

## **INDEX**

Preface .....	2
Workshop report: “Causal inference and program evaluation” by Adriaan Soetevent .....	4
NAKE teaching program 2001-2002 .....	16
Program NAKE Day 2001 .....	18
Workshop report: "Rationality and social norms" by Katrin Oltmer .....	19

## Preface

I have just returned from my holiday and I am trying to get used to a normal work schedule again. There is not much time to gradually settle down however, as there is a lot of work to be done for the new academic year that is about to start. Luckily, I managed to complete the NAKE teaching program of Utrecht courses before the holidays, due to the quick and positive response of the fellows. Thanks! The program is included in this *NAKE Nieuws*. As you will see, it is a quite elaborate program again, so there should be something to everybody's taste. Actually, the number of courses is quite a bit higher than usual. The reason for this is that, as already announced in the previous issue, NAKE is going to offer a number of courses on (empirical) finance. In order to make it worthwhile for students especially interested in this subject to travel to Utrecht, we concentrated these courses: there will be three courses on finance in block 2 and also three in block 4. Moreover, we will offer additional courses on organisation theory in block 2. I hope that these courses attract a large number of students so that we can make them a regular part of the NAKE program.

As usual, this issue also includes some workshop reports. This time the winning authors are **Katrin Oltmer** (VU), who summarises Kaushik Basu's lectures on "Rationality and Social Norms" and **Adriaan Soetevent** (RUG) who surveys Guido Imbens' course on "Causal Inference and Program Evaluation". Notice that this are still reports of the December 2000 workshop in Maastricht. For the reports of the brilliant lectures during the workshop at the Vrije Universiteit Amsterdam in June 2001 you will have to wait for the next issue. Two students received their NAKE diploma during this workshop: **Mark Sanders** (UM) and **Theo Leers** (Tilburg University). Congratulations with this achievement!

The NAKE Research Day is also imminent: Friday October 12, 2001 is the day to mark in your calendar. Elsewhere in this issue you find the program for this day. Just like last year, the NAKE Day will be combined with the Tinbergen lecture organised by the Royal Netherlands Economic Association. Jacques Drèze (CORE) and Frank Vandenbroucke (federal minister of social affairs in Belgium) will deliver this lecture. Location will be the Dutch Central Bank in Amsterdam again. Notice that the annual general meeting of the members of NAKE will take place over lunch during the NAKE Day. You can register for the NAKE Day either electronically (see the NAKE homepage) or with the registration form in this *NAKE Nieuws*.

I hope to see you in Amsterdam!

Best regards,

Lex Meijdam

# Causal inference and program evaluation

Guido Imbens

Report by Adriaan R. Soetevent, University of Groningen\*

## Abstract

This report surveys the contents of the lectures given by Guido Imbens at the 29th NAKE workshop which took place in Maastricht from December 11-15, 2000. Imbens gave five inspiring lectures about how treatment effects can be measured in a non-experimental environment, which is the typical case in economics. This report provides a review of these lectures. In Section 1 the notational apparatus is introduced; Section 2 briefly reviews the techniques used to assess treatment effects in classical experiments; Section 3 discusses how treatment effects can be measured in a non-experimental setting when the assumption of unconfoundedness is made. The plausibility of this assumption is discussed in Section 4. Section 5 concludes with some deviations from the unconfoundedness assumption.

## 1 Some notation

The population contains  $N$  units, and elements of the population are usually indexed by  $i = 1, \dots, N$ . The set of treatments is denoted by  $T_i \in T$ , e.g. in case of a program of receiving a labor market training:  $T_i \in \{\text{training, no training}\}$ .  $Y_i(t), \forall i = 1, \dots, N, t \in T_i$  denotes the potential outcome for individual  $i$ , conditioned on receiving treatment  $t$ . The covariates or pure treatment variables are assigned the variable  $x_i, i = 1, \dots, N$ . Our aim is to estimate the *causal effect* which is measured by comparison of the potential outcomes. Basically, the absolute difference between treatments (measuring  $y_i(t) - y_i(t')$ ) as well as the relative difference (measuring  $y_i(t)/y_i(t')$ ) can be compared. In this lectures, the focus will be on absolute differences. However, this does not affect the generality of the results.

The variables that can be observed are  $x_i, T_i$  and  $y_i(T_i)$  for each unit  $i = 1, \dots, N$ . This has the important implication that one observes only one potential outcome per unit  $i$ . For example, for someone who went to the university, you can never observe the potential earnings he would have earned had he not gone to the university. Central in this course is the question how the

---

\*Faculty of Economics, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands, Ph: +31 - (0) 50 - 363 37 66; E-mail: a.r.soetevent@eco.rug.nl.

causal effect can still be measured in this instance and which assumptions have to be imposed to this purpose.

New in the treatment is the notation in potential outcomes. Till the mid of the 80s, notation was done in observable variables<sup>1</sup>.

One must be aware that causal effects can never be observed at the unit level, and so the focus is on *average* causal effects:

$$\tau = \frac{1}{N} \sum_{i=1}^N [y_i(1) - y_i(0)]. \quad (1)$$

## 2 Randomized experiments

Although the focus in this course is on non-experimental studies, this section briefly reviews the assumptions underlying the measurement of treatment effects in classical randomized experiments. This will prove to be a useful starting point.

A *classical randomized experiment* in case of a binary treatment is defined by the following assumptions on the assignment mechanism.  $P(T; x, y(0), y(1))$  is the probability of the assignment vector  $T$  as a function of  $x$  and potential outcomes for which we have the following:

1.  $0 < P(T_i) < 1 \quad \forall i$ ;
2.  $P(T_i | x, y(0), y(1))$  does not depend on  $x_j, y_j(0), y_j(1) \quad \forall j \neq i$ ;
3.  $P(T; x, y(0), y(1)) = P(T; x, y(0)', y(1)')$  if  $y_i(T_i) = y_i(T_i)'$ , which means that assignment does not depend on unobserved potential outcomes;
4.  $P(\cdot)$  is known.

### Examples of randomized experiments

- An example of a completely randomized experiment is to choose randomly  $M$  out of  $N$  units to receive a certain treatment. For this assignment mechanism  $P(T | x, y(0), y(1)) = 1 / \binom{N}{M}$  if  $\sum_{i=1}^N T_i = M$ ;
- The Bernoulli trial. Flip for each unit a coin with  $P(\text{heads}) = p$ . If heads occurs, the subject receives the treatment, otherwise it is assigned to the control group.

---

<sup>1</sup>Implicit in all that follows is the *no interference* assumption: the treatment of unit  $i$  does not affect the outcome for unit  $j$ . An example of a situation in which this assumption is violated is when in a country all people would go to university. In that case the effect on earnings of joining university is different for the first and the last person who join.

## Fisher exact tests

Fisher employed the following null hypothesis:

$$H_0 : y_i(0) = y_i(1) \quad \forall i$$

$$H_a : y_i(0) \neq y_i(1) \quad \text{for some } i$$

Note that these hypotheses do not read as:

$$H_0 : \frac{1}{N} \sum [y_i(1) - y_i(0)] = 0$$

$$H_a : \frac{1}{N} \sum [y_i(1) - y_i(0)] \neq 0$$

Neyman thought that the second hypothesis was the more interesting one. Under Fisher's  $H_0$ , all potential outcomes are known.

Lets fix a statistic  $S = S(y, T, x)$ , for example  $S = \frac{\sum y_i T_i}{M} - \frac{\sum y_i (1 - T_i)}{N - M} = \bar{y}_1 - \bar{y}_0$ , which denotes the average difference between the treatment and control group. In a randomized experiment, the distribution of  $S$  is known given Fisher's  $H_0$ . Thus, a vector  $\tilde{T}$  can be drawn,  $\tilde{y}$  can be recalculated and with these,  $\tilde{S} = S(\tilde{y}, \tilde{T}, x)$  can be calculated. The  $p$ -value of  $S$  can be calculated and if this value is below a certain critical value  $\alpha$ ,  $H_0$  is rejected.

