Instrumental Variables

Yona Rubinstein

July 2016

The Limitation of Panel Data

- So far we learned how to account for selection on time invariant factors using fixed effects.
- Indeed over-weight people are more likely to be on diet than others. And indeed accounting for their time invariant factors conditions out some of the bias.
- Yet even over-weight people choose whether and when to practice diet or not.
- The causal interpretation of the before-after comparisons, that is fixed effect estimates, requires the diet regimes to be independent of developments in persons weight.
- If people start a diet when they gain weight and stop when they already lost enough then fixed effect estimates might be exaggerating the impact of diet on weight.
- Hence, the causal interpretation of fixed effect models is not robust to selection on shocks and/or time varying unobservables.

When Do We Start a Diet?

- Why people go on diet in a particular day?
- It might be not more than a coincidence. The particular timing in that case mimics a randomized trial.
- Yet, it might also reflect particular circumstances. For example a recent weight gain or a particular "need" to lose a few pounds.
- In that the case exact timing is no longer independent of subject's weight (gain). And the exact timing does not mimic a randomized trial, even holding subjects' fixed effects constant.
- If people start a diet when they gain a few extra pounds, above and beyond their "long term" weight, then the regression coefficient of weight on diet might overstate (understate) the causal impact of keeping diet on weight.

The Causal Model

- To illustrate let's consider the following model:

$$Y_{it} = \beta_0 + \beta_D D_{it} + U_{it}, \tag{1}$$

  where $D_{it} = 1$ if person i is on diet and 0 otherwise.
- Let β_D denote the causal impact of diet on weight.
- Let's allow the person-specific error term to reflect time invariant factors shaping person's i "long term" weight (θ_i) and person-time specific fluctuations around her/his mean:

$$U_{it} = \theta_i + \varepsilon_{it}, \tag{2}$$

  where ε_{it} reflects the latter.
- Equation (1) is a version of the linear causal model. The error-time term in this equation (ε_{it}) is the part of potential outcome left over after controlling for θ_i.
- This error term is uncorrelated with diet by assumption. If this assumption turns out to be correct, the population regression of Y_{it} on D_{it} and θ_i generates the causal impact of diet on weight.

The Concern

- The problem we initially want to tackle is how to estimate the β_D when the reasons that people start diet on a particular timing are unobserved.
- The key concern is that conditional on persons fixed effects (θ_i), persons' diet status in time t is correlated with the fluctuations (ε_{it}) in their weight.
- In that case, the regression coefficient, without controlling for person fixed effect is:

$$\beta_D^{OLS} = \beta_D + (\theta_1 - \theta_0) + (\varepsilon_1 - \varepsilon_0), \tag{3}$$

where:

$$\begin{aligned} \varepsilon_0 &= E\left(\varepsilon_{it}^0 | D_{it} = 0\right) = E\left(\varepsilon_{it} | D_{it} = 0\right); \\ \varepsilon_1 &= E\left(\varepsilon_{it}^1 | D_{it} = 1\right) = E\left(\varepsilon_{it} | D_{it} = 1\right). \end{aligned}$$

- Controlling for person fixed effect we eliminate part of the bias $(\theta_1 - \theta_0)$ but not necessarily all. The fixed effect estimator equals to:

The Bias

- The problems are that: the timing is not random!
- If people go on diet when they gain weight and stop their diet when they lose weight then:

$$\begin{aligned} \varepsilon_0 &< 0 \\ \varepsilon_1 &> 0 \end{aligned}$$

- In that case:

$$\beta_D^{FE} = \beta_D + (\varepsilon_1 - \varepsilon_0) > \beta_D. \tag{5}$$

- If people start a diet when they gain weight then the regression coefficient β_D^{FE} under estimates the causal impact of diet on weight.
- *Instrumental Variables* method (hereafter IV) can be used to estimate β_D.

