CONTRIBUTI DI RICERCA CRENOS

REDUCING BIAS IN A MATCHING ESTIMATION OF ENDOGENOUS TREATMENT EFFECT

Maria Gabriella Campolo
Antonino Di Pino
Edoardo Otranto

WORKING PAPERS

2018/05

CENTRO RICERCHE ECONOMICHE NORD SUD
(CRENOS)
UNIVERSITÀ DI CAGLIARI
UNIVERSITÀ DI SASSARI

CRENOS was set up in 1993 with the purpose of organising the joint research effort of economists from the two Sardinian universities (Cagliari and Sassari) investigating dualism at the international and regional level. CRENoS' primary aim is to improve knowledge on the economic gap between areas and to provide useful information for policy intervention. Particular attention is paid to the role of institutions, technological progress and diffusion of innovation in the process of convergence or divergence between economic areas. To carry out its research, CRENoS collaborates with research centres and universities at both national and international level. The centre is also active in the field of scientific dissemination, organizing conferences and workshops along with other activities such as seminars and summer schools.
CRENoS creates and manages several databases of various socio-economic variables on Italy and Sardinia. At the local level, CRENoS promotes and participates to projects impacting on the most relevant issues in the Sardinian economy, such as tourism, environment, transports and macroeconomic forecasts.

www.crenos.it
info@crenos.it

CRENOS – CAGLIARI
VIA SAN GIORGIO 12, I-09100 CAGLIARI, ITALIA
TEL. +39-070-6756406; FAX +39-070- 6756402

CRENOS - SASSARI
VIA MURONI 23, I-07100 SASSARI, ITALIA
TEL. +39-079-213511

Title: REDUCING BIAS IN A MATCHING ESTIMATION OF ENDOGENOUS TREATMENT EFFECT

First Edition: April 2018

Reducing Bias in a Matching Estimation of Endogenous Treatment Effect

Maria Gabriella Campolo[*]
University of Messina

Antonino Di Pino[*]
University of Messina

Edoardo Otranto[*]
University of Messina and CRENoS

Abstract

The traditional matching methods for the estimation of treatment parameters are often affected by selectivity bias due to the endogenous joint influence of latent factors on the assignment to treatment and on the outcome, especially in a cross-sectional framework. In this study, we show that the influence of unobserved factors involves a cross-correlation between the endogenous components of propensity scores and causal effects. A correction for the effects of this correlation on matching results leads to a reduction of bias. A Monte Carlo experiment and an empirical application using the LaLonde's experimental data set support this finding.

Keywords: endogenous component of propensity scores, endogenous treatment, Propensity Score Matching, State-Space Model.

Jel classification: C25, C31, C35

[*] Dipartimento di Economia, Università degli Studi di Messina, 98122, Messina, Italy. E-mail: mgcampolo@unime.it.

[*] E-mail: dipino@unime.it.

[*] E-mail: eotranto@unime.it.

1. Introduction

The problems deriving from the use of observational data to infer the causal effects of a treatment are considerable (LaLonde 1986, Winship and Morgan 1999). In particular, by examining cross-sectional methods for estimating causal effects, we can observe how estimation methods such as instrumental variables, dummy endogenous variable models and regression discontinuity design, as well as techniques based on propensity score, have to take into account the endogeneity of the selection into treatment. The most frequent remedy adopted in this circumstance is to identify the decision of a subject to undergo treatment using observed covariates (e.g. Heckman et al. 1997; Heckman and Navarro-Lozano 2004). In this context, matching methods based on propensity score are often criticized for not sufficiently specifying the factors determining the assignment to treatment. As a consequence, a bias due to endogeneity occurs in the estimation of the causal effects. The aim of this study is to suggest how to improve the traditional matching procedure, based on propensity score, in order to reduce bias due to endogeneity of treatment.

A typical assumption of models with endogenous treatment effects is based on the hypothesis that the decision of a subject to receive a certain treatment depends on the difference in the outcomes potentially gained by the subject under the two alternative regimes of treatment and control, respectively (see, e.g., Winship and Morgan 1999). Starting from this assumption, the decision of a subject to undergo the treatment is endogenous with respect to the potential outcome. Fixing the notation, for a data set of n subjects, let y_{Ti} be the outcome for i-th subject, where T is an indicator assuming value 0 for untreated subjects and 1 for treated subjects; moreover $\mathbf{Z}$ denotes the set of k observed variables determining the choice of the treatment regime. The non-random selection of the units into the treatment regime, due to the endogeneity of treatment, means that unobserved covariates, not included in $\mathbf{Z}$, influence both the propensity of a subject to undergo treatment and the outcomes. As a consequence, matching estimation of the treatment effect, based on the comparison of treated and untreated units with the same propensity score, is biased (e.g., Heckman and Navarro-Lozano 2004).

A natural solution to reduce the bias, as the detection of new statistically significant covariates in the treatment choice equation, could fail; in fact Heckman and Navarro-Lozano (2004) show that this happens when these variables are not exogenous with respect to the outcome. As a consequence, the econometric distinction between exogeneity and endogeneity plays a crucial role in choosing the appropriate conditioning set in matching estimation problems. There are not standard (and simple) solutions for this problem, and in fact in applied frameworks we see greater success of estimation procedures based on Instrumental Variables or Control Functions (which explicitly model the omitted conditioning variables) than matching methods.

In this study, we try to circumvent the problem of misspecification of the selection equation in matching methods based on propensity score, assuming that the potentially omitted endogenous factors can be represented by a stochastic component of the propensity score correlated with the causal effects of the treatment. This implies that the causal effect of each subject is correlated with the causal effect of another subject with a similar propensity score; moreover the stochastic component is autocorrelated, as causal effects relative to similar propensity scores will be more similar. In order to assess this endogenous relationship, we model the causal effects by adopting a sort of state-space model (see, for example, Harvey

1990), where common latent factors are detected in correspondence with the endogenous stochastic component of the propensity score sorted in an ascending (or descending) order. State-space models are generally adopted for time series; the extension to this framework can be obtained by substituting the ordering of the observations in terms of dating with the order in terms of increasing propensity score. The predictions of these components are used as correction terms in the matching procedure. The estimation method proposed, called State-Space Corrected Matching (*SSCM*), is based on the Kalman filter (see Harvey 1990) and possesses the attractive characteristic of not imposing identification conditions on the probability of undergoing the treatment as in randomized experiments.

We verify the performance of this method comparing its bias with respect to the bias occurring with traditional propensity score matching (cf., among others, Rosenbaum and Rubin 1983) both by Monte Carlo experiments and an application on real data. In the Monte Carlo experiment we generate data in a cross-sectional context, adopting a two-regime model whose data generation process (*DGP*) is affected by endogeneity. In doing this, we set non-null correlations between the errors of the selection equation and the stochastic component of the outcome in both regimes. Applying our correction method we obtain a marked reduction of bias in the estimated average treatment effect for the treated (*ATT*) in comparison with the traditional Propensity Score Matching estimator (*PSME*). The application is made using the LaLonde (1986) experimental data set on the effect of a training program[1]; in this experiment we obtain a markedly better performance of the balancing statistics by measuring the reduction of bias in covariates after matching. The results are particularly interesting if a one-to-one criterion without replacement (whatever the level of caliper) is adopted. Both the Monte Carlo experiments and the empirical application show that the reduction of bias using *SSCM* occurs even in the condition of misspecification of the set of covariates conditioning selection into treatment[2]. In addition, the simulation results show that our *SSCM* method reduces bias in the *ATT* estimation even if a markedly observed heterogeneity in covariates occurs.

The paper is structured as follows. In the next section we briefly review the recent literature on the problem of endogeneity in the matching method as an evaluation estimator. In Section 3 we describe the new procedure and introduce the *SSCM* estimator. Section 4 shows the results of the Monte Carlo experiments, while, in Section 5, we describe the empirical application based on the LaLonde experimental and non-experimental data set, comparing the performance of *SSCM* estimation method and traditional *PSME* in terms of balancing statistics. Section 6 concludes with some final remarks.

2. Endogenous treatment in a matching estimation: a brief review

In this section, we briefly discuss how several evaluation methods face the problem of the empirical identification of the choice to undergo treatment in order to avoid the selectivity effect in estimating treatment parameters. A common characteristic of the conventional estimation methods which differ for matching is that an acceptable randomization of the

[1] This data set is available in the public domain by accessing to the link: http://users.nber.org/~rdehejia/data/nswdata2.html

[2] Incomplete specification of the set of covariates such as the non-experimental dataset compared with the LaLonde data (see for example Dehejia and Wahba, 1999 and 2002).

selection into treatment is only obtained through the imposition of limitations to the specification of the model and to the variables conditioning the choice. On the contrary, the matching procedures do not impose any exclusion restriction on covariates conditioning the selection, but suffer from the influence of latent variables correlated to both the selection and the outcome.

In particular, in order to identify the selection equation, the conventional econometric evaluation models (e.g., Instrumental Variables and Control Functions), make a distinction between the set of covariates that explain the outcome equations, **X**, and the set of covariates that explain the selection equation, **Z**. The two sets are not necessarily mutually exclusive, but contain some different variables. In particular, in order to identify the selection into the treatment, variables explaining the choice to undergo the treatment, but not correlated with the outcomes, are included in **Z**, but not in **X**.

Instead, matching procedure does not rely on exclusion restrictions. In general, matching methods do not distinguish between the variables in **Z** and in **X**. However, as a consequence of the absence of exclusion restrictions, the exposure to treatment, *T*, may be considered as not independent of the potential outcomes. A common assumption to circumvent this problem is that the variables included in **Z** can produce a randomization of *T* with respect to the outcomes (Rosenbaum and Rubin 1983). This assumption is also known as selection on observables, and requires that all variables relevant to the probability of receiving treatment may be observed and included in **X**. This allows the untreated units to be used to construct an unbiased counterfactual for the treatment group.

Several critical observations have been directed towards this approach by analysts (cf., among others, Heckman and Navarro-Lozano 2004, Heckman 2008). In particular, Heckman and Navarro-Lozano (2004) showed that only a part of the variables that guarantee the conditional independence of the treatment are generally included in **Z** (only the observed covariates), and this circumstance may induce bias in the estimated treatment parameters.