Neyman instead focuses on estimation and inference and tries to estimate  $\tau$  (see Neyman (1990)). The estimate  $\hat{\tau}$  of  $\tau$  for the completely randomized experiment for example, looks like,

$$\hat{\tau} = \frac{1}{M} \sum_{i=1}^N y_i T_i - \frac{1}{N - M} \sum_{i=1}^N y_i (1 - T_i). \quad (2)$$

This estimate is unbiased for  $\tau$ , since  $E(T_i) = M/N$  and<sup>2</sup>

$$E(\hat{\tau}) = \frac{1}{M} \sum_{i=1}^N y_i(1) \frac{M}{N} - \frac{1}{N - M} \sum_{i=1}^N y_i(0) \frac{N - M}{N} = \tau \quad (3)$$

The variance of  $\hat{\tau}$  is calculated as  $\frac{S_0^2}{N - M} + \frac{S_1^2}{M} - \frac{S_{01}^2}{N}$  with

$$S_0^2 = \frac{1}{N - 1} \sum_{i=1}^N (y_i(0) - \overline{y(0)})^2$$

$$S_1^2 = \frac{1}{N - 1} \sum_{i=1}^N (y_i(1) - \overline{y(1)})^2$$

$$S_{01}^2 = \frac{1}{N - 1} \sum_{i=1}^N (y_i(1) - y_i(0) - \tau)^2$$

---

<sup>2</sup>Note that the only random element here is the assignment mechanism.

It is the last term that causes problems in practice since not both  $y_i(1)$  and  $y_i(0)$  are observed. For this reason,  $S_{01}^2$  is not estimable.

Opposed to experimental studies, observational studies in economics normally have a non-random character. Therefore, the next section deals with this type of non-experimental studies. It is shown that inferences can still be made if an additional assumption is invoked: the assumption of unconfoundedness, which ensures that there is enough detail in the covariates.

### 3 Non-experimental studies

In this section will be shown how data from observational studies can be rearranged such that techniques for analyzing randomized experiments can also be applied in this instance. In order to compare outcomes with similar values of the covariates  $x$ , strata have to be defined. The central problem is thus answering the question how strata should be defined.

In the previous section, randomized experiments with a sharply defined null hypothesis were considered. Under this  $H_0$ , all of the potential outcomes could be filled in and exact distributions and  $p$ -values could be calculated. However, in this section non-experimental studies will be treated. In a non-experimental setting, it is usually not possible to control for the covariates. Whereas in experimental situations one can obtain a control and treatment group which are homogenous with respect to the observable characteristics  $x$ <sup>3</sup>, this is not possible in non-experimental studies since it is likely that the decision to be assigned to a treatment is in this case not independent from the observable characteristics. To give an example, in a study which evaluates the effect of a labor training program, all individuals which are given the treatment will have suffered a period of unemployment, whereas this does not have to be the case for the individuals in the control group.

To be able to estimate the causal effects of a treatment from data of observational studies an additional assumption has to be made. This assumption is called the *unconfoundedness* assumption, which says that the probability of assignment to a treatment does not depend on the potential outcomes conditional on observed covariates<sup>4</sup>:

$$P(T; x, y(0), y(1)) = P(T; x) \text{ or } T \perp y(0), y(1)|x.$$

Combined with the independence of different units, the unconfoundedness assumption implies

$$\begin{aligned} \tau(x) &= E(y(1) - y(0)|x) = E(y(1)|T = 1, x) - E(y(0)|T = 0, x) \\ &= E(y|T = 1, x) - E(y|T = 0, x), \end{aligned}$$

---

<sup>3</sup>With random assignment, homogeneity of the control and treatment group with respect to the unobservable characteristics is also guaranteed if the size of the groups is sufficiently large.

<sup>4</sup>Note that this can be a very bad assumption, e.g. in the case when the functionary, who has to decide whether or not an unemployed person will receive a job training program, has a lot more information about the potential earnings  $y(T_i)$  than is contained in the covariates.

which can be estimated. The second equality is due to unconfoundedness. It follows then that

$$\tau = E(\tau(x)) \quad (4)$$

can be estimated by averaging  $\hat{\tau}(x)$  over the distribution of  $x$ .

The unconfoundedness assumption has no testable implications; you can never test whether  $f(y(1)|T = 1, x)$  is equal to  $f(y(1)|T = 0, x)$  since the data contain no information over the value of the latter function. Therefore, any method for estimating  $y(1) - y(0)$  must compare units with  $T = 1$  with units for which  $T = 0$ . Unconfoundedness validates comparisons between different units with identical values for  $x$ . Any alternative compares units having  $T = 1, X = x$  to units having  $T = 1, X = x' \neq x$ . A reason for this might be, that you want to remove a bias caused by unobservable differences by means of a difference in the observable characteristics. An example of this is the method of instrumental variables, studied in Section 5.1.

### 3.1 Inference and estimation for a single covariate

Four different methods for estimating the average causal effect  $\tau$  can be discerned, when the units are characterized by a single observable covariate  $x \in \mathbb{R}^1$ :

**blocking** The sample is divided  $k$  subsamples on basis of  $x$ . As in the random experiment, within each subsample  $\tau_k$  is estimated as  $\hat{\tau}_k = \bar{y}_{1k} - \bar{y}_{0k}$  and subsequently  $\tau$  is estimated as  $\hat{\tau} = \frac{1}{K} \sum_{k=1}^K \frac{N_k}{N} \hat{\tau}_k$ .

**matching** For each treated unit (an  $i$  with  $T_i = 1, x_i = x$ ), the closest control unit ( $j$  with  $T_j = 0, \min_j \{|x_j - x_i|\}$ ) is located. The estimation made is  $\hat{\tau}_i = y_i - y_j$ . This is done for all treated and control units and  $\hat{\tau} = \frac{1}{N} \sum_{i=1}^N \hat{\tau}_i$ .

**weighting** This method blows up the control variables if there are relatively few controls. If the probability of receiving a treatment given the covariate is denoted as  $P(x) = E(T|x) = Pr(T = 1|x)$ ,  $\tau$  is estimated by  $\hat{\tau} = \frac{1}{N} \sum_{i=1}^N \frac{y_i T_i}{P(x_i)} - \frac{y_i (1-T_i)}{1-P(x_i)}$ <sup>5</sup>.

**curve fitting** With this method a curve is modelled for both the treatment and the control group. For example:  $E(y(1)|x) = \beta_0 + \beta_1 x$  and  $E(y(0)|x) = \alpha_0 + \alpha_1 x$ . Then  $\alpha$  and  $\beta$  are estimated by least squares and  $\tau$  is estimated by  $\hat{\tau} = \frac{1}{N} \sum_{i=1}^N (\hat{\beta}_0 + \hat{\beta}_1 x_i - \hat{\alpha}_0 - \hat{\alpha}_1 x_i) = \frac{1}{N} \sum_{i=1}^N \hat{y}_{1i} - \hat{y}_{0i}$ . A danger of this method is that it may rely too much on extrapolation.

In empirical work, the researcher has to feel uncomfortable if one of these methods would give very different results than the others. Further on, blocking, matching and curve fitting all have a counterpart in non-parametric analysis: step functions, nearest neighbors and series estimators respectively. All four methods focus on the relation between the exogenous variables  $T$  and  $x$ .

---

<sup>5</sup>Note that  $E(\frac{y_i T_i}{P(x_i)} | x_i) = E(\frac{y_i(1)}{P(x_i)} | T_i, x_i) \cdot Pr(T_i = 1 | x_i) = E(y_i(1) | T_i, x_i) = E(y_i(1) | x_i)$ . The last equality follows from unconfoundedness. Similarly we have  $E(\frac{y_i(1-T_i)}{1-P(x_i)} | x_i) = E(y_i(0) | x_i)$ .

