Measuring Inequality by Counting 'Complaints': Theory and Empirics

by

Kurt Devooght Katholieke Universiteit Leuven

Discussion Paper No. DARP 59 January 2002 Distributional Analysis Research Programme
The Toyota Centre
Suntory and Toyota International
Centres for Economics and
Related Disciplines
London School of Economics
Houghton Street
London WC2A 2AE

Distributional Analysis Research Programme

The Distributional Analysis Research Programme was established in 1993 with funding from the Economic and Social Research Council. It is located within the Suntory and Toyota International Centres for Economics and Related Disciplines (STICERD) at the London School of Economics and Political Science. The programme is directed by Frank Cowell. The Discussion Paper series is available free of charge and most papers are downloadable from the website. To subscribe to the DARP paper series, or for further information on the work of the Programme, please contact our Research Secretary, Sue Coles on:

Telephone: UK+20 7955 6678 Fax: UK+20 7955 6951 Email: s.coles@lse.ac.uk

Web site: http://sticerd.lse.ac.uk/DARP

© Author: **Kurt Devooght**

All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

ABSTRACT

This paper examines how people assess inequality of an income distribution and how inequality could be measured. We start from philosophical analysis of L. Temkin who distinguishes nine plausible aspects of inequality. His approach is based on the concept of 'complaints' or distances between incomes. We examine the Temkin approach by means of the questionnaire-experimental method pioneered by Amiel and Cowell. We want to find out whether the aspects of equality have any plausibility for student-respondents and if so, whether there are any aspects which are more appealing than others. Both the numerical and verbal responses show that a considerable majority of the respondents might be influenced by the intuitions included in the combination of the Weighted Additive Principle and the Average view of complaints. The questionnaire results also shed some light towards the Sequence and provide support for the poll result of the Temkin reports.

ACKNOWLEDGEMENT

I would like to thank Erik Schokkaert, Frank Cowell, and Leima Davidovitz for various suggestions which have led to improvements in this paper. I am very grateful for helpful comments from seminar participants. I remain responsible for any remaining errors. This paper was written while the author enjoyed the hospitality of STICERD, LSE, London and he acknowledges the financial support provided by the Training and Mobility of Researchers Programme of the European Communities, grant #ERBFMRXCT98024.

Keywords: Inequality, complaints, transfer principle, Temkin

JEL classification: D63

Address for correspondence:
Kurt Devooght
Center for Economics and Ethics
K.U. Leuven
Naamsestraat 69
3000 Leuven, Belgium
email: kurt.devooght@econ.kuleuven.ac.be

Measuring Inequality by Counting 'Complaints': Theory and Empirics. Kurt DEVOOGHT

1. Introduction.

Since the seminal contributions of Kolm [1969] and Atkinson [1970] on inequality measurement the literature in this field has ever been expanding. One major reason for the continuing interest in this topic, apart from the practical relevance for economic policy, is undoubtedly the complexity of the concept of inequality itself. The economist's concept of inequality, however, has often been not much more than a small set of core axioms on which most of the income-distribution analysis has been built. Pioneering experimental research by Amiel and Cowell [1999] has shown that the economist's concept of inequality does not simply coincide with the perceptions ordinary people hold on inequality. For instance, most economists assume that multiplying all incomes by the same scalar leaves inequality unaltered but no more than half of the respondents of the Amiel and Cowell questionnaire agree with it. The fact that respondents give different answers to the same question may mean that they hold different concepts of inequality or that inequality has various meanings. Inequality apparently is complex and multifaceted.

This complexity should not leave economists in dismay. I want to take the research of Amiel and Cowell one step forward. One possible way to do so is to link the concept of inequality to the broader context of justice and equity and to explore the interactions and trade-offs between the ideal of equality and other moral ideals. I do not proceed this line of thinking here. An alternative way to continue is to explore in more detail the concept of inequality that people do have in mind. I mentioned already that people's concept of inequality poorly matches the economist's concept but little is known about how people really think. What do people have in mind if they are asked to assess the inequality of an income distribution? The main aim of this paper is to shed some light on the various meanings or aspects of inequality people endorse. I will do this by means of the questionnaire-experimental method pioneered by Amiel and Cowell.

To pursue this aim one needs to start from a theoretical, philosophical analysis which explores the various ways in which one might assess an income distribution's inequality. In other words: which ways of looking at inequality can we logically think of? I opted for the highly elaborated, but among economists poorly known, analysis of Larry Temkin [1986, 1993] which is based on the concept of complaints.

Temkin's approach is for two more reasons interesting. First, Amiel and Cowell [1999] have shown that many respondents disagree with the transfer principle. This result causes serious problems to the conventional theory of inequality measurement since the transfer principle is a major building block for many theories and measures of inequality. In an attempt to make this empirical result understandable, Amiel and Cowell suggest that people may rather think about inequality in terms of distances [see also Broome, 1989]. Temkin now provides a philosoph-

ical framework in which he explores various ways to measure the size of individual complaints or -which amounts to the same- distances and to calculate overall inequality. I hope that this framework might lead economists on their way explaining the empirical rejection of the transfer principle. Second, Temkin [1993] presents poll results of a very small population (about 10 persons) and concludes that the results lend support to his theoretical insights. Are his conclusions still valid for a much larger population? To answer this question I designed a questionnaire which is as close as possible to the one Temkin presents.

The next section summarizes the most important features of Temkin's philosophical approach to inequality. Section 3 sketches the set-up and the design of the questionnaire whereas section 4 presents the main findings of the experiment. Section 5 concludes.

2. Philosophical Background.

The approach Temkin advocates to measure inequality basically consists in a two-step procedure of which the first step measures complaints at an individual level according to a particular view of complaints (2.2). The logical second step then aggregates the individual complaints according to a particular principle of equality to obtain the overall inequality in a society or (reference) group (2.3). But first we have to define the concept of individual complaints (2.1).

2.1. The Notion of Individual Complaints.

The Temkin approach to inequality is founded on the concept of indi-

vidual complaints. Someone has a complaint regarding inequality if his income is lower than the income of someone else through no fault or choice of his own¹. To say that the worse-off have a complaint is to recognize that the mere existence of involuntary income inequality is bad, unjust or unfair. For the remainder of this paper I assume that none of the income differences is resulting from personal fault or voluntary choice. I also assume that egalitarians are concerned with income inequality alone (and not welfare inequality) so that worse off or better off simply means poorer or richer.

2.2. The Size of Individual Complaints.

For any situation, then, in which some have less than others through no fault or choice of their own, the question arises how one should determine the seriousness or size of different people's complaints. Temkin argues that there seem to be three plausible answers. Each answer is determined by the choice of a particular reference level of income with which individual incomes are compared to establish the size of individual complaints.

On the first view, the reference level would be the average income of the income distribution. This answer regards an unequal world as a deviation from the situation that would have obtained if income had been distributed equally. Only those below the average have a complaint

^{1&}quot; Throughout this book I use the expression 'through no fault of their own' as shorthand for the expression 'through no fault or choice of their own'. Thus, 'fault' need not mean moral fault. It might include any instance of voluntary choice or responsibility." [Temkin, 1993, 13-14]. In this formulation 'own fault' is reduced to the concept of voluntariness as it appears in the proposals of Arneson [1989] and Cohen [1989, 1990] as opposed to Dworkin [1981a, b].

because it seems particularly unfair that in a world of equally deserving people someone has less than average income - the level one would be at if fate had treated each person equally. On this view, the size of an individual complaint of someone below the average is simply the difference between the average and his income whereas those above the average have nothing to complain about regarding inequality. Temkin calls this the relative to the average view of complaints (AVE)².

One might also plausibly regard an unequal world as a deviation from the situation in which each person is as well off as the best-off person. The size of someone's complaint will then depend on the difference between his income and the income of the richest person. Anyone with an income below the income of the best-off person will have a complaint because it will appear to be unfair or unjust for all of them to be worse off than the best-off person through no fault or choice of their own. So even the second best-off person will have a complaint, though his complaint may be small relative to the complaints of others. Temkin has labelled this second view the relative to the best-off person view (BOP).

There is a third way to measure the size of someone's complaint. This way accepts that all but the very best-off have a complaint, but contends that the size of someone's complaint depends on how he fares relative to all of the others who are better off than he - henceforth, the relative to all those better off view of complaints (ATBO). This view may be

² The use of the average as reference income has recently been challenged by another philosopher. Hartkamp [1999] argues it is morally wrong if some persons are worse off than all could have been simultaneously. The reference income i.e. the income at which all could be simultaneously well off, only coincides with the average if the redistribution could be performed costly (without redistribution costs, incentive problems, monitoring costs etc.).

defended as follows. It is bad for someone to be worse off than another through no fault of his own. This is why any person who is in such a position will have a complaint. But if it is bad to be worse off than one person through no fault of your own, it should be even worse to be worse off than two people through no fault of your own. And, in general, the more people there are who are better off than someone (and the larger the gap between them), the more that person should have to complain about regarding inequality. Therefore, to determine the size of someone's complaint one must compare his level to those of all who are better off than he, and not only to the level of the very best-off person. Someone's complaint is then simply the sum of the differences between his income and the incomes of all those better off.

Temkin's notion of inequality allows us to make judgements about whether, and the extent to which, particular individuals have a complaint regarding inequality. This kind of judgement is about how bad the inequality in a situation is from the standpoint of particular individuals in that situation. It yields statements in the sense of 'individual 1 has a complaint of x in situation A while individual 2 has a complaint of y in situation B'. But we finally want to consider judgements about how situation A and situation B compare regarding inequality. Temkin suggests that we can compare situations regarding inequality by combining the notion of individual complaints with the principles of equality embraced by egalitarians.

