Universität St.Gallen

Identification of Dynamic Treatment Effects by Instrumental Variables

Ruth Miquel

May 2002 Discussion paper no. 2002-11

Department of Economics University of St. Gallen

Editor:	Prof. Jörg Baumberger University of St. Gallen Department of Economics Bodanstr. 1 CH-9000 St. Gallen Phone ++41 71 224 22 41 Fax ++41 71 224 28 85 Email joerg.baumberger@unisg.ch
Publisher:	Forschungsgemeinschaft für Nationalökonomie an der Universität St. Gallen Dufourstrasse 48 CH-9000 St. Gallen Phone ++41 71 224 23 00 Fax ++41 71 224 26 46
Electronic Publication:	www.fgn.unisg.ch/public/public.htm

Identification of Dynamic Treatment Effects by Instrumental Variables[1]

Ruth Miquel

Author's address: Ruth Miquel
SIAW
Dufourstrasse 48
9000 St. Gallen
Tel. ++41 71 224 23 29
Fax ++41 71 224 22 98
Email ruth.miquel@unisg.ch
Website www.siaw.unisg.ch/lechner/miquel

[1] I am grateful to Michael Lechner, Markus Frölich, Conny Wunsch and Gauthier Lanot for very helpful comments. A previous version of this paper has been presented at the "7th Colloque des jeunes économètres", Mont st. Odile. All remaining errors are mine.

Abstract

This paper considers the problem of the identification of causal effects using instrumental variables. We are interested in the effects of some treatments on certain outcomes. First, we consider that a participation in a treatment or a program is only possible one time but we have the choice between more than one program. Under a monotonicity condition and an exclusion restriction, pair-wise Local Average Treatment Effects are identifiable. Second, we consider the case where only one program is available but more than one participation is possible, leading to a comparison of sequences of participations (or sequences of programs). In this framework a problem of endogeneity appears: the outcome after one period, affected by the participation in this period, can influence the participation in the next period. Under different versions of the monotonicity condition and the exclusion restriction, identification of the causal effects of sequences of programs are investigated. The introduction of a second period implies a loss of identification for some effects of interest even without any endogeneity problem.

Keywords

Compliers, Local Average Treatment Effect, dynamic treatment regimes,
nonparametric identification, instruments

JEL Classification

C40

1) Introduction

The estimation and identification of causal effects is often encountered in conjunction with the evaluation of some treatments or programs, as for example vocational training programs. The literature dealing with this issue increased a lot since the last twenty years. The aim is to see if the application of a program is beneficial to a particular population. A summary of the principal methods and problems associated with program evaluation can be found in Angrist and Krueger (1999), Heckman, LaLonde and Smith (1999) and Lechner and Pfeiffer (2001). The evaluation problem is usually examined with Rubin's framework of potential outcomes (Roy 1951, Rubin 1974). Assuming that two possible states of nature exist, the outcome of an individual is described by two potential outcomes, each depending on or "defined" by one state of nature. The observed or realized outcome is one of these potential outcomes according to the realization of the state of nature or, in other words, according to the program applied. Then the effect of the program is computed as the difference between the potential outcomes. For example, for the evaluation of training programs we want to compute the effect of participating in the program compared to not participating. Thus, we need to compute the difference of the two potential outcomes defined in the state of nature "participation" and in the state of nature "no participation". As already mentioned, only one outcome is observable. Therefore, we have to state some hypotheses to be able to construct this difference and to estimate the causal effect of the program.

One possible approach is to use instrumental variables (if available). In this framework different effects are estimated depending on the assumptions made. Imbens and Angrist (1991, 1994) and Angrist, Imbens and Rubin (1996) have proved that the effect identified with instrumental variable restrictions is the Local Average Treatment Effect (LATE), an observable effect for an unobservable population: the compliers or the individuals for whom a change in the instrument's value implies a change in the treatment status. A second approach using instrumental variables is the Local Instrumental Variables method (LIV)[1] introduced by Heckman and Vytlacil (Heckman 1997, Heckman and Vytlacil 1999, 2000 a-b). With this method they extend the linear IV analysis and they are able to estimate a wider class of treatment parameters. A huge literature applies or extends these methods, among others Abadie (2000, 2002), Abadie, Angrist and Imbens (1998) Heckman, Tobias and Vytlacil (2000), Manski and Pepper (2000), Vytlacil (2002). Although these authors present very interesting results, for example that LATE allows the estimation of quantiles, LATE and LIV are equivalent, all these papers work in a static framework. They do not study the impact of multiple participations over time and the influence of previous program participations. In fact, the effects of the programs attended by the individuals in the past have an influence on the choice of the next program and this influence cannot be handled in such a framework.

Although the dynamic character of the problem can be treated in a parametric framework (panel data models like duration models, van den Berg, 2000, Arellano and Honoré, 2002), few work has been done in a nonparametric framework. The origin can be found in the series of papers from Robins (1986, 1989, 1997) and Robins, Greenland and Hu (1999) in Biometrics. These papers use counterfactual outcomes in a dynamic framework using some sequential randomization to identify the effects of interest. A more complete work on the identification problem of average treatment effects in a counterfactual dynamic framework can be found in Lechner and Miquel (2001). The effects of sequences of programs are identified only for certain populations of participants when the dynamic aspect of the problem is taken into account (the outcome of one period, itself influenced by the participation in this period, affects the participation in the next period). But until now, no work has been done in the area of IV methods in a dynamic framework.

This paper presents an extension of the static case (one-period case) where now many programs or treatment are available (instead of only one program) and it presents the necessary conditions to identify the effects of sequences of programs (many periods). The identified effects look like some Local Average Treatment Effects. In the static framework, the LATE is the effect of a program for the compliers, a population that changes its participation with a change in the value of an instrumental variable. In the multi-period framework we need to redefine which persons are compliers and some other similar populations. With these definitions, we establish the assumptions needed for the identification of the LATEs. Unfortunately, we are not able to identify all effects of interest whether we are in a nondynamic case where the outcome of the previous period does not affect the participation in the second period, or not. The paper is restricted to three periods. The more general case (with T periods) will generate no other problems or results as those presented here but will be more difficult to handle due to the large number of sequences of programs and possible instruments.

The next section of the paper introduces the notation used in the paper. Section 3 recalls the results of the one-period static case. Section 4 extends the one-period case with one program

[1] This method directly estimates the Marginal Treatment Effect.

to a one-period case with multiple programs. Section 5 provides the assumptions and identification results in a multi-period framework with one instrument available. Section 6 presents different sets of assumptions to identify the effects in a multi-period case with two instruments available. These sets of conditions are differentiated depending on the sequentiality and dynamics introduced in the problem. The last section concludes.

2) Notation

One of the difficulties arising when we work in a dynamic framework is a trade-off between the degree of complexity of the notation required to correctly describe the problem with all ist facets and the need of clarity. The notation used in this paper is based on the one introduced in Lechner and Miquel (2001). Without imposing any restricting on the problem we can simplify it to the case of three discrete time periods denoted with the subscript $t = 0,1,2$. For each period only one training program exists. So each period the choice is between participating or not participating in this program. (This general framework will be modified a little in the next section.) Three different types of variables are available.

The first type consists of instrumental variables. These variables fulfill particular exclusion restrictions, which we will discuss later. These exclusion restrictions vary depending on the dynamic framework used. We limit us to the case of binary variables, i.e. to the case where the instrument variables are dummy variables. Moreover, the number of available instruments varies depending on the case studied. We can dispose of one or two instruments (and in the remainder of this section we treat the two-instruments case). We relate those variables to the time period. The first instrument will have the subscript 1 and the second will have the subscript 2 because the second one is always available only in the second period. The random vector of the instrumental variables in all periods is represented by $\underline{Z}_2 = (Z_1, Z_2)$, $Z_t \in \{0,1\}$. A particular value of the instrument at time *t* is denoted z_t.[2]

The second type of variables relates to the participation in the program or treatment. $\underline{S}_2(Z) = (0, S_1(Z_1), S_2(Z_2))$ represents the history of potential participations. In each period the participation indicator takes either of two values: 0, when the person does not participate in the program and 1, when she does, $S_t(Z_t) \in \{0,1\}$. In the first period nobody participates in the program, hence $S_0(Z_0) = 0$. We can represent the dependence of the participation

[2] For all variables a value is represented by a lower case letter and the random variable by an upper case letter.

indicator on the instrument using such notation because we make an additional assumption. Three different dependencies are possible and will be investigated[3]: firstly, the participation in both periods depends only on the first instrument (only one instrument is available); secondly, each participation indicator depends on one instrument (the participation in the first period depends on the first instrument and the participation in the second period depends only on the second instrument), and finally, the participation in the second period depends on both instruments. In the second type of dependencies, the participation in the second period depends only on the second instrumental variable, Z_2, although the first instrument is still available. The first period participation indicator depends only on the first instrument because in the first period only this instrument is available. Due to the dependence of the participation indicator on the instrument, $S_t(Z_t)$ represents a potential participation, the potentiality being defined in terms of the instrument. $S_t(z_t)$ is still a random variable in the special world in which the instrumental variable takes the particular value z at time t. The observed participation indicator in one period is denoted by S_t. Sequences of programs are simply denoted by a sub-vector of the history: $\underline{S}_1(\underline{Z}_1) = (0, S_1(Z_1))$, $\underline{S}_2(\underline{Z}_2) = (0, S_1(Z_1), S_2(Z_2))$. We have to notice that the notation used for the instrumental variable $\underline{Z}_t$ is a misuse of the notation. The "underline notation" depicts the history of a variable up to a certain time. Using $\underline{S}_t(\underline{Z}_t)$ we are not saying that the dependence applies to all the history of Z up to period t, but that each element of this history has an influence on a corresponding element of the history of the participations.[4] We also neglect the subscript for individuals. The participation indicator for an individual depends only on the value of her own instrumental variable. The instrument for the other individuals has no influence on her participation: we implicitly assume that the SUTVA assumption from Rubin (1974) holds.[5]

The third category of variables represents outcomes. The impact of the treatments is modeled using the concept of potential variables. Each sequence of treatments defines a different world. In each of these worlds a random outcome exits. Thus, each of these outcomes is regarded as a potential outcome, as only one of the states of the world is realized. In the world characterized by a particular value of the sequence $\underline{s}_t$, the potential outcome at time τ ($\tau \geq t$) is $Y_\tau^{\underline{s}_t}$. The observed outcome at τ is represented by Y_τ.

[3] The second type of dependences is used for the presentation of the notations.

[4] Note that for the participation indicator, $\underline{S}_t(.)$ represents the history of participations up to period t.

[5] We use a modified version (for the multi-period case) of the Stable Unit Treatment Value Assumption: a) if $\underline{Z}_{2i} = \underline{Z}_{2i}^{'}$ then $\underline{S}_{2i}(Z) = \underline{S}_{2i}(Z')$ and b) if $\underline{Z}_{2i} = \underline{Z}_{2i}^{'}$ and $\underline{S}_{2i} = \underline{S}_{2i}^{'}$ then $Y_{\tau i}^{\underline{s}_2} = Y_{\tau i}^{\underline{s}_2^{'}}$.

In our three-time period framework, the following sequences of treatments are possible:

$$\underline{S}_1 \in \{(0,1),(0,0)\} \text{ ; } \underline{S}_2 \in \{(0,1,1),(0,0,1),(0,1,0),(0,0,0)\} \text{ ,}$$

implying six potential outcomes:

$$Y_\tau^{(0,1)}, Y_\tau^{(0,0)}, Y_\tau^{(0,1,1)}, Y_\tau^{(0,0,1)}, Y_\tau^{(0,1,0)}, Y_\tau^{(0,0,0)} .$$

To simplify the exposition the effects are summarized in the following notation:

$$\theta_\tau^{\underline{s}_t, \underline{\tilde{s}}_t}(M=m) = E(Y_\tau^{\underline{s}_t} - Y_\tau^{\underline{\tilde{s}}_t} \mid M=m)$$

where M represents functions of the participation indicators.[6]

The relations between the observable outcome, Y_τ and the potential outcomes defined in Lechner and Miquel (2001) are valid in our framework. The following equations hold in the case of three time periods. To simplify the notation we suppress the first period participation indicator (it is zero for everybody) in the remainder of the paper.

$$\begin{aligned} Y_\tau &= S_1(z_1)Y_\tau^{1\cdot} + (1-S_1(z_1))Y_\tau^{0\cdot} \\ &= S_1(z_1)S_2(z_2)Y_\tau^{11} + S_1(z_1)(1-S_2(z_2))Y_\tau^{10} + (1-S_1(z_1))(1-S_2(z_2))Y_\tau^{00} + (1-S_1(z_1))S_2(z_2)Y_\tau^{01}. \end{aligned}$$

In the following we will denote period $t=0$ "start period" and period $t=1$ "first period". In the start period nothing happens and in the first and second period participation becomes possible.

3) The static case using the multi-period notation

This part restates the "static case" using the multi-period notation. We consider only the first period, hence, the participation in the program is only specified for this period. The participation for the second period remains unspecified. To identify the LATE we need the following assumptions using Dawid's (1979) notation of independence $\coprod$:

ASSUMPTION 1: (STATIC INDEPENDENCE CONDITION)

$$(S_1(z_t), Y_\tau^{\underline{s}_t}), (S_1(1), S_1(0)) \coprod Z_1 \text{ , } \forall \underline{s}_t, z_t .$$

[6] For example $M = \underline{S}_2$ and $m = (0,1,1)$.

ASSUMPTION 2: (ONE PERIOD MONOTONICITY CONDITION):

$$p\left[S_t(z_t) \geq S_t(z_t')\right]=1 \text{ or } p\left[S_t(z_t) \leq S_t(z_t')\right]=1 \ ; \ z_t, z_t' \in \{0,1\}, \ z_t \neq z_t'.$$[7]

These assumptions are those stated by Imbens and Angrist (1994). The variables are indexed only with a time subscript. The static independence condition (SIC), which is the first part of condition 1 in Imbens and Angrist (1994), defines an exclusion restriction. The joint distribution of the potential outcome and the potential participation indicator is independent of the instrument in the first period. Also, the joint distribution of the potential participations is independent of the instrument in the first period. The monotonicity condition (OMC), which is the second condition in the above mentioned paper, ensures a monotonic influence of the instrument on the participation. If it is more likely to participate when $S_t = z_t$ than when $S_t = z_t'$, then each person who participates when $S_t = z_t'$ should also participate when $S_t = z_t$. In the terminology of Imbens and Angrist this hypothesis excludes the subpopulation of defiers (if the instrument corresponds to an assignment in a program, these individuals will do the opposite of it). Both assumptions are untestable.

Theorem 1: UNDER ASSUMPTION 1 (SIC) AND 2 (OMC) WITH $t=1$, THE EFFECT

$\theta_\tau^{(1.),(0.)}(S_1(1) - S_1(0) = 1)$ IS IDENTIFIED.