### 3.2 Inference and estimation for the multiple covariate case

When  $x$  is a vector instead of a scalar, the problem becomes intricate and what is needed is a tool to reduce the dimensionality of the problem. Such a tool is provided by Rosenbaum and Rubins (1983) *propensity score*:  $P(x) = Pr(T = 1|x)$ , which is the conditional probability of receiving the treatment of interest. This score is useful because of the implication that if

$$T \perp y(0), y(1)|x,$$

it is also true that

$$T \perp y(0), y(1)|P(x).$$

*Proof:*

$$\begin{aligned} Pr(T = 1|P(x)) &= E(T|P(x)) = E(E(T|x, P(x))|P(x)) \\ &= E(P(x)|P(x)) = P(x), \text{ and} \end{aligned} \tag{5}$$

$$\begin{aligned} Pr(T = 1|y(0), y(1), x, P(x)) \\ &= E[E(T|y(0), y(1), x, P(x))|y(0), y(1), P(x)] \\ &= E(P(x)|y(0), y(1), P(x)) = P(x). \end{aligned} \tag{6}$$

Therefore  $x \perp T|P(x)$  and matching on  $P(x)$  implies matching stochastically on  $x$ . The strategy to be followed is to estimate the average causal effect  $\tau$  using the methods of blocking etc., with the subsamples constructed with use of the propensity score.

#### The Dehejia and Wahba study, 1999

The practical usefulness of the propensity score is nicely illustrated in the Dehejia and Wahba (1999) study in which the results of Lalonde (1986) are further investigated. Lalonde compares earnings of males who participated in a labor market training program with both the earnings of males in an experimental control group and with a number of other comparison groups. Whereas the comparison to the experimental control group leads to credible estimates of the training effect, comparison to the other constructed comparison groups leads to ridiculous estimates. In view of this result, Lalonde's conclusion is that in order to get reliable estimates, experimental data are indispensable.

Dehejia and Wahba however point out that there are large differences in the values of the covariates in Lalonde's treatment group and those of the non-experimental comparison groups. Subsequently they show that the earning differences between the treatment and various comparison groups become much more credible when implementation methods based on the propensity score are used. The estimates of the average effect of the training program are now comparable with Lalonde's estimates using an experimental control group. Moreover, the variation between the different implementation methods decreases, such that the results are more robust.

Table 1: Comparison of observables in the PSID and CPS control groups.

	PSID	CPS
black	0.25	0.07
married	0.87	0.71
earnings '74	19,000	14,000

## 4 Plausibility of the unconfoundedness assumption

Hitherto, the obtained results were based on the validity of the unconfoundedness assumption. In this section the focus is on the validity of this assumption. Further, the sensitivity of the results to violations of the assumptions is assessed by means of sensitivity analysis. An extreme version of this analysis is the bounds approach.

### 4.1 Bounds

The bounds approach follows from the observation that

$$\begin{aligned}
 \tau &= E(y(1) - y(0)) = E(y(1)) - E(y(0)) \\
 &= E(y(1)|T = 1)P(T = 1) + E(y(1)|T = 0)P(T = 0) \\
 &\quad - E(y(0)|T = 1)P(T = 1) - E(y(0)|T = 0)P(T = 0),
 \end{aligned} \tag{7}$$

and noting that the only terms that are not estimable are  $E(y(1)|T = 0)$  and  $E(y(0)|T = 1)$ . With the other terms, bounds can be constructed. Suppose for example that  $y(0), y(1) \in \{0, 1\}$ . Then the bounds

$$\begin{aligned}
 \tau &\geq E(y|T = 1)P(T = 1) - P(T = 1) - E(y|T = 0)P(T = 0) = \underline{\tau} \\
 \tau &\leq E(y|T = 1)P(T = 1) + P(T = 0) - E(y|T = 0)P(T = 0) = \bar{\tau},
 \end{aligned}$$

are obtained. Note that in this way, never an interval narrower than one is obtained, since  $\bar{\tau} - \underline{\tau} = 1$ . The bounds always include zero, such that the sign of the effect cannot be obtained.

This result may not be very useful, but the advantage is that the unconfoundedness assumption is completely avoided.

### 4.2 Multiple control groups

When multiple control groups are available, these can be used to assess the validity of the unconfoundedness assumption. For example, in the Lalonde/Dehejia-Wahba study, the PSID and CPS control groups are available of which some observable characteristics are listed in Table 1.

The table shows that these control groups are very different in observable characteristics. So possibly, they are also very different in the unobservables. If this is likely, but the results are similar, credibility is added to the results.

It is therefore advisable to choose alternative control groups with plausibly different biases. For a training program one could for example choose controls from the same local labor market and controls from a different local labor market; or controls observed in the same time period and a historical control group.

### 4.3 Sensitivity to inclusion of observables

Another legitimate question is to ask whether the results are sensitive to the inclusion of (sets of) covariates. In the example of the training program, the question is whether lagged earnings matter, or the time interval (yearly, quarterly) over which earnings are reported. The Dehejia and Wahba article for example brings about that two years of earnings are needed to measure the training effect adequately. Card and Sullivan (1988) show in another study, that the labor market history of individuals has to be taken into account to measure the effect of a training program properly.

## 5 Deviations from unconfoundedness

In this last section, the unconfoundedness assumption is relaxed. In the first subsection, the technique of instrumental variables is considered as an alternative to assuming unconfoundedness. In the second subsection, attention is paid to treatments which can on more than two values.

### 5.1 Instrumental variables

The instrumental variables approach is used when the topic of interest concerns the effect of  $x$  on  $y$  but  $x$  is determined endogenously. In that case  $x$  is correlated with the disturbances and the way to circumvent the associated problems is to look for an instrument  $z$ , which has no *direct* effect on  $y$ , but only an *indirect* effect through  $x$ . Well known examples of this approach are e.g. Angrist (1990) and Angrist and Krueger (1991). The first study explores the influence of having served in the military on earnings and mortality later. The problem is that having served or not is likely to be dependent on unobservable characteristics of an individual, like e.g. physical condition. The notation used here is the following: The instrument is binary,  $z \in \{0, 1\}$ , as is the treatment variable  $x$ ,  $x \in \{0, 1\}$ . The outcome is denoted by  $y$ .  $x(z)$  denotes the potential outcome for the treatment:  $x(0)$  is the potential outcome of serving if one is not drafted and  $x(1)$  is the potential outcome of serving if one is drafted.  $y(z, x)$  denotes the potential outcome of the earnings where  $y(0, 0)$  is the outcome if one is not drafted and did not serve;  $y(1, 0)$  is the potential outcome if one is drafted but did not serve, etc. The assumptions made are

1. There is random assignment of the instruments,  
 $z \perp y(0, 0), y(0, 1), \dots, y(1, 1), x(0), x(1)$ ;
2.  $x$  is monotone:  $x(1) \geq x(0)$ , which means that there are no defiers, i.e. people who serve when they are not drafted, but who do not serve when they are;
3. The exclusion restriction: there is no direct effect of  $z$  on  $y$ . Therefore,  $y(x) = y(0, x) = y(1, x)$ .

The main result is that the average difference in covariates  $x$  between individuals who are drafted and individuals who are not, equals the fraction of people that complies, i.e. people who do not serve when not drafted, but serve when drafted:

$$E(x|z = 1) - E(x|z = 0) = Pr(\text{compl})$$

A second result is that the average difference in potential earnings between individuals who are drafted and individuals who are not, equals the average difference in earnings between people who served and those who did not conditional of being a complier times the probability of being a complier:

$$E(y|z = 1) - E(y|z = 0) = E(y(1) - y(0)|\text{compl}) \cdot Pr(\text{compl})$$

It then follows that

$$\frac{E(y|z = 1) - E(y|z = 0)}{E(x|z = 1) - E(x|z = 0)} = E(y(1) - y(0)|\text{compl}), \quad (8)$$

so the IV-approach measures the average effect on compliers.

Returning to the assumptions and to the example of being drafted for the military, we see that the exclusion restriction is critical. The first assumption is likely to be fulfilled. However, getting drafted,  $z = 1$ , may prompt other questions: "Do I stay in school or will I leave for Canada?" a decision which will affect potential earnings and for this reason violates assumption 3<sup>6</sup>.