Smith and Todd (2005), in order to investigate the properties of *PSME* applied to nonexperimental data, replicate the well-known experiment of LaLonde (1986) and Dehejia and Wahba (1999, 2002) using experimental treated units of the US National Supported Work (*NSW*) with nonexperimental comparison units from the Current Population Survey (*CPS*) and the Population Survey of Income Dynamic (*PSID*), respectively. They found that, performing traditional *PSME* in a cross-sectional data set, an effective reduction of bias is obtained only if the comparison between treatment and control group should satisfy the following criteria: (i) both treatment and untreated units must come from the same data sources (i.e., the same surveys or the same type of administrative data) so that the sample characteristics are measured in an analogous way, (ii) treated and untreated subjects must reside in the same geographic area, and (iii) the data must contain a rich set of variables that affect both the selection into treatment and the outcomes. If the data fail to satisfy these criteria, the performance of the cross-sectional propensity matching estimators diminishes greatly. In addition, Smith and Todd found that the results obtained applying cross-sectional *PSME* are strongly affected by the different specification of the set of covariates and by the influence of time-invariant latent factors such as geographic mismatch and differences in the measurement of the dependent variable. In order to reduce bias due to time-invariant latent factors, they suggest adopting the Difference-in-Differences matching (*DIDM*) estimator introduced in Heckman et al. (1997) and Heckman et al. (1998) subordinately, however, to

the condition that the data set is characterized by a panel structure. The result obtained by Smith and Todd is that the difference in time-invariant latent variables affecting both matching and outcomes is better controlled adopting a *DIDM* estimator, based on the comparison of the difference in outcome over time (before and after the selection into treatment) between treated and untreated units.

Applying estimation methods alternative to matching, analysts found that a condition close to the randomization of the selection may be achieved if the individuals belonging to the treatment and comparison group solely differ with respect to the variable determining the participation status (and to the variables correlated with it); while other variables determining heterogeneity between treated and untreated should be considered as confounder variables. However, methods adopted to control the selection for heterogeneity lead to impose limitations in the use of data and of covariates. This is the case of the approach suggested by Carneiro et al. (2003) and Aakvik et al. (2005), in which unobserved covariates are modeled, as in a factors model, such as latent components common to both outcome equations and selection equation. This approach implies that an across-correlation between the errors terms of the outcome equations is determined by latent factors, and the estimation procedure would apply factor analysis methods. However, since individuals cannot be observed jointly in both treatment-regimes, a distribution of 'counterfactuals' should be preliminarily provided to perform the factor analysis. This preliminary step should be generally supported by the introduction of exclusion restrictions to prevent pairwise comparisons between observed units and counterfactuals being affected by selectivity.

Another example is given by the Local Average Treatment Effect (*LATE*) estimator (cf. Imbens and Angrist 1994, Angrist et al. 1996, among others). The *LATE* approach provides a consistent estimate of average treatment effect only for a subgroup of the population, the so-called compliers. It does not measure the effect of the treatment for everyone (*ATE*). In addition, the status of "complier" is determined using the instruments available. This implies that different instruments will give a different *LATE*, so a different set of subjects classified as compliers will provide a different *LATE*.

The heterogeneity of the subjects belonging to both treatment and comparison group implies that treatment effects may be heterogeneous for different categories of subjects. The Marginal Treatment Effect (*MTE*) approach provides a strategy to isolate the effect of the treatment with respect to other confounding factors (e.g. Carneiro et al. 2011). The *MTE* detects how much the individual's outcome increases when there is a small increase in the propensity score or, equivalently, how much higher the outcome of an individual, that is on the margin of treatment, can be expected to be by inducing him/her to undergo the treatment, via the instruments included in **Z**. The presence of instruments in the selection equation ensures that the reason for this increase is not due to the motivation of the choice.

The rationale of this method is to compare marginal participants to marginal non-participants. In this context, the term 'marginal' refers to those subjects (treated and untreated) falling in a small neighbor of the threshold for selection. This approach implies that the choice for T_i=1 or T_i=0 of the subjects located near the threshold for selection is purely random.

Such as in the MTE estimation, also with the Regression Discontinuity approach (e.g. Hahn et al. 2001), the identification of treatment effects is made possible by comparing subjects arbitrarily close to the cut-off point z_0 (the cut-off point is function of one or more

covariates, **Z**) in which they can receive or cannot receive the treatment. The *LATE* at z_0 is identified for the subgroup of subjects for whom a treatment function (parametrically or nonparametrically specified) changes discontinuously at z_0.
From this brief review of the more frequently adopted evaluation methods, alternative to matching, we can observe that every attempt to obtain a randomization of the selection leads to a loss of generality of the significance of the estimated treatment parameters.
Unlike the most commonly used approaches, based on the randomization of the selection, in this study we suggest that it is possible, with a propensity score matching approach in a cross-sectional framework, to reduce the bias due to endogeneity without imposing limitations on the sample or adopting exclusion restrictions. This purpose can be reached by providing a stochastic specification and a proper estimation of the relationship between the causal effects of treatment and probability to undergo the treatment, as illustrated in next section.

3. Model specification

The most innovative aspect of this analysis is the individuation of an autoregressive process that characterizes, jointly, individual propensity scores and causal effects. Another important novelty, strictly linked to the previous one, is given by the correction term derived as the state variable of a State-Space model that identifies the endogenous components of the causal effects.
To better explain the endogenous relationship between causal effect and propensity to undergo treatment, we start to consider the potential outcome gained by choosing one of the two treatment statuses as a relevant (endogenous) determinant of the decision to undergo treatment. In particular, we specify the model assuming that the difference between the expected outcomes, y_{1i} and y_{0i}, obtainable, respectively, under the regimes T_i=1 (if the subject belongs to the treatment group) and T_i=0 (if the subject belongs to the comparison group), determines, at least in part, the choice of the regime.
Let us consider a Probit (or Logit) model, where the (latent) propensity to undergo treatment of the *i*-th subject, T_i^*, depends linearly on the covariates in **Z**:

$$T_i^* = \mathbf{z}_i'\boldsymbol{\beta} + \nu_i \qquad (1)$$

where $\mathbf{z}_i'$ is the *i*-th row of the matrix **Z**, $\boldsymbol{\beta}$ is a vector of unknown coefficients and ν_i is a zero mean random disturbance with unit variance. If $T_i^* > 0, T_i = 1$ (the subject underwent treatment), otherwise $T_i = 0$ (the subject did not undergo treatment).
As a consequence of the endogeneity of T_i^* with respect to the causal effects, $\Delta_i = y_{1i} - y_{0i}$, we can suppose that the propensities T_i^*, sorted in ascending or descending order, are autocorrelated, and the same holds for the causal effects Δ_i. In practice, subjects *i* and *i+1,* with contiguous propensities T_i^* and T_{i+1}^*, show similar causal effects, $\Delta_i$and Δ_{i+1}. The hypothesis of autocorrelation of causal effects and propensity score is consistent with the

"Two-Regime" Roy Model (e.g., Heckman and Honoré 1990; Carneiro et al. 2003)[3], and we specify our model according to it, adding to the selection equation (1) two further equations:

$$y_{1i} = \mu_{1i} + u_{1i} \qquad \text{if } T_i = 1;\ \text{otherwise latent} \tag{2a}$$

$$y_{0i} = \mu_{0i} + u_{0i} \qquad \text{if } T_i = 0;\ \text{otherwise latent} \tag{2b}$$

In Equations (2a) and (2b), μ_{1i} and μ_{0i} are the outcomes obtained by treated and untreated subjects, respectively, depending on the decision to undergo treatment (T = 1) or not (T = 0). The error terms u_{1i} and u_{0i} are normally distributed with zero mean and variances equal to σ_1^2 and σ_0^2 respectively. The covariances $\sigma_{1\nu}$ and $\sigma_{0\nu}$ of the disturbances of both outcome equations, u_{1i} and u_{0i}, with the disturbances of the selection equation (1), v_i, are measurements of the endogeneity of the propensity to undergo treatment, T_i^*, with respect to the outcome gained under T = 1 and T = 0.
Correlation between outcomes and propensity scores, as well as the autocorrelation of the causal effects, may be specified starting from the definition of causal effects, Δ_i, obtained as:

$$\Delta_i = y_{1i} - y_{0i} = (\mu_{1i} - \mu_{0i} + (u_{1i} - u_{0i}) \tag{3}$$

We suppose that u_{1i} and u_{0i} are both linearly related to v_i, involving a certain degree of endogeneity. Formally we have:

$$u_{1i} - u_{0i} = \sigma v_i + \varepsilon_i \tag{4}$$

Putting $\mu_{1i} - \mu_{0i} = \mu_i$, Eq. (3) can be written as a *measurement equation* of a State-Space Model, as follows (cf., among others, Harvey 1990):

$$\Delta_i - \mu_i = \sigma v_i + \varepsilon_i \tag{5}$$

In Eq. (5), ε_i is a vector of n×1 disturbance terms uncorrelated across i. The variable v_i can be considered as the state variable whose elements are not observable, but are assumed to be generated by a first-order Markov process (*transition equation*):

$$v_i = \rho v_{i-1} + \eta_i \tag{6}$$

The dependent variable of Eq. (5), $\Delta_i - \mu_i$, represents the stochastic component of the causal effect Δ_i, endogenous with respect to the decision to undergo treatment. Starting from this result, the selectivity effect due to the endogeneity of the decision to undergo treatment may be corrected by estimating σv_i in Equation (5), and using the corresponding predicted values, $\sigma \hat{v}_i$, as a correction term in the matching estimation of the causal effects. In doing

[3] Moreover, as we will show below, this hypothesis is also supported by the empirical evidence of our Monte Carlo experiment.

this, a preliminary estimation of causal effects Δ_i is obtained at a first stage by applying a traditional propensity score matching procedure. Then, at a second stage, matching is replicated using the corrected outcomes $y_{1i} - \sigma\hat{v}_i$ so as to obtain the corrected causal effects $\Delta_i - \sigma\hat{v}_i = \hat{\mu}_i$. We call this estimator the State-Space Corrected Matching (*SSCM*) estimator.

3.1 Interpretation of the Selection Equation

The endogenous selection of the units into the treatment regime, as assumed in our model, means that the observed covariates do not fully determine a specification of the selection process. In particular, we assume that the choice of the regime of the subject should also depend on the potential outcome, μ_{1i} and μ_{0i}, through the influence of the error term of the selection equation, ν_i. In order to specify how ν_i is influenced by the outcome, let us start by solving the measurement equation (5) by ν_i:

$$\nu_i = \frac{1}{\sigma}[\Delta_i - (\mu_{1i} - \mu_{0i}) + \psi_i] \tag{7}$$

with $\psi_i = -\varepsilon_i$ and $\mu_{1i} - \mu_{0i} = \mu_i$.