The proof is given in appendix A.1. Theorem 1 states, that the local average treatment effect for the compliers in the first period is identified. If we assume the instrument to be "should participate", when taking the value 1, and "should not participate", when taking the value 0, then the "compliers in the first period" are the subpopulation that follows the "assignment" of the instrument in the first period. The participation indicator S_1 will be equal to 1 and 0 respectively when Z_1 is equal to 1 and 0 respectively. Note that we do not specify the behavior of the compliers in the second period. We only want to differentiate the population on the basis of their behavior in the first period.

4) The static case with multiple programs

[7] In Imbens and Angrist (1994) this condition is presented in the following form $S_t(z_t) \geq S_t(z_t')$ or $S_t(z_t) \leq S_t(z_t')$; $z_t, z_t' \in \{0,1\}, \ z_t \neq z_t'$.

Let us stay in this one-period case with only one instrument and look at multiple alternatives faced by the individuals. Now, they have the choice to participate in one among several programs. As an example, consider only for this section that four different programs are available. The participation indicator S_1 equals zero when the individual does not participate in any program, and S_1 equals one, two or three depending on the program visited, $S_1 \in \{0,1,2,3\}$.[8] The potential participation indicators, $S_1(1), S_1(0)$, can also take four different values and also four potential outcomes, Y^0, Y^1, Y^2, Y^3, are now available. To place ourselves once again in a binary world, where only two programs are considered, we introduce news binary variables, $\tilde{S}_{kl}(1), \tilde{S}_{kl}(0), \tilde{S}_{kl}$, $k,l = 0,1,2,3$. The first two groups are potential variables (they depend on the value of the instrument Z_1) and the last one is observed.[9] When the participation choice is restricted to be between programs k and l, these variables equal 1 when $S_1(1), S_1(0), S_1$ respectively take the value k and equal 0 when $S_1(1), S_1(0), S_1$ respectively take the value l. When we compare pairs of programs, they indicate the participation in one of the programs of the pair. As we only compare two programs we can use the terminology of the one-program case and rename one of our choices "participation" and the other one "nonparticipation" in the pair considered. In the following, these variables are called "pair participation indicators" and when we mention a "pair participation" we talk about the participation in one program instead of in a second program in a pair-wise comparison (only two programs are available for the choice).

ASSUMPTION 3: (CONDITIONAL - STATIC INDEPENDENCE CONDITION)

$$(\tilde{S}_{kl}(z_1), Y_\tau^k, Y_\tau^l), (\tilde{S}_{kl}(1), \tilde{S}_{kl}(0)) \amalg Z_1 \mid S_1 = \{k,l\}, \ \forall z_1, k, l .$$

The hypotheses needed for the identification of effects are still an independence and a monotonicity condition. The first one, the conditional static independence condition (C-SIC), states the exclusion restriction which defines the instrument. The instrument Z_1 should have no influence on the joint distribution of one potential binary variable indicating a pair participation and the two potential outcomes corresponding to this pair (they are indexed by the programs in the pair), given that a participation in one of those programs considered in the

[8] The four sequences of programs defined in the notation section can be considered as four different programs, for example the sequence $(1,1)$ can be defined as program 1.

[9] They can be computed from observable quantities.

pair is actually observed. It is exactly the same condition as the one defined by Angrist and Imbens. The second condition, the conditional one-period monotonicity condition (C-OMC), affects the pair participation indicators and not the participation indicators. Nevertheless, the same interpretation can be made in terms of pair participation. Given that the choice is between participating in program k and participating in program l, the instrument has a monotonic influence on the participation.

ASSUMPTION 4: (CONDITIONAL-ONE PERIOD MONOTONICITY CONDITION):

$p\left[\tilde{S}_{kl}(z_1) \geq \tilde{S}_{kl}(z_1^{'}) \mid S_1=\{\text{k,l}\}\right]=1$ or $\text{p}\left[\tilde{S}_{kl}(z_1) \leq \tilde{S}_{kl}(z_1^{'}) \mid S_1=\{\text{k,l}\}\right]=1;\ z_1, z_1^{'} \in \{0,1\},\ z_1 \neq z_1^{'}.$

Under these assumptions all the effects comparing any pair of programs are identified for the compliers of the subpopulation participating in one of the programs considered in the pair. The compliers population is defined with respect to the pair participation and not with respect to all possible participations. These results are presented in theorem 2.

Theorem 2: UNDER ASSUMPTION 3 (C- SIC) AND 4 (C – OMC), THE EFFECTS

$\theta_{\tau}^{(k.),(l.)}(\tilde{S}_{kl}(1)-\tilde{S}_{kl}(0)=1, S_1=\{k,l\})$, $\forall k,l$ IS IDENTIFIED.

The proof is given in appendix A.2. Here, identification is obtained as easily as in the case where we have the choice to participate in only one program. Indeed, the conditioning set restores a "binary approach" to the problem and the same mechanisms apply. The results are similar to those found by Imbens (2000) and Lechner (2001) for the identification of treatment effects under the conditional independence assumption when more than one treatment is available. Identification is satisfied for subpopulations that have a restricted choice of programs: they can only participate in one of two programs.

5) The multi-period, one-instrument case

Now, we consider the first type of dependence between the participation indicators and the instrument. The behavior of the individuals is investigated for both periods, but we only observe the first instrument. Or equivalently, the instruments are perfectly correlated, $Z_1 = Z_2$. Within this context, how can we define the subpopulation of compliers? Does a population of compliers exist? More than one such subpopulation can be defined. For example, some individuals could comply only in one period and thus form a subpopulation of compliers for this period. Nevertheless, we will use the term "compliers" without reference to a time period but for the subpopulation of persons who comply in both periods. Thus, more than one effect can be defined depending on the subpopulation considered. Nontheless, identification will be possible only for some of these subpopulations. With one instrument, two different classes of effects are identified. The assumptions needed comprise an independence condition (or an exclusion restriction) and a monotonicity condition. Although the first period variables can have an influence on the second period variables (e.g. the second participation depends on the first one), the structure of the problem is very similar to the one-period, multiple-program case of the previous section. The conditions used are also in a conditional form. Instead of conditioning on a restricted participation choice (only between two programs), we condition on the behavior in one period. We can choose to fix the behavior in the first period or in the second period depending on the effects we are interested in. For example, we can look at the population that participates in the first period.

ASSUMPTION 5: (CONDITIONAL-STATIC INDEPENDENCE CONDITION):

$$(S_t(z_1), Y_\tau^{\underline{s}_2}), (S_t(z_1^{'}), S_t(z_1)) \coprod Z_1 \mid S_{t^{'}} = s_{t^{'}} \;, \;\forall\, z_1, z_1^{'}, \;\underline{s}_2, \; t \neq t^{'} .$$

ASSUMPTION 6: (CONDITIONAL-ONE PERIOD MONOTONICITY CONDITION):

$$p\left[S_t(z_1) \geq S_t(z_1^{'}) \mid S_{t^{'}} = s_{t^{'}}\right] = 1 \text{ or } p\left[S_t(z_1) \leq S_t(z_1^{'}) \mid S_{t^{'}} = s_{t^{'}}\right] = 1;\; z_1, z_1^{'} \in \{0,1\},\, z_1 \neq z_1^{'}, t \neq t^{'}.$$

Theorem 3 states that the first identifiable effects are those for the subpopulations whose participation in the first period is the same whatever the value of the instrument.

Theorem 3: UNDER ASSUMPTIONS 5 (C-SIC) AND 6 (C-OMC) WITH $t = 2$ AND $t^{'} = 1$, THE EFFECTS $\theta_\tau^{11,10}(S_1 = 1, S_2(1) - S_2(0) = 1)$ AND $\theta_\tau^{01,00}(S_1 = 0, S_2(1) - S_2(0) = 1)$ ARE IDENTIFIED.

The proof is given in appendix A.3. In this case, the first period ($t = 1$) plays the role of the start period ($t = 0$) for the subpopulations defined by their participation status in the first period. The identifiable effects compare sequences of treatments which have an identical participation in the first period. Thus, in the subpopulations defined by the participation or by the nonparticipation in the first period, the local average treatment effect (participation versus no participation in the second period) is identified for the second period compliers.

We get a similar result when we consider the subpopulation who complies in the first period and always or never participates in the second one. This time the conditions affect the first period potential participation indicators given the participation in the second period. The proof is given in appendix A.4.

Theorem 4: UNDER ASSUMPTIONS 5 (C-SIC) AND 6 (C-OMC) WITH $t=1$ AND $t^{'}=2$, THE EFFECTS $\theta_{\tau}^{11,01}(S_2=1,S_1(1)-S_1(0)=1)$ AND $\theta_{\tau}^{10,00}(S_2=0,S_1(1)-S_1(0)=1)$ ARE IDENTIFIED.

We compare sequences that have the same participation in the second period. The effect induced by the different sequences comes from the comparison of participation and nonparticipation in the second period. However, the assumptions used are not very intuitive because we condition on a variable from the future. Nevertheless, we can identify the same effects under a set of assumptions which are a little different but more intuitive.

ASSUMPTION 7: (STATIC INDEPENDENCE CONDITION-BIS):

$$(S_1(z_1),S_2(z_1),Y_{\tau}^{s_2}),(S_1(0),S_1(1),S_2(0),S_2(1))\amalg Z_1 \ , \ \forall z,\underline{s_2} \ .$$

The price to pay for the suppression of the conditioning set is a strengthening of the independence condition. Now it deals with the joint distribution of the potential outcome and the participation indicators in both periods and the joint distribution of all the participation indicators. The results of theorem 3 and 4 are restated in theorem 5 and proved in appendix A.5.

Theorem 5:

A) UNDER ASSUMPTIONS 7 (SIC_BIS), AND 6 (OMC) WITH $t=2$ AND ASSUMING THAT $p[S_1(1)=S_1(0)]=1$, THE EFFECTS $\theta_\tau^{11,10}(S_1=1, S_2(1)-S_2(0)=1)$ AND $\theta_\tau^{01,00}(S_1=0, S_2(1)-S_2(0)=1)$ ARE IDENTIFIED.

B) UNDER ASSUMPTIONS 7 (SIC_BIS), AND 6 (OMC) WITH $t=1$ AND ASSUMING THAT $p[S_2(1)=S_2(0)]=1$, THE EFFECTS $\theta_\tau^{11,01}(S_2=1, S_1(1)-S_1(0)=1)$ AND $\theta_\tau^{10,00}(S_2=0, S_1(1)-S_1(0)=1)$ ARE IDENTIFIED.

Apparently the most important change is without doubt the third assumption which excludes the defiers *and* the compliers of the first or the second period. Of course this restriction is undesirable. Moreover, the number of sequences we can compare in this context is very small. With only one instrument available, some effects can be identified if we step back into a "one period context". The possible comparisons of sequences is restricted due to the fact that only sequences with the same participation in one period can be considered to obtain identification.

6) The multi-period, two-instruments case

Nevertheless, in this two-period framework more than one instrument is available. This additional information will allow to identify more effects. As already mentioned in the notation section, we consider different combinations of the participation indicator in the second period and both instruments. In this part we assume that the participation in the second period is only influenced by the second instrument. (The second instrument does not include the information provided by the first instrument.)

6.1) Static case

Although we have more than one period, we are still in a nondynamic framework, because the first period outcome has no effect on the participation and on the outcome in the second period. As in the previous sections, we need an independence and a monotonicy assumption to obtain identification. Besides the dependence on the time index of the instrument, nothing is new and we can keep the independence assumption (SIC-BIS) of the previous section. To clarify this notion of independence, we rewrite it taking explicitly account of the time index of the instrument.

ASSUMPTION 8 (STATIC INDEPENDENCE CONDITION-BIS):

$$(S_1(z_1), S_2(z_2), Y_\tau^{\underline{s}_2}), (S_1(0), S_1(1), S_2(0), S_2(1)) \amalg (Z_1, Z_2) \ , \ \forall z_1, z_2, \ \underline{s}_2 \ .$$

Each (previously defined) joint distribution is independent of the instrument in each period. Until now the monotonicity condition has affected only one period. A priori, we do not have any other restriction on the subpopulations we considered now, thus, we need to exclude the defiers in both periods. Hence the new monotonicity condition is an extension of the one-period monotonicity to a two-period monotonicity.

ASSUMPTION 9 (TWO-PERIOD MONOTONICITY CONDITION):

$$p\left[S_1(1) \geq S_1(0)\right]=1 \text{ or } p\left[S_1(1) \leq S_1(0)\right]=1, \ p\left[S_2(1) \geq S_2(0)\right]=1 \text{ or } p\left[S_2(1) \leq S_2(0)\right]=1$$

$$\text{and } \ p\left[S_1(1) \geq S_1(0), S_2(1) \geq S_2(0)\right]=1 \text{ or } p\left[S_1(1) \leq S_1(0), S_2(1) \geq S_2(0)\right]=1,$$

$$\text{or } p\left[S_1(1) \geq S_1(0), S_2(1) \leq S_2(0)\right]=1 \text{ or } p\left[S_1(1) \leq S_1(0), S_2(1) \leq S_2(0)\right]=1 .$$

In the following, we will arbitrarily consider the case where $S_1(1) \geq S_1(0)$ and $S_2(1) \geq S_2(0)$. Then, under those assumptions, the following fourteen effects can be identified:

Theorem 6: UNDER ASSUMPTIONS 8 (SIC-BIS) AND 9 (TMC),

A) ALL LATE – EFFECTS OF TYPE $\theta^{ij,kl}((S_2(1)-S_2(0))(S_1(1)-S_1(0))=1)$,

$i, j, k, l = 1,0$, ARE IDENTIFIED.

B) THE FOLLOWING 8 EFFECTS ARE IDENTIFIED:

B.1) $\theta^{11,01}(S_2(0)(S_1(1)-S_1(0))=1), \quad \theta^{11,01}(S_2(1)(S_1(1)-S_1(0))=1),$

B.2) $\theta^{11,10}(S_1(0)(S_2(1)-S_2(0))=1), \quad \theta^{11,10}(S_1(1)(S_2(1)-S_2(0))=1),$

B.3) $\theta^{10,00}((1-S_2(0))(S_1(1)-S_1(0))=1), \quad \theta^{10,00}((1-S_2(1))(S_1(1)-S_1(0))=1),$

B.4) $\theta^{01,00}((1-S_1(0))(S_2(1)-S_2(0))=1), \quad \theta^{01,00}((1-S_1(1))(S_2(1)-S_2(0))=1).$

The theorem is proved in the appendix A.6. The effects comparing all pairs of sequences are identified for the compliers in both periods. For other subpopulations less effects are identified. We can only compare sequences which have the same participation in one period.