The Source of the Problem

- The problem we face is that **subjects** make *choices* and that these *choices* **are not** *independent* of the **outcomes of interest**.
- Subjects' choices are influenced by two type of factors: (i) the *outcome of interest*, (ii) others **independent** from the *outcome of interest*.
- Let's assume that there is a variable Z_{it}, observed by the eonometrician, that is correlated with (ii) and uncorrelated with (i).
- A mathematical representation of subjects' choices is spelled out below:

$$D_{it} = \alpha_0 + \alpha_Z Z_{it} + \alpha_U U_{it} + \epsilon_{it}, \tag{6}$$

- Where ϵ_{it} is by construction a "*left-over*" uncorrelated with Z_{it} or U_{it}.
- Clearly, D_{it} is *"contaminated"* by U_{it}, that is, as long as:

$$\alpha_U \neq 0.$$

Decomposition of the Choice / Treatment Status Equation

- Let's consider the following decomposition of the choice / treatment status equation into the three components in (6):

$$D_{it} = \alpha_0 + \underbrace{\alpha_Z Z_{it}}_{\substack{\text{the "effect"} \\ \text{of an} \\ \textit{exogenous} \\ \text{source } Z_{it} \\ \text{on } D_{it}}} + \underbrace{\alpha_U U_{it}}_{\substack{\text{the "effect"} \\ \text{of an} \\ \textit{endogenous} \\ \text{source } U_{it} \\ \text{on } D_{it}}} + \underbrace{\epsilon_{it}}_{\substack{\text{"left-over"} \\ \text{uncorrelated} \\ \text{with } Z_{it} \\ \text{or } U_{it}}}. \quad (7)$$

- We have a problem using D_{it} to identify the impact of treatment (D_{it}) as long as D_{it} is *"contaminated"* by U_{it}:

$$\alpha_U \neq 0.$$

A Solution

- An intuitive solution will be to decompose D_{it} into these parts.
- If we could condition out U_{it}, that is to have a "clean" from U_{it} measure of D_{it} we could obtain a consistent ("good") estimate of the causal impact of D on Y.
- For example, if we had access to the variable Z_{it} we could decompose D_{it} into two parts:

$$D_{it} = \underbrace{\alpha_0 + \alpha_Z Z_{it}}_{\text{"\underline{\textit{clean}}"}} + \underbrace{\alpha_U U_{it} + \epsilon_{it}}_{\text{"\underline{\textit{contaminated}}"}} \tag{8}$$

- And use only the *first part* to estimate the causal impact of D on Y.
- This is known as the <u>*Instrumental Variables / IV*</u> method.

Summarizing

- So far we learned how to account for selection on time invariant factors using fixed effects.
- Indeed over-weight people are more likely to be on diet than others. And indeed accounting for their time invariant factors conditions out some of the bias.
- Yet even over-weight people choose whether and when to practice diet or not.
- The causal interpretation of the before-after comparisons, that is fixed effect estimates, requires the diet regimes to be independent of developments in persons weight.
- If people start a diet when they gain weight and stop when they already lost enough then fixed effect estimates might be exaggerating the impact of diet on weight.
- Hence, the causal interpretation of fixed effect models is not robust to selection on shocks and/or time varying unobservables.

Back to the Basics

- The treatment effects literature is about how some outcome of interest, such as earnings, is affected by some *treatment*, for instance working in the financial industry.
- Although treatment effects must be related to *structural models*, where the outcome of interest is the left hand side variable and the treatment is a right-hand side variable, the treatment effect models have a terminology and set up all their own though.
- Notation. As in our previous notes let i index individuals and D_i denote a treatment indicator, equal to 1 if a person is treated, and equal to 0 otherwise (we can use of course other letters).
- $D_i = 1$ indicates exposure to treatment, for example if the first born child lives on a 3 children family and 0 if she lives in 2 children family.
- Y_i^0 denotes the potential outcome that would occur when person i is not treated ($D_i = 0$); Y_i^1 denotes the potential outcome that would occur when person i is not treated ($D_i = 1$).
- The "problem" these are not both observed.

Treatment Effects and Observed Outcomes

- Recall that the treatment effect for individual i is:

$$\Delta_i \equiv Y_i^1 - Y_i^0 = \beta_{Di}. \tag{9}$$

- If the treatment has the same impact on all then:

$$\beta_{Di} = \beta_D. \tag{10}$$

- Let's start with the simplest case of homogeneous/common treatment effect, that is $\beta_{Di} = \beta_D$.
- In this case the outcomes are:

$$\begin{aligned} Y_i^0 &= \beta_0 + U_i \\ Y_i^1 &= \beta_0 + \beta_D + U_i \end{aligned} \tag{11}$$

- The observed outcome will be:

$$\begin{aligned} Y_{0i} &= (\beta_0 + U_0) + \varepsilon_i \\ Y_{1i} &= (\beta_0 + U_1) + \beta_D + \varepsilon_i \end{aligned} \tag{12}$$

where ε_i is by group zero mean error term.

