

THE WORLD BANK GROUP ARCHIVES

PUBLIC DISCLOSURE AUTHORIZED

Folder Title: RPO Number 670-84 - Growth, Employment and Size Distribution of Income - Volume 3

Folder ID: 30124630

Series: Project Managers' Research Project Files

Dates: 11/06/1975-01/19/1981

Sub-Fonds: Records of the Office of the Vice President, Development Economics and Chief Economist and later Senior Vice President, Development Economics and Chief Economist (DECVP)

Fonds: Records of the Office of the Chief Economist

ISAD Reference Code: WB IBRD/IDA DEC-03-76

Digitized: 2/18/2021

To cite materials from this archival folder, please follow the following format:
[Descriptive name of item], [Folder Title], Folder ID [Folder ID], ISAD(G) Reference Code [Reference Code], [Each Level Label as applicable], World Bank Group Archives, Washington, D.C., United States.

The records in this folder were created or received by The World Bank in the course of its business.

The records that were created by the staff of The World Bank are subject to the Bank's copyright.

Please refer to <http://www.worldbank.org/terms-of-use-earchives> for full copyright terms of use and disclaimers.



THE WORLD BANK

Washington, D.C.

© International Bank for Reconstruction and Development / International Development Association or

The World Bank

1818 H Street NW

Washington DC 20433

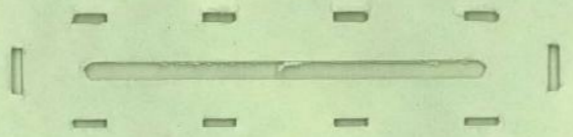
Telephone: 202-473-1000

Internet: www.worldbank.org

PUBLIC DISCLOSURE AUTHORIZED

COMPLETED PROJECT

670-84
Growth, Employment and Size Distribution
of Income Vol. III



The World Bank Group
Archives
30124630
R1986-027 Other #: 7 3554B
RPO Number 670-84 - Growth, Employment and Size Distribution of Income - Volume 3

DECLASSIFIED
WBG Archives

Mr. Hollis Chenery, VPD

Gobind Nankani, VPD

Pending Evaluation of Completed Research Projects

January 19, 1981

670.06
cc 670.29
cc 671.61
cc 670.71

cc 670.80.

cc 670.84 ✓

670.85

671.14

671.41

In FY80 and FY81 ten completed research projects were evaluated (RPO Nos. 670-06, 670-29, 670-71, 670-80, 670-84, 670-85, 670-98, 671-14, 671-41, 671-61). This brings to 25 the number of completed research projects that are yet to be evaluated. Of these 25, three were completed within the last twelve months and, as is usual, shall not be evaluated until that period is up. Of the remaining 22, 9 are to be evaluated by one or the other of the Steering Groups. The other 13 are to be evaluated by ad hoc panels in the coming months, especially the summer.

cc: Messrs. S. Acharya
E.B. Waide

GN:aa

RPO 670-84

Editorial Sub-Committee

September 25, 1980

Shankar Acharya, VPD

Evaluation Report on RPO 670-84

Attached as promised is a copy of the evaluation report by B.B. King on this research project. It has interesting things to say about the process by which the Ranis book on Taiwan came to be published. It also raises the general issue that I brought up at our last meeting regarding procedures for ensuring that adequate efforts are undertaken to take account of serious criticisms of referees.

cc. without attachment: Mr. B.B. King

OFFICE MEMORANDUM

SEP 24 1980
 J. Duloy
 DRC
 LIBRARY/FILE COPY
 (F)

TO: Mr. S. Acharya

DATE: September 23, 1980

FROM: Benjamin B. King *bank*

RPO 670-84

SUBJECT: Evaluation of RPO 670-84I. Introduction

1. The panel, consisting of Messrs. Baird, El Serafy, Hawkins, Mohan and myself met on August 15, 1980. Mr. Duloy was present from the DRC. Subsequently, a draft outline and a complete draft were circulated. The report, which follows, takes the comments of the panel and of Messrs. Duloy and Pyatt of the DRC into account.

2. This project probably cost something of the order of \$300,000.^{1/} It cannot be judged a success. It was carried out by consultants, who appear to have thought at the outset that they had received a grant for a proposal which they could thenceforward carry out with a relatively free hand. From the very beginning and throughout the project, the department chosen to sponsor the project, the DRC, were critical of its conduct - in the panel's view, rightly. The result was an acrimonious relationship between the Bank and the consultants. The final product was a book, which is so flawed that, in the panel's view, its publication under the Bank's imprimatur does the Bank no credit. Since the book illustrates much of what was wrong with the project, we shall start our report by describing its flaws (Section II). Then we shall describe the course of events (Section III). We believe that, despite the flaws in the book, the project has not been wholly without merit (Section IV). Finally, we draw our conclusions (Section V).