The persons concerned comply in one period (when the sequences differ). In the other period only the participation or the treatment under a special value of the instrument is specified. When the instrument equals 1 we cannot differentiate between the compliers and the always-takers. When it equals 0, due to the monotonicity assumption the defiers are excluded and the effect affects the never-takers. Identification is obtained because the effects can be expressed as a function of observable probabilities and expectations. For example,

$$\theta^{11,01}\left((S_2(1)-S_2(0))(S_1(1)-S_1(0))=1\right)$$
$$=\frac{E\left[Y_tS_2\mid Z_1=1,Z_2=1\right]-E\left[Y_tS_2\mid Z_1=0,Z_2=1\right]-E\left[Y_tS_2\mid Z_1=1,Z_2=0\right]+E\left[Y_tS_2\mid Z_1=0,Z_2=0\right]}{P\left[(S_2(1)-S_2(0))(S_1(1)-S_1(0))=1\right]}.$$

And the denominator equals

$$\begin{aligned}&E(S_2\mid Z_2=1,Z_1=1)-E(S_2\mid Z_2=0,Z_1=1)-E(S_2\mid Z_2=1,Z_1=1,S_1=0)(1-E(S_1\mid Z_1=1))\\&+E(S_2\mid Z_2=0,Z_1=1,S_1=0)(1-E(S_1\mid Z_1=1))-E(S_2\mid Z_2=1,Z_1=0,S_1=1)E(S_1\mid Z_1=0)\\&+E(S_2\mid Z_2=0,Z_1=0,S_1=1)E(S_1\mid Z_1=0).\end{aligned}$$

Of course this last expression muss be different from 0. We can express the effect in terms of covariances, but the link between the effect and the "classical" instrumental variable estimator is not as obvious as in the static case in Imbens and Angrist (1994):

$$\theta^{11,01}\left((S_2(1)-S_2(0))(S_1(1)-S_1(0))=1\right)$$
$$=\frac{\dfrac{\operatorname{cov}\left(Y_tS_2,Z_2\mid Z_1=1\right)}{p\left(Z_2=1\mid Z_1=1\right)\left(1-p\left(Z_2=1\mid Z_1=1\right)\right)}-\dfrac{\operatorname{cov}\left(Y_tS_2,Z_2\mid Z_1=0\right)}{p\left(Z_2=1\mid Z_1=0\right)\left(1-p\left(Z_2=1\mid Z_1=0\right)\right)}}{\left\{\begin{array}{l}\dfrac{\operatorname{cov}\left(S_2,Z_2\mid Z_1=1\right)}{p\left(Z_2=1\mid Z_1=1\right)\left(1-p\left(Z_2=1\mid Z_1=1\right)\right)}-\dfrac{\operatorname{cov}\left(S_2,Z_2\mid Z_1=1,S_1=0\right)p\left(S_1=0\mid Z_1=1\right)}{p\left(Z_2=1\mid Z_1=1,S_1=0\right)\left(1-p\left(Z_2=1\mid Z_1=1,S_1=0\right)\right)}\\-\dfrac{\operatorname{cov}\left(S_2,Z_2\mid Z_1=0,S_1=1\right)p\left(S_1=1\mid Z_1=0\right)}{p\left(Z_2=1\mid Z_1=0,S_1=1\right)\left(1-p\left(Z_2=1\mid Z_1=0,S_1=1\right)\right)}\end{array}\right\}}.$$

Is it possible to identify more effects or other effects with some additional conditions? Unfortunately, even if we limit the relation between the periods by introducing some sort of independence the number of identified effects does not increase. First, we explicitly assume that being a complier in one period is independent of being a complier in the other period:

ASSUMPTION 10: (PERIOD COMPLIER INDEPENDENCE) $S_1(1)-S_1(0) \perp\!\!\!\perp S_2(1)-S_2(0)$.

The behavior in the first period has no effect on the behavior in the second period. It is also possible that a first period complier becomes an always-taker for example and that a "first period never-taker" becomes a complier. Then, under these assumptions theorem 7 states that all the LATE effects for the compliers are identified. The proof of theorem 7 is given in appendix A.7.

Theorem 7: UNDER ASSUMPTIONS 8 (SIC-BIS), 9 (TMC) AND 10 (PCI), ALL LATE – EFFECTS OF TYPE $\theta_{\tau}^{ij,kl}((S_2(1)-S_2(0))(S_1(1)-S_1(0))=1)$, $i,j,k,l=1,0$, ARE IDENTIFIED.

Theorem 7 only deals with the compliers in both periods. In this case the estimation of the effects is simplified in the sense that the probability to be a complier in both periods (the denominator of the estimable form of the effets) is a simpler function of the observables: $\left[E(S_2\,|\,Z_2=1)-E(S_2\,|\,Z_2=0)\right]\left[E(S_1\,|\,Z_1=1)-E(S_1\,|\,Z_1=0)\right]$. Nonetheless, its expression in terms of covariances does not give the "classical" instrumental variables estimates:

$$\begin{aligned}&\theta^{11,01}\left((S_2(1)-S_2(0))(S_1(1)-S_1(0))=1\right)\\&=\frac{\operatorname{cov}\left(Y_tS_2,Z_2\,|\,Z_1=1\right)-\operatorname{cov}\left(Y_tS_2,Z_2\,|\,Z_1=0\right)}{\operatorname{cov}\left(S_2,Z_2\right)\operatorname{cov}\left(S_1,Z_1\right)}\,p\left(Z_1=1\right)\left(1-p\left(Z_1=1\right)\right).\end{aligned}$$

It may also be interesting to consider the other subpopulations with a mixed behavior. Identification is satisfied only for the effects discussed in theorem 6. The results are obtained under a modified independence assumption:

ASSUMPTION 11: (PERIOD INDEPENDENCE):

A) $S_1(1)-S_1(0)\amalg S_2(z_2)\quad z_2=0,1$.

B) $S_2(1)-S_2(0)\amalg S_1(z_1)\quad z_1=0,1$.

The potential participation indicator of one period is independent of the "complier participation" indicator of the other period.

Theorem 8: UNDER ASSUMPTIONS 8 (SIC-BIS), 9 (TMC) AND 11 (PI), THE FOLLOWING 8 EFFECTS ARE IDENTIFIED:

A) $\theta^{11,01}(S_2(0)(S_1(1)-S_1(0))=1), \quad \theta^{11,01}(S_2(1)(S_1(1)-S_1(0))=1),$

B) $\theta^{11,10}(S_1(0)(S_2(1)-S_2(0))=1), \quad \theta^{11,10}(S_1(1)(S_2(1)-S_2(0))=1),$

C) $\theta^{10,00}((1-S_2(0))(S_1(1)-S_1(0))=1), \quad \theta^{10,00}((1-S_2(1))(S_1(1)-S_1(0))=1),$

D) $\theta^{01,00}((1-S_1(0))(S_2(1)-S_2(0))=1), \quad \theta^{01,00}((1-S_1(1))(S_2(1)-S_2(0))=1).$

The proof is given in appendix A.8. Theorem 8 establishes the identification of effects of sequences with a common part: people do the same in the first period for both sequences or they do the same in the second period for both sequences. These effects are identical to those identified in section 5, with one exception. The populations for which the effects are investigated are not based on always-takers and never-takers in one period as in the previous case. In the period when they do not comply, the individuals are selected on a particular value of one potential participation indicator (the participation given one of the values of the regressor). Evidently, those potential indicators are not observed. Thus, although the sequences compared are the same as those in the multi-period, one-instrument case, the populations concerned differ. In the one-instrument case the populations are observable in one period. In the two-instrument case they are not.

Unfortunately, the identification of other possible effects cannot be obtained. Nevertheless, we can bind these effects, if the outcomes, the potential outcomes and their expectations are themselves bound. Some of the bounds are informative and some of them are informative depending on the value of different probabilities (conditional probabilities involving the realizations of the potential indicators). The results are available upon request.

Both assumptions affecting the interrelations of the periods (10 (PCI) and 11 (PI)) are not intuitive. It is possible to postulate a stronger assumption which is easier to understand and will lead to the same results.[10] Namely, we can assume strict independence of the periods. Each period, individuals decide on participation independently of the participation in the other period. But as stated before, this independence assumption does not improve identification. It

[10] This new assumption implies the two other ones.

just simplifies the computation of the probabilities to be in different subpopulations as presented in the proofs of the theorems.

6.2) Sequentiality

It is more realistic to assume that the participation is decided on at the beginning of each period, because it is often the case that a participation in a particular program is conditioned on a previous participation in another particular program (eg. before they participate in a training program, the trainees should have experienced a job search assistance). We now consider a sequential process. The participation decision is made on the basis of the observed past participations. This sequentiality emerges in the independence assumption. For the second period, independence is fulfilled given the participation and the instrument in the first period. This framework is still a static framework because the outcomes of the previous periods (or some other variables related to the outcomes) have no effect on the actual participation decision.

ASSUMPTION 12: (SEQUENTIAL STATIC INDEPENDENCE CONDITION):

$\forall z_1, z_2, \tilde{z}_1, \underline{s}_t, s_1$:

a) $(S_1(z_1), S_2(z_2), Y_\tau^{\underline{s}_2}), (S_1(0), S_1(1), S_2(0), S_2(1)) \amalg Z_1, \ \tau \geq 1$

b) $(S_1(1), S_1(0)) \amalg Z_2 \mid Z_1 = \tilde{z}_1,$

c) $(S_2(z_2), Y_\tau^{\underline{s}_2}), (S_2(0), S_2(1)) \amalg Z_2 \mid S_1 = s_1, Z_1 = \tilde{z}_1, \ \tau \geq 2.$

The independence condition concerning the first-period instrument remains unchanged. The second-period instrument has to fulfil two conditions: it has to be independent of the potential participation in the first period given the instrument in the first period, and it has to be independent of the potential participation in the second period and of the potential outcomes (the joint distribution) given the instrument and the realized participation in period one. This condition is also fulfilled for the joint distribution of the potential participation in the second period. Although this assumption is more intuitive in the sense that the decisions are made sequentially, it is neither stronger nor weaker than the static independence condition–bis. Rather they are equivalent.

Lemma 1: THE STATIC INDEPENDENCE CONDITION – BIS (SIC-BIS, ASSUMPTION 8) AND THE SEQUENTIAL STATIC CONDITION (SSIC, ASSUMPTION 12) ARE EQUIVALENT.

The proof is presented in appendix A.9. As the conditions are equivalent, the same effects are identified under the sequential form of the independence condition as under condition SIC-BIS as stated in theorem 9.

Theorem 9: UNDER ASSUMPTIONS 12 (SSIC) AND 9 (TMC),

A) ALL LATE – EFFECTS OF TYPE $\theta^{ij,kl}((S_2(1)-S_2(0))(S_1(1)-S_1(0))=1)$, $i,j,k,l=1,0$, ARE IDENTIFIED.

B) THE FOLLOWING 8 EFFECTS ARE IDENTIFIED:

B.1) $\theta^{11,01}(S_2(0)(S_1(1)-S_1(0))=1),\quad \theta^{11,01}(S_2(1)(S_1(1)-S_1(0))=1),$

B.2) $\theta^{11,10}(S_1(0)(S_2(1)-S_2(0))=1),\quad \theta^{11,10}(S_1(1)(S_2(1)-S_2(0))=1),$

B.3) $\theta^{10,00}((1-S_2(0))(S_1(1)-S_1(0))=1),\quad \theta^{10,00}((1-S_2(1))(S_1(1)-S_1(0))=1),$

B.4) $\theta^{01,00}((1-S_1(0))(S_2(1)-S_2(0))=1),\quad \theta^{01,00}((1-S_1(0))(S_2(1)-S_2(0))=1).$

The proof is direct as identification is already proved under the SIC-BIS assumption. The expression of the effects in terms of observables has already been presented in the previous section.

Moreover, assumption 12 (SSIC) implies a testable condition stated in Lemma 2.

LEMMA 2 : $S_1 \amalg Z_2 \mid Z_1$.

PROOF : AS S_1 IS A FUNCTION OF $S_1(z_1)$ AND Z_1, PART B) OF ASSUMPTION SSIC IMPLIES THIS RESULT.

This condition is a testable restriction between the observed participation in the first period and the instrument in the second period. Indeed, the sequentiality in the condition introduces a cross period relation. To identify the effects, we need to restrict this cross section relation.

6.3) Dynamics

Suppose now, the interactions between the periods are complicated. It seems credible that in addition to the influence of the realized participation in the first period, the realized outcome also influences the decision to participate in the second period. For example, if a person is still unemployed after a certain period of time after the end of the first training program, it is more likely that she participates in another program. Again, this new hypothesis affects only the independence assumption. We are not in a static framework anymore, because the endogenous outcome of the first period (or some functions of it) influences the second period variables.

Hence, assumption 12, SSIC, must be modified for the second period. Point c) now states that the independence of the potential outcomes, the potential participation indicators and the instrument is required given the observable outcome, the observable participation in period one and the value of the first-period instrument. The part of the potential variables, which is not explained by the observed variables of period one, should be independent of the second-period instrument.

ASSUMPTION 13: (DYNAMIC INDEPENDENCE CONDITION):

$$\forall z_1, z_2, \tilde{z}_1, \underline{s}_t, s_1 :$$

a) $(S_1(z_1), S_2(z_2), Y_\tau^{\underline{s}_2}), (S_1(0), S_1(1), S_2(0), S_2(1)) \amalg Z_1, \ \tau \geq 1$

b) $S_1(z_1) \amalg Z_2 \mid Z_1 = \tilde{z}_1$

c) $(S_2(z_2), Y_\tau^{\underline{s}_2}), (S_2(0), S_2(1)) \amalg Z_2 \mid Y_1 = y_1, S_1 = s_1, Z_1 = \tilde{z}_1, \ \tau \geq 2$.

As in the sequential case the monotonicity assumption 9 (two period monotonicity condition) is still required unchanged. Moreover, Lemma 2 is also valid. The independence condition presents a new cross period restriction with its dependence on the first-period outcome. Thus, we need a condition on this interaction as we needed one for the observed participation in the first period. Nevertheless, a similar result cannot be achieved directly and we need to explicitly add a testable hypothesis which is presented in assumption 14 below.

ASSUMPTION 14 : $Y_1 \amalg Z_2 \mid Z_1, S_1$.

With this additional assumption there is no loss of identification due to the dynamic nature of the problem. We can identify the same effects as in the sequential case.

Theorem 10: UNDER ASSUMPTIONS 13 (DIC), 9 (TMC) AND ASSUMPTION 14,

A) ALL LATE – EFFECTS OF TYPE $\theta^{ij,kl}((S_2(1)-S_2(0))(S_1(1)-S_1(0))=1)$, $i,j,k,l=1,0$, ARE IDENTIFIED.