We assume that both the components of the outcomes, μ_{1i} and μ_{0i}, are determined in part by the observed covariates $\mathbf{Z}$, and in part by unobserved factors, δ_{1i} and δ_{0i}. Then we can explain μ_{1i} and μ_{0i} by adopting the following linear specification:

$$\mu_{1i} = \mathbf{z}_i'\boldsymbol{\alpha}_1 + \delta_{1i} \tag{8a}$$

$$\mu_{0i} = \mathbf{z}_i'\boldsymbol{\alpha}_0 + \delta_{0i} \tag{8b}$$

where $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_0$ are coefficients measuring the partial effect of observed covariates on respectively, μ_{1i} and μ_{0i}, while δ_{1i} and δ_{0i} are latent components.

Substituting Eqs. (8a) and (8b) into Eq. (7), we obtain:

$$\nu_i = \frac{1}{\sigma}[\Delta_i - \mathbf{z}_i'(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0) - (\delta_{1i} - \delta_{0i}) + \psi_i] \tag{9}$$

Eq. (9) shows that the stochastic component of the propensity to undergo treatment, v_i, includes an endogenous variable given by the difference between the unobserved factors, δ_{1i} and δ_{0i}, of each regime. However, since the cross-sectional nature of the analysis does not allow us to observe a subject under the two regimes simultaneously, the term $\mathbf{z}_i'(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0)$and the difference $\delta_{1i} - \delta_{0i}$ cannot be identified. To this end we propose adopting a well known counterfactual procedure: the Blinder-Oaxaca decomposition (e.g. Blinder 1973; Oaxaca 1973).

In practice, by introducing the Blinder-Oaxaca decomposition, we can detect three distinct effects on v_i, and, consequently, on the probability to undergo treatment T_i^*: i) the effect of the "shift" in the coefficients, $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_0$, on the outcome due to the choice of the regime; ii) the effect of "change" in covariates, $\mathbf{z}_{1i}$ and $\mathbf{z}_{0i}$ between the regimes; and iii) a component

that measures the interaction between the first two effects. Considering the average effects of these components, we have (e.g. Jann 2008):

$$\bar{\mathbf{z}}'(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0) + \left(\bar{\delta}_1 - \bar{\delta}_0\right)' = (\bar{\mathbf{z}}_1 - \bar{\mathbf{z}}_0)'\boldsymbol{\alpha}_1 + \bar{\mathbf{z}}_0'(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0) + (\bar{\mathbf{z}}_1 - \bar{\mathbf{z}}_0)'(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0) \quad (10)$$

Where, $\bar{\mathbf{z}}$, $\bar{\mathbf{z}}_1$ and $\bar{\mathbf{z}}_0$ are the mean values of the covariates of the full sample, and of the treated and untreated groups, respectively. The term $\bar{\mathbf{z}}_0'(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0)$ measures the "shift" effect on the outcome as a consequence of choice of the regime of treatment ($T = 1$), assuming the values of the covariates, $\mathbf{z}_i$, fixed at $T = 0$. The term $(\bar{\mathbf{z}}_1 - \bar{\mathbf{z}}_0)'\boldsymbol{\alpha}_1$ measures the extent to which the differential in outcome between treated and untreated is due to difference in covariates (the so-called "endowment" effect). Assuming that the difference in unobserved factors, $\delta_{1i} - \delta_{0i}$, accounts for the fact that differences in endowments and coefficients exist simultaneously between the two regimes, we replace the mean difference of $\delta_{1i} - \delta_{0i}$, with the "Interaction Term" $(\bar{\mathbf{z}}_1 - \bar{\mathbf{z}}_0)'(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0)$, such as in the "three-fold" version of the Blinder-Oaxaca decomposition (see Jann2008, among others).
Extending the Blinder-Oaxaca decomposition to the *i*-th observation, we replace $\mathbf{z}_i'(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0)$ in the Eq. (10) and obtain:

$$\nu_i = \frac{1}{\sigma}[\Delta_i - (\mathbf{z}_{1i} - \mathbf{z}_{0i})'\boldsymbol{\alpha}_1 - \mathbf{z}_{0i}'(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0) - (\mathbf{z}_{1i} - \mathbf{z}_{0i})'(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0) + \psi_i] \quad (11)$$

Then, substituting the Eq. (11) into the Eq. (1), the selection equation can be expressed as:

$$T_i^* = \mathbf{z}_i'\boldsymbol{\beta} + \frac{1}{\sigma}[\Delta_i - (\mathbf{z}_{1i} - \mathbf{z}_{0i})'\boldsymbol{\alpha}_1 - \mathbf{z}_{0i}'(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0) - (\mathbf{z}_{1i} - \mathbf{z}_{0i})'(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0) + \psi_i] \quad (12)$$

As a result, the estimation of the selection equation can be improved by introducing, as further explanatory variables, the terms, $\mathbf{z}_{0i}'(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0)$, $(\mathbf{z}_{1i} - \mathbf{z}_{0i})'\boldsymbol{\alpha}_1$ and $(\mathbf{z}_{1i} - \mathbf{z}_{0i})'(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0)$ preliminarily obtained performing an Oaxaca-Blinder decomposition in the three-fold version. The coefficients, $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_0$, can be estimated by running two Least Squares regressions on the equations (9a) and (9b), after replacing the latent dependent variables μ_{1i} and μ_{0i} with the observed y_{1i} and y_{0i}, respectively[4].

4. Monte Carlo experiment

We propose a Monte Carlo experiment to compare the performance of the *SSCM* procedure with that of the *PSME* in terms of bias reduction under both the conditions of heterogeneous and homogeneous covariates between regimes.
For this purpose, we generate 500 data sets of 2,000 units from the Two-Regime model above in Equations (1), (2a) and (2b). The exogenous covariates Z are generated in order to reproduce the very frequent condition of heterogeneity in observed covariates between

[4] As we will show below, the terms $\mathbf{z}_{0i}'(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0)$, $(\mathbf{z}_{1i} - \mathbf{z}_{0i})'\boldsymbol{\alpha}_1$ and $(\mathbf{z}_{1i} - \mathbf{z}_{0i})'(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0)$ can be included as further explanatory variables in the right side of the measurement equation (Eq.5), in order to better specify the endogenous effects of the choice of the regime on the causal effects.

treatment and comparison group, and the condition of homogeneity in the observed covariates between regimes.

More in detail, we consider a 5x1 vector $\mathbf{z}_i$ in equation (1), where the five variables are generated as follows:

- the variable z_{1i} is generated from a Uniform distribution ranging in [0,10] (U(0,10));
- z_{2i} from a Normal distribution with mean 10 and variance 16 ($N(10; 16)$):
- $z_{3i} = z_{1i}N(-5; 16)$;
- $z_{4i} = z_{1i}z_{2i} + N(0; 4)$;
- $z_{5i} = N(-10; 16) + z_{2i}^2$.

The coefficients included in the vector $\boldsymbol{\beta}$ are: $\beta_1 = 10$; $\beta_2 = -10$; $\beta_3 = 10$; $\beta_4 = 10$; $\beta_5 = -10$.

The choice of the distributions and their parameters is well subjective, but it has the purpose of including in the explanatory variables both independent and dependent (also quasi-collinear) variables.

In order to simulate the effect of endogeneity on the estimates, we consider two different DGP, with and without endogeneity, so as to fix two distinct sets of population parameters under the condition of endogeneity and exogeneity, respectively.

In order to embed endogeneity in the selection equation (Eq.1) and in the outcome equations (2a) and (2b), we generate the random variable $\xi_i = N(0; 1) + (u_{1i} - u_{0i})$, where u_{1i} and u_{0i} are the error terms of the outcome equations (2a) and (2b), generated, respectively, as follows:

$$u_{1i} = \sigma_{1\nu}\vartheta_i + \varepsilon_{1i} \text{ and } \quad u_{0i} = \sigma_{0\nu}\vartheta_i + \varepsilon_{0i} \tag{14}$$

where ϑ_i is a $N(0;1)$ random variable, and ε_{1i} and ε_{0i} are independently generated by a $N(0;36)$ and a $N(0;16)$, respectively. Finally, we standardize ξ_i and obtain the disturbance term of the selection equation, ν_i, following a $N(0;1)$ distribution, as stated above specifying the *selection equation* (Eq.1).

Note that we decide to fix different values of $\sigma_{1\nu}$ and $\sigma_{0\nu}$ in each experiment (reported, below, in Tables 1 and 2), in order to reproduce different conditions of endogeneity. In particular, if $\sigma_{1\nu}$and $\sigma_{0\nu}$ are both positive, we obtain an unobserved heterogeneity that positively influences both the propensity to undergo the treatment and the ability of the subject to gain the outcome. The opposite occurs if one of these covariances has a negative value[5].

To reproduce the probability to undergo treatment, we generate a cumulative Normal Standard distribution, as in a *Probit* function, given by $\Phi(\beta_0 + \beta_1 z_{1i} + \ldots + \beta_k z_{ki} + \nu_i$. The response variable T_i of the *selection equation* (1) is found to be zero if the values randomly assumed by the Gaussian cdf $\Phi(...)$ are less than 0.5, while T_i is equal to 1, if $\Phi(...) \geq 0.5$. The simulated response variable T_i determines, in this experiment, the assignment of the

[5] For example, considering a two-regime model of wage for unionized and non-unionized workers, latent cultural factors may induce a worker who gains a higher wage not to join the union, even though an unionized worker should have greater economic protection.

units to the regime of treated (outcome equation (Eq. 2a)) or untreated (outcome equation (Eq. 2b)).
Regarding to the outcome equations, we include in the right-hand sides of the outcome equations (Eq. 2a and Eq. 2b) two components, μ_{1i} and μ_{0i}, exogenously generated by a *N*(15; 25) and a *N*(10; 16) random variables, respectively. Endogeneity in outcome equations (Eq. 2a and Eq. 2b) is given by the relationship between the disturbance terms u_{1i} and u_{0i} and the error term of the selection equation, v_i, as above specified.
The "Population" treatment parameter considered in our analysis is given by the Average Treatment on Treated (*ATT*): $E(y_{1i} - y_{0i} | T_i) = 1$. Applying Eq. (5), Population *ATT* is equal to $E[\mu_{1i} - \mu_{0i} + (\sigma_{1v} - \sigma_{0v})v_i + (\varepsilon_{1i} - \varepsilon_{0i}) | T_i] = 1$, and converges on different limits depending on the pre-determined values of the covariances σ_{1v} and σ_{0v}. As a consequence, setting both σ_{1v} and σ_{0v} equal to zero, we obtain the Population *ATT* in absence of endogeneity.
Table 1 shows the Population *ATT* parameters generated under different values determined for σ_{1v} and σ_{0v} in order to reproduce endogeneity.