The OLS Regression Coefficient

- By substituting equation (12) into the observed outcome we receive that (for simplicity of notation we denote Y_{Di} as Y_i):

$$Y_i = D_i \left(\beta_0 + \beta_D + U_1 + \varepsilon_i\right) + \left(1 - D_i\right)\left(\beta_0 + U_0 + \varepsilon_i\right), \qquad (13)$$

- Which leads to the following linear reduced form model:

$$Y_{Di} = \beta_0 + \underbrace{\beta_D D_i}_{\text{the treatment effect}} + \underbrace{\left(U_0 + \left(U_1 - U_0\right) D_i + \varepsilon_i\right)}_{U_i = \text{the error term}}$$

- The OLS estimator for β_1 is:

$$\beta_D^{OLS} = E\left[Y_i | D_i = 1\right] - E\left[Y_i | D_i = 0\right] = \beta_D + \left(U_1 - U_0\right). \qquad (14)$$

- Our concern is that individuals are not randomly selected on the "unobservables", and therefore $\left(U_1 - U_0\right) \neq 0$.

Instrumental Variables: the Intuition beyond the IV Approach

- An intuitive solution will be to "find" a shifter in D_i that is not correlated with the "unobservables". Let's assume that we observe a variable Z_i that:

1. Influences individuals' treatment status, which means that Z_i and D_i are correlated;
2. Is exogenous to the outcome of interest (Y), which means that U_i and Z_i are not correlated.

- Then we could decompose the variation in the treatment status D_i into two parts by estimating using OLS the following regression model:

$$D_i = \alpha_0 + \alpha_Z Z_i + V_i \tag{15}$$

- And obtain:

$$\begin{aligned} \hat{D}_i(Z_i) &= \alpha_0 + \alpha_Z Z_i \\ V_i &= D_i - \hat{D}_i(Z_i) \end{aligned}$$

The Intuition beyond the IV Approach

- Note that given the assumptions (1) and (2) in the previous slide:

$$\begin{aligned} cov\left(\hat{D}_i, V_i\right) &= 0, \\ cov\left(\hat{D}_i, U_i\right) &= 0 \end{aligned}$$

- Therefore if D_i and U_i are correlated it must be the case that U_i and V_i are correlated. In that case by estimating equation (15) we decompose the variation in the treatment status into two parts: (i) the first part $\hat{D}_i\left(Z_i\right)$ that is "clean" from the "influence" of *Us* and (ii) the second part that is "contaminated" by *Us*:

$$D_i = \underbrace{\alpha_0 + \alpha_Z Z_i}_{\substack{\text{"clean"} \\ \hat{D}_i\left(Z_i\right)}} + \underbrace{\alpha_U U_i + \epsilon_i}_{\substack{\text{"contaminated"} \\ +V_i}} \tag{16}$$

- We'll use only the *first part* to estimate the causal impact of D on Y.
- This is known as the *Instrumental Variables / IV* method.

The IV Estimator

- Let's assume that we are interested in estimating the following causal relationship:

$$Y_i = \beta_0 + \beta_D D_i + U_i. \tag{17}$$

- We can obtain a consistent estimate of β_D using an instrument Z_i that fulfills the following conditions:

1. *Correlated* with treatment status D_i:

$$cov\left(Z_i, D_i\right) \neq 0$$

2. *Uncorrelated* with any other determinants of the dependent variables Y_i, that is uncorrelated with the error term in the outcome equation:

$$cov\left(Z_i, U_i\right) = 0$$

- The latter is called an *exclusion restriction* since Z_i can be said (assumed) to be excluded from the causal model of interest in equation (17).
- In this case we obtain a consistent estimate of β_D by using the IV

The IV Method using TSLS

- Let's assume that we are interested in estimating the following causal relationship:

$$Y_i = \beta_0 + \beta_D D_i + U_i. \tag{19}$$

- Assuming that we observe a valid instrument Z_i we can obtain a consistent estimate of β_D using the following two steps procedure:
- *First stage:* decompose D_i to the "clean" and the "contaminated" elements by estimating equation the following equation:

$$D_i = \alpha_0 + \alpha_Z Z_i + V_i. \tag{20}$$

- *Second stage:* impute $\hat{D}_i = \alpha_0 + \alpha_Z Z_i$ and use that to estimate the causal relationship between D_i by estimating the following equation:

$$Y_i = \beta_0 + \beta_D \tilde{D}_i + W_i. \tag{21}$$

Does the TSLS Provide the IV Estimator?