B) THE FOLLOWING 8 EFFECTS ARE IDENTIFIED:

B.1) $\theta^{11,01}(S_2(0)(S_1(1)-S_1(0))=1),\quad \theta^{11,01}(S_2(1)(S_1(1)-S_1(0))=1),$

B.2) $\theta^{11,10}(S_1(0)(S_2(1)-S_2(0))=1),\quad \theta^{11,10}(S_1(1)(S_2(1)-S_2(0))=1),$

B.3) $\theta^{10,00}((1-S_2(0))(S_1(1)-S_1(0))=1),\quad \theta^{10,00}((1-S_2(1))(S_1(1)-S_1(0))=1),$

B.4) $\theta^{01,00}((1-S_1(0))(S_2(1)-S_2(0))=1),\quad \theta^{01,00}((1-S_1(0))(S_2(1)-S_2(0))=1).$

The proof is given in appendix A.10.

As shown in the different proofs, the effects are functions of some observable conditional expectations (an example is presented in section 6.1, the estimable form of the effects being the same in the dynamic case and in the other case). The conditioning events are all the possible values of the instrument in both periods: $\{(Z_1=1,Z_2=1),(Z_1=1,Z_2=0),$ $(Z_1=0,Z_2=1),(Z_1=0,Z_2=0)\}$. When the realizations of the instrument in both periods are correlated but not perfectly correlated, $0<p(Z_1\neq Z_2)<1$, our results still hold.

The most difficult part in applications is finding the instrument. The dynamic independence condition and the two period monotonicity condition are not more difficult to fulfil than the corresponding assumptions in the static case. For example, when investigating the effect of participating in a training program on the probability to be unemployed, the region of residence or the labor office which the individuals are assigned to could be good candidates. They can influence the participation in a training program, but have no influence on the probability to be unemployed (the outcome). However, two testable conditions exist that the instrument has to fulfil. Firstly the observed participation and the region of residence in the second period must be independent given the region of residence in period one. This rules out the moving due to a nonparticipation in a program (if the persons know that participation is easier in another region and this stays the same in the second period, they decide to move to this new region in the second period). Moreover, the outcome of period one should have no influence on the region of residence in the second period given the participation and the region of residence in period one. (The individuals should not move based on the fact that they have a bad outcome in the first period. It certainly is the most difficult testable assumption to fulfil as it is often the case that some persons move to another region because they are still unemployed and the other region seems more attractive.)

6.4) Robustness

Until now, we made an implicit exclusion restriction when we assumed the potential participation indicators in the second period to depend only on one instrument, namely the second one. But is this really necessary? If there is no such restriction, four potential participation indicators exist in the second period: $S_2(1,1), S_2(0,1), S_2(1,0), S_2(0,0)$, where the first number is the value of the instrument in the first period (or the first instrument) and the second one the value of the instrument in the second period. Now, it is extremely difficult to divide the population in the four categories of compliers, defiers, never-takers and always-takers in the second period. If we treat the instrument as an assignment to a program, its value in the first period corresponds to the assignment to a program for this period and its value in the second period corresponds to the assignment for this period. Thus, we can possibly regard the second-period compliers as those who participate if the instrument is 1 in the second period and who do not participate if the instrument is 0 in this period (independently of the value of the instrument in the first period). Another possibility is to look at a sequence of participations and a sequence of assignments. A complier would be a person who follows her sequence of assignments in all the periods. For a complier, the following pairs of events have to be true: $(S_1(1)=1, S_2(1,1)=1), (S_1(0)=0, S_2(0,1)=1), (S_1(1)=1, S_2(1,0)=0)$ and $(S_1(0)=0, S_2(0,0)=0)$.

With this new dependence, we need to restate the monotonicity condition.[11]

---

[11] Even if we have a monotonicity property for the first period variables, we do not obtain identification of all the effects. We do not need such condition for the possible identifications.

ASSUMPTION 15: (ONE PERIOD MONOTONICITY CONDITION-II):

$$p\left[S_2(z_1,1) \geq S_2(z_1,0)\right]=1 \text{ or } p\left[S_2(z_1,1) \leq S_2(z_1,0)\right]=1, \ \forall z_1.$$

Under this condition and one of the independence conditions previously made (plus assumption 14 if the independence condition is the dynamic one), only the effects that compare sequences with the same treatment in the first period are identified, but not for all subpopulations. The following effects are identified:

$$\theta_\tau^{11,10}(S_1(1)(S_2(1,1)-S_2(1,0))=1), \theta_\tau^{11,10}(S_1(0)(S_2(0,1)-S_2(0,0))=1),$$

$$\theta_\tau^{01,00}((1-S_1(1))(S_2(1,1)-S_2(1,0))=1), \theta_\tau^{01,00}((1-S_1(0))(S_2(0,1)-S_2(0,0))=1).$$

These results are proved in appendix A.11. It follows that our previous results are not robust to this change in the relation between the participation indicator in the second period and the instruments. Thus, the implicit exclusion restriction imposed on this relation is a necessary condition to obtain identification.

Before coming to the conclusions we present in table 1 a summary of the hypotheses and results obtained in this work. The last set of results presents only one combination of the monotonicity condition and an independence assumption, but other combinations are possible.

Table 1

7) Conclusion

In this article we examine the identifiability of different effects of sequences of programs using instrumental variables. More than one LATE can be defined. We consider different sets of assumptions and their implications on the identification of the parameters of interest. In a one-period framework with more than one program, identification is obtained for subpopulations defined by a pair-wise participation, i.e. they can only choose between two programs. Unfortunately, in a multi-period framework we cannot identify all possible effects even in a case with some independence condition between the periods. Nevertheless, the effects for individuals who comply each period are identified whatever the assumption set is. These sets of assumptions principally comprise a monotonicity hypothesis and some exclusion restrictions. When only one instrument is available in both periods, only few effects can be identified. These compare sequences which have the same participation in one of the two periods. Also, the affected populations are observed in the period with the same participation in both sequences and these populations comply in the other period. The effects for such individuals are not identifiable anymore when more than one instrument is available. Nonetheless, we can identify a lot of other different effects. The introduction of endogeneity in the dynamic case does not pose any additional problems. All the effects identified without this endogeneity are also identified with it. This paper covers a part of the problem of identification of dynamic treatment effects using instrumental variables. The LIV approach of Heckman and Vytlacil (1999) is not at all discussed and needs also to be translated into a dynamic context. Moreover, the IV approach allows the identification of more than some means. In a static framework the distribution (quantiles) of some treatment effects are also identified. This point is not yet investigated in the dynamic framework. It is also essential to see some applications which we intend to do this in the near future.

BIBLIOGRAPHY:

Abadie A. (2000): ”Semiparametric Instrumental Variable Estimation of Treatment Response Models”, National Bureau of Economic Research Technical Working Paper n° 260.

Abadie A. (2002): “Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models”, Journal of American Statistical Association, Vol. 97, 284-292.

Abadie A., J. D. Angrist and G. W. Imbens (2002): ”Instrumental Variables estimates of the Effect of Subsidized training on the Quantiles of Trainee Earnings”, Econometrica, Vol.70, n°1, 91-117.

Angrist J. D., G. W. Imbens and D. B. Rubin (1996): “Identification of Causal Effects Using Instrumental Variables”, Journal of American Statistical Association, Vol. 91, 444-458.

Angrist J. D. and A. B. Krueger (1999): "Empirical Strategies in Labor Economics," in O. Ashenfelter and D. Card (eds.), Handbook of Labor Economics, Vol. III A, chapter 23, 1277-1366.

Angrist J. D. and G. W. Imbens (1991):"Sources of Identifying Information in Evaluation Models", NBER Technical Working Paper, n° 117.

Arellano M. and B. Honoré (2002): “Panel Data Models: Some Recent Developments”, in J.J. Heckman and E. Leamer (eds.): Handbook of Econometrics, Vol. V, ch.53.

Dawid A.P. (1979): “Conditional Independence in Statistical Theory”, Journal of the Royal Statistical Society Series B, 41, 1-31.

Heckman J. J. (1997): "Instrumental Variables, A Study of Implicit Behavioral Assumptions", Journal of Human Resources, 32(3), 441-462.

Heckman J. J., R. J. LaLonde, and J. A. Smith (1999): "The Economics and Econometrics of Active Labor Market Programs", in O. Ashenfelter and D. Card (eds.): Handbook of Labor Economics, Vol. III A, 1865-2097, North-Holland, Amsterdam.

Heckman J. J., J. L. Tobias and E. Vytlacil (2000): “Simple Estimators for Treatment Parameters in a Latent Variables Framework with an Application to Estimating the Returns to Schooling”, National Bureau of Economic Research Working Paper 7950.

Heckman J. J. and E. Vytlacil (1999): “Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects”, Proceeding of the National Academy of Sciences 96, 4730-4734.

Heckman J. J. and E. Vytlacil (2000a): “The Relationship between Treatment Parameters within a Latent Variable Framework”, Economics Letters, 66, 33-39.

Heckman J. J. and E. Vytlacil (2000b): “Local Instrumental Variables”, in Hsiao, c., k. Morimune, and J. Powell (eds), Nonlinear Statistical Inference: Essays in Honor of Takeshi Amemiya, Cambridge: Cambridge Univrsity Press.

Imbens G. W. and J. D. Angrist (March 1994): " Identification and Estimation of Local Average Treatment Effects", Econometrica, vol. 62, n°2, 467 – 475.

Imbens G. W. (2000): “The Role of Propensity Score in Estimating Dose-Response Functions”, Biometrica, vol. 87, n°3, 706-710.

Lechner M. (2001): “Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption”, in: Lechner, M., Pfeiffer, F. (eds), Econometric Evaluation of Labour Market Policies, Heidelberg: Physica/Springer, 43-58.

Lechner M. and F. Pfeiffer, (eds., 2001), Econometric Evaluation of Active Labour Market Policies, Heidelberg: Physica.

Lechner M. and R. Miquel (2001): “A Potential Outcome Approach to Dynamic Programme Evaluation: Nonparametric Identification”, Working Paper n° 2001-07, University of St. Gallen.

Manski C. F. and J. V. Pepper (2000): “Monotone Instrumental Variables: With an Application to the Returns to Schooling”, Econometrica, Vol. 68, n°4, 997-1010.

Resnick S. I. (1999): A Probability Path, Birkhäuser Boston.

Robins J. M. (1986): "A new approach to causal inference in mortality studies with sustained exposure periods - Application to control of the healthy worker survivor effect." Mathematical Modelling, 7:1393-1512, with 1987 Errata to "A new approach to causal inference in mortality studies with sustained exposure periods - Application to control of the healthy worker survivor effect.” Computers and Mathematics with Applications, 14:917-921; 1987 Addendum to "A new approach to causal inference in mortality studies with sustained exposure periods - Application to control of the healthy worker survivor effect." Computers and Mathematics with Applications, 14:923-945; and 1987 Errata to "Addendum to ’A new approach to causal inference in mortality studies with sustained exposure periods - Application to control of the healthy worker survivor effect’." Computers and Mathematics with Applications, 18:477.

Robins J. M. (1989): "The Analysis of Randomized and Nonrandomized AIDS Treatment Trials Using a New Approach to Causal Inference in Longitudinal Studies", Sechrest, L., H. Freeman, A. Mulley (eds.), Health Service Research Methodology: A Focus on Aids, 113-159, Washington, D.C.: Public Health Service, National Center for Health Services Research.

Robins J. M. (1997): "Causal Inference from Complex Longitudinal Data. Latent Variable Modeling and Applications to Causality", in M. Berkane, ed., Lecture Notes in Statistics (120),. NY: Springer Verlag, pp. 69-117.

Robins J. M., S. Greenland and F. Hu (1999): "Estimation of the Causal Effect of a Time-varying Exposure on the Marginal Mean of a Repeated Binary Outcome", Journal of the American Statistical Association, 94, 687-700, with discussion.

Roy A.D. (1951): "Some Thoughts on the Distribution of Earnings", Oxford Economic Papers, 3, 135 – 146.

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

van den Berg G. J. (2001): "Duration Models: Specification, Identification, and Multiple Durations", in Heckman, James J., and Edward Leamer (eds.), Handbook of Econometrics, vol. V, Amsterdam: North-Holland.

Vytlacil E. (2002): “Independence, Monotonicity, and Latent Index Model: an Equivalence Result”, Econometrica, Vol 70, n°1, 331-341.

Appendix:

A.1) Proof of theorem 1:

We need to prove that $\theta_{\tau}^{(1.),(0.)}(S_1(1)-S_1(0)=1)$ is a function of observables. In the sample we can estimate the expectations of the outcome given a value of the instrument in period one. $E(S_1 \mid Z_1=1)$ and $E(S_1 \mid Z_1=0)$ are observable as well.

$$
\begin{aligned}
&E(Y_\tau \mid Z_1=1)-E(Y_\tau \mid Z_1=0) \\
&\qquad = E\left[S_1(1)Y_\tau^{(1.)}+(1-S_1(1))Y_\tau^{(0.)} \mid Z_1=1\right]-E\left[S_1(0)Y_\tau^{(1.)}+(1-S_1(0))Y_\tau^{(0.)} \mid Z_1=0\right] \\
&\qquad \underset{SIC}{=} E\left[S_1(1)Y_\tau^{(1.)}+(1-S_1(1))Y_\tau^{(0.)}\right]-E\left[S_1(0)Y_\tau^{(1.)}+(1-S_1(0))Y_\tau^{(0.)}\right] \\
&\qquad = E\left[(S_1(1)-S_1(0))(Y_\tau^{(1.)}-Y_\tau^{(0.)})\right] \\
&\qquad = p\left[S_1(1)-S_1(0)=1\right]E\left[Y_\tau^{(1.)}-Y_\tau^{(0.)} \mid S_1(1)-S_1(0)=1\right] \\
&\qquad\quad - p\left[S_1(1)-S_1(0)=-1\right]E\left[Y_\tau^{(1.)}-Y_\tau^{(0.)} \mid S_1(1)-S_1(0)=-1\right] \\
&\qquad \underset{OMC}{=} p\left[S_1(1)-S_1(0)=1\right]E\left[Y_\tau^{(1.)}-Y_\tau^{(0.)} \mid S_1(1)-S_1(0)=1\right] \\
&\qquad \underset{OMC}{=} \left\{E(S_1(Z_1) \mid Z_1=1)-E(S_1(Z_1) \mid Z_1=0)\right\}E\left[Y_\tau^{(1.)}-Y_\tau^{(0.)} \mid S_1(1)-S_1(0)=1\right]
\end{aligned}
$$

$$
\Rightarrow \quad E\left[Y_\tau^{(1.)}-Y_\tau^{(0.)} \mid S_1(1)-S_1(0)=1\right]=\frac{E(Y_\tau \mid Z_1=1)-E(Y_\tau \mid Z_1=0)}{E(S_1 \mid Z_1=1)-E(S_1 \mid Z_1=0)}.
$$

The effect equals a function of estimable expectations and thus is identified. q.e.d

A.2) Proof of theorem 2:

As in the previous proof, we need to prove that $\theta_{\tau}^{(k.),(l.)}(S_{kl}(1)-S_{kl}(0)=1, S_1=\{k,l\})$ is a function of observables.