## 5.2 Multivalued treatments

In this section is dealt with two aspects of multivalued treatments, i.e. treatments which can take on more than two values. First, instrumental variable models are discussed and next methods for unconfounded assignment.

In the binary case, (8) showed how – under some assumptions – the IV-approach could be used to estimate the average effect for compliers. How can these ideas be applied when the endogenous regressor is continuous?

---

<sup>6</sup>Note that in the common application of IV, assumptions 1 and 3 are combined in the statement that the residuals of  $y$  and  $x$  are uncorrelated with  $z$ .

To explore this problem, the example of the Fulton fish market in New York is taken, as described in Angrist, Graddy and Imbens (2000). The quantity of whiting traded is denoted by  $q$  and the price per pound is denoted by  $p$ ;  $p^e$  is the equilibrium price. The weather conditions *at sea* as described by wave height and wind speed, are the instrumental variables  $z$ . These affect the supply but not the demand of whiting. The assumptions made are:

1. The weather is independent of the demand and supply functions,  $z_t \perp q_t^D(p, z), q_t^S(p, z_t)$ . This implies that  $z_t \perp p_t^e(z)$ , and this can be interpreted as that weather is as good as randomly assigned.
2.  $q_t^D(p, z)$  is assumed independent of  $z$ , the exclusion restriction<sup>7</sup>.
3.  $q_t^D(p)$  is non-decreasing in  $p$  for all  $z$ ;  
 $q_t^S(p, z)$  is non-decreasing in  $p$  for all  $z$ ;  
 $q_t^S(p, z)$  is monotone in  $z$  for all  $p$ .

If there are two possible weather conditions, good or bad:  $z \in \{0, 1\}$ , (8) estimates the demand elasticity with  $y$  replaced by  $q$  and  $x$  by  $z$ .

Now assume that there are three weather conditions:  $z \in \{fair, mixed, storm\}$ . Moreover, let  $p_t^e(f) = 1 \forall t$ ,  $p_t^e(m) \forall t$  and  $Pr(p_t^e(s) = z) = Pr(p_t^e(s) = 3) = 1/2$ . The interest is in  $\beta^* = q^D(3) - q^D(2)$ . Which information is in the data about  $\beta^*$ ? With the data, the following two pieces of information can be estimated:

$$\beta^{f/m} = \frac{E(q|z = m) - E(q|z = f)}{E(p|z = m) - E(p|z = f)} = E[q_t^D(2) - q_t^D(1)|p_t^e(m) = 2] \quad (9)$$

$$\beta^{m/s} = \frac{E(q|z = s) - E(q|z = m)}{E(p|z = s) - E(p|z = m)} = E[q_t^D(3) - q_t^D(2)|p_t^e(m) = 3]. \quad (10)$$

So there is choice in which equation to estimate. However, (9) gives the right averages but works with wrong prices, and (10) gives the wrong average, whilst using the right prices. There is no way out of this dilemma since  $z$  is the only instrument you have and the data are not going to tell which values of the instruments should be used.

In summary, the IV-approach gets much more complicated in case of continuous treatments. In the remainder, attention is paid to the case with unconfoundedness and more than two treatments.

Assume that  $t \in \{1, \dots, K\}$  and denote the potential outcome by  $y(t)$ . In Section 3.2, it was shown how the propensity score could be used to reduce the dimension of the problem when there are multiple covariates. The question is if there exists a function  $h(x)$  analogous to the propensity score, such that:

$$T \perp \{y(t)\}_{t=1}^K | X \Rightarrow T \perp \{y(t)\}_{t=1}^K | h(x)$$

<sup>7</sup>Note that this assumption is not valid if the weather on the sea is highly correlated with the weather on the shore.

Unfortunately, it turns out that it is not possible to find a scalar valued solution for  $h(x)$ . If  $h$  is allowed to be vector valued, a solution can be found: let  $h(t, x) = Pr(T = t|x)$ , then  $T \perp \{y(t)\}_{t=1}^K | \{h(t, x)\}_{t=1}^K$ . This solution however gives not much dimension reduction if  $\dim(K)$  is large relative to  $\dim(x)$ .

What can be done to solve this problem, is to weaken the unconfoundedness assumption: no full independence on the potential outcomes is needed. Specify  $D(t) = 1\{T = t\}$ , then

$$D(t) \perp y(t)|x \Rightarrow D(t) \perp y(t)|h(t, x).$$

In this assumption, only on potential outcome is considered at a time and the independence is on  $D(t)$ , not on  $T$ .

## References

- [1] J. Neyman, On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9, *Statistical Science* 5(4), 465–480 (1990).
- [2] P. Rosenbaum and D. Rubin, The Central Role of the Propensity Score in Observational Studies for Causal Effects, *Biometrika* 70(1), 41–55 (1983).
- [3] R. Dehejia and S. Wahba, Causal Effects in Non-Experimental Studies: Re-Evaluating the Evaluation of Training Programs, *Journal of the American Statistical Association* 94(448), 1053–1062 (1999).
- [4] R. Lalonde, Evaluating the Econometric Evaluations of Training Programs, *American Economic Review* 76(4), 604–620 (1986).
- [5] D. Card and D. Sullivan, Measuring the Effect of Subsidized Training Programs on Movements In and Out of Employment, *Econometrica* 56(3), 497–530 (1988).
- [6] J. Angrist, Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence Form Social Security Administrative Records, *American Economic Review* 80, 313–335 (1990).
- [7] J. Angrist and A. Krueger, Does Compulsory School Attendance Affect Schooling and Earnings, *Quarterly Journal of Economics* 106, 979–1014 (1991).
- [8] K. G. Angrist J. and G. Imbens, The Interpretation of Instrumental Variables Estimators in Simultaneous Equations Models with an Application to the Demand for Fish, *Review of Economic Studies* 67, 499–527 (2000).

## NAKE teaching program 2001-2002

### Block 1: 7 September - 19 October, excl. 12 October 2001<sup>1</sup>

***10.30 - 12.30***

01.44 The economics of household behaviour Kooreman

***13.00 - 14.45***

01.05 Environmental problems and policy: A theoretical introduction Folmer

01.46 Optimisation methods in econometrics Magnus

***15.00 - 17.00***

01.15 Social economics: Heterodox approaches to economic theory Brenner

01.20 Topics in applied microeconomics van Damme

### Block 2: 26 October - 30 November 2001

***10.30 - 12.30***

01.83 Simulation-based econometric methods van Soest

01.87 Finance Palomino

***13.00 - 14.45***

01.12 Fiscal policy in open economies Bovenberg

01.35 Theory of incomplete markets Herings

01.53 Testing the validity of asset pricing models with applications  
to stock picking and performance measurement Nijman &  
Melenberg

***15.00 - 17.00***

01.03 General equilibrium theory van der Laan &  
Talman

01.06 Growth and environment Smulders &  
Withagen

01.30 Nonlinear time series models in empirical finance Franses & van Dijk

---

<sup>1</sup> 12 October 2001 is the NAKE researchday 2001 at the Dutch Central Bank Amsterdam

**Block 3: 25 January - 1 March 2002****10.30 - 12.30**

01.62	Regulation and privatisation	Ruys
01.85	The empirics of economic growth	Florax & de Groot

**13.00 - 14.45**

01.**	Open economy macro	Kool
01.89	Economic Organization Theory	Hendrikse

**15.00 - 17.00**

01.**	Organisation	van Ees
01.77	Generalized method of moments and latent variables	Wansbeek & Meijer

**Block 4: 15 March - 24 April 2002****10.30 - 12.30**

01.21	Asymptotic theory for integrated processes	Boswijk
01.42	Market micro-structure	De Jong

**13.00 - 14.45**

01.57	Social choice theory	Storcken & Peters
01.84	Bayesian Analysis of Dynamic Econometric Models using Simulation Techniques	van Dijk (Herman)
01.86	Issues in Banking: Banks 'Specialness', Loan Pricing and Availability, and Firm-Bank Relationships	Ongena

**15.00 - 17.00**

01.27	Advanced industrial organisation	Furth & van Cayseele
01.88	Asset pricing	Verbeek

The course descriptions can be found on the NAKE homepage. You can register electronically through the NAKE homepage.