2.3. Aggregating Individual Complaints.

Until now I have explored how much an individual in a particular situation may complain about regarding inequality. From now on I will try to determine what the egalitarian's views actually are. Egalitarians sometimes hold conflicting views on how a society's overall inequality should be assessed. These firmly considered theoretical principles of equality in combination with the various views of complaints will allow us to make judgements about how situations compare regarding inequality.

One of the principles of equality that has strong appeal is the Maximin Principle of equality (MP)³. It maintains that egalitarians should especially be concerned with the complaint of the worst-off person or the complaints of a selected worst-off group (e.g. the 10 % poorest). It expresses the fact that our moral sensibilities are especially attuned to the lot of the worst-off. On this view, it would be unjust for society to benefit the rich if instead it could benefit the poor even if it could assist the rich more easily and efficiently than the poor. A society should devote most of its efforts and resources to its worst-off members. Consequently, the complaints of the poorest should be the only focus of egalitarians. A maximin principle of equality may also include a tie-breaking clause that if the worst-off individuals in two different distributions fare the same, then one goes on comparing the complaints of the second worst-off persons, the third worst-off persons, and so on, until the tie is broken. This procedure is known as lexicographic maximin or leximin.

³This principle is an application of the maximin principle of justice [Rawls, 1971] to the ideal of equality though both should not be confused.

Another principle that seems relevant is what can be referred to as the Additive Principle of equality (AP). According to this principle a situation's overall inequality is measured by summing up each of the complaints that its individuals have, and the larger the sum, the worse the inequality is. It captures the view that it is bad for one person to be in such position that he has a complaint and the corresponding view that it should be even worse if, in addition to the first person with his complaint, there is a second person who has a complaint.

An additive principle might seem preferable to a maximin principle insofar as it is concerned with the complaints of all those who have a complaint, and not just with the complaints of some arbitrarily selected worst-off group. It seems, indeed, implausible that we should be concerned with the complaints of the people at that level, but shouldn't be concerned at all with the complaints of those who are just above that level. Yet a maximin principle might seem preferable to an additive principle insofar as it is concerned with the distribution rather than merely the sum total of complaints. This suggests that a principle that combines these two elements would have great appeal.

One such principle is the Weighted Additive Principle of equality (WAP). It measures inequality by summing up people's complaints, after first attaching extra weight to them in such a way that the larger someone's complaint is the more weight is attached to it. What the weights should be based on or how the weighting scheme should look like is not discussed by Temkin. However, this principle gives expression to both the view that we should be especially concerned with the worst-off and

the view that we should be concerned with all complaints. It would give us a way of capturing the intuition that a world in which 4 people have complaints of 20 would be worse than a world in which 20 people have complaints of 4, ceteris paribus. This is a plausible intuition, but one that neither the additive principle nor the maximin principle captures.

It should be stressed that neither the maximin, the additive nor the weighted additive principle yield an ordering in itself. The principles provide only a judgement about how good or bad a situation is regarding inequality in combination with one of the three discussed views of complaints. Each principle of equality can now be combined with respectively the relative to the average, the relative to the best-off person, and the relative to the all those better off views of complaints which renders nine plausible positions⁴. For instance, AP+BOP makes the sum-total of all complaints relative to the best-off person. Each position thus allows us to measure the overall inequality of a distribution and to compare the degrees of inequality of different distributions.

In this section I have been examining the question of how Temkin judges one situation to be worse than another regarding inequality. He provides arguments for nine plausible positions which might influence our egalitarian judgements. Each position represents a different aspect of inequality and might be considered as a particular way of measuring a situation's inequality. It should be stressed that Temkin argues for

⁴ Temkin also describes three other plausible positions: gratuitousness, deviation to the median and social inequality. I will not deal with these positions in this paper because the first is rather obscure, the second is very similar to AP+AVE (and identical when the median equals the mean) and the third requires additional cultural and institutional information to rank income distributions.

the plausibility of each aspect presented, not that each aspect is equally plausible. Each aspect represents certain plausible views that cannot easily be dismissed. Moreover, many of these positions are fundamentally incompatible, resting as they do on contrary views. It simply cannot be true, for instance, both that everybody but the best-off person has a complaint and that only those below the average have a complaint. One might insist that what the conflict between these aspects illustrates is just how complex and multifaceted the notion of inequality is. Thus far Temkin.

3. Experimental Settings.

In this section, I like to present the set-up of the questionnaire. Temkin [1993, chp 5] refers to results of a poll taken in 1984 but only ten of the forty who received the poll completed it. Because Temkin does not give a complete report of his questionnaire we are not able to build an identical setting. I therefore rely on the design used by Amiel and Cowell [1992, 1999] since it is readily available and well thought-out.

The questionnaire, which can be found in Appendix A, starts with some explanation and an introductory story. Part one is a numerical part which consists of 14 numerical questions to find out in an indirect way how respondents compare situations regarding inequality. Focusing on two diagrams at a time, respondents were asked to indicate whether they consider the situation in diagram A less unequal than the situation in diagram B or consider situation B less unequal than situation A or are indifferent between both situations regarding inequality. There were

in fact three different versions of the numerical part each consisting of 14 questions, some of them appearing in all three versions⁵. This enabled us to collect answers on a total of 32 different questions. To be comparable with Temkin, these questions included exactly the same 15 questions Temkin presented his respondents with. The three versions were distributed randomly among the respondents.

This setting differs from both Temkin's poll and Amiel and Cowell's questionnaires. It is different from Temkin's poll in that I allow only the three above indicated answers for each A and B to be compared instead of asking respondents to characterize situation A's inequality by choosing one of eleven possible statements⁶. In order to simplify our questionnaire, I furthermore assume that the incomes in the income distributions represent personal incomes instead of group incomes. The distinct income vectors chosen have never been included in an Amiel and Cowell questionnaire. This enables us to supplement the already existing results of the Amiel-Cowell approach with relevant data. I also try another layout by using a graphical presentation of persons positioned on an income line -as suggested by the figures in Amiel and Cowell [1999]-

⁵ Questionnaire **A** has 6 questions which compare two diagrams where the second diagram contains at least one person who has more income than in the first diagram (income increase). The next 8 questions deal which two diagrams where the second is obtained from the first by a rank-preserving transfer as to test the transfer principle. Questionnaire **B** has 6 questions on income decrease i.e. comparing two diagrams where the second diagram contains at least one person who has less income than in the first diagram, 6 questions on the so-called increasing Sequence i.e. a series of unbalanced income increases and 2 questions on the transfer principle. Questionnaire **C** starts with the same 6 questions as series B, has 6 questions on the decreasing Sequence and concludes with 2 questions on the transfer principle. The last two questions of each questionnaire (on the transfer principle) are the same. Only questionnaire A is given in appendix A. The various income distributions used in questionnaire B and C are given in the tables in section 4.

⁶ For instance, to say that distribution A is worse than distribution B the respondents had to choose one of the following four statements: "definitely worse and univocally so"; "definitely worse though I feel some slight pull in the other direction"; "somewhat worse and univocally so" and "somewhat worse though I feel some pull in the other direction" [Temkin 1993, 55]

instead of a simple vector presentation.

The adequate choice of the numerical questions was crucial for our main objective. I wanted to test whether respondents think along the lines of one of the aspects of inequality Temkin described when asked to compare two simple income distributions. The questionnaire is designed in such a way that each aspect of inequality describes a unique path of answers to the numerical questions. Let us explain. Recall that we deal with nine aspects of inequality. For any two income distributions, each aspect of inequality yields an answer to the question how these distributions compare regarding inequality. For instance, question A10 asked to compare situation A = (9, 8, 7, 4) with situation B = (9, 8, 6, 5). The maximin principle of equality combined with the best-off person view of complaints ranks situation B as strictly less unequal than situation A (Table 3.1.). On the other hand, two aspects (AP+BOP and AP+AVE) imply that the inequality in both situations is the same i.e. is a tie (T).

Table 3.1.	ΜP			ΑP			WAP		
	ВОР	AVE	ATBO	ВОР	AVE	ATBO	вор	AVE	ATBO
A 1 0	В	В	В	T	T	В	В	В	В

The same can be done for each other question. How each aspect of inequality ranks the two situations w.r.t. inequality has been calculated for each question and a complete overview of what the aspects of inequality imply is given in Appendix B.

Most of the calculations are straightforward except for the aspects based on the Weighted Additive Principle. First of all, Temkin seems to be inconsistent when it comes to count the size of complaints according to the average view of complaints which might give problems when AVE is combined with WAP. Temkin [1993, 20-23] argues extensively that only those who are worse off than the average have a complaint. However, he applies another procedure to measure AVE complaints in his numerical examples where he allows both the incomes below and above the average to count⁷. I believe that from a philosophical point of view, this ambiguity is of a major importance but will not deal with it here⁸. For the remainder of this paper I have chosen to include both deviations below and above the average mainly because this is what Temkin actually does in his numerical examples and supports in chapter 9 of his book. This particular choice yields the same inequality ordering of income distributions as counting deviations below the average only except in case inequality is measured by WAP+AVE where both approaches may differ.