$$
\begin{aligned}
&E(Y_\tau \mid Z_1=1, S_1=\{k,l\})-E(Y_\tau \mid Z_1=0, S_1=\{k,l\}) \\
&\qquad = E\left[\tilde{S}_{kl}(1)Y_\tau^{(k.)}+(1-\tilde{S}_{kl}(1))Y_\tau^{(l.)} \mid Z_1=1, S_1=\{k,l\}\right] \\
&\qquad\qquad\qquad - E\left[\tilde{S}_{kl}(0)Y_\tau^{(k.)}+(1-\tilde{S}_{kl}(0))Y_\tau^{(l.)} \mid Z_1=0, S_1=\{k,l\}\right] \\
&\underset{C-SIC}{=} E\left[\tilde{S}_{kl}(1)Y_\tau^{(k.)}+(1-S_{kl}(1))Y_\tau^{(l.)} \mid S_1=\{k,l\}\right]-E\left[\tilde{S}_{kl}(0)Y_\tau^{(k.)}+(1-\tilde{S}_{kl}(0))Y_\tau^{(l.)} \mid S_1=\{k,l\}\right]
\end{aligned}
$$

$$= E\left[(\tilde{S}_{kl}(1)-\tilde{S}_{kl}(0))(Y_\tau^{(k.)}-Y_\tau^{(l.)}) \mid S_1=\{k,l\}\right]$$

$$= p\left[\tilde{S}_{kl}(1)-\tilde{S}_{kl}(0)=1 \mid S_1=\{k,l\}\right] E\left[Y_\tau^{(k.)}-Y_\tau^{(l.)} \mid \tilde{S}_{kl}(1)-\tilde{S}_{kl}(0)=1, S_1=\{k,l\}\right]$$
$$- p\left[\tilde{S}_{kl}(1)-\tilde{S}_{kl}(0)=-1, S_1=\{k,l\}\right] E\left[Y_\tau^{(k.)}-Y_\tau^{(l.)} \mid \tilde{S}_{kl}(1)-\tilde{S}_{kl}(0)=-1, S_1=\{k,l\}\right]$$

$$\underset{C-OMC}{=} p\left[\tilde{S}_{kl}(1)-\tilde{S}_{kl}(0)=1 \mid S_1=\{k,l\}\right] E\left[Y_\tau^{(k.)}-Y_\tau^{(l.)} \mid \tilde{S}_{kl}(1)-\tilde{S}_{kl}(0)=1, S_1=\{k,l\}\right]$$

$$\underset{C-OMC}{=} \left\{E(\tilde{S}_{kl}(1) \mid Z_1=1, S_1=\{k,l\}) - E(\tilde{S}_{kl}(0) \mid Z_1=0, S_1=\{k,l\})\right\}$$
$$E\left[Y_\tau^{(k.)}-Y_\tau^{(l.)} \mid \tilde{S}_{kl}(1)-\tilde{S}_{kl}(0)=1, S_1=\{k,l\}\right]$$

$$\Rightarrow \quad E\left[Y_\tau^{(k.)}-Y_\tau^{(l.)} \mid \tilde{S}_{kl}(1)-\tilde{S}_{kl}(0)=1, S_1=\{k,l\}\right] = \frac{E(Y_\tau \mid Z_1=1, S_1=\{k,l\}) - E(Y_\tau \mid Z_1=0, S_1=\{k,l\})}{E(\tilde{S}_{kl} \mid Z_1=1, S_1=\{k,l\}) - E(\tilde{S}_{kl} \mid Z_1=0, S_1=\{k,l\})} \cdot$$

$$E\left[Y_\tau^{(k.)}-Y_\tau^{(l.)} \mid \tilde{S}_{kl}(1)-\tilde{S}_{kl}(0)=1, S_1=\{k,l\}\right] = \frac{E(Y_\tau \mid Z_1=1, S_1=\{k,l\}) - E(Y_\tau \mid Z_1=0, S_1=\{k,l\})}{p(S_1=k \mid Z_1=1, S_1=\{k,l\}) - p(S_1=k \mid Z_1=0, S_1=\{k,l\})}$$

$$= \frac{E(Y_\tau \mid Z_1=1, S_1=\{k,l\}) - E(Y_\tau \mid Z_1=0, S_1=\{k,l\})}{\frac{p(S_1=k \mid Z_1=1)}{p(S_1=k \mid Z_1=1)+p(S_1=l \mid Z_1=1)} - \frac{p(S_1=k \mid Z_1=0)}{p(S_1=k \mid Z_1=0)+p(S_1=l \mid Z_1=0)}}$$

The effect equals a function of estimable expectations and probabilities and thus is identified. q.e.d

A.3) Proof of theorem 3:

We only prove the theorem for the effect $\theta_\tau^{11,10}(S_1=1, S_2(1)-S_2(0)=1)$. The proof for the second effect, $\theta_\tau^{01,00}(S_1=0, S_2(1)-S_2(0)=1)$, follows exactly the same steps with $E(Y_\tau \mid S_1=0, Z_1=z_1) = E(S_2(z_1)Y_\tau^{01} + (1-S_2(z_1))Y_\tau^{01} \mid S_1=0, Z_1=z_1)$ as starting point.

$$E(Y_\tau \mid S_1=1, Z_1=1) - E(Y_\tau \mid S_1=1, Z_1=0)$$
$$= E\left[S_2(1)Y_\tau^{11} + (1-S_2(1))Y_\tau^{10} \mid S_1=1, Z_1=1\right] - E\left[S_2(0)Y_\tau^{11} + (1-S_2(0))Y_\tau^{10} \mid S_1=1, Z_1=0\right]$$

$$\underset{C-SIC}{=} E\left[S_2(1)Y_\tau^{11} + (1-S_2(1))Y_\tau^{10} \mid S_1 = 1\right] - E\left[S_2(0)Y_\tau^{11} + (1-S_2(0))Y_\tau^{10} \mid S_1 = 1\right]$$

$$= E\left[(S_2(1) - S_2(0))(Y_\tau^{11} - Y_\tau^{10}) \mid S_1 = 1\right]$$

$$\underset{C-OMC}{=} p\left[S_2(1) - S_2(0) = 1 \mid S_1 = 1\right] E\left[Y_\tau^{11} - Y_\tau^{10} \mid S_1 = 1, S_2(1) - S_2(0) = 1\right] \; .$$

We also have:

$$p\left[S_2(1) - S_2(0) = 1 \mid S_1 = 1\right] = 1 - p\left[S_2(1) = S_2(0) = 0 \mid S_1 = 1\right] - p\left[S_2(1) = S_2(0) = 1 \mid S_1 = 1\right]$$

$$\underset{C-SIC}{=} 1 - p\left[S_2(1) = S_2(0) = 0 \mid S_1 = 1, Z_1 = 1\right] - p\left[S_2(1) = S_2(0) = 1 \mid S_1 = 1, Z_1 = 0\right]$$

$$= 1 - p\left[S_2 = 0 \mid S_1 = 1, Z_1 = 1\right] + p\left[S_2(1) = 0, S_2(0) = 1 \mid S_1 = 1, Z_1 = 1\right]$$
$$- p\left[S_2 = 1 \mid S_1 = 1, Z_1 = 0\right] + p\left[S_2(1) = 0, S_2(0) = 1 \mid S_1 = 1, Z_1 = 0\right]$$

$$\underset{C-OMC}{=} 1 - p\left[S_2 = 0 \mid S_1 = 1, Z_1 = 1\right] - p\left[S_2 = 1 \mid S_1 = 1, Z_1 = 0\right]$$

$$= p\left[S_2 = 1 \mid S_1 = 1, Z_1 = 1\right] - p\left[S_2 = 1 \mid S_1 = 1, Z_1 = 0\right].$$

Thus $E\left[Y_\tau^{11} - Y_\tau^{10} \mid S_1 = 1, S_2(1) - S_2(0) = 1\right] = \dfrac{E(Y_\tau \mid S_1 = 1, Z_1 = 1) - E(Y_\tau \mid S_1 = 1, Z_1 = 0)}{E(S_2 \mid S_1 = 1, Z_1 = 1) - E(S_2 \mid S_1 = 1, Z_1 = 0)}$. q.e.d

A.4) Proof of theorem 4:

Similarly we only prove the theorem for one effect, $\theta_\tau^{11,01}(S_2 = 1, S_1(1) - S_1(0) = 1)$, the identification of the other effect following the same scheme.

$$E(Y_\tau \mid S_2 = 1, Z_1 = 1) - E(Y_\tau \mid S_2 = 1, Z_1 = 0)$$
$$= E\left[S_1(1)Y_\tau^{11} + (1-S_1(1))Y_\tau^{01} \mid S_2 = 1, Z_1 = 1\right] - E\left[S_1(0)Y_\tau^{11} + (1-S_1(0))Y_\tau^{01} \mid S_2 = 1, Z_1 = 0\right]$$

$$\underset{C-SIC}{=} E\left[S_1(1)Y_\tau^{11} + (1-S_1(1))Y_\tau^{01} \mid S_2 = 1\right] - E\left[S_1(0)Y_\tau^{11} + (1-S_1(0))Y_\tau^{01} \mid S_2 = 1\right]$$

$$= E\left[(S_1(1) - S_1(0))(Y_\tau^{11} - Y_\tau^{01}) \mid S_2 = 1\right]$$

$$\underset{C-OMC}{=} p\left[S_1(1)-S_1(0)=1 \mid S_2=1\right] E\left[Y_\tau^{11}-Y_\tau^{01} \mid S_2=1, S_1(1)-S_1(0)=1\right]$$

$$\underset{C-OMC}{=} \left\{p\left[S_1=1 \mid S_2=1, Z_1=1\right]-p\left[S_1=1 \mid S_2=1, Z_1=0\right]\right\} E\left[Y_\tau^{11}-Y_\tau^{01} \mid S_2=1, S_1(1)-S_1(0)=1\right]$$

$$\Rightarrow \quad E\left[Y_\tau^{11}-Y_\tau^{01} \mid S_2=1, S_1(1)-S_1(0)=1\right]=\frac{E(Y_\tau \mid S_2=1, Z_1=1)-E(Y_\tau \mid S_2=1, Z_1=0)}{E(S_1 \mid S_2=1, Z_1=1)-E(S_1 \mid S_2=1, Z_1=0)}.$$

q.e.d

A.5) Proof of theorem 5:

A) We only prove that $\theta_\tau^{11,10}(S_1=1, S_2(1)-S_2(0)=1)$ is identified.

$$E(S_1 Y_\tau \mid, Z=1)-E(S_1 Y_\tau \mid, Z=0)$$
$$=E\left[S_1(1)S_2(1)Y_\tau^{11}+S_1(1)(1-S_2(1))Y_\tau^{10} \mid Z=1\right]-E\left[S_1(0)S_2(0)Y_\tau^{11}+S_1(0)(1-S_2(0))Y_\tau^{10} \mid Z=0\right]$$

$$\underset{SIC-bis}{=} E\left[S_1(1)S_2(1)Y_\tau^{11}+S_1(1)(1-S_2(1))Y_\tau^{10}\right]-E\left[S_1(0)S_2(0)Y_\tau^{11}+S_1(0)(1-S_2(0))Y_\tau^{10}\right]$$

$$=E\left[(S_1(1)S_2(1)-S_1(0)S_2(0))(Y_\tau^{11}-Y_\tau^{10})\right]+E\left[(S_1(1)-S_1(0))Y_\tau^{10}\right]$$

$$\underset{S_1(1)=S_1(0)}{=} E\left[(S_1(1)S_2(1)-S_1(0)S_2(0))(Y_\tau^{11}-Y_\tau^{10})\right]$$

$$\underset{OMC}{=} E\left[Y_\tau^{11}-Y_\tau^{10} \mid S_1(1)S_2(1)-S_1(0)S_2(0)=1\right] p\left[S_1(1)S_2(1)-S_1(0)S_2(0)=1\right]$$

$$=E\left[Y_\tau^{11}-Y_\tau^{10} \mid S_1=1, S_2(1)-S_2(0)=1\right] p\left[S_2(1)-S_1(0)=1 \mid S_1=1\right] p(S_1=1).$$

The end of the proof is similar to the end of the previous theorems.

B) The proof is identical to the one of point A) with the periods reversed and thus is not presented.

A.6) Proof of theorem 6:

To simplify the proof of this theorem and the following ones we present intermediate results in lemma A1.

- *Lemma A1: under Assumption SIC-bis, the following equalities hold*

(0.1) $E\left[Y_\tau S_2 \mid Z_1=1, Z_2=1\right] - E\left[Y_\tau S_2 \mid Z_1=0, Z_2=1\right] = E\left[S_2(1)(S_1(1)-S_1(0))(Y_\tau^{11}-Y_\tau^{01})\right]$

(0.2) $E\left[Y_\tau S_2 \mid Z_1=1, Z_2=0\right] - E\left[Y_\tau S_2 \mid Z_1=0, Z_2=0\right] = E\left[S_2(0)(S_1(1)-S_1(0))(Y_\tau^{11}-Y_\tau^{01})\right]$

(0.3)
$$\begin{aligned}&E\left[Y_\tau S_2 \mid Z_1=1, Z_2=1\right] - E\left[Y_\tau S_2 \mid Z_1=1, Z_2=0\right]\\ &\quad = E\left[(Y_\tau^{11}-Y_\tau^{01})S_1(1)(S_2(1)-S_2(0))\right] + E\left[Y_\tau^{01}(S_2(1)-S_2(0))\right]\end{aligned}$$

(0.4)
$$\begin{aligned}&E\left[Y_\tau S_2 \mid Z_1=1, Z_2=1\right] - E\left[Y_\tau S_2 \mid Z_1=0, Z_2=0\right]\\ &\quad = E\left[(Y_\tau^{11}-Y_\tau^{01})(S_1(1)S_2(1)-S_1(0)S_2(0))\right] + E\left[Y_\tau^{01}(S_2(1)-S_2(0))\right]\end{aligned}$$

(0.5)
$$\begin{aligned}&E\left[Y_\tau S_2 \mid Z_1=1, Z_2=0\right] - E\left[Y_\tau S_2 \mid Z_1=0, Z_2=1\right]\\ &\quad = E\left[(Y_\tau^{11}-Y_\tau^{01})(S_1(1)S_2(0)-S_1(0)S_2(1))\right] - E\left[Y_\tau^{01}(S_2(1)-S_2(0))\right]\end{aligned}$$

Similar results are available for the conditional expectations of $Y_\tau S_1$, $Y_\tau(1-S_1)$, $Y_\tau(1-S_2)$ and Y_τ. They are not presented here but are available upon request.

PROOF: AS DEFINED IN ABADIE (2002, PROOF OF LEMMA 2.1), ASSUMPTION SIC-BIS IMPLIES THAT $Y^{jk}S_i(m),\ j,k,i,m=0,1,$ ARE INDEPENDENT OF Z_i. THE RESULTS FOLLOW BY DIRECT CALCULATIONS.