## - **Rationality and Social Norms** -

### **Workshop Report on the Lectures of Kaushik Basu**

**by Katrin Oltmer**

#### **1. INTRODUCTION**

Social norms and cultural aspects are important and influential factors for an individual's economic decision making. Traditional economic analysis, where rational and utility maximising behaviour is one of the core ingredients, often fails to incorporate these factors into its models and theories. The main concerns of Basu's lectures are the importance of factors beyond standard economic theory and the challenge of the rationality assumption that lies at the basis of all standard economic analysis.

Basu stated that game theory could be an appropriate tool to model social factors that have been left out in former analysis. Game Theory is the study of interactive decision making, where situations in which a decision maker's behaviour affects not only his or her own gains and losses, but also those of other decision makers, is analysed. Before the rise of game theory, rationality has been represented by maximisation of the utility function. In cases where an individual does something that is not in accordance with the utility function, the utility function has to be made more complex. However, if all actions carried out by an individual under different circumstances (e.g. at different times or places) would be comprised by the utility function, such that all possible actions could explain utility maximising behaviour, the utility function would not say anything more at the end. In Basu's words it would result into a tautological situation.

This report is divided into two parts. Part 1 covers some ideas of rationality and Part 2 deals with the role of social norms in individuals' decision-making problems. Part 1 goes into the inconsistency of choice and gives some illustrative examples of game theoretical problem solving. Part 2 describes different categories of social norms, conditions for the survival of social norms, the problem of peer pressure, and a three persons interaction case. The report closes with some concluding remarks.

<b>PART 1: IDEAS OF RATIONALITY</b>
-------------------------------------

## 2. QUESTIONS ON THE CONSISTENCY OF CHOICE

### 2.1 *Internal Consistency of Choice*

Rational behaviour of economic agents is the core assumption in economic modelling. For example, a rationally behaving individual maximises his/her utility function satisfying the axioms of rationality. In his lecture, Basu refers to the *Axiom of Chernoff* and the *Weak Axiom of Revealed Preferences (WARP)*, which are imposed in order to safeguard the internal consistency of choice. The two axioms are explained in the following:

#### *The Axiom of Chernoff*

The *Axiom of Chernoff* requires that an alternative  $x$  that is chosen from a set of alternatives  $S$  and belongs to a subset  $T$  of  $S$  must be chosen from  $T$  as well. Formalisation of this axiom gives the following expression:

If  $x \in C(S)$  &  $x \in T \subseteq S$ , then  $x \in C(T)$ ,

where,  $C(S)$  is the choice function or choice set of  $S$ . Sen (1993) calls this axiom the 'basic contraction consistency' or 'property  $\alpha$ '.

#### *The Weak Axiom of Revealed Preference (WARP)*

The *WARP*, which was first suggested by Paul Samuelson in 1947, demands that if  $x$  is chosen from a choice set  $C(S)$  when  $y$  is available in  $S$  as well, there cannot be any set  $T$  containing both alternatives for which  $y$  is chosen from a choice set  $C(T)$  and  $x$  is not. The formal representation of *WARP* is as follows:

If  $x \in C(S)$  &  $y \in S \setminus C(S)$ , then there exist no  $T \in K$ , such that  $x \in T$  &  $y \in C(T)$ ,

where,  $K$  represents all alternatives.

These two axioms satisfy the internal consistency of choice. However, Basu (referring to Sen (1993)) pointed out that once some external factors to choice, such as social norms and values on which the choice is based, are brought into the analysis, complying with the two axioms above may become difficult. Consider the following pair of choices, which clearly violate the *Axiom of Chernoff* and the *WARP*:

$$\text{A) } C(\{x, y\}) = \{x\}$$

$$\text{B) } C(\{x, y, z\}) = \{y\}$$

An illustration of this pair of choices is the cinema example. A person, say Peter, meets a distant acquaintance at the cinema, who is inviting him for a cup of tea at his home after the film is over. Now, Peter has the choice between accepting the invitation (x), and going home (y) (case A). Since Peter considers the distant acquaintance to be a nice person, he chooses to accept the invitation (x). The following week at the cinema, Peter meets again the distant acquaintance, who now is offering him the choice over having a cup of tea (x), going home (y), or having some cocaine (z) (case B). This time, Peter prefers to go home (y). Peter's choices in case A (taking x from {x, y}) and in case B (taking y from {x, y, z}), clearly rejects the *Axiom of Chernoff* and the *WARP*. The reason for Peter choosing y and rejecting x in case B is that adding z to the menu reveals information about the underlying situation, namely something about the distant acquaintance, which in turn influences Peter's preferences over the alternatives offered. Sen (1993) calls this phenomenon the 'epistemic value of the menu'.

Basu even went a step further by arguing that adding the z is not necessarily needed in order to change choice. Imagine that the second time Peter meets the distant acquaintance, the choice offered to him is the same as in case A. However, while the distant acquaintance is inviting Peter for a cup of tea, he is pulling some cocaine out of his pocket. In this case, Peter might prefer to go home above having a cup of tea with the distant acquaintance although no additional choice has been added to the menu. The question Basu is raising here is whether a choice function does exist at all. The objective of Sen's paper (1993) has the objective that a choice function is retained.

## ***2.2 Inconsistency of Dynamic Choice***

The dynamic inconsistency of human behaviour, as Basu presented it in this part of the lecture, is a very recognisable phenomenon in everyday life. He started with the example of George Akerlof in India who is supposed to send a parcel to his friend Joseph Stiglitz in the US (Akerlof, 1994). In short, each morning for over eight months Akerlof woke up and decided that the *next* morning he would go to the post office to send the parcel. This happened until a few months before his own departure, and he decided to include the parcel in the large shipment of another friend of him who was returning to the US at the same time as himself. In the situation just described, Akerlof overvalued the costs of sending the parcel on the current day relative to the costs of any other day, which causes him to put off sending the parcel. Basu formalised this problem as follows.

$N$  = capital and time left for own departure ( $N-1$ ,  $N-2$ , etc.),

$x$  = utility of one more day of having the parcel (Stiglitz's utility),

$X$  = utility of having the parcel from day 0 (the day of departure) onwards (Akerlof's utility),

$c$  = costs of mailing,

$p$  = salience premium (overvaluation of sending the parcel today) and  $p > 0$ ,

$Z$  = overall or combined utility (Akerlof's plus Stiglitz's utility).

The salient premium let the costs of sending the parcel today be  $[c(1 + p)]$ , whereas the costs of sending the parcel tomorrow are  $c$ . It is assumed that  $[pc > x]$  and  $[xN > pc]$ . The overall utility of sending the parcel today is hence

$$a) Z_n = X + xn - c(1 + p)$$

and that of sending the parcel tomorrow is

$$b) Z_{n-1} = X + x(n-1) - c.$$

Subtracting equation a) from equation b) then gives

$$c) Z_{n-1} - Z_n = cp - x,$$

and according to the first assumption [ $cp - x > 0$ ]. In equation c), the  $n$  has vanished, which means that it does not matter on which day the calculation is made, it would always result into postponement (procrastination).

What would be the result of a comparison between the utility of sending the parcel today and sending the parcel on day 0? In this case, the utility of sending the parcel today would be

$$d) Z_N = X + xN - c(1 + p)$$

And those of sending the parcel on day 0

$$e) Z_0 = X - c.$$

Subtracting again equation e) from equation d) gives

$$f) Z_n - Z_0 = xN - cp$$

and according to the second assumption [ $xN - pc > 0$ ], which shows that sending the parcel today is preferred above sending the parcel on day 0. This result is the opposite of the first result, where postponing was preferred above sending the parcel today.