Second, for the aspects based on the Weighted Additive Principle one had of course to specify the weighting function in order to rank the alternatives. We attach extra weight to larger complaints by requiring that the weighting function is strictly increasing in its argument and strictly convex in some interval. The latter condition excludes the (non-weighted)

^{7&}quot; Also, recall that to measure a situation regarding AP & AVE one must decide whether on the relative to the average view of complaints deviations above the average should count as well deviations below the average. I assumed they should in my examples." [Temkin, 1993, 58 fn]

⁸Rabinowicz [1999, p 20-24] also points to this fundamental ambiguity in Temkin's book. Much in Temkin's approach, e.g. the BOP and ATBO versions of individual complaints, gives the impression that the badness of pairwise inequality is a personal disvalue or harm. The badness of pairwise inequality is said to be a personal disvalue if it is a disvalue for a person, more precisely for the person who is worse off in a given relationship. But then Temkin's treatment of AVE-complaints is inappropriate since those above the average are not worse off but do have a complaint according to Temkin's examples. On the other hand, Temkin emphasizes (1993, ch. 9) that the disvalue of inequality is essentially impersonal i.e. the badness of pairwise inequality accrues to the relation between two individuals. If the latter is true, then, there seems something inappropriate in framing an impersonal conception of the badness of pairwise inequality in terms of personal complaints.

AP-based inequality measures. The additive inequality measures can be considered as weighted additive if all weights are set equal. The rationale why I do not impose strict convexity on the whole of the weighting function is because I wanted to follow Temkin's text as close as possible. He indeed allows that some complaints do not receive extra weight. This specification also ensures that complaints are treated symmetrically which is another of Temkin's implicit requirements. Symmetry should be understood as the requirement that a complaint C of someone below the average should receive the same weight as the same complaint of someone above the average. I then checked for all possible weighting functions satisfying these conditions. Since WAP in fact represents a class of inequality measures some cases are undecided.

Now, the 14 questions (in each of the three versions) are chosen in such a way that each aspect of inequality describes a unique pattern of judgements on these questions. If someone judges all 14 cases in the same way as one particular aspect of inequality then we are sure that his judgements are in at least some cases different from the judgements of any other aspect of inequality. This enables us to differentiate between the aspects of inequality. Whether someone who judges all 14 cases in the same way as an aspect of inequality really thinks along that line is plausible but far from sure. He might have other ideas in mind, although we know that it is not one of the other eight aspects of inequality (then

⁹ This is indicated in the tables in appendix B by a question mark. The implications in appendix B may differ from the implications Temkin [1993] gives for the same questions. The differences are due to a more consistent use of the formulas and the weighting functions for which Temkin has no uniform outlook.

he should have given other answers to some questions). To be sure we should ask him explicitly.

Therefore I have added a second part in which is investigated in a verbal and more direct way how respondents might think about comparing distributions regarding inequality. This verbal part consists of two questions in which the first verbal question (Q15) asks the respondents to indicate which view of measuring the size of individual complaints corresponds most closely to their own view: BOP, AVE, ATBO or none. The second question (Q16) then asks them to indicate the way in which those individual complaints should subsequently be aggregated to obtain the overall inequality in Alfaland. They were provided with the choice between the MP, AP, WAP or none¹⁰. If they choose a particular aspect of inequality in the verbal part and have judged the cases according to this position, then this is a quite strong suggestion that they really think according to that aspect of inequality. For instance, if our respondent after 14 choices in line with AP+AVE in the verbal part indicates that the way one should think about inequality is AP+AVE, then, we have a strong case to defend that we know how this respondent, under the given experimental conditions, thinks about inequality.

Finally, part three asks the respondents to answer some personal questions (age, male/female, subject of study, ...). The questionnaires were filled out anonymously during an economics class by 185 first year Business School students at the Catholic University of Leuven, Belgium.

 $^{^{10}}$ In both questions respondents were also given the opportunity to add their own proposal or any comment.

4. Experimental Results.

4.1. Verbal Part.

Let us start with the results of the verbal questions. This part of the questionnaire matches directly the aspects of inequality Temkin describes. It might give us an answer to the question whether the aspects of inequality have any plausibility for the respondents at all and whether there are one or more aspects which are really more appealing than others. If so, it might reveal something of the students' attitudes towards inequality when provided with the particular framework of complaints.

Table 4.1.:Verbal Part (percentages, # between brackets)							
$\mathbf{Q16} \backslash \mathbf{Q15}$	BOP AVE		ATBO	None	Total		
AP	1.1 (2)	8.7 (16)	0 (0)	0.5 (1)	10.3 (19)		
MP	0.5 (1)	2.2 (4)	0.5 (1)	0 (0)	3.2 (6)		
WAP	1.6 (3)	58.9 (109)	8.7 (16)	8.1 (15)	77.3 (143)		
None	0 (0)	5.4 (10)	2.2 (4)	1.6 (3)	9.2 (17)		
Total	3.2 (6)	75.2 (139)	11.4 (21)	10.2 (19)	100 (185)		

The results given in Table 4.1. are quite remarkable. Almost 60 % of the respondents chooses both WAP and AVE as the aspect of inequality that suits best their own attitudes towards inequality¹¹. Given the wide range of possible combinations it is remarkable that one result clearly stands out. Asked to give their view about the size of the inequality an arbitrary person is entitled to complain about (Q15), even 75 % of the

¹¹ It could be added that verbal support for WAP and AVE is considerably higher among the female students (68%) than among the male student population (52%).

respondents selects that answer which measures the size of an individual complaint as the (absolute) distance between one's income and the average income. The choice of the average as reference income is not surprising because at least students in economics should immediately see that giving anyone the average income is the only feasible outcome of an egalitarian (re)distribution. On the other hand, more than 75 % states that the weighted additive principle is the proper way to aggregate individual complaints in order to obtain a measure of the overall inequality (Q16). Again, this is not totally unexpected since the weighted additive principle takes an intermediate stance in between the more extreme positions of the maximin principle and the additive principle.

One of Temkin's objectives, however, was to show that the other aspects of inequality have attractive features as well and are not merely constructions of mind. Is this claim sustainable? Support for aspects of inequality other than WAP+AVE is small or even non-existent (AP+ATBO). The only aspects which receive some considerable support are AP+AVE (8.7 %) and WAP+ATBO (8.7 %). It should not surprise anyone that both combinations include either WAP or AVE. Looking at Q15 and Q16 separately, the picture is clear: AVE strictly dominates the three other alternatives in question 15 and WAP does the same in question 16. On the other hand, one should acknowledge that some of Temkin's logical positions clearly do not have much appeal to our respondents: BOP (3.2 %) and MP (3.2 %) for instance. About 10 % on Q15 and 9 % on Q16 chooses for answer D (of which 1.6 % selected D in both Q15 and Q16) which means that they did not find a suitable formulation among the

given alternatives. This indicates that the vast majority of the students were satisfied with the specified alternatives.

This conclusion might be supported by another observation. Many students made use of the opportunity to give a comment or their own proposal (20.0 % on Q15 and 7.6 % on Q16). The majority of the comments are reasonable and useful. This means that the students understood the questionnaire in the way intended. Many comments (9 respondents) suggest that their answers are only conditional and should depend on the reasons behind the inequality in Alfaland (e.g. personal characteristics of the individuals). For those students considerations of justice interfere with mere inequality considerations. This does not surprise me and I designed the questionnaire in such a way as to avoid these expected interferences (by giving no information about the origins of the inequality). As mentioned in the introduction, investigating the interactions between the ideal of equality and other ideals as justice and equity is one way to learn more about how people compare unequal distributions. I however opted to pursue that goal by exploring the concept of inequality itself and wanted the respondents to report an as much as possible 'unspoiled' attitude towards inequality and inequality alone.

Seven respondents argue that the size of the 'group' of the poor and of the 'group' of the rich matters (although the questionnaire never mentioned 'groups'). It is considered as fairly offensive if a huge number is rich while few remain very poor. This highlights that thinking in terms of groups is apparently deep-rooted. In the comments on Q16, five students argue that they perceive the aggregation of individual com-

plaints as something between MP and AP. They apparently did not understand that WAP exactly does what they want. This adds in an indirect way to the overwhelming agreement with the weighted additive principle (148/185 = 80.0 %). Only three students make complaints about the wording of the questionnaire: they say that having a complaint does not mean that people actually do complain. They stress that complaints should be considered as objective states of affairs rather than merely subjective expressions of distress. It is not clear to them that the questionnaire does consider complaints as objective features of income distributions¹².

4.2. Numerical Part: Synchronic view.

The majority of the respondents consider WAP+AVE as the aspect of inequality that best suits their own attitude towards inequality measurement. But do they also choose the answer to the numerical question along this line of reasoning? Are the outcomes of the numerical part consistent with the results of the verbal part? We can answer this question if we consider the results of the numerical questions in a synchronic manner i.e. if we examine the responses of one individual at a time. Indeed, each aspect of inequality implies a particular outcome on each of the 14 numerical questions. The questionnaire is designed such that each aspect of inequality describes a slightly different pattern of answers on the 14

¹² Single comments include (undeveloped) proposals of what can be labelled as a relative measure (divide by total income), a weighted average (instead of weighted sum), a headcount of those below a reference income, the sum of the distances to the mode, an inequality measure including functionings or primary goods (besides income), and a comment that not inequality is important but sufficiency (people should have sufficient income to lead a decent life; compare with Frankfurt [1987]).

questions. The question then is: to what extent do respondents choose the answers implied by the aspect of inequality they verbally endorse?

I first focus on those who supported WAP+AVE verbally. Table 4.2. gives the percentages of students who have chosen the by WAP+AVE predicted ranking for different percentages of questions. In the absence of a determinate function for how much (extra) weight the larger complaints should be given, there is no simple algorithm for comparing distributions according to WAP+AVE. Our specification of the weighting function yields an unambiguous ordering in 8 of the 14 cases. The other six questions cannot be solved without additional specification of the weighting function. Any answer on these questions can be accounted for simply by slightly changing the weighting scheme within the conditions I imposed on it earlier. For reasons of comparability, however, I have omitted these questions here and report only on the consistency with the eight questions for which WAP+AVE yields an unambiguous ordering.