Table1. Population *ATT* parameters derived from *DGP*

DGP	$\sigma_{1v}(\rho_{1v})^*$	$\sigma_{0v}(\rho_{0v})^*$	Population *ATT*	Population *ATT* with observed heterogeneity
Endogeneity (1)	5.4(0.9)	2.4(0.6)	6.12	6.15
Endogeneity (2)	5.4(0.9)	-2.4(-0.6)	9.80	9.80
Endogeneity (3)	5.4(0.9)	0.8(0.2)	8.47	8.47
Endogeneity (4)	5.4(0.9)	-0.8(-0.2)	8.47	8.47
No Endogeneity	0.0 (0.0)	0.0 (0.0)	5.00	5.00

Note: * Taking into account the variances of u_1 and u_0, the corresponding correlation coefficients are approximately given by the values shown in brackets.

Moreover, in order to consider heterogeneity in covariates, we replicate the experiment with some different specifications in the covariates $\mathbf{Z}$ with respect to the previous scheme. For $T_i = 1$, the variable z_{1i} is generated from a *U*(0;13) and the variable z_{2i} from a *N*(13,4), while the variables z_{3i}, z_{4i} and z_{5i} change according to the previous scheme. As a consequence of heterogeneity in covariates, negligible changes in the values of the population *ATT* occur, as reported in the last column of Table 1.
The differences between the estimated *ATT* obtained under endogeneity and the "unbiased" population *ATT* value (equal to 5) quantify the effect of endogeneity simulated by the *DGP*. Hence we can evaluate the bias of the estimated *ATT* parameters, obtained by applying matching methods in different conditions of endogeneity.

4.1 Comparison of Matching Estimators
The proposed *SSMC* estimator is compared to the *PSME* on the simulated data.
For the sake of clarification, we provide a brief description of the steps needed to apply the estimation methods.

i) *SSCM*:

Step 1: let us start to run the Blinder-Oaxaca decomposition of the outcome, using the treatment dummy T_i in order to indicate the choice of regime. The outcomes, y_{1i} and y_{0i}, are pooled to form the dependent variable, as well as the covariates of the selection, in order to generate the five explanatory variables. The decomposition is replicated for the dummy 1- T_i (equal to 0 for treated, and equal to 1 for untreated).

Step 2: two new variables, named "*Split*" and "*Endowment*", are generated. The variable *Split* measures the "shift" effect on the outcome as a consequence of the choice of the regime given by $\mathbf{z}'_{0i}(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0)$, while the variable *Endowment* measures the effect of difference in covariates between regimes, given by $(\mathbf{z}_{1i} - \mathbf{z}_{0i})'\boldsymbol{\alpha}_1$ (see, above, Section 3.1).

Step 3: the propensity score matching is performed using the estimated propensity scores in order to obtain a preliminary estimation of the causal effects $\tilde{\Delta}_i = y_{1i} - y_{0c}$ (where y_{0c} is the counterfactual of y_{1i} belonging to the comparison group and characterized by the same propensity score of the *i*-th unit). The propensity to undergo the treatment is assumed to be conditional to the five above generated exogenous covariates and the variable *Split*.

The estimated propensity scores are then sorted in ascending order, and the estimated causal effects are indexed and ordered accordingly to the estimated propensity scores.

Step 4: A Maximum Likelihood estimation of the State-Space model is performed adopting $\tilde{\Delta}_i = y_{1i} - y_{ci}$ as a dependent variable and the variables *Split* and *Endowment* as covariates in the measurement equation; while the transition equation is specified as in equation (7).[6] The predicted values, $\widehat{\Delta}_i$, are the estimates of endogenous components of the causal effects, $\sigma\nu_i$ (cf., above, Sect. 3).

Step 5: subtracting $\widehat{\Delta}_i$, from y_{1i}, we obtain the corrected outcomes $\hat{y}_i$. A matching procedure is then replicated to link the corrected outcomes $\hat{y}_i$ with the respective counterfactuals in order to obtain the causal effects and the treatment parameters.

ii) *PSME*.

A matching procedure is performed, using the propensity scores estimated by *Probit*[7]. The estimation of the causal effects is given by $\widehat{\Delta}_i = y_{1i} - y_{0c}$ (where y_{0c} indicates a counterfactual of y_{1i}). The estimated propensity to undergo the treatment is assumed to be conditional to the five above generated exogenous covariates. The estimated causal effects allow us to compute the treatment parameters.

4.2 Comparing Simulation Results of Matching Procedures

We summarize in Tables 2 and Table 3 the estimated *ATT* values obtained by embedding different endogeneity conditions into the *DGP*. Note that, computing the bias with respect to the population *ATT* value (set to 5), the *SSCM* estimator performs better than the *PSME* procedure. The bias resulting from the application of *SSCM* is markedly smaller than of the

[6] The different step of the estimator are performed using STATA 14 packages.

[7] The STATA 14 package used to perform matching is PSMATCH2 (Leuven and Sianesi 2003). Performing PSMATCH2, a "one to one" linkage without replacement with a caliper equal to 0.05 is imposed. In addition the "Common Support" condition is ensured.

one resulting from *PSME*. We can observe, in particular, that, if we reproduce the "more common" endogeneity conditions (characterized by covariances, σ_{1v} and σ_{0v}, with the same sign) in the *DGP*, the confidence intervals obtained by the *SSCM* estimates include the population *ATT* value. In the less frequent case, in which the propensity to undergo treatment is endogenously affected in the two regimes with opposite sign, confidence intervals of the *SSCM* estimates do not include the population parameter. However the percentage of bias of *SSCM* estimation does not exceed 15% in absolute value.
While the mean of the estimated *ATT* using the *SSCM* procedure is not influenced by the presence of heterogeneity in covariates between the regimes, standard errors and confidence intervals are found to be markedly increased with respect to the case of no heterogeneity in covariates.

Table 2. Estimated ATT parameters without heterogeneity in observed covariates. Population *ATT* value = 5

DGP	*SSCM*			*PSME*		
Endogeneity	***ATT***	*95% CI*		***ATT***	*95% CI*	
σ_{1v} 5.4; σ_{0v} 2.4	4.974	*4.93*	*5.018*	7.996	*7.97*	*8.022*
σ_{1v} 5.4; σ_{0v} -2.4	4.32	*4.279*	*4.362*	6.814	*6.782*	*6.846*
σ_{1v} 5.4; σ_{0v} 0.8	4.983	*4.939*	*5.027*	7.571	*7.534*	*7.607*
σ_{1v} 5.4; σ_{0v} -0.8	4.729	*4.688*	*4.77*	7.572	*7.536*	*7.608*
	% BIAS*	*St.Dev.*	*t***	***% BIAS****	*St.Dev.*	*t***
σ_{1v} 5.4; σ_{0v} 2.4	-0.51%	*0.022*	222.16	59.92%	*0.013*	*607.170*
σ_{1v} 5.4; σ_{0v} -2.4	-13.59%	*0.021*	205.72	36.28%	*0.016*	*414.280*
σ_{1v} 5.4; σ_{0v} 0.8	-0.35%	*0.022*	222.48	51.41%	*0.018*	*410.660*
σ_{1v} 5.4; σ_{0v} -0.8	-5.42%	*0.021*	225.71	51.45%	*0.018*	*412.000*

Note: * = % of *Bias* [(Est. *ATT*-5)/5]%; ** = t-ratio: *ATT*/*St.Dev.*)

Table 3. Estimated *ATT* parameters with heterogeneity in observed covariates. Population *ATT* value = 5

DGP	*SSCM*			*PSME*		
Endogeneity	***ATT***	*95% CI*		***ATT***	*95% CI*	
σ_{1v} 5.4; $\boldsymbol{?}_{0v}$ 2.4	5.059	*3.482*	*6.527*	7.941	6.551	9.599
σ_{1v} 5.4; $\boldsymbol{?}_{0v}$ -2.4	4.246	*2.879*	*5.681*	6.888	5.548	8.339
σ_{1v} 5.4; $\boldsymbol{?}_{0v}$ 0.8	5.128	*3.356*	*6.722*	7.982	6.566	9.500
σ_{1v} 5.4; $\boldsymbol{?}_{0v}$ -0.8	4.757	*3.433*	*6.083*	7.599	5.993	9.045
	% BIAS*	*St.Dev.*	*t***	***% BIAS****	*St.Dev.*	*t***
σ_{1v} 5.4; σ_{0v} 2.4	1.19%	*0.529*	9.56	58.82%	*0.498*	15.94
σ_{1v} 5.4; σ_{0v} -2.4	-15.07%	*0.512*	8.29	37.76%	*0.491*	14.03
σ_{1v} 5.4; σ_{0v} 0.8	2.53%	*0.517*	9.93	59.63%	*0.479*	16.66
σ_{1v} 5.4; σ_{0v} -0.8	-4.86%	*0.523*	9.09	51.99%	*0.484*	15.71

Note: * = % of *Bias* [*ATT*-5)/5]%; ** = *t*-ratio: [*ATT*/*St.Dev.*)

Tables 2 and Table 3 report the statistics on *ATT* and *Bias* only, while in the Appendix, simulation results are reported in detail, including the statistics measuring balancing in covariates between treated and untreated cases after matching. These statistics allow us to evaluate the extent to which the adopted matching procedure reduces differences in covariates between treated and untreated units (cf. Rubin 2001; Haviland et al. 2007). In addition, we also report the mean of estimated coefficients measuring dependence between causal effects and propensity score values, and the mean of the coefficients measuring the autoregressive component of the causal effects, resulting by the estimation of the Transition Equation of the *state space* model (see, above, Eq. 7). In the following table (Table 4), we summarize the description of the coefficients and indicators provided by the Monte Carlo experiments.