This inconsistency of dynamic choice does also violate the WARP condition. E.g., on day 2, there is the choice between sending the parcel today, tomorrow, or on day 0. According to the result in equation c), sending the parcel tomorrow is always preferred above sending it today. Hence, the choice on day 2 is [ $c(\{2,1,0\}) = 1$ ]. However, on day 1, the choice is only between sending the parcel today or tomorrow. Again, according to equation c), the choice on day 1 is [ $c(\{1,0\}) = 0$ ]. These two choices clearly violate WARP.

The above example describes a situation involving repeated decisions with time inconsistent behaviour, which means that people do things that they would never have thought of in the beginning. Referring to the example above, at the time when Stiglitz was leaving India, Akerlof would have never considered keeping the parcel until his own departure. This inconsistency of dynamic choice arises through the ability to take small decisions at subsequent points in time, of which each is close to

utility maximisation and hence resulting in small losses only. However, the cumulative effect of a series of repeated suboptimal choices may be quite large (Akerlof, 1994).

Another example of time inconsistent behaviour, Basu mentioned in his lecture, is the smoker example. A smoker, who is taking the decision to smoke on a single day, prefers to smoke rather than not to smoke. However, if the smoker is asked whether to smoke or not to smoke during the rest of his life, the smoker chooses not to smoke.

### 3. SOME EXAMPLES OF GAME THEORETICAL PROBLEM SOLVING

#### 3.1 The Game of Hex

In his short introduction to game theory, Basu started with an example of a parlor game, namely the *Game of Hex*<sup>1</sup>. The *Game of Hex* is a two-player strategy game played on an  $N \times N$  rhombus of hexagonal cells. It is an example for a finite game with perfect information. According to the *Theorem of Zermelo*, every finite game with perfect information has a pure strategy Nash equilibrium that can be derived through backward induction (Mas-Colell *et al.*, 1995).

The two players, Black and White, take turns in marking the hexagons. The objective of the game is to create an unbroken path of either Black or White marked hexagons across the board (Black from top to bottom, White from left to right). Figure 1 shows an example of Hex Game in a 4x4 rhombus. In this game, winning is clearly defined. Either Black or White wins, whereas the first mover tends to win.

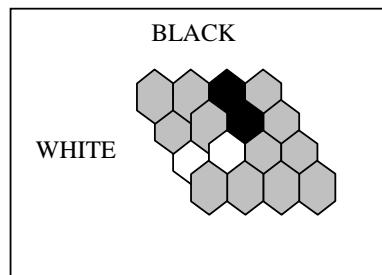


Figure 1: The Game of Hex in a 4x4 Rhombus.

<sup>1</sup> The Hex Game was invented in 1942 by Piet Hein, a Danish poet and mathematician, and it has been played on bathroom tiles (<http://members.iex.net/~rfinn/gameshlf/abstract/hex/hex.htm> and <http://homepages.enterprise.net/drking/hexagons/hex/index.html>, 30/12/2000)

With the help of this parlor game example Basu demonstrated the paradoxical situation that gametheory is able to show that there is a winning strategy, but cannot tell us what this strategy is (or even which of the players has a the wining strategy as in the game of chess). In this example, White starts the game. Assume that Black has a winning strategy, meaning that Black has a fully defined plan that specifies every move in every possible distinguishable situation, such that winning is secured. Black has written down his winning strategy in a book. Now, if White would take the book, also White would have a winning strategy. The only problem White encounters is the positioning of the first stone. Black, on the other hand, only has to react on White's first move, and Black's winning strategy includes all potential first moves of White. The solution to White's problem would be, if White would regard its first stone not as a white stone but as, for instance, a grey stone. After Black reacted on White's first move with the *grey* stone, White can react on Black's move (this time with a white stone) according to the winning strategy written down by Black. Summarising, if Black has a winning strategy, White has a winning strategy. Since White has a first-mover-advantage, White wins, so Black cannot have a winning strategy. As there has to be a winner in this game, white must have a winning strategy. It is not possible to derive this strategy however.

### **3.2 The Hawk-Dove-Game**

Furthermore, Basu's review on game theory comprised examples of the *Prisoner's Dilemma* and the *Hawk-Dove-Game*. In the famous *Prisoner's Dilemma*, the dominant strategy is not the best outcome for the players, meaning that self-interested, rational behaviour is not leading to a socially optimal result (Mas-Colell *et al.*, 1995).

In the *Hawk-Dove-Game*<sup>2</sup>, playing 'dove' is the peaceful way of playing, and playing 'hawk' the aggressive one. The two players are competing against a fixed amount of resources. If both players would play 'dove', the resources would be divided equally among the two players. If one player would play 'hawk' and the other 'dove', the hawkish player would gain all the resources and the dovish player would get nothing. Finally, if both players would play 'hawk', the outcome of both players (and hence

---

<sup>2</sup> The *Hawk-Dove-Game* is widely discussed in the biological literature and was mainly inspired by Maynard Smith in 1982.

also the overall outcome) would be negative, because their aggressive behaviour would lead to a destruction of resources.

Basu illustrated the *Hawk-Dove-Game* with the Cuban Missile Crisis between the USA and the Soviet Union in 1962, one of the most important conflicts of the Cold War period<sup>3</sup>. Both 'players', the USA and the Soviet Union, were threatening to play 'hawk', which in any case would have had disastrous consequences, in the worst case a declaration for a third world war. Certainly, the negative outcome in terms of loss of resources would have been uncountable, especially with the kind of weapons involved in this crisis. Basu stated that the crisis was resolved by the Soviet Union going dovish and retreating their nuclear missiles from Cuba, since they did not want to take the risk of a nuclear disaster. However, as it have been found out many years after the crisis, the actual agreement between the USA and the Soviet Union by that time was, that also the USA would retreat their weapons from Turkey. Seemingly, the fact that in the end also the USA adopted a dovish strategy was not supposed to reach the public in those years.

### ***3.3 The Traveller's Dilemma***

Basu's Traveller's Dilemma tells the story about two travellers who bought identical antiques on their vacations. On the return trip, their luggage got lost, and the airline offers them compensation for their lost luggage. A representative of the airline asked both travellers separately the amount they paid for the antique. Since the airline is afraid of excessive claims, the representative informs the travellers about the following condition: The airline is willing to compensate any claim between \$2 and \$100. Each of the travellers will get the amount that is equal to the *lower claim* of the two claims declared. If the two claims differ, the airline will reward the person with the lower claim with \$2 for her honesty. The person with the higher claim will be penalised with \$2 for her dishonesty. If both travellers announce the same claim, they will get exactly the amount they declared (Capra *et al.* 1999). In a pay-off table, this game theoretical problem can be described as follows.

---

<sup>3</sup> The reason for the Cuban Missile Crisis was the stationing of Soviet nuclear missiles in Cuba as a reaction on the presence of US nuclear missiles in Turkey.

		<b>CLAIM OF TRAVELLER B</b>						
		<b>2</b>	<b>3</b>	<b>4</b>	<b>....</b>	<b>98</b>	<b>99</b>	<b>100</b>
<b>CLAIM OF TRAVELLER A</b>	<b>2</b>	2,2	4,0	4,0	....	4,0	4,0	4,0
	<b>3</b>	0,4	3,3	5,1	....	5,1	5,1	5,1
	<b>4</b>	0,4	1,5	4,4	....	6,2	6,2	6,2
	:	:	:	:	:	:	:	:
	:	:	:	:	:	:	:	:
	<b>98</b>	0,4	1,5	2,6	....	98,98	100, 96	100,96
	<b>99</b>	0,4	1,5	2,6	....	96,100	99,99	101,97
	<b>100</b>	0,4	1,5	2,6	....	96,100	97,101	100,100

*Table 1: Pay-Off Table of the Traveller's Dilemma.*

The solution to this game, which means the only rational choice for both travellers, is to claim \$2. This solution is a Nash-equilibrium and also a unique equilibrium. For example, if both travellers would think about to declare 100 in the beginning, than traveller A would argue that her final pay-off would be larger if she would declare 99, since she would get the \$2 reward. However, since rationality is common knowledge in game theory, traveller B knows that traveller A behaves rational by declaring 99. Player B would hence declare 98, because this choice would maximise her final pay-off. Continuing with this reasoning would bring both travellers back to declare 2. This procedure is called *backward induction*.