The results are encouraging. Ten respondents or 9.2 % are completely consistent which did surprise us provided that they had to be consistent on a considerable amount of questions. Almost 90 % are in at least two thirds (63%) of the questions consistent which may be considered as a very high percentage. I am inclined to believe that these results lend support for the claim that the majority of the respondents have outspoken views about how to measure inequality and do apply these ideas reasonably well when asked to measure inequality.

Table 4.2.: WAP+AVE consistency (8 questions)

9.2 % (10/109) are in 100 % of the questions consistent

40 % (44/109) are in at least 88 % of the questions consistent

61 % (67/109) are in at least 75 % of the questions consistent

89 % (97/109) are in at least 63 % of the questions consistent

97 % (106/109) are in at least 50 % of the questions consistent

100 % (109/109) are in at least 38 % of the questions consistent

Some doubts may be raised about these conclusions. (i) The mere fact that only three possible answers per question were possible makes it necessary to control the results for random answers. If someone answers all questions randomly then the number of consistent answers should average 2.67 (8/3). The actual consistency rate - the total number of consistent answers divided by the total number of respondents - is 5.97. It is needless to say that the consistency rate in our questionnaire is considerably higher than in the case of random answers.

- (ii) What does remain of the consistency if we assume that the respondents have a relative measure in mind i.e. when they compare the distributions relative to total income? WAP+AVE as it stands is indeed not scale invariant. Various checks on the results show that a relative WAP+AVE measure only on few occasions yields a different outcome than its absolute counterpart and these cases do not change the consistency significantly.
- (iii) Questions might be raised about the role of the reference income. The tables in appendix B give the implications of WAP-measures taking

both the differences below and above the average into account. Recall that there may be a difference between what WAP+AVE predicts if one takes complaints below the average only and what it implies if one takes distances both below and above the average. An eminent example is Q14: the former predicts a tie (which no one of those who verbally endorse WAP+AVE has chosen) whereas the latter predicts T or B (often selected)¹³. This is an interesting fact: in this case students seem to look at dispersion at both sides of the mean.

(iv) Is it not the case that the answer patterns of the respondents are better described by a pattern implied by another aspect of inequality than WAP+AVE? Respondents might be consistent with WAP+AVE but possibly are even more in line with the answers implied by another aspect of inequality. This is not the case as table 4.3. indicates¹⁴: 63.3 % of those who verbally endorse WAP+AVE are most in line with its implications as well. Their answers come in percentage terms closer to the implications of WAP+AVE than to any other aspect of inequality. Only 10 % is inconsistent in that they are in percentage terms more compatible with another aspect than with WAP+AVE. For 26.6 % we can not say: their answer is equally in line with the implications of WAP+AVE and one or more other aspects. The particular questions are

 $^{^{13}\,\}mathrm{WAP} + \mathrm{AVE}$ implies outcome B if the weighting function is strictly increasing.

¹⁴Table 4.3. has been calculated as follows. For each respondent, I compared his pattern of answers with the pattern of answers implied by each of the nine aspects of inequality and calculated the percentage of 'correct' answers to each of those nine aspects. I obtained nine percentages for each respondent. If a respondent has endorsed WAP+AVE verbally then there are three possibilities: (i) the respondent scores the highest percentage on the implications of WAP+AVE in which case we call him consistent; (ii) the respondent scores his highest percentage on more than one aspect in which case we don't know whether he is consistent; (iii) or the highest percentages is obtained by another aspect then WAP+AVE in which case he is clearly inconsistent. Similar calculations have been performed for those who have not endorsed WAP+AVE (Others) verbally.

not sufficient to differentiate between them.

The objection does not hold for those who have chosen WAP+AVE verbally but it remarkably does apply to those who have not chosen WAP+AVE in the verbal questions. They score in the numerical questions almost always better on the pattern of WAP+AVE than on their own pattern. Almost 48 % is most in line with the predictions of WAP+AVE although they have selected a different aspect of inequality in the verbal part. For 31.6 % we are unable to differentiate between different patterns. The consistency of other aspects of inequality is considerably lower in favour of being in line with WAP+AVE. We cannot be sure that people think in terms of WAP+AVE but we are at least sure that the answers implied by this aspect of inequality are very attractive and of particular interest¹⁵.

Table 4.3.					
Verbal	%				
WAP+AVE	63.3	Consistent			
[n=109]	26.6	Consistent or inconsistent			
	10.1	Inconsistent			
Other	47.4	Most in line with WAP+AVE.			
[n=76]	31.6	Also in line with WAP+AVE			
	21.0	Not in line with WAP+AVE			

(v) Finally, there is at least one reason for a cautious interpretation

 $^{^{15}\,\}mathrm{Table}$ 4.3. also reveals that only 27 respondents (10.1 * 109/100 + 21*76/100) out of the total population of 185 (= 14.6 %) definitely reject WAP+AVE because their answers to the numerical questions are more in line with another aspect of inequality.

of the results. I am well aware that there is probably too little differentiation between the questions. Some questions imply the same outcome irrespective the chosen aspect of inequality (e.g. A3, A6, A9) which, if that particular outcome is chosen, does not learn us anything about which aspect of inequality the respondent might have in mind. If respondents choose this answer, they are consistent with any of the nine aspects of inequality making conclusions about consistency pointless. Similarly, the really 'reasonable' choices in questions A7-A14 are limited to B or T again making consistency interpretations awkward. These observations, however, do not impair the fact that each aspect of inequality does describe a different pattern of answers than each other aspect of inequality although the difference between some may be minimal (one or two questions only). The similarities are mainly created by the fixed set of questions self-imposed by our aim to replicate the Temkin poll. Further research should incorporate questions which differentiate more widely between the implications of the various aspects of inequality.

All things considered, I am rather confident in these results. My confidence is based on the observation that even before they are presented with the verbal question most respondents already have chosen answers which are in line with what they later answer to the verbal question. Even before the verbal question was raised and they were asked to think systematically in terms of complaints (distances), the respondents 'spontaneously' select the answers one particular aspect of inequality implies. That 60% of them then also picks out the WAP+AVE verbal description is very significant. Notice that the students had no option to change

their answers and only very few made visible corrections -for whatever reason- to their responses.

It may be interesting to have a closer look at the questions on which the respondents were not consistent? Are there any cases where most of the respondents who verbally endorse WAP+AVE do not choose the implied answer? Yes there are. An elaborated discussion of those questions is postponed to section 4.3.2 on the Sequence.

In all series, only very few responded in a consistent way to question A14 (see table 4.4. below) concerning a transfer in a hypothetical situation consisting of five persons. In series C the consistent answer is even never chosen. The reason for this could be the difficult nature of the question. It is the only question where the respondents were asked to compare two situations with five persons instead of four. But there might be a more fundamental problem: question 14 required respondents to compare situations regarding a Pigou-Dalton transfer. Questions concerning the transfer principle are dealt with in section 4.3.1.

4.3. Numerical Part: Diachronic view.

An alternative way to deduce valuable information from the empirical results is to consider the results in a diachronic manner i.e. to look at all the answers on one question at a time. I first deal with the questions involving a Pigou-Dalton transfer, next with questions on the Sequence and finally I compare some of our results with those of Temkin's poll.

4.3.1. The Transfer Principle.

Our questionnaire includes in total 8 questions dealing with a Pigou-Dalton transfer. Our results are comparable to earlier results on the transfer principle reported by Amiel and Cowell [1992, 1998, 1999]: empirical agreement with the transfer principle, which was widely considered as being uncontroversial, is in most cases remarkably low. Depending on the case, they found agreement with the transfer principle being between 38 % and 78 %. The highest percentage was reached in the case where a small amount was transferred from the richest to the poorest person. The agreement with the transfer principle in our population (Table 4.4.) is in general higher than in the Amiel and Cowell study (between 42.7 % and 92.4 %) although we should recall that both experiments are too different to be directly comparable: they compared other income distributions and used a simple vector presentation 16.

¹⁶The different results may point to the relevance of the presentation of the income distribution graphic presentations may induce different outcomes than simple vector presentations.

ŗ.	Table 4.4.: Transfer Principle						
% (n=92)	Q.	A more equal	B more equal	A =B			
A = (4, 6, 8, 10);	A. 7	42.4	50.0	7.6			
B = (4, 6, 9, 9)							
A = (4, 6, 8, 10);	A. 8	35.9	56.5	7.6			
B = (4, 7, 8, 9)							
A = (4, 6, 8, 10);	A. 9	1.1	91.3	7.6			
B = (5, 6, 8, 9)							
A = (4, 7, 8, 9);	A.10	6.5	92.4	1.1			
B = (5, 6, 8, 9)							
A = (4, 6, 8, 10);	A. 11	42.4	43.5	14.1			
B = (4, 7, 7, 10)							
A = (4, 6, 8, 10);	A. 12	22.8	73.9	3.3			
B = (6 , 6, 6 , 10)							
% (185)							
A = (4, 6, 8, 10);	A.B.C.13	29.2	63.2	7.6			
B = (5, 5, 8, 10)							
A = (2, 5, 5, 6, 7);	A.B.C.14	53.0	42.7	4.3			
B = (2, 5, 6 , 6, 6)							

Can we find any pattern in the way respondents answer the questions on the transfer principle? There are at least some interesting lines to point at. The agreement is high if the poorest person is at the receiving end of the transfer. On the other hand, if the richest is the transferor, agreement decreases with the rank of the transferee: the lower his rank, the more agree that the transfer decreases inequality. Many consider it particularly offensive if the transfer makes the group of the richest larger (A7, 14). Finally, attitudes seem to be scattered most if neither the poorest nor the richest are involved in the transfer: some think it increases inequality, others think it decreases it, while even few believe inequality remains unchanged (A11).