Table 4. Description of the coefficients and indicators provided by the Monte Carlo experiments

RHO(T; Δ)	Correlation coefficient between estimated propensity score and causal effects
Estimated ATT_{SSCM} (ATT_{PSME})	Estimated ATT parameter obtained performing the matching Estimator
Unbiased Pop. *ATT*	ATT generated by DGP without endogeneity
Biased Pop. *ATT*	ATT generated by DGP under endogeneity
Estim. Transition coeff	Coefficient of the Transition equation measuring the autoregressive effect in ordered causal effects
Shift coeff.	Coefficient estimated in the "measurement equation", corresponding to the explanatory variable "Split" provided by the Blinder-Oaxaca decomposition
Endowment coeff.	Coefficient estimated in the "measurement equation", with the introduction as regressor of the difference in covariates provided by the Blinder-Oaxaca decomposition
Mean bias after matching	Standardized mean difference between treatment and control units after matching
Median bias after matching	Median difference between treatment and control units after matching
BAFT	"Rubins' B" indicator: The absolute standardized difference of the means of the linear index of the propensity score in the treated and non-treated (matched).
RAFT	"Rubin's R" index: The ratio of variances of the propensity score index between treated and non-treated (matched).
No. of units on the common support	No. of observations belonging to the treatment group or to the comparison group who have an estimated propensity score equal to that of one or more observations belonging to the opposite group.

Analyzing the statistics, reported below in Appendix, obtained by the Monte Carlo experiments, we observe that the *RHO*(*T*; Δ) coefficient, measuring, in the *SSCM* procedure, the correlation between the estimated propensity score and the causal effects, is generally higher if heterogeneity in covariates has been embedded in *DGP*. We found that, with heterogeneous covariates, *RHO*(*T*; Δ) ranges between 14% and 15%. Instead *RHO*(*T*; Δ) ranges between 2% and 5% if covariates are not affected by heterogeneity. In addition, we generally observe significant estimated values of the coefficient of the transition equation, ρ, measuring the autoregressive effect in ordered causal effects. Estimated values are close to -0.44 in all the Monte Carlo experiments.
The impact of the endogenous change of regime and of the difference in covariates across regimes in the *measurement equation* (Eq. 7), are measured by the "Shift" and the "Endowment" coefficient, respectively. Note that the Endowment coefficient is generally higher (ranging between 0.73 and 0.76) if covariates are not affected by heterogeneity. It ranges between 0.52 and 0.56 in the case of heterogeneous covariates across regimes. On the contrary, the "Shift" coefficient is higher if heterogeneity in covariates occurs (between 1.39 and 1.59) than the opposite case of homogeneous covariates (between -0.06 and 0.15).
The *BAFT* and *RAFT* indicators provide, in all the experiments, values compatible with a satisfactory balancing using both *SSCM* and *PSME* estimators[8]. In general, a better balancing is obtained setting the experiment with homogeneous covariates between regimes.

5. Application using experimental vs. nonexperimental data sets on training programs

In this section we report the results of the comparison of our *SSCM* method and the traditional Propensity Score Matching (*PSME*) both applied to the experimental treated units from the US National Support Work (*NSW*) paired with the nonexperimental untreated comparison units drawn from the Current Population Survey (*CPS*) and the Population Survey of Income Dynamic (*PSID*), respectively.
The data we use, obtained from LaLonde (1986), are from the *NSW* labor market experiment in which participants were randomized between a treatment group (on-the-job training lasting between nine months and a year) and control groups. The outcomes are given by the annual earnings gained after the experiment by the participants belonging to the two groups.
Following LaLonde (1986) and Dehejia and Wahba (1999, 2002), we use the experimental *NSW* data paired with the untreated units drawn from the *PSID* and the *CPS* dataset, respectively. In this way, we compare the difference in the propensity score distribution between treated units and nonexperimental counterfactuals applying, respectively, our *SSCM* method and the traditional *PSME*. Balancing score statistics are also adopted in order to evaluate the effectiveness of each estimator in reducing bias in covariates before and after matching.
In Table 5 we report the sample characteristics of the treatment group and the two comparison groups. Marked differences characterize the comparison between *NSW* and *CPS*

[8] Rubin (2001) recommends that *BAFT* be less than 25 and that *RAFT* be ranged between 0.5 and 2 for the samples to be considered sufficiently balanced.

data and between *NSW* and *PSID* data. In particular, the age of the subject and the real earnings before treatment and proportion of no degree differ between treatment and comparison groups. Note that these differences may influence the decision of the participants to undergo the training program.
Table 6 shows the estimated *ATT* obtained comparing experimental treated units with experimental and non-experimental control units, respectively. Matching results are obtained using both *PSME* and *SSCM* applying a different caliper (0.05, 0.01 and 0.005) without replacement and under three different matching criteria (number of matches equal to one, four and eight, respectively).

Table 5. Description, sample means and Standard Errors of covariates for *NSW*, *CPS* and *PSID* participants

	NSW - treated		*PSID* – untreated		*CPS* – untreated	
Variable Description	Mean	*Std. Err.*	Mean	*Std. Err.*	Mean	*Std. Err.*
Age	25.82	*0.526*	34.85	*0.209*	33.23	*0.087*
Years of schooling	10.35	*0.148*	12.12	*0.062*	12.03	*0.023*
Proportion of Black	0.84	*0.027*	0.25	*0.009*	0.07	*0.002*
Proportion of Hispanic	0.06	*0.017*	0.03	*0.004*	0.07	*0.002*
Proportion of nodegree	0.71	*0.034*	0.31	*0.009*	0.30	*0.004*
Proportion of Married	0.19	*0.029*	0.87	*0.007*	0.71	*0.004*
Real Earnings in 1974	2096	*359*	19429	*269*	14017	*76*
Table 5 - continued						
Real Earnings in 1975	1532	*237*	19063	*272*	13651	*73*
Real Earnings in 1978	6349	*578*	21554	*312*	14847	*76*
Proportion of individuals Black and Unemployed before training	0.60	*0.0361*	0.01	0.002	0.01	*0.001*
Sample size	185		2490		15992	

Table 6. *ATT* estimation. Comparison between *NSW* treated units and *CPS* and *PSID* untreated units, respectively

NSW treated and untreated (experimental)						
	No of matches=1		No of matches=4		No of matches=8	
Caliper	*SSCM*	*PSME*	*SSCM*	*PSME*	*SSCM*	*PSME*
0.005	2387	1964	2259	2093	2147	2241
0.01	1524	1939	1853	2051	2122	2139
0.05	1453	1590	1972	2059	2039	2313
NSW treated and CPS untreated						
	No of matches=1		No of matches=4		No of matches=8	
Caliper	*SSCM*	*PSME*	*SSCM*	*PSME*	*SSCM*	*PSME*
0.005	1255	855	1015	1594	983	1633
0.01	2065	1027	1656	1611	1703	1626
0.05	1410	1587	1189	1567	1422	1609
NSW treated and PSID untreated						
	No of matches=1		No of matches=4		No of matches=8	
Caliper	*SSCM*	*PSME*	*SSCM*	*PSME*	*SSCM*	*PSME*
0.005	2558	400	-106	515	70	319
0.01	161	-96	-305	658	-530	470
0.05	-1925	-214	1671	2517	1879	2337

Table 7. Mean bias after matching. Comparison between NSW treated units and CPS and PSID untreated units, respectively

NSW treated and CPS untreated						
	No of matches=1		No of matches=4		No of matches=8	
Caliper	*SSCM*	*PSME*	*SSCM*	*PSME*	*SSCM*	*PSME*
0.005	7.52	37.17	7.20	7.48	6.01	7.06
0.01	8.09	41.69	6.32	7.77	5.84	6.73
0.05	10.54	35.17	6.91	7.83	6.59	6.77
NSW treated and PSID untreated						
	No of matches=1		No of matches=4		No of matches=8	
Caliper	*SSCM*	*PSME*	*SSCM*	*PSME*	*SSCM*	*PSME*
0.005	13.76	27.32	6.69	4.54	6.64	4.02
0.01	11.51	26.14	6.21	6.82	7.44	6.24
0.05	10.48	23.27	10.01	7.22	8.97	6.25

Table 8. Rubin's *RAFT* index: The ratio of variances of the propensity score between treated and non-treated after matching. Comparison between *NSW* treated units and *CPS* and *PSID* untreated units, respectively

NSW treated and CPS untreated						
	No of matches=1		No of matches=4		No of matches=8	
Caliper	*SSCM*	*PSME*	*SSCM*	*PSME*	*SSCM*	*PSME*
0.005	0.74	2.71	0.80	0.80	0.66	0.772
0.01	1.04	8.46	0.86	0.80	0.79	0.767
0.05	0.75	9.68	0.81	0.79	0.79	0.770
NSW treated and PSID untreated						
	No of matches=1		No of matches=4		No of matches=8	
Caliper	*SSCM*	*PSME*	*SSCM*	*PSME*	*SSCM*	*PSME*
0.005	1.00	2.82	1.17	0.98	0.91	1.01
0.01	0.79	1.91	0.90	0.95	1.06	0.98
0.05	0.50	2.55	1.07	0.76	1.17	0.76

Comparing balancing statistics between experimental treated units and non-experimental untreated units, we can observe how the application of the *SSCM* method leads to improving bias reduction with respect to the use of the *PSME* procedure, especially if the criterion "one-to-one" (number of matches equal to 1) without replacement is applied. In particular, the results reported in Tables 7 and 8 show how the standardized mean difference in covariates and the ratio between the variances of the propensity score between treated and untreated after matching are markedly smaller when the one-to-one criterion is adopted, whatever the caliper level.

We compute also the differences in estimated propensity scores after matching between treated and untreated units. These results are reported below in the graphs of Figures 1, 2, 3 and 4, where the treated units are ordered on the basis of the propensity scores in ascending order (horizontal axis), while on the vertical axis, the estimated propensity score of the treated is reported jointly with the estimated propensity score of the corresponding counterfactual. Figures 1 and 2 report the propensity scores linked using the one-to-one criterion without replacement, while in Figures 3 and 4 the treated and untreated units are linked adopting replacement.