- What do we actually do when we estimate equation (21)? To address this question let's write it explicitly:

$$Y_i = \beta_0 + \beta_D (\alpha_0 + \alpha_Z Z_i) + W_i. \tag{22}$$

- Note that the OLS estimator of β_D in the equation above is, the IV/TSLS estimator is:

$$\beta_D^{IV} = \frac{\alpha_Z cov(Y_i, Z_i)}{\alpha_Z^2 var(Z_i)} = \frac{cov(Y_i, Z_i)}{\alpha_Z var(Z_i)}. \tag{23}$$

- Yet, since α_Z, estimated in the first stage, is:

$$\alpha_Z = \frac{cov(D_i, Z_i)}{var(Z_i)}. \tag{24}$$

- Then the IV/TSLS estimator equals:

$$\beta_D^{IV} = \frac{cov(Y_i, Z_i) / var(Z_i)}{cov(D_i, Z_i) / var(Z_i)} = \frac{cov(Y_i, Z_i)}{cov(D_i, Z_i)}. \tag{25}$$

Are We Sure that the IV Formula is Correct?

- What we actually do is estimating the following equation:

$$Y_i = \beta_0 + \beta_D \left(\alpha_0 + \alpha_Z Z_i\right) + V_i. \tag{26}$$

- And that in this case:

$$\beta_D = \frac{\alpha_Z cov\left(Y_i, Z_i\right)}{\alpha_Z^2 var\left(Z_i\right)} = \frac{cov\left(Y_i, Z_i\right)}{\alpha_Z var\left(Z_i\right)}. \tag{27}$$

- Yet, since:

$$\alpha_Z = \frac{cov\left(D_i, Z_i\right)}{var\left(Z_i\right)}. \tag{28}$$

- Then:

$$\beta_D^{IV} = \frac{cov\left(Y_i, Z_i\right) / var\left(Z_i\right)}{cov\left(D_i, Z_i\right) / var\left(Z_i\right)} = \frac{cov\left(Y_i, Z_i\right)}{cov\left(D_i, Z_i\right)}. \tag{29}$$

The IV Formula in the Binary Case

- Consider a case where both the treatment and the instrument take a binary form:

$$D_i = (0,1)$$
$$Z_i = (0,1)$$

- We already proved that the regression coefficient of a binary on a continuos variable equals to the gap in means, that is:

$$\frac{cov(Y_i, Z_i)}{var(Z_i)} = E(Y|Z=1) - E(Y|Z=0)$$
$$\frac{cov(D_i, Z_i)}{var(Z_i)} = E(D|Z=1) - E(D|Z=0)$$

- The IV estimator in this case equals to the gap in mean outcomes divided by the gap in mean treatment conditional on Z:

$$\beta_D^{IV} = \frac{E(Y|Z=1) - E(Y|Z=0)}{E(D|Z=1) - E(D|Z=0)}. \tag{30}$$

The Effect of Schooling on Earnings: Angrist and Krueger 1991

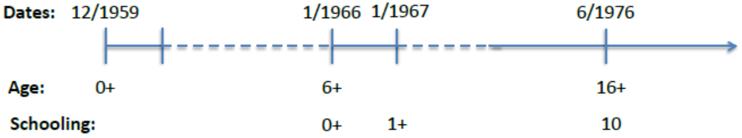
- One of the most studied public policies is the effect of schooling on labor market outcomes.
- In the reduced form sense we are interested in estimating the causal impact of schooling on earnings:

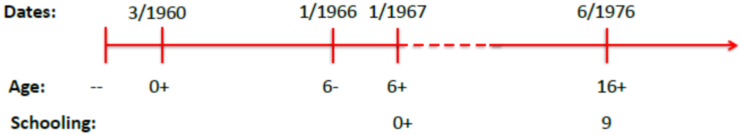
$$Y_i = \beta_0 + \beta_S S_i + U_i, \tag{31}$$

  where Y measures earnings (log hourly, for instance) and S stands for schooling.
- A key concern is is that years of schooling may be *endogenous*, with pre-schooling levels of ability affecting both schooling choices and earnings given education levels.
- Angrist and Krueger (1991), exploit variation in schooling levels that arise from differential impacts of compulsory schooling laws to estimate β_S.