According to Temkin, this is completely in line with the egalitarian's intuition and he argues that it is inherent to the transfer principle itself that the relative agreement with it depends on the rank of the transferor and the transferee. He takes for granted that an egalitarian's first choice would be for a transfer to go from the richest to the poorest person. "... it seems intuitively clear that the egalitarian will want transfers to go from the 'highest' [richest] possible group to the 'lowest' [poorest] possible group. This intuitive claim can be derived from (a corollary of) the Pigou-Dalton condition. Suppose we could transfer n units of welfare from the first group [richest] to either the third or fourth [poorest]. According to PD, transferring n units from the first group to the third group would improve inequality, so too would transferring n units from the third group to the fourth. It follows that the inequality would be better if the n units were first transferred to the third group and then transferred to the fourth, than if it were simply transferred to the third. But the distribution resulting from the two transfers is exactly the distribution resulting from the direct transfer of n units from the first group to the fourth. Hence, PD implies that it would be better to transfer the welfare to the fourth group than to the third. Similar reasoning could be adducted to show that according to PD an egalitarian's preference ranking about transfers ... should for welfare to be transferred from the first, second, third, and fourth groups, respectively, and to the fourth, third, second, and first groups, respectively." [Temkin, 1993, 72]

This is not exactly what the empirical results tell us. Temkin's argument can not account for the disagreement with the transfer principle since he perceives each Pigou-Dalton transfer as inequality decreasing though the degree of decrease may vary with the rank of transferee and transferor. We may restate his argument, retaining his proposed egalitarian's preference ranking, but conceding that some egalitarians may consider a Pigou-Dalton transfer as not inequality reducing. If one further assumes, as Temkin implicitly does, that the egalitarian's preference ranking is expressed in the ranking of transfers according to the empirical agreement with the inequality-reducing effect of it¹⁷, then our questionnaire results are consistent with the egalitarian's intuition Temkin describes, except for question A10. This particular question A10 elicits that the corollary of the Pigou-Dalton condition is not convincingly supported by our results. Question A8 is about a transfer from the richest to the second poorest, question A10 takes this newly evolved situation as the starting situation for a transfer from the second poorest to the poorest. The distribution resulting from the two transfers is exactly the distribution emerging from a transfer from the richest to the poorest as in question A9. According to Temkin, the move from the richest to the

¹⁷The egalitarian's intuition that one PD transfer is more inequality reducing than another transfer, should then mean the same as more respondents believe that the first PD transfer is inequality reducing than respondents think that the second PD is inequality reducing.

poorest in question A9 should be better than the transfers in both question A8 and A10. This is not exactly the case because agreement in A10 is higher than in A9 -although not significantly. More research and a cautious interpretation meanwhile is certainly needed.

Nevertheless, this result suggests that the transfer from the second poorest to the poorest is the most important step. This brings us to a remarkable result. Question 13 and question A10 are very similar is this respect: a transfer of one unit from the second poorest to the poorest person. The results, however, are very different: 92.4% consider the transfer in A10 as inequality reducing whereas only 63.2% does the same in 13. Further research showed us that the lower percentage in question 13 does not depend on the series of questions the respondent previously had to answer. Of those who completed series A 68.5% answered that B was more equal than A which implies that the result has nothing to do with the previous questions. The explanation has to be found in the difference between the questions. The transfer in question 13 removes the difference between the poorest and the second poorest completely whereas the transfer in A10 bridges the gap only partially. It seems that some people are particularly offended by an increase of the group of the poorest or by getting the impression that a group clearly leaps behind the rest or has been excluded. This suggest that the number of poorest or the gap between the poorest and the richer in income terms may be important in comparing income distributions.

Let us return to the problem of those who verbally stated that WAP+AVE is the right way to measure inequality but on some numerical questions

involving a Pigou-Dalton transfer responded in an inconsistent way. On all questions on the transfer principle WAP+AVE (with the above described weighting structure) implies that B is more equal than A (what is what the transfer principle states) expect for question A7, 13 and 14 where it is indecisive. But a large number of respondents are inconsistent with the transfer principle and the implications of WAP+AVE on question A 11 and even more on question 14. Provided that students think according to WAP+AVE, one way to explain this result is to weaken the restrictions we imposed on the weighting function. These students have probably a different weighting scheme in mind than the researcher had. If we loosen the restrictions that each complaint should be treated symmetrically much can be explained - if not everything. It should be noted that then larger complaints may receive lower weights than smaller complaints and in this sense does not comply with how Temkin defines the weighting structure of WAP+AVE. Asymmetric weighting however is compatible with how we phrased WAP+AVE in Q15 and Q16. What at first glance seem to be answers which are inconsistent with the implications of WAP+AVE and the transfer principle may be completely reasonable and consistent with a different weighting scheme than we had imposed thus far. This suggests that WAP+AVE may provide a rationale for the rejection of the transfer principle in some specific cases.

For purely illustrative reasons I designed two weighting schemes which imply some of the so-called inconsistent answers. It should be clear that I do not intend to propose new -ad hoc- inequality measures since I am fully aware that the available information is highly insufficient. The first

scheme takes Temkin's suggestion serious that people may take income ranks into account. We may assume for instance that those who behave inconsistently with the transfer principle have a scheme in mind which divides each individual complaint by the income rank of that individual. Complaints of richer persons receive less weight than equal complaints of their poorer counterparts.

$$I' = \sum_{i=1}^{n} \frac{1}{i} |x_i - \bar{x}| \tag{4.1}$$

Applied to question 14, this inequality measure ranks A more equal than B (just like 53 % of our respondents do). Apart from complaints, additional information is smuggled into the inequality measure i.c. information about the income ranks.

Without additional information, the theory has furthermore difficulties to explain the behaviour of the respondents who rank A above B regarding inequality in question A11 (and to a lesser extent question A8). Both questions have in common that the transfer goes from someone above the average to someone below the average making the gap between the poorest and the second poorest larger. This suggests that the poorest becomes deprived, isolated or even excluded. This may seem particularly offensive to some. This intuition can be included if the weight given to the largest complaint not only depends on the size of the complaint but also on the distance between the poorest and the second poorest, as in

the following inequality measure.

$$I'' = (1 + x_2 - x_1)(x_1 - \bar{x}) + \sum_{i=2}^{n} \frac{1}{i} |x_i - \bar{x}|$$
(4.2)

This scheme contains the same features as the previous one except that it gives more weight to the complaint of the poorest depending on the income difference between the two poorest persons: the higher the income distance, the higher the weight given to the complaint of the poorest. If that difference is nihil then this scheme reduces to the previous one. Complaints of the same size may receive a different weight. This measure belonging to the class of WAP+AVE measures ranks B more equal than A in questions A9, A10, A12 and 13 whereas it gives a ranking contradicting the transfer principle in A7, A8, A11 and 14.

We can conclude. The empirical results highlight that the transfer principle is not likely to be considered a monolithic principle respondents either accept or reject. The transfer principle receives in some cases almost unanimous agreement whereas in other situations the majority rejects that a transfer from a rich to a poor person is inequality decreasing. The aspect of inequality most people verbally endorse namely WAP+AVE can account for this asymmetric assessment of rank-preserving transfers. I designed two simple weighting schemes -based on additional information on income ranks and distances- to show that case-dependent agreement with the transfer principle can easily be explained by WAP+AVE. It is remarkable that respondents exactly select an aspect of inequality that gives room for rejecting the transfer princi-

ple under reasonable assumptions about the weighting function which all together makes the empirical rejection of the transfer principle witnessed by Amiel and Cowell more understandable.

4.3.2. The Sequence.

Other deviations from consistency occur in the responses to question B11-12 (see Table 4.5.) and C11-12 (see Table 4.6.). These questions are part of a series of questions dealing with the so-called Sequence. To understand what is meant by the Sequence imagine a population of npeople who are alike in virtually every respect. All have the same income. Now let one individual in the population drop to a lower level of income. All except one enjoy the same high income, but the one exception has a lower income. Now let the remaining n-1 individuals start dropping to this lower income level one by one. Go until all are on the lower income level and no one has been left behind with the original high income level. The question then is: What happens to inequality along this 'sequence' of events? What happens to inequality in this story of -what has been called-unbalanced immiserisation has been discussed by Temkin [1986, 1993. Amiel and Cowell [1999, 78-79] examine the reverse story, one of unbalanced enrichment: all are at the same low income level and one by one jumps to the higher income level until all are in that position. I included both stories in the questionnaire to test for framing effects and to add to Amiel and Cowell's results¹⁸.

¹⁸ For an extensive theoretical and empirical analysis of the Sequence, see Amiel and Cowell [1994].