A class room experiment of the Traveller's Dilemma game showed that in reality people would actually choose numbers in the 90s and not 2, the rational outcome of the theoretical game. Also intuitively, choosing 2 seems to be a strange outcome if you had the chance to get higher payments instead. The empirical question is hence why in reality people claim differently from what game theory would predict. The travellers' assumption of the opponent behaving rational is crucial for the theoretical outcome of the game. If one traveller would make an irrational move, she would give a sign to the other traveller that her believes about rationality is wrong, and the two travellers might start to co-operate.

Another important factor that influences the outcome of the Traveller's Dilemma is the level of the penalty/reward, which was investigated by Capra *et al.* (1999).

Theoretically, the Nash-equilibrium of the Traveller's Dilemma is independent of the penalty/reward value. In a repeated game experiment, Capra *et al.* test the influence of high, intermediate, and small penalty/reward values on the rational behaviour of the players. The penalty/reward values are not announced as single number but rather as fuzzy categories. They found that small penalty/reward values result into an irrational (non-Nash equilibrium) outcome of the game. On the other hand, if the penalty/reward values are high people move closer to a rational, Nash-equilibrium outcome.

## **PART II: SOCIAL NORMS**

### **4. DIFFERENT CATEGORIES OF SOCIAL NORMS**

In his introduction to the effects and the functioning of social norms, Basu started with the First Theorem of Welfare Economics. Roughly speaking, this theorem tells us that if all individuals in a society are free to choose, society will reach an optimal state where all resources are allocated in an economically efficient way. However, since the First Theorem of Welfare Economics is a pure theoretical and mathematical statement only, it does not always succeed in describing real world situations. Basu argued that the reasons why individuals do not act according to what is technically possible are social norms. In other words, social norms are the factors that prevent individuals from behaving rationally, or, considering the game theoretical examples in the previous part of this report, the factors that let 'real life games' having different outcomes than theoretical games. There are three different categories of social norms:

- Rationality-limiting norms
- Equilibrium selection norms
- Preference-changing norms

#### ***4.1 Rationality-limiting Norms (RLN)***

RLN are those norms that limit our feasible set of actions we encounter in everyday life. Strictly speaking, social norms are a subset of this feasible set of actions. The feasible set of actions is so incredible large that calculating every action would not be possible, such that social norms facilitate decisions making in everyday life. An example for RLN is queue-breaking. A person in e.g. Great Britain who finds a queue at any particular box-office can either join the queue or break the queue. Breaking the queue would certainly be the most time-saving (and hence the most rational) option. However, since the social norm in this situation is to join queues, breaking queues would not be accepted (and might give trouble). For this person, the social norm of joining queues has become an automatism, which means that other options are not even considered when finding the queue.

#### ***4.2 Equilibrium Selection Norms (ESN)***

ESN occur in situations where two equilibria are possible and where the choice of one particular equilibrium depends on the norm. Examples for the ESN are driving on the left or right side of the way and the acceptance of torn currency notes. Two equilibria are possible in the driving example, driving on the left or driving on the right side of the way. If a person lives in a country where driving on the right side of the way is the norm, he does certainly not consider driving on the left instead. Torn currency notes, even if they are 'fixed' again, are not accepted in India. Hence, a person does not accept a torn currency note from another person, because she knows that she cannot make use of it either.

#### ***4.3 Preference-changing Norms (PCN)***

For the explanation of the PCN, Basu referred to a paper by Lindbeck *et al.* (1999), who modelled the combined influence of social norms and economic incentives on individual behaviour with respect to the utilisation of unemployment benefits. The paper argues that in most modern welfare states the social norm is to earn one's income from own work and not to live off other people. This means that the preference of people living in this kind of societies is to work rather than not to work. In this case, the social norm creates stigma costs, because people living on unemployment benefits feel embarrassed and not accepted by their working fellow

citizens. However, these stigma costs are not exogenously given but are rather dependent on the number of people adhering to the social norm. To put it differently, the larger the share of people living on unemployment benefits, the less embarrassing it is and the lower are the stigma costs. The intensity or strength of the social norm is decreasing, which in turn induce changing preferences with respect to the utilisation of unemployment benefits.

The model presented in the paper is looking for an equilibrium wage, at which people with lower wages choose not to work and those with higher wages choose to work, and which determines the intensity of the social norm. People have two choices. Firstly, they have an economic, namely whether to work or to be unemployed. Secondly, they have a political choice, which is to determine the size of welfare transfers. Mathematically, this problem can be described as follows.

$$u[(1 - t)w] = u(T) + \mu - v(x),$$

where,  $u$  = utility,  $w$  = wage rate,  $t$  = tax rate,  $T$  = level of unemployment benefit,  $\mu$  = joy of leisure,  $v$  = stigma costs,  $x$  = number of people unemployed. The equation on the left-hand side gives the utility of working people and the equation on the right-hand side the utility of unemployed people.

Solving the model, Lindbeck *et al.* found indeed two stable equilibria as it was intuitively suggested. The first (second) equilibrium describes the situation where unemployment is low (high) and hence the share of people living on unemployment benefits is low (high), and where the stigma costs are high (low). The multiplicity of equilibria indicates that societies with the same tax rates, level of unemployment benefits, wage rates and preferences may get caught in different equilibria, in which different levels of intensities of the social norm persist.

Having described the different definitions of social norms, the next paragraph will go into the survival power of different social norms and the emergence of rights.

## 5. THE SURVIVAL OF SOCIAL NORMS

Introducing the theory of the survival of social norms, Basu gave an illustration about a trap-setters society. In a trap-setters society, the person who set the trap has the right to take the animal caught in the trap. If this right would not exist, the two

possibilities for the trap-setter would be either to wait next to the trap until an animal is caught, or not to set traps at all. Since these two options are not very appealing, the trap-setters society might not even exist if there would not be the right of taking the animal. The social norm of not taking an animal from another person's trap has evolved into the right for the trap-setter. In other words, the fact that the trap-setters society does exist shows that the social norm of not taking the animal has had the strongest survival power.

The trap-setters story shows an example where the strategy of not taking the animal from another person's trap is an *Evolutionary Stable Strategy*. This term is drawn from the biological literature (in particular from Maynard Smith), to which Basu is referring in this part of his lecture. Next to the concept of *Evolutionary Stable Strategy*, he discussed the principle of *Evolutionary Stable Norm*.

### ***5.1 Evolutionary Stable Strategies (ESS)***

An ESS is a strategy such that, if all members of a population adopt the ESS, no mutant strategy could invade the population under influence of natural selection. In the following the ESS is described in a game theoretical context.

Consider a normal form game between two players [ $N = \{1,2\}$ ] and a finite number of strategies [ $S = \{1,\dots,m\}$ ]. The strategies can be ordered such that strategy 1 notifies a very aggressive strategy and strategy  $m$  a very dovish one. Accordingly, the strategies in-between are decreasingly aggressive and increasingly dovish. The pay-off function is given by  $[P(p,q)]$ , where  $[p,q \in S]$ <sup>4</sup>. The pay-off for the two players are symmetrical. Hence, defining the first player's pay-off is sufficient. The strategies are predetermined, which means that aggressive players will always be aggressive and dovish players will always be dovish. Now, strategy  $p \in S$  is an ESS, if it is immune against all  $q \in S \setminus \{p\}$ , which means that strategy  $p$  is immune against strategy  $q$ , if

- i)  $P(p, p) > P(q, p)$  or
- ii)  $P(p, p) = P(q, p) \ \& \ P(p, q) > P(q, q)$

The rule given under i) says that strategy  $p$  does always perform better against itself than any other strategy. The rule given under ii) notes that it performs equally well

---

<sup>4</sup> In the original, biological context, the pay-off function indicates fitness. The higher the fitness of a species, the more rapidly they are reproducing.

against itself as another and at the same time performs better against the other strategy than the strategy does against itself. The ESS is always a Nash-equilibrium. In fact, ESSs are a subset of all possible Nash-equilibria in the game.