Explaining in Figure 1 the results obtained applying *SSCM* and *PSME*, respectively, we found that the *SSCM* estimated propensity scores of both treated and untreated are paired at a higher level of propensity scores than using *PSME*. This implies that, performing *SSCM*, matching results are supported by a higher conditional probability to undergo the treatment than the traditional method. We can explain this result because, unlike traditional matching methods, the *SSCM* procedure, introducing a correction term for endogenous latent factors, increases the components explaining propensity score with respect to the traditional *PSME*. In addition, as showed by both Figures 1 and 2, propensity scores of treated and untreated are closer when performing *SSCM*.

On the contrary, if we replicate the matching estimation by introducing replacement, no significant differences are found between the *SSCM* and *PSME* estimated propensity score (Figures 3 and 4). Other comparisons not reported here for the sake of brevity, show how

SSCM and *PSME* generally lead to similar results in propensity score estimation if the number of matches increases (for example, one-to-four or one-to-eight). In addition, we observe that, by increasing the number of matches, no marked differences in propensity score estimates occur by adopting replacement or not.

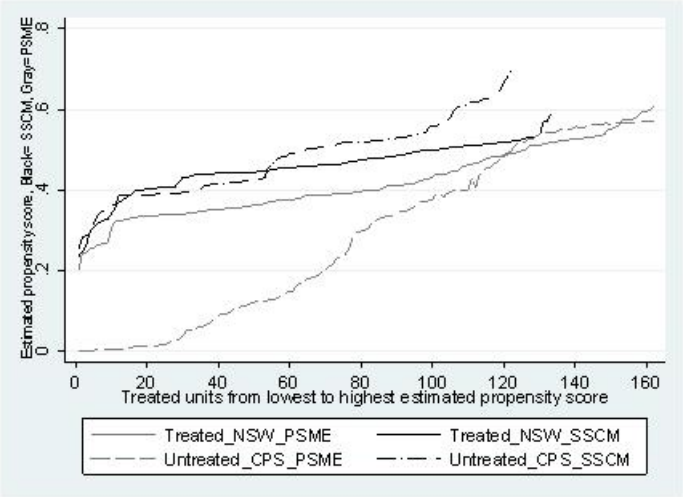

Figure 1. *SSCM* and *PSME* propensity scores comparison. No of matches=1 without replacement. Caliper: 0.05. *NSW-CPS*

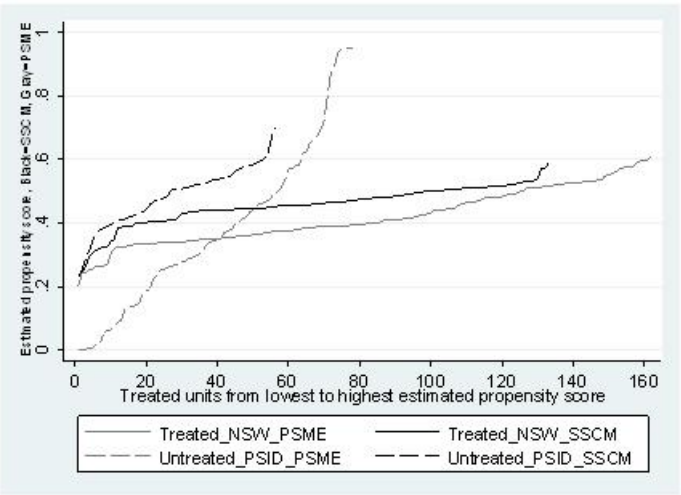

Figure 2. *SSCM* and *PSME* Propensity scores comparison. No of matches =1 without replacement. Caliper: 0.05. *NSW-PSID*

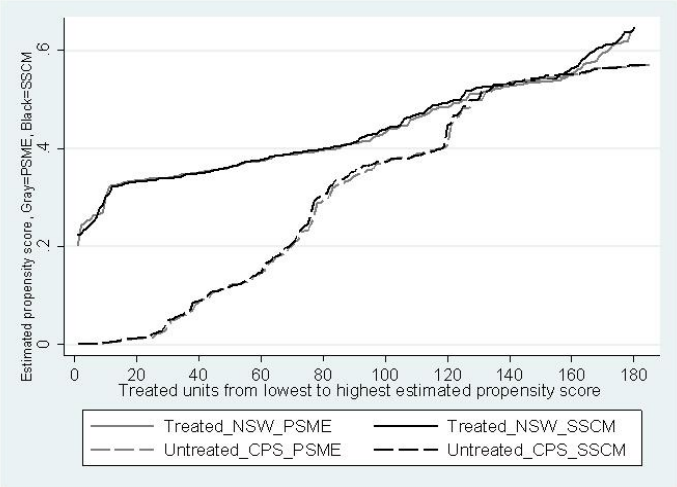

Figure 3. – *SSCM* and *PSME* Propensity scores comparison. No of matches =1 with replacement. Caliper: 0.05. *NSW-CPS*

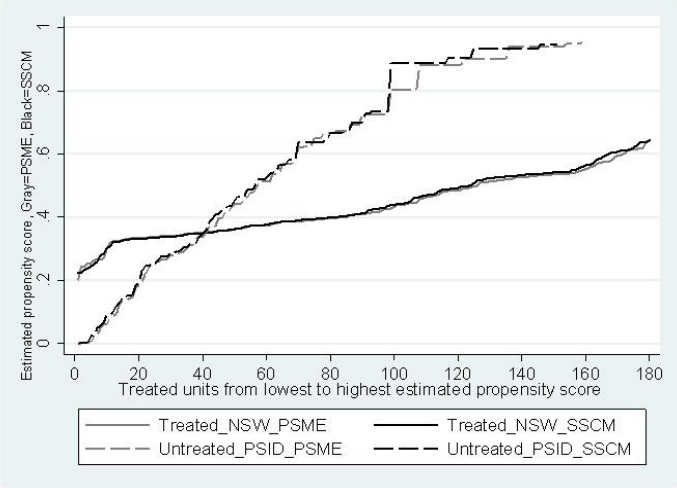

Figure 4. *SSCM* and *PSME* Propensity scores comparison. No of matches =1 with replacement. Caliper: 0.05. *NSW-PSID*

6. Concluding remarks

The aim of this study is to improve the propensity-score matching approach so that estimation results do not overly suffer from the influence of the endogeneity of treatment. In doing this, the main innovation introduced here is given by the specification of the endogenous relationship between the individual propensity score (individual probability to undergo the treatment) and the individual causal effect of treatment. In practice, we start by

assuming that the probability that a subject undergoes treatment endogenously depends, at least in part, on the potential effect of the treatment. This implies that two subjects with the same (or similar) propensity score should expect similar results in terms of causal effects. This allows us to consider the causal effect of treatment on each subject such as correlated with the causal effect of another subject with similar propensity score.
A consequence of this assumption is that the causal effects, ordered by their correspondent propensity scores, are autocorrelated via their endogenous component. As an empirical verification of this assumption, we apply a state-space model to estimate the autocorrelated endogenous component of the causal effects, so as to use the result of this estimate as a correction term. In particular, the results of the Monte Carlo experiments here reported confirm that, simulating endogeneity of the selection into treatment in a Two-Regime model, the predicted components of causal effects can be successfully used, at a second stage of the estimation procedure, to correct the matches outcomes.
As the results of our empirical analysis show, this method allows us to reduce the selectivity bias in matching without imposing, to the data or the model, any restriction usually adopted to reproduce a condition comparable to randomization.
At this stage of our research, we have deepened the characteristics of the *SSCM* estimator only by means of Monte Carlo experiments and empirical tests. However the inferential properties must still be investigated. This will be the next aim of this research.
Considering our study at an early stage, the field of application currently investigated is limited to the comparison of propensity-score based methods between treated and untreated units in a cross-sectional context. For this reason, we assume as a basic model a Two-Regime model with endogenous treatment in which the outcome value of each unit can be observed only under one of the two regimes (treatment or control). Possible extensions could deal with a broader evaluation of the application of the *SSCM* method compared to other cross-sectional matching methods not based on the propensity score estimation such as Nearest Neighbor, and Bias Corrected matching (cf. Abadie and Imbens 2011). In addition, we plan to apply the *SSCM* procedure also in a longitudinal or panel context so as to compare the results with those obtained from the application of the Difference-in Differences Propensity Score matching.

References

Aakvik, A., James J. H., and Vytlacil E. J. (2005). "Estimating Treatment Effects for Discrete Outcomes When Responses To Treatment Vary: an Application to Norwegian Vocational Rehabilitation Programs", *Journal of Econometrics,* 125(1-2), 15-51.

Abadie, A., and Imbens, G. (2011), "Bias-Corrected Matching Estimators for Average Treatment Effects", *Journal of Business & Economic Statistics*, 29(1), 1-11.

Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996), "Identification of Causal Effects Using Instrumental Variables", *Journal of the American statistical Association*, *91*(434), 444-455.

Blinder, A. S. (1973), "Wage Discrimination: Reduced Form and Structural Estimates", *Journal of Human resources*, 436-455.

Carneiro, P., Hansen, K. T., and Heckman, J. J. (2003), "Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice", *International Economic Review*, 44 (2), 361-422.

Carneiro, P., Heckman, J. J., and Vytlacil, E. J. (2011), "Estimating Marginal Returns to Education", *American Economic Review*, 101, 2754–2781.

Dehejia, R. H., and Wahba S. (1999), Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs", *Journal of the American Statistical Association*, 94(448), 1053-1062.

Dehejia, R. H., and Wahba S. (2002), "Propensity Score-Matching Methods for Nonexperimental Causal Studies", *The Review of Economics and Statistics*, 84(1), 151-161.

Hahn, J., Todd, P., and Van der Klaauw, W. (2001), "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design", *Econometrica*, 69(1), 201-209.

Harvey, A. C. (1990), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge: Cambridge University Press.

Haviland, A., Nagin, D. S., and Rosenbaum, P. R. (2007), "Combining Propensity Score Matching and Group-based Trajectory Analysis in an Observational Study", *Psychological methods*, *12*(3), 247-267.

Heckman, J. J. (2008), "The Principles Underlying Evaluation Estimators with an Application to Matching", *Annales d'Economie et de Statistique*, (91/92), 9-73.

Heckman, J. J., and Honore, B. E. (1990), "The Empirical Content of the Roy Model", *Econometrica, 58*, 1121-1149

Heckman, J. J., Ichimura, H., and Todd, P. (1997), "Matching As an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme", *Review of Economic Studies,* 64 (4), 605–654.