Table 4.5.: Sequence Unbalanced Enrichment									
% (n=46)	Q.	A more equal	B more equal	A =B					
A = (1, 1, 1, 1); B	B.7	95.7	4.3	0					
= (1, 1, 1, 10)									
A = (1, 1, 1, 10); B	B.8	32.6	65.2	2.2					
= (1, 1, 10, 10)									
A = (1, 1, 10, 10); B	B.9	56.5	41.3	2.2					
= (1, 10 , 10, 10)									
A = (1, 10, 10, 10); B	B.10	2.2	97.8	0					
= (10 , 10, 10, 10)									
A = (1, 1, 1, 1); B	B.11	6.5	26.1	67.4					
= (10, 10, 10, 10)									
A = (1, 1, 1, 10); B	B.12	21.7	21.7	56.6					
= (1, 10, 10, 10)									

Table 4.6.: Sequence Unbalanced Immiserisation									
% (n=47)	Q.	A more equal	B more equal	A =B					
A = (10, 10, 10, 10);	C.7	100	0	0					
B = (1, 10, 10, 10)									
A = (1, 10, 10, 10);	C.8	36.2	57.4	6.4					
B = (1, 1, 10, 10)									
A = (1, 1, 10, 10);	C.9	74.5	19.1	6.4					
B = (1, 1, 1, 10)									
A = (1, 1, 1, 10);	C.10	6.4	93.6	0					
B = (1, 1, 1, 1)									
A = (10, 10, 10, 10);	C.11	29.8	0	70.2					
B = (1, 1, 1, 1)									
A = (1, 10, 10, 10);	C.12	46.8	8.5	44.7					
B = (1, 1, 1, 10)									

The picture for unbalanced enrichment (Table 4.5) and unbalanced immiserisation (Table 4.6) is pretty comparable. Any change to a perfect equal distribution induces inequality (B7 and C7). Any income change resulting in a perfect equal distribution reduces inequality (B10 and C10). Almost all respondents share these insights. Any change of income towards an equal division of the population between low income and high income is considered as equality improving by a majority of respondents (B8, C8) and any income change departing from such an equal division is judged inequality inducing by most (B9, C9). When we take the results of questions 7-10 together and look how the participants judge the

inequality along the sequence, the picture is clear. In the model of enrichment 47.8 % say that inequality first increases, then decreases, increases again and finally decreases (answer pattern ABAB)¹⁹. In the model of immiserisation even 55.3 % assess the (reversed) sequence of events in the same way. No aspect of inequality allows for this combination except some measures of the WAP+AVE class.

The consistency however is less remarkable once we consider question B11-12 and C11-12. Question 11 compares two perfectly equal income distributions where one simply can be obtained by multiplying each individual income of the other by the same scalar²⁰. Under all conventional definitions of equality as well as under all aspects of inequality these two distributions are perfectly equal. But about 30 % do not agree with this view, mostly to provide support for the distribution with the highest total income. They say that an equal distribution is more 'equal' than another equal distribution. One explanation which may make this result understandable is to assume that respondents link the ideal of equality to other moral ideals as there are justice, fairness, sufficiency, or freedom and responsibility. If people are inconsistent on this question they may believe that higher total income overrides inequality considerations and the richest distribution deserves to be labelled more equal. Efficiency arguments are then a constitutive part of inequality comparisons or even of the concept of inequality itself. Exploring this is an alternative way

¹⁹ Amiel and Cowell [1994] report that 42% of the respondents answers on similar questions to B8 and B9 that inequality first decreases and then increases. This is remarkably similar to our result although we should say that their questionnaire did not include the counterparts of B7 and B10 and that the income distance between the poor and the rich was smaller (the poor had 5 instead of 1).

²⁰ Question B11 tests in fact the property of scale invariance: 67.4% thinks that multiplying a distribution with the same scalar, keeps the inequality unchanged.

one could follow to learn more about the concept of inequality which is a direction I have excluded from the outset of this paper²¹. Maybe we should not go that far and can the results be explained by a common psychological attitude. It may be that respondents tend to look for complicated answers to simple questions. Posing a very simple question to students often triggers an attitude of suspicion, disbelief and distrust forcing them to give not their first but second plausible answer. They expect academics to ask them difficult questions with complicated answers and they might consider seemingly easy questions to be trick questions.

Question 12 leaves us with another inconsistent result. In question B12 two jump to a higher level of income leaving only one person at the lowest income level. Most find that this move keeps inequality unchanged (56.6%). A considerable number (21.7%) however thinks that A is more equal than B and exactly the same number believes that the opposite is true. If we perform the sequence of events in reverse order (C12), if two drop from the higher income level to the lower one, leaving only one at the top income level, the assessment of the inequality is totally different. Only 44.7% believes that this move does not alter the inequality while even more think that the situation where more are at the higher level is better regarding inequality (46.8%). The results on question B12 are pretty much in line with the results on this kind of symmetry of which

 $^{^{21}}$ Some may now think that Pareto improvements -at least someone gets more without anyone getting less- interfere with inequality comparisons on other occasions as well. This is not the case. Take question B7 for instance. Only few (4.3%) state that they believe that the unequal distribution is more 'equal' than the equal one. They trade-off equality for extra income, even if the income has to go to one only. But the vast majority does not allow that this Pareto improvement interferes with their inequality comparison. For empirical test results of the Pareto principle see Schokkaert and Devooght [1999].

Amiel and Cowell report (see Table 4.7). They suggested that a possible explanation for the symmetry phenomenon would be that respondents evaluate distributions in terms of income differences or distances. In fact, this symmetry is only implied by the following aspects of inequality: AP+AVE, AP+ATBO, WAP+AVE, and WAP+ATBO. Only the best-off person view of complaints and/or the maximin principle of equality deviate from the symmetry.

The results on C12 -which are clearly at odds with B12 and the Amiel and Cowell results- require special attention. A closer look at all the results on the Sequence reveals that framing effects -the phenomenon that answers depend on the way certain pieces of information are presented in the questions- may play a role. In particular question 8, 9, and 12 should be scrutinized in this respect because the other questions involve equal distributions where the answers are relatively straightforward. The difference between the questions in series B and series C is only the order in which distribution A and distribution B appear. In principle this randomization should not influence the results. But there are differences between thus corresponding questions. The results on question B8 and its reverse C9 reveal considerable differences (65.2% versus 74.5% say that the situation where two are at the lower level and two at the higher level is more equal) and similar differences are apparent between the results on questions C12 and B12²². This might be understandable if students

²² There is another way to look at the results. Question B12 is the combination of question B8 and B9. Indeed, in B12 two jump to the higher level of income whereas B8 starts from the same situation but only one jumps to the higher level and subsequently another one jumps to that level in B9 to end with the same final situation as in B12. The results show that there seems to be two forces at work: the first change (B8) seems to pull most respondents in the direction of an inequality decrease whereas the second change (B9) pulls most in the direction of a inequality increase. When

consider that distribution A is first and B emerges as a modification of A, although nothing in the questionnaire explicitly told them to think in such way. If this is true than respondents see series B7-B12 as questions of unbalanced enrichment and C7-C12 as questions of unbalanced immiserisation. It is well known from psychological literature and empirical research on justice²³ that respondents deal asymmetrically with gains and losses. In our context (questions 8, 9 and 12) students seem to be more keen to state that levelling down incomes decreases inequality than that they are keen to approve that the reverse changes increase inequality. This observation should make us aware of the possibility that the results in the other sections are subject to framing as well and should urge us to more specific research and cautious and qualified conclusions meanwhile.

Table 4.7.: Amiel and Cowell $[1994]^{24}$								
% A more equal B more equal A = I								
A = (5, 5, 5, 10); B = (5, 10, 10, 10)	23	22	53					
A = (5, 5, 5, 10); B = (5, 10, 10, 10)	19	18	62					

4.3.3. The Temkin Poll Reconsidered.

Finally, I would like to compare the results of our study with the results

asked to consider both changes together (B12) most indicate that both forces cancel each other out (A=B). The same analysis can be done for the relevant questions in series C but here the results are different. Although the majority still believes that the first change decreases inequality, much more think that the second change increases it. It is not surprising that then asked to combine both changes (C12) 46.8% states that inequality has decreased. There seems to be a stronger pull in that direction in series C than in series B.

²³See Schokkaert and Overlaet [1989].

²⁴ The first row of this table is taken from a questionnaire where the sequence started from a situation in which one was at the highest income and consequently one by one jumped to the highest income until only one poor remained. The second row is from a questionnaire where the sequence started from (and ended in) a perfectly equal situation as in our questionnaire series B.

of Temkin's poll. The cases are exactly the same - except for the presentation (diagrams versus persons positioned on a line). In the first six questions A1-6 the respondents were asked to compare two income distributions where the second could be derived from the first by adding a small amount of income to the income of one or more persons - like manna falling from heaven. A majority regards such a Pareto improvement only in two cases as inequality decreasing (A1 and A3, see Table 4.8). Those cases have the common characteristic that the extra money is assigned to (one of) the poorest members of the distribution²⁵. In all other cases the Pareto improvement is considered as inequality increasing. This is an interesting point since Schokkaert and Devooght [1999] have found that when respondents are asked for an all-things-considered judgement most of the respondents do not consider a Pareto improvement as a desirable change, most probably because they also take inequality considerations into account. The actual results confirm that students have strong opinions about what happens to inequality when additional money is allotted to one and it should not surprise us when the bad of increased inequality overrides the good of a Pareto improvement.

The results are quite similar to the limited results of Temkin except for question A2 and A4. To understand the relevance of this divergent result I should explain what Temkin's poll aimed at. He hoped to find support for the claims that inequality is complex, many plausible ways

 $^{^{25}}$ If, for instance, the additional income does not go to the poorest (as in A1) but to the second poorest (as in A4) the results change dramatically. This result is consistent with our comments on questions A10 and 13 which pointed to the fact that respondents seem to be very concerned with the group of the poorest members in society.

exist to compare distributions regarding inequality and that his model of aspects of inequality has intuitive plausibility. For example, where there is virtual unanimity among the aspects of inequality, unanimity among poll respondents should strengthen his claims or where there is almost no agreement among the aspects of inequality, the respondents too, should be sharply divided. Question A2 and A4 were eminent examples of the latter: both the aspects of inequality and the respondents are sharply divided. But this larger scale experiment shows an opposite tendency. Aspects of inequality are sharply divided on question A2 but respondents remarkably seem to share a common view. On the other hand, the majority of the aspects of inequality agree that A is more equal than B²⁶ in question A4 but there is disagreement among respondents about how A compares to B. I believe that this proves that the aspects of inequality do not have equal plausibility to the respondents. If the aspects are divided, respondents may still think that only one aspect is really plausible and where the majority of the aspects favours a particular answer, respondents may still believe that a different answer implied by only one or few aspects is much more plausible. This is exactly what seems to happen in A2 and A4 respectively and is completely consistent with our earlier observation that one aspect of inequality, WAP+AVE, seem to have great plausibility to the respondents.

