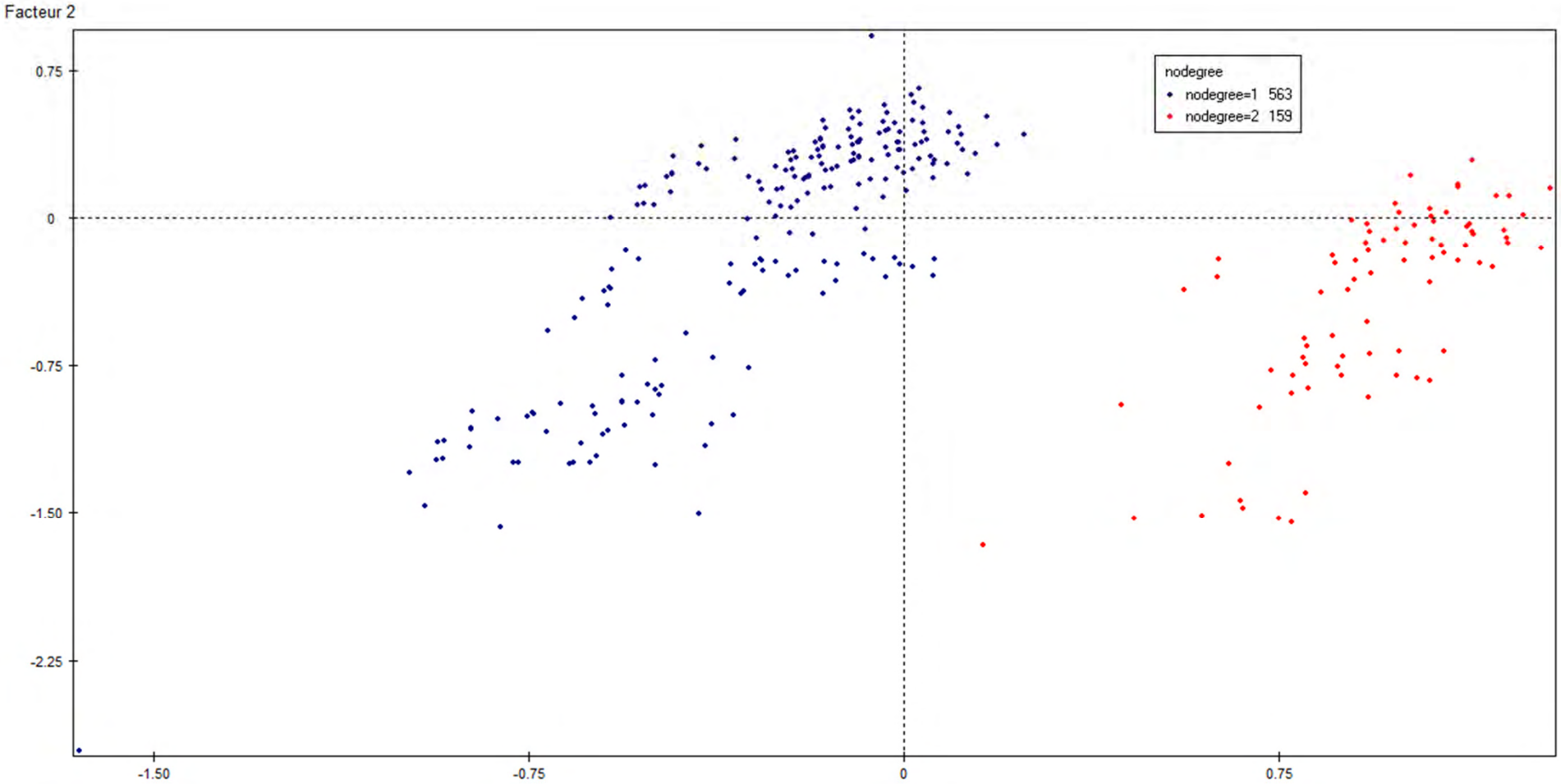
Conf interval for LaLonde data

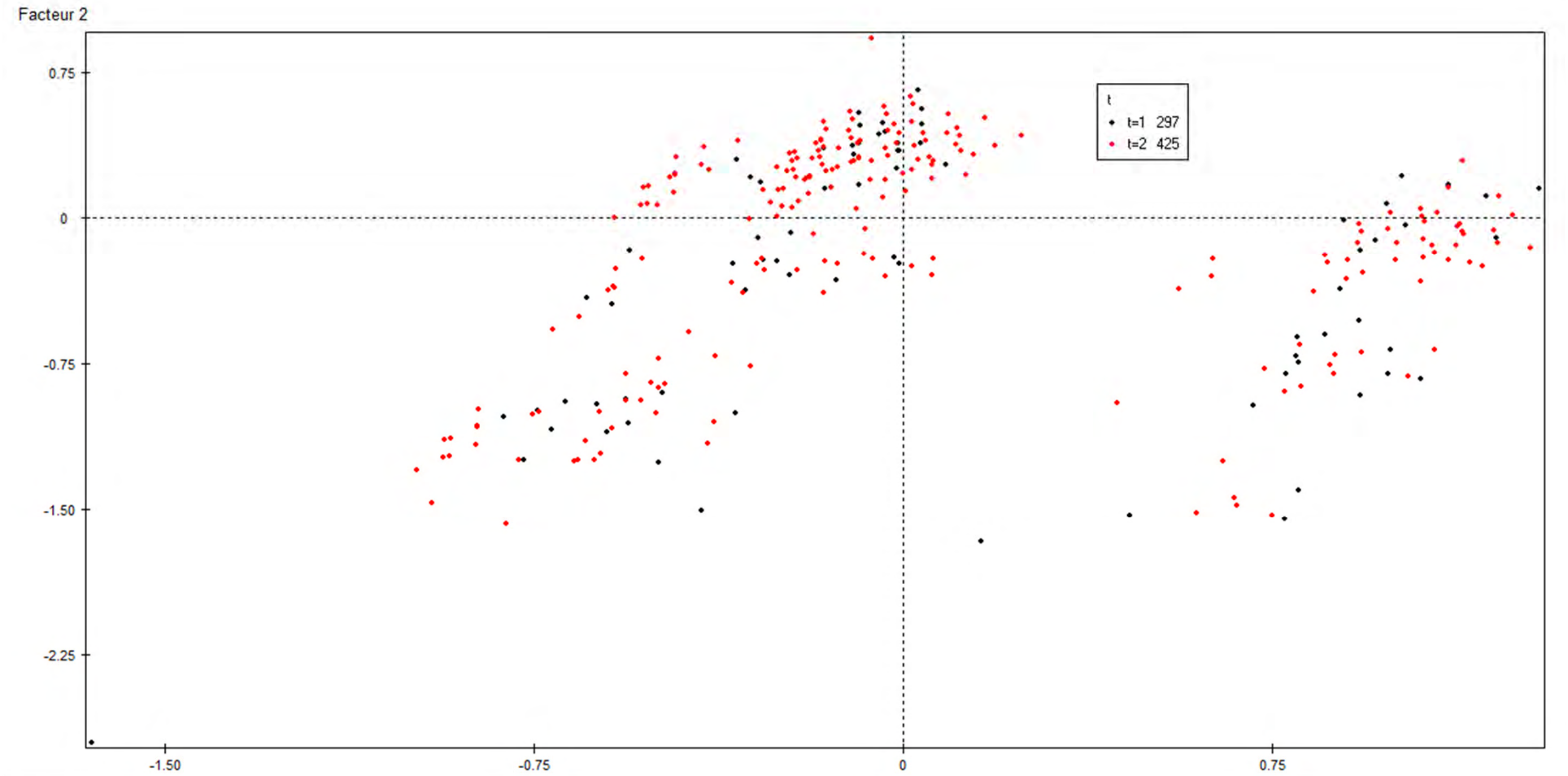
$$I_{between} \in (0, \frac{\chi^2_{(T-1)(J-1),\alpha}}{nQ})$$

$$I_{between} \in \left(0; \frac{\chi^2_{(1)(33),\alpha=0.05}}{722 * 7}\right) = 0; \frac{49.77}{5054} = 0; 0.009379$$