We will prove the identification of only one effect of theorem 6. The proofs for the identification of the other five effects follow exactly the same steps using different combinations of the results of lemma A1 as starting point.

To identify the effect $\theta^{11,00}((S_2(1)-S_2(0))(S_1(1)-S_1(0))=1)$, substracting the sum of (0.2) and $E\left[Y_\tau(1-S_1) \mid Z_1=1, Z_2=1\right] - E\left[Y_\tau(1-S_1) \mid Z_1=1, Z_2=0\right]$ from the sum of (0.1) and $E\left[Y_\tau(1-S_1) \mid Z_1=0, Z_2=1\right] - E\left[Y_\tau(1-S_1) \mid Z_1=0, Z_2=0\right]$, we get:

$$E\left[(S_2(1)-S_2(0))(S_1(1)-S_1(0))(Y_\tau^{11}-Y_\tau^{00}+Y_\tau^{00}-Y_\tau^{01})\right]\text{-}E\left[(S_2(1)-S_2(0))(S_1(1)-S_1(0))(Y_\tau^{00}-Y_\tau^{01})\right]$$
$$=E\left[Y_\tau S_2 \mid Z_1=1, Z_2=1\right]-E\left[Y_\tau S_2 \mid Z_1=0, Z_2=1\right]-E\left[Y_\tau S_2 \mid Z_1=1, Z_2=0\right]+E\left[Y_\tau S_2 \mid Z_1=0, Z_2=0\right]$$
$$-E\left[Y_\tau(1-S_1) \mid Z_1=1, Z_2=1\right]+E\left[Y_\tau(1-S_1) \mid Z_1=1, Z_2=0\right]+E\left[Y_\tau(1-S_1) \mid Z_1=0, Z_2=1\right]$$
$$-E\left[Y_\tau(1-S_1) \mid Z_1=0, Z_2=0\right]$$

$$E\left[(S_2(1)-S_2(0))(S_1(1)-S_1(0))(Y_\tau^{11}-Y_\tau^{00})\right]$$
$$\Rightarrow \quad =E\left[Y_\tau(S_2+S_1-1) \mid Z_1=1, Z_2=1\right]-E\left[Y_\tau(S_2+S_1-1) \mid Z_1=0, Z_2=1\right]$$
$$-E\left[Y_\tau(S_2+S_1-1) \mid Z_1=1, Z_2=0\right]+E\left[Y_\tau(S_2+S_1-1) \mid Z_1=0, Z_2=0\right]$$

$$E\left[Y_\tau^{11}-Y_\tau^{00} \mid (S_2(1)-S_2(0))(S_1(1)-S_1(0))=1\right]$$
$$\underset{TMC}{\Rightarrow} \quad =\frac{E\left[Y_\tau(S_2+S_1-1) \mid Z_1=1, Z_2=1\right]-E\left[Y_\tau(S_2+S_1-1) \mid Z_1=0, Z_2=1\right]}{p\left[(S_2(1)-S_2(0))(S_1(1)-S_1(0))=1\right]}$$
$$\frac{-E\left[Y_t(S_2+S_1-1) \mid Z_1=1, Z_2=0\right]+E\left[Y_t(S_2+S_1-1) \mid Z_1=0, Z_2=0\right]}{p\left[(S_2(1)-S_2(0))(S_1(1)-S_1(0))=1\right]};$$

It remains to be proved that $p\left[(S_2(1)-S_2(0))(S_1(1)-S_1(0))=1\right]$ is observable. We do it only for the complier probability of the effects in A). The proof for the other probabilities is similar and will be presented in detail in the proof of theorem 10 (in this case, it is a little more complicated due to the dynamic relation but the steps are similar).

$$p\left[S_2(1)-S_2(0)=1\right] \underset{TMC}{=} p\left[(S_2(1)-S_2(0))(S_1(1)-S_1(0))=1\right]$$
$$+p\left[(S_2(1)-S_2(0))=1 \mid S_1(1)=S_1(0)=1\right]p\left[S_1(1)=S_1(0)=1\right]$$
$$+p\left[(S_2(1)-S_2(0))=1 \mid S_1(1)=S_1(0)=0\right]p\left[S_1(1)=S_1(0)=0\right]$$

$$p\left[(S_2(1)-S_2(0))(S_1(1)-S_1(0))=1\right]=p\left[S_2(1)-S_2(0)=1\right]$$
$$\Rightarrow \quad -p\left[(S_2(1)-S_2(0))=1 \mid S_1(1)=S_1(0)=1\right]p\left[S_1(1)=S_1(0)=1\right]$$
$$-p\left[(S_2(1)-S_2(0))=1 \mid S_1(1)=S_1(0)=0\right]p\left[S_1(1)=S_1(0)=0\right]$$

So we have to prove that each of the terms on the right hand side of the equation is observable.

a) $p\left[S_2(1)-S_2(0)=1\right] \underset{TMC}{=} 1-p\left[S_2(1)=S_2(0)=0\right]-p\left[S_2(1)=S_2(0)=1\right]$

$$\underset{SIC-bis}{=} 1-p\left[S_2(1)=S_2(0)=0 \mid Z_2=1\right]-p\left[S_2(1)=S_2(0)=1 \mid Z_2=0\right]$$

$$\underset{TMC}{=} 1-p[S_2=0 \mid Z_2=1]-p[S_2=1 \mid Z_2=0]$$

$$= p[S_2=1 \mid Z_2=1]-p[S_2=1 \mid Z_2=0]$$

$$\Rightarrow \quad p[S_2(1)-S_2(0)=1] = E[S_2 \mid Z_2=1]-E[S_2 \mid Z_2=0].$$

b) $p[S_2(1)-S_2(0)=1 \mid S_1(1)=S_1(0)=1] \underset{SIC-bis}{=} p[S_2(1)-S_2(0)=1 \mid S_1(1)=S_1(0)=1, Z_1=0]$,

$$\begin{aligned} p[S_2(1)-S_2(0)=1 \mid S_1(1)=S_1(0)=1, Z_1=0] \underset{TMC}{=} & 1-p[S_2(1)=S_2(0)=0 \mid S_1=1, Z_1=0] \\ & -p[S_2(1)=S_2(0)=1 \mid S_1=1, Z_1=0] \end{aligned}$$

$$\begin{aligned} \underset{TMC}{\Leftrightarrow} p[S_2(1)-S_2(0)=1 \mid S_1=1, Z_1=0] \underset{SIC-bis}{=} & 1-p[S_2(1)=S_2(0)=0 \mid Z_2=1, S_1=1, Z_1=0] \\ & -p[S_2(1)=S_2(0)=1 \mid Z_2=0, S_1=1, Z_1=0] \end{aligned}$$

$$\underset{C}{=} 1-p[S_2=0 \mid Z_2=1, S_1=1, Z_1=0]-p[S_2=1 \mid Z_2=0, S_1=1, Z_1=0]$$

$$= p[S_2=1 \mid Z_2=1, S_1=1, Z_1=0]-p[S_2=1 \mid Z_2=0, S_1=1, Z_1=0]$$

$$= E[S_2 \mid Z_2=1, S_1=1, Z_1=0]-E[S_2 \mid Z_2=0, S_1=1, Z_1=0]$$

thus
$$\begin{aligned} & p[S_2(1)-S_2(0)=1 \mid S_1(1)=1, S_1(0)=1] \\ & = E[S_2 \mid Z_2=1, S_1=1, Z_1=0]-E[S_2 \mid Z_2=0, S_1=1, Z_1=0]. \end{aligned}$$

c) Following the proof in b), we get:

$$\begin{aligned} p[S_2(1)-S_2(0)=1 \mid S_1(1)=0, S_1(0)=0] = & E[S_2 \mid Z_2=1, S_1=0, Z_1=1] \\ & -E[S_2 \mid Z_2=0, S_1=0, Z_1=1]. \end{aligned}$$

d) $p[S_1(1)=S_1(0)=1] \underset{SIC-bis}{=} p[S_1(1)=S_1(0)=1 \mid Z_1=0] \underset{C}{=} p[S_1=1 \mid Z_1=0] = E[S_1 \mid Z_1=0]$.

e) $p[S_1(1)=S_1(0)=0] \underset{SIC-bis}{=} p[S_1(1)=S_1(0)=0 \mid Z_1=1] \underset{TMC}{=} p[S_1=0 \mid Z_1=1] = 1-E[S_1 \mid Z_1=1]$.

All the expectations are observable in the sample, therefore our effect is identified.

A.7) Proof of theorem 7:

The only difference from the previous proof or the only part where the independence assumption plays a role is in the proof of the observability of the probability to be a complier in both periods.

$$PCI, TMC \Rightarrow p\left[(S_2(1)-S_2(0))(S_1(1)-S_1(0))=1\right]=p\left[(S_2(1)-S_2(0))=1\right]p\left[(S_1(1)-S_1(0))=1\right]$$

thus following the proof of theorem 3 for each probability, we have :

$$\begin{aligned}
&E\left[Y_\tau^{11}-Y_\tau^{00} \mid (S_2(1)-S_2(0))(S_1(1)-S_1(0))=1\right] \\
&\quad=\frac{E\left[Y_\tau(S_2+S_1-1)\mid Z_1=1,Z_2=1\right]-E\left[Y_\tau(S_2+S_1-1)\mid Z_1=0,Z_2=1\right]}{\left\{E(S_2\mid Z_2=1)-E(S_2\mid Z_2=0)\right\}\left\{E(S_1\mid Z_1=1)-E(S_1\mid Z_1=0)\right\}} \\
&\qquad\qquad\frac{-E\left[Y_\tau(S_2+S_1-1)\mid Z_1=1,Z_2=0\right]+E\left[Y_\tau(S_2+S_1-1)\mid Z_1=0,Z_2=0\right]}{\left\{E(S_2\mid Z_2=1)-E(S_2\mid Z_2=0)\right\}\left\{E(S_1\mid Z_1=1)-E(S_1\mid Z_1=0)\right\}}.
\end{aligned}$$

q.e.d.

A.8) Proof of theorem 8:

As for theorem 6, we only present the proof of point A) of theorem 8. The identification of points B), C) and D) is proved exactly in the same way. The proof is based on the result of LEMMA A1.

A.1) The starting point is equation (0.1)

$$E\left[S_2(1)(S_1(1)-S_1(0))(Y_\tau^{11}-Y_\tau^{01})\right]=E\left[Y_\tau S_2\mid Z_1=1,Z_2=1\right]-E\left[Y_\tau S_2\mid Z_1=0,Z_2=1\right]$$

$$\underset{TMC}{\Rightarrow}\; E\left[Y_\tau^{11}-Y_\tau^{01}\mid S_2(1)(S_1(1)-S_1(0))=1\right]=\frac{E\left[Y_\tau S_2\mid Z_1=1,Z_2=1\right]-E\left[Y_\tau S_2\mid Z_1=0,Z_2=1\right]}{p\left[S_2(1)(S_1(1)-S_1(0))=1\right]}$$

$$\underset{PI}{\Rightarrow}\; E\left[Y_\tau^{11}-Y_\tau^{01}\mid S_2(1)(S_1(1)-S_1(0))=1\right]=\frac{E\left[Y_\tau S_2\mid Z_1=1,Z_2=1\right]-E\left[Y_\tau S_2\mid Z_1=0,Z_2=1\right]}{p\left[S_2(1)=1\right]p\left[(S_1(1)-S_1(0))=1\right]}$$

$$\underset{TMC}{\Rightarrow}\; E\left[Y_\tau^{11}-Y_\tau^{01}\mid S_2(1)(S_1(1)-S_1(0))=1\right]=\frac{E\left[Y_\tau S_2\mid Z_1=1,Z_2=1\right]-E\left[Y_\tau S_2\mid Z_1=0,Z_2=1\right]}{E\left[S_2\mid Z_2=1\right]\left\{E\left[S_1\mid Z_1=1\right]-E\left[S_1\mid Z_1=0\right]\right\}}.$$

All expectations on the right hand side of the equality are estimable in the sample.

A.2) The starting point is equation (0.2):

$$E\left[S_2(0)(S_1(1)-S_1(0))(Y_\tau^{11}-Y_\tau^{01})\right]=E\left[Y_\tau S_2 \mid Z_1=1, Z_2=0\right]-E\left[Y_\tau S_2 \mid Z_1=0, Z_2=0\right],$$

and then the proof ends in the same way as the previous point.

A.9) Proof of lemma 1

We prove that the assumptions SSIC and SIC-Bis are equivalent.

First, we prove that SSIC implies SIC-Bis:

It holds that for three random variables A,B,C, $(A,B)\coprod C \Leftrightarrow A\coprod C$ and $B\coprod C \mid A$. Thus given that if $Z_1=z_1$ then $S_1=S_1(z_1)$, b) and c) with $S_1(z_1)=A$ and $Z_2=C$ imply that $(S_1(\tilde{z}_1),S_2(z_2),Y_\tau^{S_2}),(S_1(1),S_1(0),S_2(0),S_2(1))\coprod Z_2 \mid Z_1=\tilde{z}_1$. For the same reason this result and a) imply that $(S_1(\tilde{z}_1),S_2(z_2),Y_\tau^{S_2}),(S_1(1),S_1(0),S_2(0),S_2(1))\coprod (Z_2,Z_1)$. We can repeat the same steps with z_1 instead of $\tilde{z}_1$ as value of the conditioning variable Z_1. We get $(S_1(z_1),S_2(z_2),Y_\tau^{S_2}),(S_1(1),S_1(0),S_2(0),S_2(1))\coprod (Z_2,Z_1)$ and this proves the first equivalence.

Second, we prove that SIC-Bis implies SSIC:

$$(S_1(z_1),S_2(z_2),Y_\tau^{S_2}),(S_1(0),S_1(1),S_2(0),S_2(1))\coprod (Z_1,Z_2)$$

$$\Rightarrow \begin{cases} (S_1(z_1),S_2(z_2),Y_\tau^{S_2}),(S_1(0),S_1(1),S_2(0),S_2(1))\coprod Z_1 \\ (S_1(z_1),S_2(z_2),Y_\tau^{S_2}),(S_1(0),S_1(1),S_2(0),S_2(1))\coprod Z_2 \mid Z_1=z_1 \end{cases}$$

Then $(S_1(z_1),S_2(z_2),Y_\tau^{S_2}),(S_1(0),S_1(1),S_2(0),S_2(1))\coprod Z_2 \mid Z_1=z_1$

$$\Rightarrow \begin{cases} (S_2(z_2),Y_\tau^{S_2})\coprod Z_2 \mid Z_1=z_1, S_1(z_1) \\ S_1(z_1)\coprod Z_2 \mid Z_1=z_1, \end{cases}$$

$\Rightarrow$ $(S_2(z_2),Y_\tau^{S_2})\coprod Z_2 \mid Z_1=z_1, S_1=s_1$.

The same arguments for the joint distribution of the potential participation indicators give part b) of SSIC.