 $^{^{26}}$ A correct calculation of the implications of the aspects of inequality yields that the aspects of inequality are less divided than Temkin reports.

Table 4.8: Income Increase									
% (n=92)	Q.	A more equal	B more equal	A = B					
A = (4, 4, 6, 8);	A.1	6.5 [10]	88.0 [80]	5.4 [10]					
B = (4, 6, 6, 8)									
A = (4, 6, 8, 10);	A.2	82.6 [50]	17.4 [37.5]	0 [12.5]					
B = (4, 8, 10, 10)									
A = (4, 6, 8, 10);	A.3	10.9 [10]	88.0 [80]	1.1 [10]					
B = (6, 6, 8, 10)									
A = (4, 6, 8, 10);	A.4	65.2 [30]	32.6 [60]	2.2 [10]					
B = (4, 8, 8, 10)									
A = (4, 6, 8, 10);	A.5	82.6 [80]	15.2 [20]	2.2 [0]					
B = (4, 6, 10 , 10)									
A = (4, 6, 8, 10);	A.6	93.5 [100]	4.3 [0]	2.2 [0]					
B = (4, 6, 8, 12)									

The Temkin poll results are between square brackets [].

Table 4.9. epitomizes the results on the parallel questions where a misfortune befalls some, but not all, members of society. We can think of a lump-sum tax levied on some incomes and wasted by the government (an assumption to avoid any redistributive effects). The majority of the respondents judge such an income decrease as inequality increasing except in the case the tax is imposed on the richest member of the distribution. The main features apparent in Temkin's results seem to be maintained by our results. But again one case stands out: students are sharply divided on question 3. Most of the aspects imply that B is more

equal than A but only 44 % of the students (and even only 33.4% in the Temkin poll) is in agreement with this judgement. The majority (54 %) goes for the reverse judgement. The only aspects implying this ranking are the aspects involving the BOP view of complaints. Together with the conclusion on question 12, this supports Temkin's claim that the BOP view of complaints can not simply be dismissed - although this claim is at odds with the results on the verbal question (only 3.2 % selects BOP)²⁷.

Table 4.9.: Income Decrease										
% (n=93)	Q.	A more equ	al	B more	A =B					
A = (4, 6, 8, 10);	B.C.1	96.8 [10	00]	2.1	[0]	1.1	[0]			
B = (2 , 6, 8, 10)										
A = (4, 6, 8, 10);	B.C.2	89.2 [8	80]	10.8	[20]	0	[0]			
B = (4, 4, 8, 10)										
A = (4, 6, 8, 10);	B.C.3	53.8 [66	.6]	44.1	[33.4]	2.1	[0]			
B = (4, 6, 6, 10)										
A = (4, 6, 8, 10);	B.C.4	17.2	[0]	82.8	[100]	0	[0]			
B = (4, 6, 8, 8)										
A = (4, 4, 6, 8);	B.C.5	76.3		23.7		0				
B = (4, 4, 4, 8)										
A = (4, 6, 8, 10);	B.C.6	76.3		19.4		4.3				
B = (4, 4, 6, 10)										

The Temkin poll results, where available, are between square brackets [].

In sum, these results are a necessary refinement of the Temkin poll but

 $^{^{27}\}mbox{"}$ Given the empirical responses ..., it seems clear that BOP influences people's egalitarian judgements and, in some cases, at least, to a large extent. The question is whether it sometimes influences our judgements more than it should." [Temkin, 1993, 69]

basically do not provide enough new insights to deny his conclusion: "To an extent, then, the poll's empirical results and ... [the] theoretical results are mutually supportive. The implications of the latter suggest that the poll has (largely) succeeded in eliciting people's egalitarian judgements. In turn, the former seems to provide further support ... for the theoretical claims." [Temkin, 1993, 64].

5. Conclusions.

Measuring inequality by counting complaints may eventually seem to be not that extravagant as it may have sounded initially. I have given a glimpse of the philosophical framework Temkin provides for building an approach to inequality on the concept of complaints and for measuring inequality accordingly. Basically, complaints acknowledge that it is unjust or unfair to be worse off than someone else through no fault or choice of your own. A two-step procedure has been put forward to measure inequality of which the first step measures each individual's complaint according to a particular view of complaints. The logical second step aggregates the individual complaints according to a particular principle of equality to obtain the overall inequality in a society. This procedure generates nine possible positions or inequality measures which he argued have great plausibility since they express intuitions which influence our egalitarian judgements, although they are not all equally appealing.

I then moved on to submit Temkin's ideas to an examination by means of the questionnaire-experimental method. I wanted to find out whether the aspects of equality have any plausibility for our student-respondents and if so, whether there are aspects which are more appealing than others. The design of our questionnaire enabled us to compare our results with those of Temkin's poll. The results were encouraging (4.3.3). From the many comments which almost all were compatible with the model came up that the students felt comfortable with the theoretical setting. Only a few wanted to change the whole approach. We furthermore found little to deny Temkin's conclusion that the empirical results and the theoretical insights were mutually supportive. This all lifted our confidence in the ability of the theoretical model of Temkin to structure people's attitudes towards inequality comparisons.

Such a model was exactly what we needed to pursue the main objective of our research: to acquire some insights in the way respondents measure and compare the inequality of different distributions. The results occurred to be quite striking. Both the verbal responses (4.1) and the synchronic reading of the numerical results (4.2) show that a considerable majority of the respondents might be influenced by the intuitions included in the combination of the Weighted Additive Principle and the relative to the Average view of complaints (WAP+AVE). It is of great importance for our conclusion that most of the students' judgements in the numerical part did not contradict their responses in the verbal part as this is proof of the consistency of their answers. These empirical results may guide empirical researchers in their choice of a theoretically attractive and widely acceptable measure as well as future theoretical research -for example on how the weighting scheme should look like or how other aspects as BOP play a role. Research on how to link the Temkin-

measures with existing inequality measures is promising as well. It can be shown that the Generalized Entropy measure for c=2 belongs to the WAP+AVE class of measures and we know that the GE class of measures is renowned for possessing many interesting properties (as scale and population invariance, transfer principle, decomposability). Measures ordinally equivalent to this particular GE measure include the variance and the coefficient of variation. Besides, AP+ATBO represents the intuition behind the well-known Gini coefficient but the verbal results made clear that our respondents have completely neglected this aspect of inequality, leaving the case for the Gini coefficient far less convincing than many economists still believe.

A diachronic reading of the numerical results (4.3) added to the conclusion that the WAP+AVE measures have some very appealing features. The WAP+AVE measures proved to be helpful in explaining the large empirical rejection of the sacrosanct transfer principle of which Amiel and Cowell report (which are confirmed by our results (4.3.1)). I have shown that some weighting schemes accept some Pigou-Dalton transfers as inequality decreasing whereas they consider other transfers as inequality increasing. It is quite remarkable that most respondents select an aspect of inequality (WAP+AVE) that gives room for rejecting the principle in some cases while accepting it in others what is exactly what they do in the numerical questions. This confirms the suggestion of Amiel and Cowell that people might think in terms of distances and I may now even add that especially distances relative to the average seem to be most important of all (although I suggested that the distance between the poorest

and the second poorest in some cases may be relevant as well).

Some words of caution are in place here. To know more about how people compare income distributions we should ask them explicitly by means of interview techniques for instance. Our research might be valuable in this respect since it may provide the direction in with questions should be modelled to acquire insights in people's egalitarian judgements. Improvements and further research are certainly required and not in the least repeating the questionnaire study with various (also non-student) populations and with a more elaborated choice of numerical questions to differentiate even further between what the various aspects of inequality imply. More research on possible framing effects is needed as well. We should finally keep in mind that the results are limited by the framework imposed on the questionnaire in casu that of individual complaints as the basis for inequality measurement. But for those who accept this framework, and a lot can be said in favour of it, these results can not be easily dismissed.

References.

Amiel, Y. & Cowell, F.A., Measurement of Income Inequality: Experimental Tests by Questionnaire, in Journal of Public Economics 47 (1992), 3-26.

Amiel, Y. & Cowell, F.A., Inequality Changes and Income Growth. In Eichhorn, W. (ed.), Models and Measurement of Inequality, Berlin, 1994, 3-26.

Amiel, Y. & Cowell, F.A., Distributional Orderings and the Transfer Principle: A Re-examination, in Research on Economic Inequality 8 (1998), 195-215.

Amiel, Y. & Cowell, F.A., Thinking About Inequality, Cambridge U.P., 1999.

Arneson, R.J., Equality and Equal Opportunity for Welfare, in Philosophical Studies 56 (1989), 77-93.

Atkinson, A.B., On the Measurement of Inequality, in Journal of Economic Theory 2 (1970), 244-263.

Broome, J., What's the Good of Inequality? In J. D. Hey (eds.), Current Issues in Microeconomics, Macmillan, 1989, 236-262.

Cohen, G.A., On the Currency of Egalitarian Justice, in Ethics 99 (1989), 906-944.