Heckman, J. J., Ichimura, H., Smith, J., and Todd, P. (1998), "Characterizing Selection Bias Using Experimental Data", *Econometrica,* 66 (5), 1017–1098.

Heckman, J. J., and Navarro-Lozano, S. (2004), "Using Matching, Instrumental Variables, and Control Functions to Estimate Economic Choice Models", *The Review of Economics and Statistics*, 86(1), 30-57.

Imbens, G., and Angrist, J. D. (1994), "Identification and Estimation of Local Average Treatment Effects", *Econometrica* 62(2), 467–475.

Jann B. (2008), "The Blinder–Oaxaca Decomposition for Linear Regression Models", *The Stata Journal*, 8(4), 453-479.

LaLonde, R. (1986), "Evaluating the Econometric Evaluations of Training Programs", *American Economic Review,* 76 (4), 604-620.

Leuven, E., and Sianesi, B. (2003), "PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing". Available at http://ideas.repec.org/c/boc/bocode/s432001.html.

Oaxaca, R. (1973), "Male–Female Wage Differentials Iin Urban Labor Markets", *International Economic Review*, 14, 693–709.

Rosenbaum, P., and Rubin, D. B. (1983), "The Central Role of the Propensity Score In Observational Studies For Causal Effects", *Biometrika*, 70(1), 41-55.

Rubin, D. B. (2001), "Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation", *Health Services and Outcomes Research Methodology*, 2(3–4), 169–188.

Smith, J. A., and Todd P. E. (2005), "Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators?", *Journal of Econometrics*, 125, 305–353.

Winship, C., and Morgan, S. L. (1999), "The Estimation of Causal Effects from Observational Data", *Annual Review of Sociology*, 25, 659-706.

APPENDIX

Reducing Bias in a Matching Estimation of Endogenous Treatment Effect.

Results of the Monte Carlo Experiments

Simulated data without heterogeneity in observed covariates

Table A1. No heterogeneity in covariates. No. of Reps.=500. Simulated endogeneity: σ_{1v} =5.4; σ_{0v} = 2.4

		SSCM		
	Mean	*Std. Dev.*	Confidence Interval 95%	
RHO(T; Δ)	0.0266	*0.0012*	0.0243	0.0289
Estimated ATT_{SSCM}	**4.9744**	***0.0224***	**4.9304**	**5.0184**
Unbiased Pop. ATT	***4.9903***	*0.0091*	*4.9725*	*5.0081*
Biased Pop. *ATT*	6.1230	*0.0137*	6.0960	6.1499
Estim. Transition coeff.	-0.4426	*0.0009*	-0.4443	-0.4409
Shift coeff.	-0.0061	*0.0185*	-0.0425	0.0304
"Endownment" coeff.	0.7529	*0.0018*	0.7492	0.7565
Mean bias after matching	2.6819	*0.0583*	2.5673	2.7965
Median bias after matching	2.5889	*0.0681*	2.4550	2.7228
BAFT	8.7095	*0.1517*	8.4114	9.0076
RAFT	1.0047	*0.0073*	0.9904	1.0191
No. of units on common support	1882			
		PSME		
	Mean	*Std. Dev.*	Confidence Interval 95%	
RHO(T; Δ)	0.0602	*0.0007*	0.0588	0.0615
Estimated ATT_{PSME}	**7.9958**	***0.0132***	**7.9700**	**8.0217**
Unbiased Pop. ATT	***4.9988***	*0.0064*	4.9862	5.0113
Biased Pop. *ATT*	6.1460	*0.0098*	6.1267	6.1653
Mean bias after matching	2.0514	*0.0347*	1.9833	2.1194
Median bias after matching				
BAFT	7.4477	*0.0873*	7.2763	7.6191
RAFT	1.0288	*0.0048*	1.0193	1.0383
No. of units on common support	1995			

Table A2. No heterogeneity in covariates. No. of Reps.=500. Simulated endogeneity: σ_{1v} =5.4; σ_{0v} = 0.8

		SSCM		
	Mean	*Std. Dev.*	Confidence Interval 95%	
RHO(T; Δ)	0.0245	*0.0011*	0.0223	0.0267
Estimated ATT_{SSCM}	**4.9825**	***0.0224***	**4.9385**	**5.0265**
Unbiased Pop. ATT	***5.0073***	*0.0084*	4.9908	5.0238
Biased Pop. *ATT*	7.2388	*0.0140*	7.2114	7.2663
Estim. Transition coeff.	-0.4419	*0.0008*	-0.4436	-0.4403
Shift coeff.	0.0287	*0.0170*	-0.0047	0.0620
" Endownment " coeff.	0.7594	*0.0017*	0.7560	0.7628
Mean bias after matching	2.4950	*0.0565*	2.3841	2.6060
Median bias after matching	2.3444	*0.0642*	2.2182	2.4705
BAFT	8.4844	*0.1457*	8.1983	8.7705
RAFT	1.0012	*0.0068*	0.9879	1.0146
No. of units on common support	1891			
		PSME		
	Mean	*Std. Dev.*	Confidence Interval 95%	
RHO(T; Δ)	0.0502	*0.0009*	0.0483	0.0520
Estimated ATT_{PSME}	**7.5705**	***0.0184***	**7.5342**	**7.6067**
Unbiased Pop. ATT	***5.0070***	*0.0088*	*4.9898*	*5.0242*
Biased Pop. *ATT*	8.4698	*0.0154*	8.4395	8.5002
Mean bias after matching	2.1077	*0.0488*	2.0118	2.2036
Median bias after matching	1.9453	*0.0550*	1.8373	2.0533
BAFT	7.6179	*0.1249*	7.3725	7.8633
RAFT	1.0128	*0.0062*	1.0007	1.0249
No. of units on common support	1995			

Table A3. No heterogeneity in covariates. No. of Reps.=500. Simulated endogeneity: σ_{1v} =5.4; σ_{0v} = -2.4

		SSCM		
	Mean	*Std. Dev.*	Confidence Interval 95%	
RHO(T; D)	0.0208	*0.0011*	0.0187	0.0229
Estimated ATT_{SSCM}	**4.3204**	***0.0210***	**4.2791**	**4.3616**
Unbiased Pop. ATT	***4.9968***	*0.0083*	*4.9805*	*5.0132*
Biased Pop. *ATT*	9.7883	*0.0179*	9.7532	9.8234
Estim. Transition coeff.	-0.4427	*0.0008*	-0.4443	-0.4411
Shift coeff.	0.1542	*0.0176*	0.1196	0.1888
" Endownment " coeff.	0.7332	*0.0018*	0.7296	0.7367
Mean bias after matching	2.7085	*0.0691*	2.5728	2.8442
Median bias after matching	2.5320	*0.0701*	2.3942	2.6698
BAFT	9.0678	*0.1807*	8.7127	9.4229
RAFT	1.0058	*0.0072*	0.9916	1.0200
No. of units on common support	1901			
		PSME		
	Mean	*Std. Dev.*	Confidence Interval 95%	
RHO(T; D)	0.0441	*0.0008*	0.0425	0.0457
Estimated ATT_{PSME}	**6.8141**	***0.0164***	**6.7818**	**6.8464**
Unbiased Pop. ATT	***4.9971***	*0.0079*	4.9816	5.0126
Biased Pop. *ATT*	9.7787	*0.0145*	9.7502	9.8072
Mean bias after matching	2.0963	*0.0454*	2.0071	2.1856
Median bias after matching	1.9439	*0.0504*	1.8449	2.0429
BAFT	7.3877	*0.1087*	7.1742	7.6011
RAFT	1.0223	*0.0057*	1.0110	1.0336
No. of units on common support	1995			

Table A4. No heterogeneity in observed covariates. No. of Reps.=500. Simulated endogeneity: σ_{1v} =5.4; σ_{0v} = -0.8

			SSCM	
	Mean	*Std. Dev.*	Confidence Interval 95%	
RHO(T; Δ)	0.0225	*0.0012*	0.0202	0.0248
Estimated ATT_{SSCM}	**4.7288**	***0.0210***	**4.6876**	**4.7700**
Unbiased Pop. ATT	***5.0020***	*0.0093*	*4.9838*	*5.0203*
Biased Pop. *ATT*	8.4783	*0.0152*	8.4485	8.5081
Estim. Transition coeff.	-0.4419	*0.0008*	-0.4435	-0.4402
Shift coeff.	0.1168	*0.0164*	0.0846	0.1491
Endownment coeff.	0.7482	*0.0017*	0.7449	0.7516
Mean bias after matching	2.6499	*0.0601*	2.5317	2.7680
Median bias after matching	2.5703	*0.0677*	2.4374	2.7033
BAFT	8.7016	*0.1570*	8.3931	9.0100
RAFT	0.9952	*0.0066*	0.9822	1.0081
No. of units on common support	1895			
			PSME	
	Mean	*Std. Dev.*	Confidence Interval 95%	
RHO(T; Δ)	0.0502	*0.0009*	0.0484	0.0520
Estimated ATT_{PSME}	**7.5723**	***0.0184***	**7.5362**	**7.6084**
Unbiased Pop. ATT	***5.0074***	*0.0088*	*4.9902*	*5.0247*
Biased Pop. *ATT*	8.4712	*0.0154*	8.4409	8.5015
Mean bias after matching	2.1056	*0.0489*	2.0095	2.2016
Median bias after matching	1.9435	*0.0550*	1.8353	2.0516
BAFT	7.6145	*0.1251*	7.3687	7.8603
RAFT	1.0120	*0.0061*	1.0000	1.0241
No. of units on common support	1995			

Simulated data introducing heterogeneity in observed covariates

Table A5. Heterogeneity in observed covariates. No. of Reps.=500. Simulated endogeneity: σ_{1v} =5.4; σ_{0v} = 2.4