A.10) Proof of theorem 10:

If the relations of Lemma A1 are still identified under DIC, the identification of the effects of interest is guaranteed. Thus, we just prove that the first relation of Lemma A1 is valid under the dynamic independence assumption.

$$
\begin{aligned}
&E\left[Y_\tau S_2 \mid Z_1=1, Z_2=1\right]-E\left[Y_\tau S_2 \mid Z_1=0, Z_2=1\right] \\
&\quad =E\left[S_1(1)S_2(1)Y_\tau^{11}+(1-S_1(1))S_2(1)Y_\tau^{01} \mid Z_1=1, Z_2=1\right] \\
&\qquad -E\left[S_1(0)S_2(1)Y_\tau^{11}+(1-S_1(0))S_2(1)Y_\tau^{01} \mid Z_1=0, Z_2=1\right]
\end{aligned}
$$

$$
\begin{aligned}
&=\underset{Y_1,S_1|Z_1,Z_2}{E}\left\{E\left[S_1(1)S_2(1)Y_\tau^{11}+(1-S_1(1))S_2(1)Y_\tau^{01} \mid Z_1=1, Z_2=1, Y_1, S_1\right] \mid Z_1=1, Z_2=1\right\} \\
&\quad -\underset{Y_1,S_1|Z_1,Z_2}{E}\left\{E\left[S_1(0)S_2(1)Y_\tau^{11}+(1-S_1(0))S_2(1)Y_\tau^{01} \mid Z_1=0, Z_2=1, Y_1, S_1\right] \mid Z_1=0, Z_2=1\right\}
\end{aligned}
$$

$$
\begin{aligned}
&\underset{DIC}{=}\underset{Y_1,S_1|Z_1,Z_2}{E}\left\{E\left[S_1(1)S_2(1)Y_\tau^{11}+(1-S_1(1))S_2(1)Y_\tau^{01} \mid Z_1=1, Y_1, S_1\right] \mid Z_1=1, Z_2=1\right\} \\
&\quad -\underset{Y_1,S_1|Z_1,Z_2}{E}\left\{E\left[S_1(0)S_2(1)Y_\tau^{11}+(1-S_1(0))S_2(1)Y_\tau^{01} \mid Z_1=0, Y_1, S_1\right] \mid Z_1=0, Z_2=1\right\}
\end{aligned}
$$

$$
\begin{aligned}
&\underset{\substack{Lemma\,1\ and\\ assumption\,14}}{=}\underset{S_1|Z_1}{E}\ \underset{Y_1|Z_1,S_1}{E}\left\{E\left[S_1(1)S_2(1)Y_\tau^{11}+(1-S_1(1))S_2(1)Y_\tau^{01} \mid Z_1=1, S_1, Y_1\right] \mid Z_1=1, S_1\right\} \\
&\quad -\underset{S_1|Z_1}{E}\ \underset{Y_1|Z_1,S_1}{E}\left\{E\left[S_1(0)S_2(1)Y_\tau^{11}+(1-S_1(0))S_2(1)Y_\tau^{01} \mid Z_1=0, S_1, Y_1\right] \mid Z_1=0, S_1\right\}
\end{aligned}
$$

$$
\begin{aligned}
&=E\left[S_1(1)S_2(1)Y_\tau^{11}+(1-S_1(1))S_2(1)Y_\tau^{01} \mid Z_1=1\right] \\
&\quad -E\left[S_1(0)S_2(1)Y_\tau^{11}+(1-S_1(0))S_2(1)Y_\tau^{01} \mid Z_1=0\right]
\end{aligned}
$$

$$
\underset{DIC}{=}E\left[S_1(1)S_2(1)Y_\tau^{11}+(1-S_1(1))S_2(1)Y_\tau^{01}\right]-E\left[S_1(0)S_2(1)Y_\tau^{11}+(1-S_1(0))S_2(1)Y_\tau^{01}\right]
$$

$$
=E\left[(S_1(1)-S_1(0))S_2(1)Y_\tau^{11}-(S_1(1)-S_1(0))S_2(1)Y_\tau^{01}\right]
$$

which proves that the first equation of lemma A1 is still valid.

For the effects in A) the last step consists of showing that $p\left[(S_2(1)-S_2(0))(S_1(1)-S_1(0))=1\right]$ is observable. We have :

$$
\begin{aligned}
p\left[S_2(1)-S_2(0)=1\right] \underset{TMC}{=}\ & p\left[(S_2(1)-S_2(0))(S_1(1)-S_1(0))=1\right] \\
&+p\left[(S_2(1)-S_2(0))=1 \mid S_1(1)=S_1(0)=1\right]p\left[S_1(1)=S_1(0)=1\right] \\
&+p\left[(S_2(1)-S_2(0))=1 \mid S_1(1)=S_1(0)=0\right]p\left[S_1(1)=S_1(0)=0\right]
\end{aligned}
$$

$$p\left[(S_2(1)-S_2(0))(S_1(1)-S_1(0))=1\right]=p\left[S_2(1)-S_2(0)=1\right]$$

$$\Rightarrow \qquad -p\left[(S_2(1)-S_2(0))=1\,|\,S_1(1)=S_1(0)=1\right]p\left[S_1(1)=S_1(0)=1\right]$$

$$-p\left[(S_2(1)-S_2(0))=1\,|\,S_1(1)=S_1(0)=0\right]p\left[S_1(1)=S_1(0)=0\right]$$

So we have to prove that each of the terms on the right hand side of the equation is observable.

a) $p\left[S_2(1)-S_2(0)=1\right]\underset{TMC}{=}1-p\left[S_2(1)=S_2(0)=0\right]-p\left[S_2(1)=S_2(0)=1\right]$

$$\underset{DIC}{=}1-p\left[S_2(1)=S_2(0)=0\,|\,Z_1=1\right]-p\left[S_2(1)=S_2(0)=1\,|\,Z_1=1\right]$$

$$=1-\underset{Y_1,S_1|Z_1=1}{E}\left[p\left[S_2(1)=S_2(0)=0\,|\,Y_1,S_1,Z_1=1\right]\right]-\underset{Y_1,S_1|Z_1=0}{E}\left[p\left[S_2(1)=S_2(0)=1\,|\,Y_1,S_1,Z_1=1\right]\right]$$

$$\underset{DIC}{=}1-\underset{Y_1,S_1|Z_1=1}{E}\left[p\left[S_2(1)=S_2(0)=0\,|\,Y_1,S_1,Z_1=1,Z_2=1\right]\right]$$

$$-\underset{Y_1,S_1|Z_1=0}{E}\left[p\left[S_2(1)=S_2(0)=1\,|\,Y_1,S_1,Z_1=1,Z_2=0\right]\right]$$

$$\underset{\substack{Lemma\,1\\ Ass.14}}{=}1-p\left[S_2(1)=S_2(0)=0\,|\,Z_2=1,Z_1=1\right]-p\left[S_2(1)=S_2(0)=1\,|\,Z_2=0,Z_1=1\right]$$

$$\underset{TMC}{=}1-p\left[S_2=0\,|\,Z_2=1,Z_1=1\right]-p\left[S_2=1\,|\,Z_2=0,Z_1=1\right]$$

$$=p\left[S_2=1\,|\,Z_2=1,Z_1=1\right]-p\left[S_2=1\,|\,Z_2=0,Z_1=1\right]$$

$$\Rightarrow \qquad p\left[S_2(1)-S_2(0)=1\right]=E\left[S_2\,|\,Z_2=1,Z_1=1\right]-E\left[S_2\,|\,Z_2=0,Z_1=1\right].$$

b) $p\left[S_2(1)-S_2(0)=1\,|\,S_1(1)=S_1(0)=1\right]\underset{DIC}{=}p\left[S_2(1)-S_2(0)=1\,|\,S_1(1)=S_1(0)=1,Z_1=0\right],$

$$p\left[S_2(1)-S_2(0)=1\,|\,S_1(1)=S_1(0)=1,Z_1=0\right]\underset{TMC}{=}1-p\left[S_2(1)=S_2(0)=0\,|\,S_1=1,Z_1=0\right]$$

$$-p\left[S_2(1)=S_2(0)=1\,|\,S_1=1,Z_1=0\right]$$

$$\underset{TMC}{\Leftrightarrow}p\left[S_2(1)-S_2(0)=1\,|\,S_1=1,Z_1=0\right]\underset{DIC}{=}1-\underset{Y_1|S_1,Z_1}{E}p\left[S_2(1)=S_2(0)=0\,|\,Z_2=1,Y_1,S_1=1,Z_1=0\right]$$

$$-\underset{Y_1|S_1,Z_1}{E}p\left[S_2(1)=S_2(0)=1\,|\,Z_2=0,Y_1,S_1=1,Z_1=0\right]$$

$$\underset{\substack{TMC\\Ass.14}}{=} 1 - p\left[S_2 = 0 \mid Z_2 = 1, S_1 = 1, Z_1 = 0\right] - p\left[S_2 = 1 \mid Z_2 = 0, S_1 = 1, Z_1 = 0\right]$$

$$= p\left[S_2 = 1 \mid Z_2 = 1, S_1 = 1, Z_1 = 0\right] - p\left[S_2 = 1 \mid Z_2 = 0, S_1 = 1, Z_1 = 0\right]$$

$$= E\left[S_2 \mid Z_2 = 1, S_1 = 1, Z_1 = 0\right] - E\left[S_2 \mid Z_2 = 0, S_1 = 1, Z_1 = 0\right]$$

thus $$\begin{aligned} &p\left[S_2(1) - S_2(0) = 1 \mid S_1(1) = 1, S_1(0) = 1\right] \\ &\qquad = E\left[S_2 \mid Z_2 = 1, S_1 = 1, Z_1 = 0\right] - E\left[S_2 \mid Z_2 = 0, S_1 = 1, Z_1 = 0\right]. \end{aligned}$$

c) Following the proof in b), we get:

$$\begin{aligned} p\left[S_2(1) - S_2(0) = 1 \mid S_1(1) = 0, S_1(0) = 0\right] = {}& E\left[S_2 \mid Z_2 = 1, S_1 = 0, Z_1 = 1\right] \\ & - E\left[S_2 \mid Z_2 = 0, S_1 = 0, Z_1 = 1\right]. \end{aligned}$$

d) $p\left[S_1(1) = S_1(0) = 1\right] \underset{DIC}{=} p\left[S_1(1) = S_1(0) = 1 \mid Z_1 = 0\right] \underset{TMC}{=} p\left[S_1 = 1 \mid Z_1 = 0\right] = E\left[S_1 \mid Z_1 = 0\right].$

e) $p\left[S_1(1) = S_1(0) = 0\right] \underset{DIC}{=} p\left[S_1(1) = S_1(0) = 0 \mid Z_1 = 1\right] \underset{TMC}{=} p\left[S_1 = 0 \mid Z_1 = 1\right] = 1 - E\left[S_1 \mid Z_1 = 1\right].$

All the expectations are observable in the sample, therefore our effect is identified.

For the effects in B) the last step consists in showing that $p\left[S_2(1)(S_1(1) - S_1(0)) = 1\right]$ is identified.

$$\begin{aligned} p\left[S_2(1)(S_1(1) - S_1(0)) = 1\right] = p\left[S_2(1) = 1\right] - {}& p\left[S_2(1) = 1, S_1(1) = S_1(0) = 1\right] \\ & - p\left[S_2(1) = 1, S_1(1) = S_1(0) = 0\right] \end{aligned}$$

a) $p\left[S_2(1) = 1\right] \underset{DIC}{=} p\left[S_2(1) = 1 \mid Z_1 = z_1\right]$

$$= \underset{Y_1, S_1 \mid Z_1 = z_1}{E}\left[p\left[S_2(1) = 1 \mid Y_1, Z_1 = z_1, S_1\right]\right]$$

$$\underset{DIC}{=} \underset{Y_1, S_1 \mid Z_1 = z_1}{E}\left[p\left[S_2(1) = 1 \mid Z_1 = z_1, Y_1, S_1, Z_2 = 1\right]\right]$$

$$\underset{\substack{Lemma 1\\Ass.14}}{=} p\left[S_2(1) = 1 \mid Z_1 = z_1, Z_2 = 1\right]$$

$$= p\left[S_2 = 1 \mid Z_1 = z_1, Z_2 = 1\right] = E\left[S_2 \mid Z_1 = z_1, Z_2 = 1\right].$$

b) $p\left[S_2(1) = 1, S_1(1) = S_1(0) = 1\right] = p\left[S_2(1) = 1 \mid S_1(1) = S_1(0) = 1\right] p\left[S_1(1) = S_1(0) = 1\right]$

$$\underset{DIC}{=} p\left[S_2(1) = 1 \mid S_1(1) = S_1(0) = 1, Z_1 = 0\right] p\left[S_1(1) = S_1(0) = 1 \mid Z_1 = 0\right]$$

$$\underset{C}{=} p\left[S_2(z_2) = 1 \mid S_1 = 1, Z_1 = 0\right] p\left[S_1 = 1 \mid Z_1 = 0\right]$$

$$\underset{DIC}{=} \underset{Y_1|S_1,Z_1}{E} p\left[S_2(1) = 1 \mid Z_2 = 1, Y_1, S_1 = 1, Z_1 = 0\right] E\left[S_1 \mid Z_1 = 0\right]$$

$$\underset{Ass.14}{=} p\left[S_2 = 1 \mid Z_2 = 1, S_1 = 1, Z_1 = 0\right] E\left[S_1 \mid Z_1 = 0\right]$$

$$= E\left[S_2 \mid Z_2 = 1, S_1 = 1, Z_1 = 0\right] E\left[S_1 \mid Z_1 = 0\right].$$

c) $p\left[S_2(1) = 1, S_1(1) = S_1(0) = 0\right] = p\left[S_2(1) = 1 \mid S_1(1) = S_1(0) = 0\right] p\left[S_1(1) = S_1(0) = 0\right]$

$$\underset{DIC}{=} p\left[S_2(1) = 1 \mid S_1(1) = S_1(0) = 0, Z_1 = 1\right] p\left[S_1(1) = S_1(0) = 0 \mid Z_1 = 1\right]$$

$$\underset{TMC}{=} p\left[S_2(1) = 1 \mid S_1 = 0, Z_1 = 1\right] p\left[S_1 = 0 \mid Z_1 = 1\right]$$

$$\underset{DIC}{=} \underset{Y_1|S_1,Z_1}{E} p\left[S_2(1) = 1 \mid Z_2 = 1, Y_1, S_1 = 0, Z_1 = 1\right]\left[1 - E\left(S_1 \mid Z_1 = 1\right)\right]$$

$$\underset{Ass.14}{=} p\left[S_2 = 1 \mid Z_2 = 1, S_1 = 0, Z_1 = 1\right]\left[1 - E\left(S_1 \mid Z_1 = 1\right)\right]$$

$$= E\left[S_2 \mid Z_2 = 1, S_1 = 0, Z_1 = 1\right]\left[1 - E\left(S_1 \mid Z_1 = 1\right)\right]. \qquad \text{q.e.d.}$$