Cohen, G.A., Equality of What? On Welfare, Goods and Capabilities, in Recherches Economiques de Louvain 56 (1990), 357-382.

Dworkin, R., What is Equality? Part 1: Equality of Welfare, in Philosophy and Public Affairs 10 (1981a), 185-246.

Dworkin, R., What is Equality? Part 2: Equality of Resources, in Philosophy and Public Affairs 10 (1981b), 283-345.

Frankfurt, H., Equality as a Moral Idea, in Ethics 98 (1987), 21-43.

Hartkamp, S.F., Equality; A Moral Realistic View. Towards A Simple Measure of Inequality, 1999, PhD., Groningen.

Kolm, S.Ch., The Optimal Production of Social Justice. In J. Margolis and H. Guitton (eds.), Public Economics, Macmillan, London, 1969, 145-200.

Rabinowicz, W., The Size of Inequality and Its Badness. Some Reflections around Temkin's Inequality, mimeo, 1999.

Rawls, J., A Theory of Justice, Oxford, 1971.

Schokkaert, E. & Devooght, K., Responsibility-Sensitive Fair Compensation in Different Cultures, 1999, DARP 46,1999, STICERD, LSE.

Schokkaert, E. & Overlaet, B., Moral Intuitions and Economic Models of Distributive Justice, in Social Choice and Welfare 6 (1989), 19-31.

Temkin, L.S., Inequality, in Philosophy and Public Affairs 15 (1986), 99-121.

Temkin, L.S., Inequality, Oxford, 1993.

Appendix A: The Questionnaire.

In this appendix we give the complete version of questionnaire A. Questionnaire B and Questionnaire C are exactly the same, except for different income distributions A and B which are given in the text (see 4.2.-4.3.).

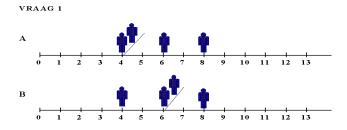
QUESTIONNAIRE INEQUALITY

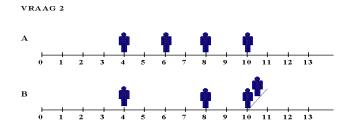
This questionnaire concerns people's attitude to inequality. We would be interested in your view, based on hypothetical situations. Because it is about attitudes there are no "right" answers. Some of the possible answers correspond to assumptions consciously made by economists: but these assumptions may not be good ones. Your responses will help to shed some light on this, and we would like to thank you for your participation. The questionnaire is anonymous.

Alfaland consists of 4 persons who are identical in every respect other than the income they earn. Two economic policy proposals A and B are being considered for implementation in Alfaland next year. It is known that - apart from their impact on personal incomes - the two policies would have the same effect on the population. The impact upon the persons' incomes would depend upon the particular state of the Alfaland economy at the time the policy (A or B) is to be introduced.

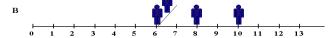
In each of questions 1 to 14 two alternative lists of incomes A and B (in Alfaland local currency) are given. Each of these pairs represents the outcomes of the A-policy and the B-policy on incomes of the four persons

in each of 14 different situations in which Alfaland might find itself next year. In each case please state which policy you consider would result in lower inequality in Alfaland by circling A or B. If you consider that the two policies will result in the same inequality then circle both A and B.

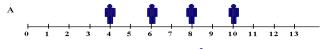






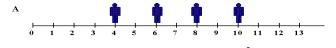


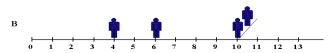
VRAAG 4

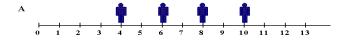


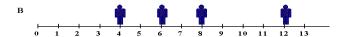


VRAAG 5

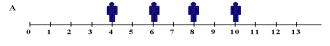


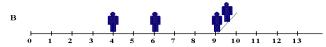




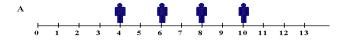


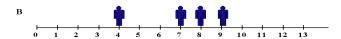
VRAAG 7

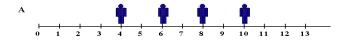


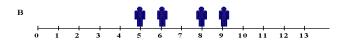


VRAAG 8



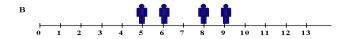




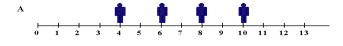


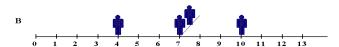
VRAAG 10

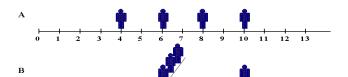




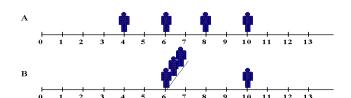
VRAAG 11

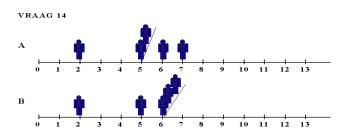






VRAAG 13





Question 15:

Let us have a closer look at the hypothetical situation of Alfaland.

Some economists assume that perfectly identical persons (as in Alfaland) which incomes differ have a complaint regarding inequality. Those economists however do not know exactly how to measure the size of such a complaint. That's what this question is all about. We offer you some alternative proposals how one might measure the size of complaints. Please circle the letter alongside the view that corresponds most closely to your own. Feel free to add your own proposal or any comment that explains the reason for your choice.

- A) The inequality an arbitrary person in Alfaland is entitled to complain about is the difference between his income and the highest income.
- B) The inequality an arbitrary person in Alfaland is entitled to complain about is the difference between his income and the average income.
- C) The inequality an arbitrary person in Alfaland is entitled to complain about is the sum of the differences between his income and all higher incomes.
 - D) None of the above.

Own proposal or comment:

Question 16:

Those economists assume furthermore that the overall inequality in Alfaland is determined in some way or another by the size of the individual complaints. We offer you again some alternative proposals. Please circle the letter alongside the view that corresponds most closely to your own. Feel free to add your own proposal or any comment that explains the reason for your choice.

A) The total inequality in Alfaland is equal to the sum of all the

complaints.

- B) The total inequality in Alfaland is equal to the size of the largest complaint.
- C) The total inequality in Alfaland is equal to the weighted sum of all the complaints where the larger complaints are given a higher weight.
 - D) None of the above

Own proposal or comment:...

Thanks

Appendix B: What the Aspects Imply.

The tables below indicate which judgement each aspect of inequality implies: 'A' means that a particular aspect of inequality considers distribution A having less inequality than distribution B; 'T' (tie) means that it considers equality in both distributions to be equal; 'B(t)' means that if we apply leximin as tie-breaker when maximin implies a tie, B has less inequality than A and '?' means that any outcome is possible depending on the weighting structure one imposes. Notice that I calculated WAP+AVE relative to all the deviations from the average; not only those below the average!

A	MP			AP				WAP	
	ВОР	AVE	ATBO	ВОР	AVE	ATBO	ВОР	AVE	ATBO
A1	B(t)	A	A	В	В	В	В	?	В
A2	B(t)	A	A	В	Т	Т	В	?	A
A3	В	В	В	В	В	В	В	В	В
A4	B(t)	A	A	В	В	В	В	?	В
A5	B(t)	A	A	В	A	A	В	?	A
A6	A	A	A	A	A	A	A	A	A
A7	В	Т	B(t)	В	Т	В	В	?	В
A8	В	B(t)	B(t)	В	В	В	В	В	В
A9	В	В	В	В	В	В	В	В	В
A10	В	В	В	Т	Т	В	В	В	В
A11	B(t)	B(t)	B(t)	Т	В	В	?	В	?
A12	В	В	В	Т	В	В	?	В	В
A13	В	В	В	Т	Т	В	?	?	?
A14	В	Т	B(t)	В	Т	В	В	T or B	?

В		MP			AP			WAP	
	ВОР	AVE	ATBO	ВОР	AVE	ATBO	ВОР	AVE	ATBO
B1	A	A	A	A	A	A	A	A	A
B2	A(t)	В	В	A	A	A	A	?	A
В3	A(t)	В	В	A	В	В	A	?	В
B4	В	В	В	В	В	В	В	В	В
В5	A(t)	В	В	A	Т	В	A	A	?
В6	A(t)	В	В	A	Т	Т	A	?	A
В7	A	A	A	A	A	A	A	A	A
В8	B(t)	A	A	В	A	A	В	?	A
В9	B(t)	A	A	В	В	В	В	?	В
B10	В	В	В	В	В	В	В	В	В
B11	Т	Т	Т	Т	Т	Т	Т	Т	Т
B12	B(t)	A	A	В	Т	Т	В	Т	Т
B13	В	В	В	Т	Т	В	?	?	?
A14	В	Т	B(t)	В	Т	В	В	T or B	?

С	MP			AP			AP WAP		
	ВОР	AVE	ATBO	ВОР	AVE	ATBO	ВОР	AVE	АТВО
С1	A	A	A	A	A	A	A	A	A
C2	A(t)	В	В	A	A	A	A	?	A
С3	A(t)	В	В	A	В	В	A	?	В
C4	В	В	В	В	В	В	В	В	В
С5	A(t)	В	В	A	Т	В	A	A	?
С6	A(t)	В	В	A	Т	Т	A	?	A
С7	A	A	A	A	A	A	A	A	A
С8	A(t)	В	В	A	A	A	A	?	A
С9	A(t)	В	В	A	В	В	A	?	В
C10	В	В	В	В	В	В	В	В	В
C11	Т	Т	Т	Т	Т	Т	Т	Т	Т
C12	A(t)	В	В	A	Т	Т	A	Т	Т
C13	В	В	В	Т	Т	В	?	?	?
C14	В	Т	B(t)	В	Т	В	В	T or B	?