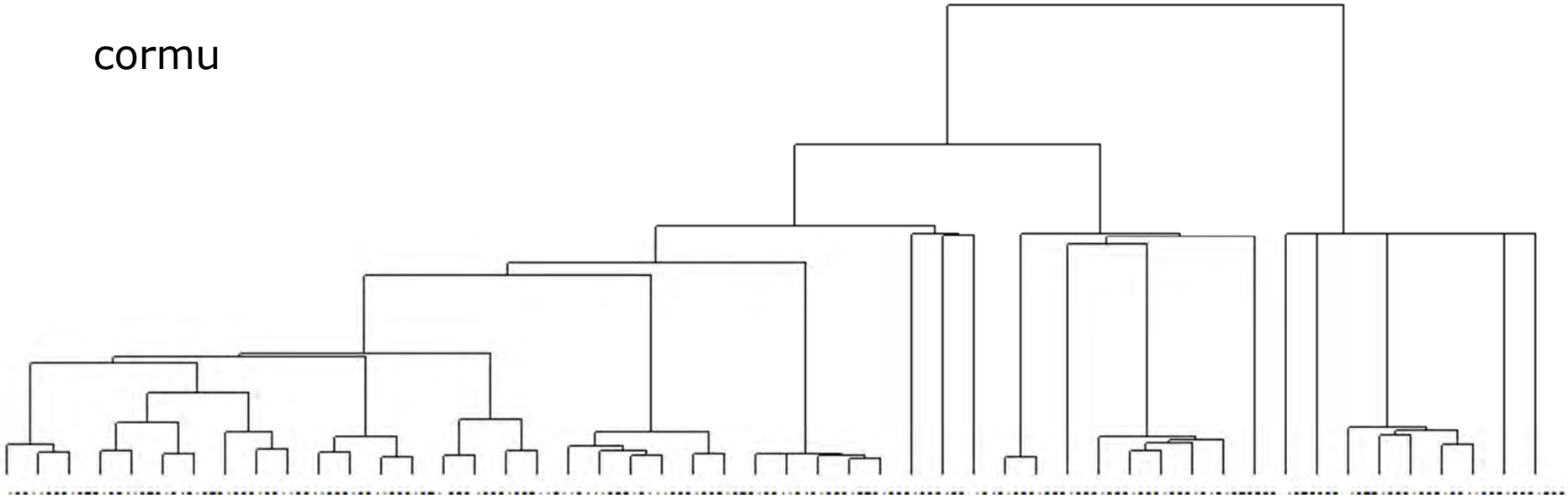
$$I_{between} \in \left(0; \frac{\chi^2_{(1)(33),\alpha=0.10}}{722 * 7}\right) = 0; \frac{47.40}{5054} = 0; 0.00865$$

$$I_{between} = \frac{1}{Q}\sum_t\sum_j \frac{b_{jt}^2}{k_t k_j} - 1 = \left(\frac{1}{Q}\sum_t \frac{1}{k_t}\sum_j \frac{b_{jt}^2}{k_{.j}}\right) - 1 \quad = \quad 0.00579$$