A.11) Proof of the results of section 6.4:

We only prove the result for the first effect under the static independence condition (SIC–BIS). Using the other independence conditions leads to the same structure for the proof, except the additional use of $\underset{S_1|Z_1}{E}(.)$ and $\underset{Y_1|Z_1,Z_2,S_1}{E}(.)$. The identification of the other effects is proved in the same way replacing S_1 by $1-S_1$ or replacing $Z_1 = 1$ by $Z_1 = 0$ or both.

$$E(Y_\tau S_1 \mid Z_1=1, Z_2=1) - E(Y_\tau S_1 \mid Z_1=1, Z_2=0)$$
$$= E(S_1(1)S_2(1,1)Y_\tau^{11} + S_1(1)(1-S_2(1,1))Y_\tau^{10} \mid Z_1=1, Z_2=1)$$
$$- E(S_1(1)S_2(1,0)Y_\tau^{11} + S_1(1)(1-S_2(1,0))Y_\tau^{10} \mid Z_1=1, Z_2=0)$$

$$\underset{SIC-BIS}{=} E(S_1(1)S_2(1,1)Y_\tau^{11} + S_1(1)(1-S_2(1,1))Y_\tau^{10} - S_1(1)S_2(1,0)Y_\tau^{11} - S_1(1)(1-S_2(1,0))Y_\tau^{10})$$

$$= E(S_1(1)(S_2(1,1)-S_2(1,0))(Y_\tau^{11}-Y_\tau^{10}))$$

$$\underset{OMC-II}{=} E(Y_\tau^{11}-Y_\tau^{10} \mid S_1(1)(S_2(1,1)-S_2(1,0))=1)\,p(S_1(1)(S_2(1,1)-S_2(1,0))=1)$$

$$\Rightarrow\; E(Y_\tau^{11}-Y_\tau^{10} \mid S_1(1)(S_2(1,1)-S_2(1,0))=1) = \frac{E(Y_\tau S_1 \mid Z_1=1, Z_2=1) - E(Y_\tau S_1 \mid Z_1=1, Z_2=0)}{p(S_1(1)(S_2(1,1)-S_2(1,0))=1)}.$$

The identification of the denominator has already been proved in the other theorems. We just have to replace the monotonicity condition used by the OMC-II.

Table 1: Summary of Assumptions and identified effects

Assumptions (I: independence, M: monotonicity, A: additional)	Effects identified
I: $(S_1(z_t), Y_\tau^{S_t}), (S_1(1), S_1(0)) \coprod Z_1$, $M: p\left[S_1(z_1) \geq S_1(z_1')\right]=1$ or $p\left[S_1(z_1) \leq S_1(z_1')\right]=1$	$\theta_\tau^{(1.),(0.)}(S_1(1)-S_1(0)=1)$
I: $(\tilde{S}_{kl}(z_1), Y_\tau^k, Y_\tau^l), (\tilde{S}_{kl}(1), \tilde{S}_{kl}(0)) \coprod Z_1 \mid S_1 = \{k,l\}$ $M: p\left[\tilde{S}_{kl}(z_1) \geq \tilde{S}_{kl}(z_1') \mid S_1 = \{\mathrm{k,l}\}\right]=1$ or $\mathrm{p}\left[\tilde{S}_{kl}(z_1) \leq \tilde{S}_{kl}(z_1') \mid S_1 = \{\mathrm{k,l}\}\right]=1$	$\theta_\tau^{(k.),(l.)}(\tilde{S}_{kl}(1)-\tilde{S}_{kl}(0)=1, S_1=\{k,l\})$
I: $(S_t(z_1), Y_\tau^{S_2}), (S_t(z_1'), S_t(z_1)) \coprod Z_1 \mid S_{t'} = s_{t'}$ $M: p\left[S_2(z_1) \geq S_2(z_1') \mid S_1 = s_1\right]=1$ or $p\left[S_2(z_1) \leq S_2(z_1') \mid S_1 = s_1\right]=1$	$\theta_\tau^{11,10}(S_1=1, S_2(1)-S_2(0)=1)$ $\theta_\tau^{01,00}(S_1=0, S_2(1)-S_2(0)=1)$
I: $(S_t(z_1), Y_\tau^{S_2}), (S_t(z_1'), S_t(z_1)) \coprod Z_1 \mid S_{t'} = s_{t'}$ $M: p\left[S_1(z_1) \geq S_1(z_1') \mid S_2 = s_2\right]=1$ or $p\left[S_1(z_1) \leq S_1(z_1') \mid S_2 = s_2\right]=1$	$\theta_\tau^{11,01}(S_2=1, S_1(1)-S_1(0)=1)$ $\theta_\tau^{10,00}(S_2=0, S_1(1)-S_1(0)=1)$
I: $(S_1(z_1), S_2(z_1), Y_\tau^{S_2}), (S_1(0), S_1(1), S_2(0), S_2(1)) \coprod Z_1$, A: $p\left[S_1(1)=S_1(0)\right]=1$ $M: p\left[S_2(z_1) \geq S_2(z_1')\right]=1$ or $p\left[S_2(z_1) \leq S_2(z_1')\right]=1$	$\theta_\tau^{11,10}(S_1=1, S_2(1)-S_2(0)=1)$ $\theta_\tau^{01,00}(S_1=0, S_2(1)-S_2(0)=1)$
I: $(S_1(z_1), S_2(z_1), Y_\tau^{S_2}), (S_1(0), S_1(1), S_2(0), S_2(1)) \coprod Z_1$, A: $p\left[S_2(1)=S_2(0)\right]=1$ $M: p\left[S_1(z_1) \geq S_1(z_1')\right]=1$ or $p\left[S_1(z_1) \leq S_1(z_1')\right]=1$	$\theta_\tau^{11,01}(S_2=1, S_1(1)-S_1(0)=1)$ $\theta_\tau^{10,00}(S_2=0, S_1(1)-S_1(0)=1)$
I: $(S_1(z_1), S_2(z_2), Y_\tau^{S_2}), (S_1(0), S_1(1), S_2(0), S_2(1)) \coprod (Z_1, Z_2)$ M: $p\left[S_1(1) \geq S_1(0)\right]=1$ or $p\left[S_1(1) \leq S_1(0)\right]=1$, $p\left[S_2(1) \geq S_2(0)\right]=1$ or $p\left[S_2(1) \leq S_2(0)\right]=1$, $p\left[S_1(1) \geq S_1(0), S_2(1) \geq S_2(0)\right]=1$ or $p\left[S_1(1) \leq S_1(0), S_2(1) \geq S_2(0)\right]=1$, or $p\left[S_1(1) \geq S_1(0), S_2(1) \leq S_2(0)\right]=1$ or $p\left[S_1(1) \leq S_1(0), S_2(1) \leq S_2(0)\right]=1$	$\theta^{ij,kl}((S_2(1)-S_2(0))(S_1(1)-S_1(0))=1)$ $\theta^{11,01}(S_2(0)(S_1(1)-S_1(0))=1), \theta^{11,01}(S_2(1)(S_1(1)-S_1(0))=1),$ $\theta^{11,10}(S_1(0)(S_2(1)-S_2(0))=1), \theta^{11,10}(S_1(1)(S_2(1)-S_2(0))=1),$ $\theta^{10,00}((1-S_2(0))(S_1(1)-S_1(0))=1), \theta^{10,00}((1-S_2(1))(S_1(1)-S_1(0))=1),$ $\theta^{01,00}((1-S_1(0))(S_2(1)-S_2(0))=1), \theta^{01,00}((1-S_1(1))(S_2(1)-S_2(0))=1).$

Assumptions (I: independence, M: monotonicity, A: additional)	**Effects identified**
I: $(S_1(z_1),S_2(z_2),Y_\tau^{s_2}),(S_1(0),S_1(1),S_2(0),S_2(1))\coprod(Z_1,Z_2)$ M: $p[S_1(1)\ge S_1(0)]=1$ or $p[S_1(1)\le S_1(0)]=1$, $p[S_2(1)\ge S_2(0)]=1$ or $p[S_2(1)\le S_2(0)]=1$ $p[S_1(1)\ge S_1(0),S_2(1)\ge S_2(0)]=1$ or $p[S_1(1)\le S_1(0),S_2(1)\ge S_2(0)]=1$, or $p[S_1(1)\ge S_1(0),S_2(1)\le S_2(0)]=1$ or $p[S_1(1)\le S_1(0),S_2(1)\le S_2(0)]=1$, A: $S_1(1)-S_1(0)\coprod S_2(1)-S_2(0)$	$\theta_\tau^{ij,kl}((S_2(1)-S_2(0))(S_1(1)-S_1(0))=1)$
I: $(S_1(z_1),S_2(z_2),Y_\tau^{s_2}),(S_1(0),S_1(1),S_2(0),S_2(1))\coprod(Z_1,Z_2)$ M: $p[S_1(1)\ge S_1(0)]=1$ or $p[S_1(1)\le S_1(0)]=1$ $p[S_2(1)\ge S_2(0)]=1$ or $p[S_2(1)\le S_2(0)]=1$ $p[S_1(1)\ge S_1(0),S_2(1)\ge S_2(0)]=1$ or $p[S_1(1)\le S_1(0),S_2(1)\ge S_2(0)]=1$, or $p[S_1(1)\ge S_1(0),S_2(1)\le S_2(0)]=1$ or $p[S_1(1)\le S_1(0),S_2(1)\le S_2(0)]=1$ A: $S_1(1)-S_1(0)\coprod S_2(z_2)$ $z_2=0,1$, $S_2(1)-S_2(0)\coprod S_1(z_1)$ $z_1=0,1$	$\theta^{11,01}(S_2(0)(S_1(1)-S_1(0))=1)$, $\theta^{11,10}(S_1(0)(S_2(1)-S_2(0))=1)$, $\theta^{11,01}(S_2(1)(S_1(1)-S_1(0))=1)$, $\theta^{11,10}(S_1(1)(S_2(1)-S_2(0))=1)$, $\theta^{10,00}((1-S_2(0))(S_1(1)-S_1(0))=1),\theta^{10,00}((1-S_2(1))(S_1(1)-S_1(0))=1)$, $\theta^{01,00}((1-S_1(0))(S_2(1)-S_2(0))=1),\theta^{01,00}((1-S_1(1))(S_2(1)-S_2(0))=1)$.
I: $(S_1(z_1),S_2(z_2),Y_\tau^{s_2}),(S_1(0),S_1(1),S_2(0),S_2(1))\coprod Z_1$ $(S_1(1),S_1(0))\coprod Z_2\mid Z_1=\tilde z_1$ $(S_2(z_2),Y_\tau^{s_2}),(S_2(0),S_2(1))\coprod Z_2\mid S_1=s_1,Z_1=\tilde z_1$ M: $p[S_1(1)\ge S_1(0)]=1$ or $p[S_1(1)\le S_1(0)]=1$, $p[S_2(1)\ge S_2(0)]=1$ or $p[S_2(1)\le S_2(0)]=1$, $p[S_1(1)\ge S_1(0),S_2(1)\ge S_2(0)]=1$ or $p[S_1(1)\le S_1(0),S_2(1)\ge S_2(0)]=1$, or $p[S_1(1)\ge S_1(0),S_2(1)\le S_2(0)]=1$ or $p[S_1(1)\le S_1(0),S_2(1)\le S_2(0)]=1$	$\theta^{ij,kl}((S_2(1)-S_2(0))(S_1(1)-S_1(0))=1)$ $\theta^{11,01}(S_2(0)(S_1(1)-S_1(0))=1),\theta^{11,01}(S_2(1)(S_1(1)-S_1(0))=1)$, $\theta^{11,10}(S_1(0)(S_2(1)-S_2(0))=1),\theta^{11,10}(S_1(1)(S_2(1)-S_2(0))=1)$, $\theta^{10,00}((1-S_2(0))(S_1(1)-S_1(0))=1),\theta^{10,00}((1-S_2(1))(S_1(1)-S_1(0))=1)$, $\theta^{01,00}((1-S_1(0))(S_2(1)-S_2(0))=1),\theta^{01,00}((1-S_1(1))(S_2(1)-S_2(0))=1)$.
I: $(S_1(z_1),S_2(z_2),Y_\tau^{s_2}),(S_1(0),S_1(1),S_2(0),S_2(1))\coprod Z_1, S_1(z_1)\coprod Z_2\mid Z_1=\tilde z_1$ $(S_2(z_2),Y_\tau^{s_2}),(S_2(0),S_2(1))\coprod Z_2\mid Y_1=y_1,S_1=s_1,Z_1=\tilde z_1$ M: $p[S_1(1)\ge S_1(0)]=1$ or $p[S_1(1)\le S_1(0)]=1$, $p[S_2(1)\ge S_2(0)]=1$ or $p[S_2(1)\le S_2(0)]=1$, $p[S_1(1)\ge S_1(0),S_2(1)\ge S_2(0)]=1$ or $p[S_1(1)\le S_1(0),S_2(1)\ge S_2(0)]=1$, or $p[S_1(1)\ge S_1(0),S_2(1)\le S_2(0)]=1$ or $p[S_1(1)\le S_1(0),S_2(1)\le S_2(0)]=1$ A: $Y_1\coprod Z_2\mid Z_1,S_1$	$\theta^{ij,kl}((S_2(1)-S_2(0))(S_1(1)-S_1(0))=1)$ $\theta^{11,01}(S_2(0)(S_1(1)-S_1(0))=1),\theta^{11,01}(S_2(1)(S_1(1)-S_1(0))=1)$, $\theta^{11,10}(S_1(0)(S_2(1)-S_2(0))=1),\theta^{11,10}(S_1(1)(S_2(1)-S_2(0))=1)$, $\theta^{10,00}((1-S_2(0))(S_1(1)-S_1(0))=1),\theta^{10,00}((1-S_2(1))(S_1(1)-S_1(0))=1)$, $\theta^{01,00}((1-S_1(0))(S_2(1)-S_2(0))=1),\theta^{01,00}((1-S_1(1))(S_2(1)-S_2(0))=1)$.
M: $p[S_2(z_1,1)\ge S_2(z_1,0)]=1$ or $p[S_2(z_1,1)\le S_2(z_1,0)]=1$ I: $(S_1(z_1),S_2(z_1,z_2),Y_\tau^{s_2}),(S_1(0),S_1(1),S_2(z_1,0),S_2(z_1,1))\coprod(Z_1,Z_2)$	$\theta_\tau^{11,10}(S_1(1)(S_2(1,1)-S_2(1,0))=1),\theta_\tau^{11,10}(S_1(0)(S_2(0,1)-S_2(0,0))=1)$, $\theta_\tau^{01,00}((1-S_1(1))(S_2(1,1)-S_2(1,0))=1),\theta_\tau^{01,00}((1-S_1(0))(S_2(0,1)-S_2(0,0))=1)$.