II. What is Wrong with the Book?

3. The purpose of the book can be briefly stated in the authors' own words. They note that previous writers, in their work, 'discern an inverse U-shaped relation between growth and equity. They conclude that..... the distribution of income must first worsen before it can improve..... Taiwan is one exception..... Taiwan's family distribution of income in the 1950s was not very different from the unfavorable levels most LDCs seem to be prey to in the early years of their transition effort. But that distribution has substantially improved during two decades of rapid growth. This "deviant" record should therefore be of interest to academicians and policy makers.' They wanted to explore, with particular emphasis on the years 1964 to 1972, how, and if possible, why it was that Taiwan achieved this unusual result. There can be little question that the subject is an interesting one.

^{1/} Including staff-time of about 1 man-year which may well be an underestimate. The dollars are approximately 1975 vintage.

4. Unfortunately, the authors saddled themselves, if we can so put it, with an albatross. This was the methodology they used and which they clung to, despite observations of various commentators, at different stages of the work, which cast doubt on its adequacy. To do justice to the debate on this methodology, would require something much longer than this report. Nevertheless, we feel that a minimum description of it and its alleged faults is necessary. We will, therefore, touch on the main points of the arguments without going into details.

5. The central unit of observation used by the authors is family income distribution (FID). The Gini coefficient is used as the index of inequality. The thrust of the methodology is to decompose this Gini coefficient into component parts as follows:

$$G = \sum G_i \phi_i$$

The Gini coefficient (G) in this equation is equal to the sum of the Gini coefficient (G_i) of different factor incomes (agricultural, wage, property) multiplied by the weights (ϕ_i) of these factor incomes in total income. The authors then go on to explain changes in the Gini for total income by changes in its six component parts (the three factor Ginis and the three factor shares): or, to be more precise, they combine these changes in a particular way and describe the three combinations as three "effects".

6. The heart of the difficulty lies in the fact that the above equation cannot be true except in very special cases. No such precise decomposition exists or can exist. However, two similar decompositions do exist. In one case the Gini coefficient for the factor income can be replaced by something known as the "concentration ratio" or pseudo-Gini, as the authors call it. In the other the true Gini can be replaced by what we shall call the estimated factor income (EFI) Gini. A note on these two is given in Annex 1. Since the authors give prominence to the EFI Gini by discussing it at length in Chapter 3, the reader might reasonably draw the conclusion that this is the basis for their contention that the error introduced by using the true Gini is sufficiently small that it can be neglected. Their case would then rest on the validity of this contention in the sense that each EFI Gini is a good approximation of the corresponding true factor income Gini and that this also applies to changes in each from year to year.

7. The principal evidence that the authors bring to bear to support their contention is on pages 92 and 93. Here, they give figures for the total error which is indeed quite low. However, it is somewhat difficult to reconcile this with other figures given on pages 102 and 103,

which suggest that the change between 1964 and 1968 for the true total Gini is positive, while the estimated change is negative.1/

8. It is not clear how the overall error was calculated, since there are certain residual incomes for which no information is given. In any event, the overall error is not the only error to be considered. This overall error is the sum of three component errors, one for each of the three factor incomes. There is solid evidence that these individual errors are much larger, but offset each other. This casts considerable doubt on the validity of the identification of the EFI Gini with the factor income Gini, as calculated by the authors.2/

9. We would, however, regard these approximations as not the principal threat to the validity of the argument. It is entirely possible that if they were properly documented and if the consequent limitations to the conclusions on changes from year to year were properly stated, they would not necessarily destroy the major part of the analysis.

10. A more serious deficiency in the analysis arises because the authors use grouped data, arranged by deciles of the population according to total income. A Gini coefficient for total income constructed from decile averages is different from a Gini coefficient constructed from individual incomes of the total population. But the direction of the error introduced is well known and it is bounded. This in itself is then not too serious a problem.