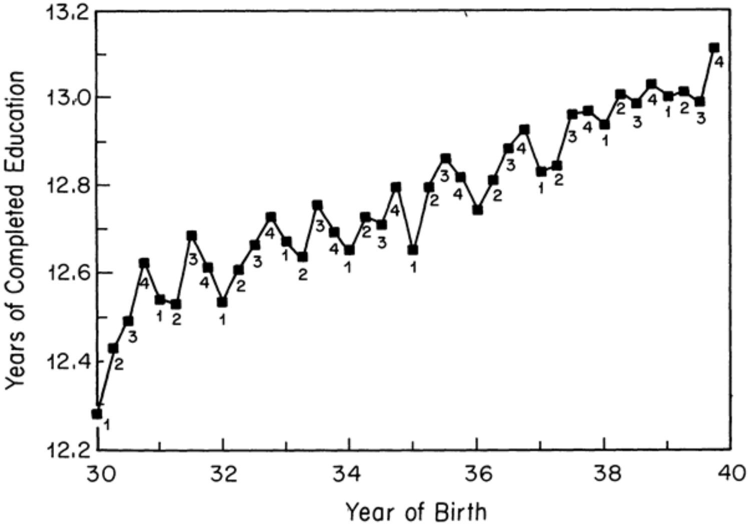
Angrist and Krueger 1991

- In the US a child is entitled to drop out of school once he/she turns 16. School districts typically require a child to have turned six by January 1st of the year the student enters school. Therefore they accumulate different lengths of schooling at the time they turn 16.





Angrist and Krueger 1991: First Stage



Angrist and Krueger 1991: First Stage

Angrist and Kruger Estimated the first stage using the following specfication:

$$S_i = \alpha_0 + \alpha_{Z1} Z_1 + \alpha_{Z2} Z_2 + \alpha_{Z3} Z_3 + V_i, \tag{32}$$

Outcome variable	Birth cohort	Mean	Quarter-of-birth effect[a] I	II	III	F-test[b] [P-value]
Total years of education	1930–1939	12.79	−0.124	−0.086	−0.015	24.9
			(0.017)	(0.017)	(0.016)	[0.0001]
	1940–1949	13.56	−0.085	−0.035	−0.017	18.6
			(0.012)	(0.012)	(0.011)	[0.0001]

Angrist and Krueger 1991: Second Stage for Men Born 1930-1939

Angrist and Kruger Estimated the fsecond stage using the following specfication:

$$Y_i = \beta_0 + \beta_S \hat{S}_i + V_i, \tag{33}$$

Independent variable	(1) OLS	(2) TSLS	(3) OLS	(4) TSLS
Years of education	0.0711	0.0891	0.0711	0.0760
	(0.0003)	(0.0161)	(0.0003)	(0.0290)
Race (1 = black)	—	—	—	—
SMSA (1 = center city)	—	—	—	—
Married (1 = married)	—	—	—	—
9 Year-of-birth dummies	Yes	Yes	Yes	Yes
8 Region-of-residence dummies	No	No	No	No
Age	—	—	−0.0772	−0.0801
			(0.0621)	(0.0645)
Age-squared	—	—	0.0008	0.0008
			(0.0007)	(0.0007)
χ^2 [dof]	—	25.4 [29]	—	23.1 [27]

The IV and the Wald Estimator

- Table 1 shows average years of education and average log earnings for individual born in the first and fourth quarter, using the 1980 census born between 1930 to 1939.