		SSCM		
	Mean	*Std. Dev.*	Confidence Interval 95%	
RHO(T; Δ)	0.1567	*0.0370*	0.0608	0.2684
Estimated ATT_{SSCM}	**5.0594**	***0.5291***	**3.4824**	**6.5270**
Unbiased Pop. ATT	***5.0105***	*0.2105*	4.3976	5.5818
Biased Pop. *ATT*	6.1694	*0.3100*	5.3431	7.0190
Estim. Transition coeff.	-0.4614	*0.0236*	-0.5444	-0.3922
Shift coeff.	1.4269	*0.3220*	0.6882	2.8712
Endownment coeff.	0.5249	*0.0387*	0.4261	0.6429
Mean bias after matching	4.9963	*1.7333*	0.8394	12.3551
Median bias after matching	4.8836	*2.3745*	0.5446	14.7685
BAFT	21.3282	*4.4243*	8.8634	35.3954
RAFT	1.1155	*0.2168*	0.4998	1.7896
No. of units on common support	1934			
		PSME		
	Mean	*Std. Dev.*	Confidence Interval 95%	
RHO(T; Δ)	0.2113	*0.0285*	0.1208	0.2990
Estimated ATT_{PSME}	**7.9409**	***0.4981***	**6.5505**	**9.5988**
Unbiased Pop. *ATT*	***5.0153***	*0.1974*	4.4988	5.6266
Biased Pop. *ATT*	6.1529	*0.3106*	5.1768	7.0055
Mean bias after matching	5.7745	*1.9577*	1.1594	13.2100
Median bias after matching	6.1974	*2.4247*	0.2901	14.9395
BAFT	22.1471	*4.1054*	5.7511	37.5376
RAFT	1.1680	*0.1905*	0.6644	1.7616
No. of units on common support	1968			

Table A6. Heterogeneity in observed covariates. No. of Reps.=500. Simulated endogeneity: σ_{1v} =5.4; σ_{0v} = 0.8

		SSCM		
	Mean	*Std. Dev.*	Confidence Interval 95%	
RHO(T; Δ)	0.1677	*0.0386*	0.0459	0.2816
Estimated ATT_{SSCM}	**5.1267**	***0.5165***	**3.3563**	**6.7218**
Unbiased Pop. ATT	***5.0107***	*0.2081*	4.3993	5.5940
Biased Pop. *ATT*	7.2626	*0.3269*	6.2737	8.0730
Estim. Transition coeff.	-0.4589	*0.0226*	-0.5227	-0.3942
Shift coeff.	1.3980	*0.3116*	0.7341	2.7256
Endownment coeff.	0.5202	*0.0382*	0.4229	0.6226
Mean bias after matching	4.8514	*1.6374*	1.3317	10.4524
Median bias after matching	4.7744	*2.4592*	0.5058	14.6848
BAFT	21.0182	*4.3332*	10.0042	36.3935
RAFT	1.1165	*0.2172*	0.4042	1.6645
No. of units on common support	1927			
		PSME		
	Mean	*Std. Dev.*	Confidence Interval 95%	
RHO(T; Δ)	0.2162	*0.0268*	0.1464	0.2988
Estimated ATT_{PSME}	**7.9816**	***0.4790***	**6.5655**	**9.4996**
Unbiased Pop. ATT	***5.0035***	*0.2056*	4.3844	5.5943
Biased Pop. *ATT*	7.2141	*0.3304*	6.2855	8.0881
Mean bias after matching	5.8488	*2.0170*	0.7382	13.0318
Median bias after matching	6.1902	*2.4461*	0.5725	14.2794
BAFT	21.6758	*4.1721*	10.4950	36.4430
RAFT	1.1946	*0.2042*	0.6493	1.7908
No. of units on common support	1969			

Table A7. Heterogeneity in observed covariates. No. of Reps.=500. Simulated endogeneity: σ_{1v} =5.4; σ_{0v} = -2.4

	SSCM			
	Mean	*Std. Dev.*	Confidence Interval 95%	
RHO(T; Δ)	0.1390	*0.0375*	-0.0020	0.2672
Estimated ATT_{SSCM}	**4.2464**	***0.5121***	**2.8788**	**5.6813**
Unbiased Pop. ATT	***5.0051***	*0.2042*	4.3985	5.6521
Biased Pop. *ATT*	9.8091	*0.3589*	8.7284	11.0152
Estim. Transition coeff.	-0.4572	*0.0240*	-0.5364	-0.3768
Shift coeff.	1.5859	*0.3853*	0.7016	3.3063
Endownment coeff.	0.5619	*0.0379*	0.4526	0.7021
Mean bias after matching	5.0343	*1.8335*	0.7780	12.2944
Median bias after matching	4.9884	*2.4593*	0.3312	13.2262
BAFT	20.5180	*4.4625*	7.2853	36.2346
RAFT	1.1321	*0.2230*	0.4343	1.9061
No. of units on common support	1943			
	PSME			
	Mean	*Std. Dev.*	Confidence Interval 95%	
RHO(T; Δ)	0.2001	0.0281	0.1293	0.2913
Estimated ATT_{PSME}	**6.8882**	**0.4908**	**5.5479**	**8.3394**
Unbiased Pop. *ATT*	***5.0095***	0.2065	4.4344	5.6268
Biased Pop. *ATT*	9.7874	0.3633	8.7151	10.7534
Mean bias after matching	5.8729	2.0717	1.0809	13.0339
Median bias after matching	6.2542	2.4963	0.7634	13.8850
BAFT	21.2310	4.4589	10.5499	34.9559
RAFT	1.1906	0.1927	0.6864	1.8085
No. of units on common support	1971			

Table A8. Heterogeneity in observed covariates. No. of Reps.=500. Simulated endogeneity: σ_{1v} =5.4; σ_{0v} = -0.8

		SSCM		
	Mean	*Std. Dev.*	Confidence Interval 95%	
RHO(T; Δ)	0.1602	*0.0385*	0.0562	0.2763
Estimated ATT -SSCM	**4.7568**	***0.5233***	**3.4332**	**6.0831**
Unbiased Pop. ATT	***5.0122***	*0.2050*	4.4301	5.6090
Biased Pop. *ATT*	8.4813	*0.3355*	7.3660	9.5244
Estim. Transition coeff.	-0.4573	*0.0234*	-0.5320	-0.3910
Shift coeff.	1.4295	*0.3139*	0.6726	2.4584
Endownment coeff.	0.5350	*0.0408*	0.4236	0.6628
Mean bias after matching	4.9532	*1.7404*	1.3584	10.1984
Median bias after matching	4.8650	*2.3882*	0.3790	11.9008
BAFT	20.7891	*4.6601*	8.6572	37.0917
RAFT	1.1283	*0.2153*	0.5115	1.7224
No. of units on common support	1931			
		PSME		
	Mean	*Std. Dev.*	Confidence Interval 95%	
RHO(T; Δ)	0.2122	0.0273	0.1462	0.3230
Estimated ATT_{PSME}	**7.5994**	**0.4838**	**5.9927**	**9.0451**
Unbiased Pop. ATT	***5.0102***	0.2049	4.4301	5.6047
Biased Pop. *ATT*	8.4679	0.3483	7.3660	9.3284
Mean bias after matching	5.9958	2.1700	1.1903	15.7593
Median bias after matching	6.4759	2.6393	0.7870	18.5810
BAFT	21.4835	4.6189	8.5545	38.2082
RAFT	1.1834	0.1941	0.5498	1.8284
No. of units on common support	1969			

Ultimi Contributi di Ricerca CRENoS

I Paper sono disponibili in: http://www.crenos.it

18/04 *William Addessi, Bianca Biagi, Maria Giovanna Brandano,* "How tourist flows are affected by the introduction of the euro?"

18/03 *Luc Bauwens, Edoardo Otranto*, "Nonlinearities and Regimes in Conditional Correlations with Different Dynamics"

18/02 *Massimiliano Bratti, Maurizio Conti, Giovanni Sulis,* "Employment Protection, Temporary Contracts and Firm-provided Training: Evidence from Italy"

18/01 *Luca De Benedictis, Vania Licio, Anna Maria Pinna, "*The long-term effects of the historical Roman road network: Trade costs of Italian provinces"

17/11 *Massimo Del Gatto, Carlo S. Mastinu,* "A Huff model with heterogeneous retailers fits well in Southern Italy"

17/10 *Sara Calligaris, Massimo Del Gatto, Fadi Hassan, Gianmarco I.P. Ottaviano, Fabiano Schivardi,* "The Productivity Puzzle and Misallocation: an Italian Perspective"

17/09 *Michele Battisti, Filippo Belloc. Massimo Del Gatto,* "Technology-specific Production Functions"

17/08 *Oliviero A. Carboni, Giuseppe Medda*, "Do Investment and Innovation Boost Export? An Analysis on European Firms"

17/07 *Edoardo Otranto, Massimo Mucciardi,* "Clustering Space-Time Series: A Flexible STAR Approach"

17/06 *Simone Franceschini, Gerardo Ettore Marletto, "*The dynamics of social capital during public participation: new knowledge from an on-going monitoring"

17/05 *Luca G. Deidda, Ettore Panetti,* "Banks' Liquidity Management and Systemic Risk"

17/04 *Luca Frigau, Tiziana Medda, Vittorio Pelligra,* "From the Field to the Lab An Experiment on the Representativeness of Standard Laboratory Subjects"

17/03 *William Addessi, Manuela Pulina,* "Sectoral Composition of Consumption Expenditure: A Regional Analysis"

17/02 *Claudio Detotto, Marta Meleddu, Marco Vannini,* "Cultural identity and willingness to protect and preserve art"

17/01 *Adriana Di Liberto, Laura Casula,* "Teacher assessments versus standardized tests: is acting "girly" an advantage?"

16/16 *Massimo Del Gatto, Carlo S. Mastinu,* "Sequentiality and Distance(s) in Cross-Border Mergers and Acquisitions: Evidence from Micro Data"

16/15 *Angelo Antoci, Simone Borghesi, Gianluca Iannucci*, "Green licenses and environmental corruption: a random matching model"

16/14 *Vittorio Pelligra, Tommaso Reggiani, Daniel John Zizzo,* "Responding to (un)reasonable requests"

16/13 *Pasqualina Arca Gianfranco Atzeni, Luca Deidda,* "Asset exemption in entrepreneurs' bankruptcy and the informative role of collateral"

16/12 *Miguel Casares, Luca Deidda, Jose E. Galdon-Sanchez*, "Loan production and monetary policy"

16/11 *Manuela Pulina, Valentina Santoni,* "An analysis on the Italian agricultural firms: effects of public subsidies"

16/10 *Tiziana Medda, Vittorio Pelligra, Tommaso Reggiani,* "Does Experience Affect Fairness, Reciprocity and Cooperation in Lab Experiments?"

16/09 *Bianca Biagi, Maria Giovanna Brandano, Manuela Pulina,* "The effect of tourism taxation: a synthetic control approach"

www.crenos.it