11. The same is by no means true for the individual factor incomes. Because individual factor incomes are ordered by total income, group or decile averages may conceal extreme variations; for example, some of the factor incomes in each decile are intrinsically likely to be zero and the number of such zero incomes is likely to change from year to year. The possibilities for error are thus great. Indeed, calculations from ungrouped or computerized data carried out after the main work on the project was completed, but well before the publication of the book, demonstrate that factor income Ginis calculated from decile data are very different from true Ginis, the latter in some cases being three or four times the former. It is hardly putting it too strongly that the story that is told is about a statistical artifact rather than about the real world.3/

1/ This discrepancy led to the somewhat anomalous "Finding 3.4" on p. 102 that one "highly favorable" effect overwhelmed (sic) a "highly unfavorable" effect to cause the slight worsening of income distribution.

2/ In fact, the problem of offsetting errors would not arise, if the authors based their case on the Pseudo-Gini. If they did it is hard to see why they put the analysis of the Pseudo-Gini in Chapter 9, near the end of the book, and that on the EFI in Chapter 3, near the beginning.

3/ The authors refer to the possibility of error from using grouped data in the last eight pages of the book, pointing out that they had no choice but to use grouped data.

12. What is perhaps most disturbing is that few of these points are immediately obvious to even a well qualified reader. Most readers would not have the patience to undertake the analysis required to uncover the possible sources of error, much less to assess their magnitude. Indeed in many cases, it is impossible from the material presented to attempt an assessment of this magnitude.

13. We now turn to questions of data. Most of the statistics on which the authors rely originate from household surveys conducted by one organization, the Directorate-General of Budget, Accounting and Statistics (DGBAS) between 1964 and 1976. The fact that one organization was responsible for the majority of the data is a distinct plus. Nevertheless, household surveys such as these are subject to well-known sources of error, to which the authors do indeed refer.^{1/}

14. The authors state that the inferences they derive "should not be sensitive to small variations" (page 13). We are inclined to doubt whether they have followed their own prescription. This is especially the case because DGBAS data were not available for the city of Taipei from 1968 onwards. So surveys by the administration of the city had to be substituted. This inevitably raises questions of continuity. They would not be so serious, were it not for the fact that the authors place a great deal of emphasis on the year 1968 as a "turning point" when surplus labor ceased to become available. There is such a sharp break in 1968 in some of the data, that inevitably one is led to the suspicion that the abrupt changes ascribed to economic factors may, in fact, have been due to the simple fact of discontinuity in the data sources. Moreover, series constructed from computerized ungrouped data show little evidence of a particular break in 1968.

15. The Bank employed a highly qualified reader (Professor A. B. Atkinson) to review the book before publication. While he recognized the interest in the subject and the possible contribution that the authors were making to it, he had criticisms, both in general and in detail. He made a series of comments on two successive drafts. Most, if not all, of our criticisms can be found in his reviews. The comments which are of particular interest to us are those in which he expresses his doubt whether the "data are sufficiently robust to stand" the use to which they were put. In his second round of comments, he noted that inadequate notice had been taken of his more general criticisms on the first round. He posed three possible causes of action:

^{1/} On page 97 there is a table comparing the shares of the three factor incomes according to (a) the household surveys used and (b) the national accounts. There are some substantial differences. This phenomenon was observed in extenso in Latin America by Oscar Altimir.

- i) publish the book as it stands (with editorial revisions),
- ii) ask the authors to revise the statement of the findings so as to bring out the necessary qualifications (which would require more than simply adding a few footnotes),
- iii) wait until the computerized data are available, so that the sensitivity of the results can be checked.

He recommended the second, noting, however, that the findings might turn out to need substantial modification when richer data are available.

16. It is our opinion that the revision Professor Atkinson called for was not carried out on the second round. In fact, there is one instance where he pointed out that a numerical error had been made, both on the first round and on the second round. It has not been corrected in the final edition (page 69, line 3, third column).

17. The style of the book is opaque and often confusing to the reader. As a simple example, it is often unclear whether the authors are referring to the Gini coefficient of a particular factor income in the true sense (using grouped data) or the Gini coefficient of the estimated factor income (the EFI Gini); there is an example in the table on page 76. The authors complicate their description of the process of decomposition by introducing transfer income, which is likely to have special characteristics different from other factor incomes. This is unnecessary, because they never use transfer income subsequently. The discussion could as well have been put in an appendix. The authors' findings, which we regard as being based on misleading evidence, are presented in such a way that it is hard to extract the gist of them. The mathematical treatment is in many cases unnecessarily labored. We doubt whether anyone other than specialists will be tempted to read very far.