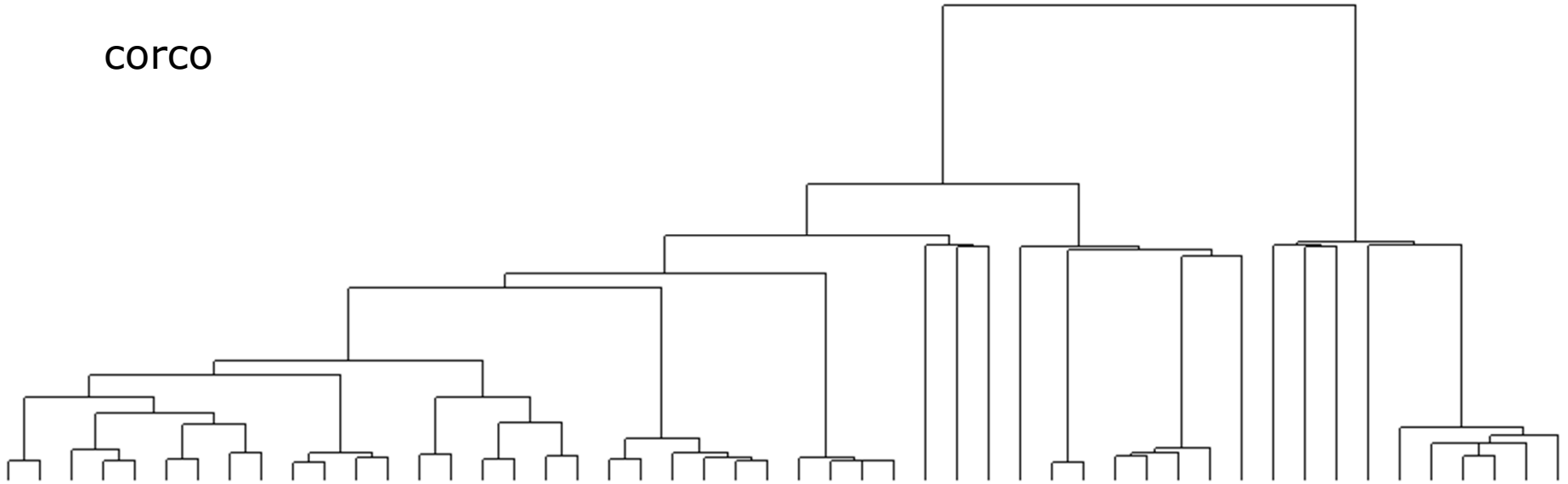

Facteur 2
0.75
0
-0.75
-1.50
-2.25
-1.50
-0.75
0
0.75
nodegree
nodegree=1 563
nodegree=2 159


Facteur 2
0.75
0
-0.75
-1.50
-2.25
-1.50
-0.75
0
0.75
t
t=1 297
t=2 425

cormu



corco



Main references (about our approach)

Camillo, F. & I. D'Attoma. (2010). A new data mining approach to estimate causal effects of policy interventions. *Expert Systems with Applications*, 37: 171-181. DOI: 10.1016/j.eswa.2009.05.072

Assessing Balance of Categorical Covariates and Measuring Local Effects in Observational Studies

To be appear in CLADAG conference Proceedings, ed. Bock and others, 2010

CAMILLO F., D'ATTOMA I., PECK L. (2009). A Promising Approach to Deal with Selection Bias: an Application to the Workfare Action. CANADIAN JOURNAL OF PROGRAM EVALUATION Vol.24 n.2 [ARTICOLO]

CAMILLO F., D'ATTOMA I. (accepted paper e in stampa). A multivariate strategy to measure and test global imbalance in observational studies. Expert System with Application – 37, 2010 pag. 171-181 (Elsevier)

furio.camillo@unibo.it

References

- Rosenbaum,P.R. and D.B. Rubin (1983), "The central role of propensity score in observational studies for causal effects", Biometria 70(1), 41-55
- Rubin,D.B (2004), "Teaching statistical inference for causal effects in experiments and observational studies", Journal of educational and Behavioral Statistics, Vol. 29, No 3, pp 343-367
- Rubin, D.B. (2005), "Causal inference using potential outcome: design, modeling, decisions", Journal of the American Statistical Association, Mar 2005; 100, 469
- Rubin, D.B, (2006), "The design versus the analysis of observational studies for causal effects: parallels with the design of Randomized Trials", Statistics in Medicine, 26: 20-36
- B.Escofier, "Analyse des correspondances multiples conditionelle", in Data Analysis and Informatics, a cura di E.Diday, pp. 333-342, North Holland, 1988
- Rubin, D.B. (1991), "Practical Implications of Modes of Statistical Inference for Causal Effects and the critical role of the assignment mechanism", Biometrics, Vol.47, No 4, pp 1213-1234
- R.H.Dehejia, S.Wahba (2002), "*Propensity score-matching methods for nonexperimental causal studies*", The Review of Economics and Statistics, 84(1): 151-161, February