Table 1: Summary Statistics, a Subset of the Angrist and Krueger 1991 Data

Variable		1ST Q	2ND, 3RD & 4TH Q		Difference
Years of Schooling		12.6881	12.7969		0.1088
Weekly Earnings (log)		5.8916	5.9027		0.0111
OLS Estimate					**0.0709**
Ratio					**0.1020**

The IV estimator is:

$$\beta_D^{IV} = \frac{cov\left(Y_i, Z_i\right) / var\left(Z_i\right)}{cov\left(S_i, Z_i\right) / var\left(Z_i\right)} = \frac{Y_4 - Y_1}{S_4 - S_1} = \frac{0.0111}{0.1088} = 0.1020. \qquad (34)$$

Compliance

- While the insturment affects choices - subjects choose based on other factors too. To discuss the implications let's consider a case where both the treatment and the instrument take a binary form, that is, $Z_i = (0, 1)$ and $D_i = (0, 1)$.
- Specifically, we could disaggregate the population into four groups

1. Non-complier $(C_i = 0)$: always take treatment;
2. Non-complier $(C_i = 0)$: never take treatment;
3. Compliers $(C_i = 1)$: take treatment;
4. Compliers $(C_i = 1)$: do not take treatment;

- What does the IV identify?

Complience, Treatment and Outcomes

- Assuming random assinment:

1. For the non-compliers $(C_i = 0)$: the mean treatment status and outcomes are the same regrdless of the instrument:

$$\begin{aligned} E\left(Y_i | Z_i = 1\right) &= E\left(Y_i | Z_i = 0\right) \\ E\left(D_i | Z_i = 1\right) &= E\left(D_i | Z_i = 0\right) \end{aligned}$$

2. For the compliers $(C_i = 1)$:

$$\begin{aligned} E\left(Y_i | Z_i = 1\right) &\neq E\left(Y_i | Z_i = 0\right) \\ E\left(D_i | Z_i = 1\right) &\neq E\left(D_i | Z_i = 0\right) \end{aligned}$$

3. Note that

$$\begin{aligned} E\left(D_i | Z_i = 1 \& C_i = 1\right) &= 1 \\ E\left(D_i | Z_i = 0 \& C_i = 1\right) &= 0 \end{aligned}$$

Average Local Treatment Effect

- What is identified by the IV? Let $\bar{C}$ denote the proportion of the compliers. In that case the IV estimator can be expressed as:

$$\beta_D^{IV} = \frac{\left[E\left(Y_i^1|C_i = 1\right) - E\left(Y_i^0|C_i = 1\right)\right]}{\left[E\left(D_i|Z_i = 1\right) - E\left(D_i|Z_i = 0\right)\right]} \cdot \bar{C}, \tag{35}$$

since $\left[E\left(Y_i|Z_i = 1\&C_i = 0\right) - E\left(Y_i|Z_i = 0\&C_i = 0\right)\right] = 0$.

- Note that:

$$\bar{C} = \left[E\left(D_i|Z_i = 1\right) - E\left(D_i|Z_i = 0\right)\right]$$

- Denote by $Y^0_{C=1}$ and $Y^1_{C=1}$ the potential outcomes the compliers then:

$$\beta_D^{IV} = Y^1_{C=1} - Y^0_{C=1} \tag{36}$$

- Hence, the IV identifies the effect on the *compliers*. This is known as the *Local Average Treatment Effect* (**LATE**).
- If the treatment effect varies over groups this should be taken into account.

The IV Estimator and Treatment Effects

- Let's now allow for *henterogenoues treatment effects.*
- Specifically, let's consider a case with two effects: one for those influenced by the "instrument", β_{DC}, and another one for those who are not influenced by the "instrument", β_{DNC}.
- If, for instance, children from high income families benefit from schooling less than children from low income families then we might find the IV estimator of the returns to schooling to be higher than the OLS estimator, even if the OLS estimator is unbiased.
- This can happen if/when children from low income families are more influenced from the instrument (mandatory schooling; income shocks, tuitions).
- In that case:

$$\beta_D^{IV} = \beta_{DC} > \beta_D = \beta_D^{OLS}.$$

- Thus, we should be careful with the interpretation of the IV estimator even when the identifying assumptions hold.

Take Home Message

- Selection into "*treatment*" is **endogenous** to the *evaluated outcomes*.
- For instance:
  - Able students are less likely to dropout from school and earn more regardless of schooling.
  - Productive firms are more likely to get external finance and have higher profits than less productive firms regardless of the external finance.
- Selection into "treatment"is also determined by other factors **external** to the subject.
- The TSLS allows to control for self-selection into "treatment" by using external factors - the *instrumental variables* - influencing subjects selection into "treatment".
- The TSLS provide a **consistent estimate** for the *casual impact* on those subject manipulated by the **instrument**.