Basu gave a numerical example of this game with three strategies and eight available resources. The strategies are the hawk-strategy (aggressive), the cookoo-strategy (intermediate), and the dove-strategy. The pay-off table of this game is as follows.

		Player 2		
		H	C	D
Player 1	H	0,0	3,1	6,2
	C	1,3	3,3	5,3
	D	2,6	3,5	4,4

*Table 2: The Pay-Off Table of the ESS Game*

In this example, the only ESS for both players is the cookoo-strategy, since it is the only strategy where one of the two rules given above holds. The explanation is given from the view-point of player 1. The rule given under i) does not hold, since the pay-off  $P(C,C) = \{3,3\}$  is not larger than the pay-off  $P(H,C) = \{3,1\}$  and neither for the pay-off  $P(D,C) = \{3,5\}$ . However, the rule given under ii) does hold for the cookoo-strategy. The pay-off for player one  $P(H,C)$ ,  $P(C,C)$ , and  $P(D,C)$  are equal. Additionally, the pay-off  $P(C,H) = \{1,3\}$  is larger than the pay-off  $P(H,H) = \{0,0\}$ , and the pay-off  $P(C,D) = \{5,3\}$  is also larger than the pay-off  $P(D,D) = \{4,4\}$ , as required by the rule given under ii).

### ***5.2 Evolutionary Stable Norm (ESN)***

The question behind an ESN is whether a population with an ESN is stable against any other country coming in, and what happens if the country coming in itself has an ESN. Which social norm will survive? Is the concept of rationality evolutionary stable or not? As the ESS, the ESN can be described in a game theoretical context.

Basu calls this game a human evolutionary game. There are two populations (players),  $[N = \{X,Y\}]$ , where X is the original population, and the two populations

both represent certain social norms. Therefore,  $[X, Y \subset S]$  and  $[S = \{1, \dots, m\}]$ . The rules for an ESN is in accordance to the rules for an ESS.

Given  $X, Y \subset S$ ,  $X$  is immune against the invasion of  $Y$  if

- i)  $N(X, X) > N(Y, X)$ , or
- ii)  $N(X, X) = N(Y, X) \ \& \ N(X, Y) \geq N(Y, Y)$ .

The ESN is also always a Nash-equilibrium. However, in equilibrium, certain kinds of non-Nash behaviours will evolutionary exist, which means that some social norms that are not part of a Nash-equilibrium will survive.

### **5.3 Peer Pressure**

The social norm of peer pressure arises in situations with high degrees of social control, and it can lead to and sustain situations where no one really wants to be. Basu explained the peer pressure problem with the example of a person who is disloyal to his king, or, using Akerlof's definition of such a person, of an outcast. A person's disloyalty consists of not giving one unit of his goods to the king. There are two types of disloyal persons. Type 0 does not give a unit to the king, and type 1 is not disloyal himself but maintains a relation with a disloyal person.

Consider a society with  $n$  goods and  $n$  persons, where each person  $i$  has  $e$  (where  $e > 1$ ) units of good  $i$ . The utility function of person  $i$  is then  $[U^i = \sqrt{x_1^i} + \dots + \sqrt{x_n^i}]$ . In a Walrasian equilibrium with all prices being equal and normalised to one, the utility for everybody in this society can be expressed as  $[U^* = n(\sqrt{(e/n)})]$ , where  $[e/n = x^i]$ , for all goods]. Now the oppressive king comes in and demands from each person one unit of their goods. He additionally declares that there are punishment costs equal to  $k \geq 0$  for persons who do not give a unit to him. With the assumption that everyone is loyal and gives one unit to the king, the new initial endowment to each person is  $e-1$ . The utility function in the new Walrasian equilibrium is  $[U^0 = n(\sqrt{((e-1)/n)})]$ , where  $U^0$  indicates the oppressive equilibrium. The expected utility for a protesting person  $U^P$  is given by  $[U^P = \sqrt{e - k}]$ . In order to find out whether the peer pressure of being loyal to the king will sustain, one has to look for parameters such that  $[n(\sqrt{((e-1)/n)})] > [\sqrt{e - k}]$ . Assume that the punishment costs are equal to zero,  $[k = 0]$ . Then,

reformulating the inequality  $[n(\sqrt{(e-1)/n})] > [\sqrt{e} - k]$  by taking roots will give  $[n(e-1) = e]$ . This result shows that the utility of the protesting person is always smaller than the utility of the loyal person, which means that no one likes to be the first person to stand up and peer pressure will maintain.

#### ***5.4 Three Person Interaction***

The three person interaction problem is related to the type 1 definition of a disloyal person. It describes a situation where the relation between two persons is dependent on whether one of the persons keeps a relation with a third person, which is not approved by the other person. Basu's example of a three person interaction was illustrated by the relations between a landlord, a labourer, and a merchant. The merchant can trade with both, the landlord and the labourer. The landlord, however, wants to employ the labourer only on the condition that he does not trade with the merchant.

*The labourer's problem* - The labourer values the good that he could obtain by trading with the merchant with  $B_2 > 0$ . The critical wage level at which the labourer would accept to work for the landlord is hence  $[(w-c)L + B_2 = 0]$ , or by rearranging  $[(w-c)L = -B_2]$ , where  $w$  is the wage rate,  $c$  a work related cost factor, and  $L$  the hours worked. The labourer would be willing to work for the landlord if the net benefits from working would compensate for the loss of not trading with the merchant.

*The landlord's problem* - The landlord has a net profit function described by  $[f(L) - wL]$ . The maximisation problem he faces is hence  $\max [f(L) - wL]$ , subject to  $[(w-c)L + B_2 \geq 0]$ . In order to attract the labourer, the landlord has to pay a wage higher than that under a situation where everyone has the choice to trade with whomever he likes. The reason for this is the  $B_2$  term in the constraint.

In a situation where everyone has the choice to trade with whomever he likes, the land lord faces a different maximisation problem, and that is  $\max [f(L) - wL]$ , subject to  $[(w-c)L \geq 0]$ . In this case, the critical wage the landlord has to offer in order to attract the labourer is only equal to the work related costs occurring to the labourer. Therefore, from a rational point of view, the landlord would be better off if he would allow the labourer to trade with the merchant.

## **6. Concluding Remarks**

Basu's lectures about rationality and social norms were very interesting and enlightening. His gift to illustrate all problems with numerous examples from all over the world gave the lectures an entertaining and amusing touch. It was shown that rationality, one of the basic assumptions in traditional economic analysis, is not present in many everyday life decision-making problems. The institution of social norms and values is very important for an individual's economic behaviour. The framework of game theory is seen as an adequate approach to model institutions such as social norms and to incorporate them into the analysis of economic decision-makings. However, it has also been pointed out that the rationality assumption is actually not as unfavourable as it was insinuated. Assuming rational behaviour helps to understand fundamental mechanisms of economic decision-making and basic determinants of complicated real world problems.

## **7. REFERENCES**

- Akerlof, G.A. (1994). Procrastination and Obedience. *American Economic Review* 81(2): 1-19.
- Capra, C.M., J.K. Goeree, R. Gomez, and C.A. Holt (1999). Anomalous Behaviour in a Traveller's Dilemma. *American Economic Review* 89(3): 678-690.
- Lindbeck, A., S. Nyberg, and J.W. Weibull (1999). Social Norms and Economic Incentives in the Welfare State. *Quarterly Journal of Economics* 114(1): 1-35.
- Mas-Colell, A., M.D. Whinston, and J.R. Green (1995). *Microeconomic Theory*. Oxford University Press.
- Sen, A. (1993). Internal Consistency of Choice. *Econometrica* 61(3): 495-521.