18. In several respects the authors were criticized by Professor Atkinson for not taking sufficient account of previous work in particular fields, for example, on earnings functions in Chapter 4 and on taxation in Chapter 6. From our own point of view, we note that, throughout, the analysis was in terms of family income. It is by now well established--in large part, through work conducted by the DRC--that family income distribution is not the same as individual income distribution; the correspondence between low-income families and low-income individuals is far from being one-to-one, for the obvious reason that families differ in size and composition. This fact might at least have been mentioned.

19. The objective of the book was to relate the particular path of growth that Taiwan took which caused or influenced the unusual phenomenon of greater equality. The investigation of this connection was more or less neglected after Chapter 2. Chapter 2 itself is rather limited in its analysis. A richer treatment would have related the trends in the disaggregation of income inequality with real economic phenomena. Such phenomena would include a description of the particular growth path; the causes of an effective land reform; the tracing of the movement of population; the mix and concentration of industry. As it stands, little insight is gained on how or why Taiwan was able to achieve growth with equity.

III. How did this come about?

20. The Bank's formal research program started in 1972 and this project was one of the early ones. Understandably, at the beginning, there was some anxiety to launch the program, and little experience to go on. This was particularly true of projects in the field of income distribution, several of which were started at this time. One proposal was submitted by the Economic Growth Center at Yale, covering seven countries that was much more ambitious than the project that finally emerged.^{1/} It was criticized strongly by the Development Research Center (DRC). Their criticism which was intended as an internal document, was passed on to the Yale Growth Center. Considerable pressure was brought to bear on the DRC to moderate their criticism in view of the high reputation of the Center and the researchers involved. The Research Committee's review panel was lukewarm in its report on the proposal and it recommended only conditional financing at best.

21. The Economic Growth Center, on its part, had previously received grants from USAID to support its research. There is evidence to suggest that they regarded the grant from the Bank as a similar sort of core support, rather than a source of funds for carrying out a particular project, in the form and content of which the Bank had a considerable interest. Indeed, the bureaucratic procedures of the Center, as it turned out, were not well adapted to coping with grants that depended on the delivery of a specific product. This misunderstanding was one source of antagonism; another was the fact that they had been shown the DRC's critical memorandum, which would probably have been worded differently had the authors known where it was going.

^{1/} The original proposal would have cost something of the order of \$1-1/2 million in current dollars.

There seems to have been an unfortunate atmosphere in which the authors felt they were being "got at" by the DRC and the latter, in its turn, felt that the authors were not taking seriously reasonable and sincere criticism. The authors at least gave the appearance of feeling that the way to deal with this situation was to appeal to higher authority. The atmosphere was not exactly ideal for carrying on a cooperative project.

22. In retrospect, it appears that the authors must have set their minds on using this particular project to try out the particular methodology they had chosen. Little else would explain their persistence in carrying on with it in the face of various kinds of adverse criticism. The Bank's interest lay, quite differently, in exposing, by whatever eclectic means was available, the root causes of the success story of Taiwan. Where any particular methodology would help, that methodology could be used. But the purpose was not to use one particular methodology or another, but to get at the truth as far as possible. Thus, there was another fundamental source of conflict.

23. The process of publication was in some ways a repetition of the process of acceptance of the proposal. The DRC, by now, had so accepted its adversary role that it agreed with the view that it should not act as judge and jury. It provided the Editorial Sub-Committee with copies of its comments, but did not pursue the matter further than that. This deprived the Sub-Committee, which was supposed to pass on the merits of the book, of valuable advice. To be sure, the Bank's reviewer (Professor Atkinson) did an excellent and conscientious job. As we have pointed out, he made many serious criticisms; but his recommendation that the book be published was conditional on his criticisms being taken into account.

24. Again, there was pressure from higher authority to hasten the process of publication. And, again, the reviewer's identity was revealed to the authors without his knowledge. In the event the necessary follow-up to determine whether Atkinson's criticisms had been taken into account fell between two stools. The Editorial Sub-Committee probably did not have the expertise or the mechanism to see that this was done. The DRC clearly felt that, once the decision in principle had been made to publish, it would be invidious for them to take the responsibility for oversight.

25. The panel felt that, in its conduct as sponsors of the project, the DRC did its best in very difficult circumstances. They pursued the task of maintaining the quality of the research conscientiously and efficiently; they put in much more staff time than is shown in the Completion Report and much more than they would ordinarily be expected

to in an alternative project of this kind. The fact that the consultants took their criticisms with an ill grace and that the final product is not one to be proud of, cannot be laid at the DRC's door.

IV. The Saving Grace

26. Despite all we have said, we feel that this project has a limited pay-off. While the proposal as such had serious faults in the way it was conducted from start to finish, it is not merely the project itself that we should be thinking about. The project was itself part of a process. Much further enquiry has been stimulated into the nature of the decomposition of indices of income distribution and into the magnitude of the difference between the analysis of grouped and ungrouped data. The search continues for a better description and explanation of the evolution of income distribution in Taiwan. One fitting outcome of the controversy is an article, written jointly by one of the authors of the book and a staff member of the DRC who criticized it. We should recognize that false steps can have their virtues.

V. The lessons

27. There is a formal list of questions which evaluation panels are asked to answer. We append ours as Annex 2. However, we prefer to treat these questions as a kind of checklist. We will also put our own conclusions in our own way. The principal ones are as follows:

- (1) The review process. This should be conducted by an independent panel, free from hierarchical pressure. In fact, this is now the case, as far as we can see.
- (2) Clarification of the Bank's role as a source of funds. The Bank should always make it clear, when collaborative research is undertaken, what its role is and what the conditions for the release of funds are. We believe that its performance in this respect is now much superior to what it was. However, to the extent that the Bank moves into special arrangements to support research activity in LDCs, the need for clarity may become more acute.
- (3) The responsibility of sponsors. Under the present system, all research proposals have to be sponsored by a particular department in the Bank. Unless this system is changed in some way, we feel that the responsibility of the sponsoring department must be respected. Part of that responsibility

September 23, 1980

lies in making the judgment as to when a difference of opinion with outside consultants on research ceases to be merely a difference of opinion and becomes a question of quality control. This responsibility should not be diluted without cause.

- (4) Anonymity of reviewers. This should be respected, whether the review is formal or informal. Otherwise, the Bank will be unable to command honest reviews.
- (5) Follow-up procedure by the Editorial Sub-Committee. When a decision to publish is made subject to the authors "taking into account reviewers' comments", we feel that a specific procedure should be set up to ensure that this is done.

Attachments

cc: Messrs. Baird
El Serafy
Hawkins
Mohan
Nankani
Duloy
Pyatt

Annex 1

The Gini coefficient

There are several formulae for the Gini coefficient. A convenient one for present purposes is as follows:

$$G = 2/nS \sum_{r=1}^n (y_r - \bar{y}) r$$

where y_r is a set of incomes (in this case) ordered so that they increase monotonically, i.e., for any $r = i, j$ and $i > j$, then $y_i \geq y_j$.

In addition $S = \sum_{r=1}^n y_r$

The concentration ratio ^{1/}

Suppose there are, as in the present case, two sets of related variables; total income y_r ; and a factor income x_{ir} forming part of each total income y_r . Then the concentration ratio for x_{ir} , in relation to total income, is defined as:

$$C_i = 2/ns_i \sum_{r=1}^n (x_{ir} - \bar{x}_i) r$$

where $s_i = \sum_{r=1}^n x_{ir}$

It should be noted that, here, the ordering depends on y_r , not x_{ir} . Clearly, if x_{ir} are large for small r and vice versa, the concentration ratio can be negative.

^{1/} Or Pseudo-Gini,; see FRK pp. 352 ff.

If there are m such factor incomes equal to total income, such that $\sum_{i=1}^m x_{ir} = y_r$, it follows that

$$\begin{aligned} \sum_{i=1}^m C_i s_i &= 2/n \sum_{r=1}^n r \sum_{i=1}^m (x_{ir} - \bar{x}_i) \\ &= 2/n \sum_r (y_r - \bar{y}) \\ &= GS \end{aligned}$$

If we denote the weight of each factor income in total income as $w_i = s_i/S$:

$$\sum w_i C_i = G$$

The Estimated Factor Income Gini

Suppose, as before, there is a set of factor incomes x_{ir} whose sum is equal to total income y_r . Suppose, further, that each factor income is regressed on total income with the result that the estimated factor income, denoted by \hat{x}_{ir} , is related to y_r as follows:

$$\hat{x}_{ir} = b_i + a_i y_r$$

Hence $\hat{x}_{ir} - \bar{x}_i = a_i (y_r - \bar{y})$

If we form the concentration ratio for \hat{x}_{ir} :

$$\begin{aligned} C_i &= 2/ns_i \sum_{r=1}^n (\hat{x}_{ir} - \bar{x}_i) r \\ &= 2a_i/ns_i \sum_{r=1}^n (y_r - \bar{y}) r \\ &= \frac{a_i S}{s_i} G \\ &= a_i G/w_i \end{aligned}$$

Note that \hat{x}_{ir} increases or decreases monotonically with r , according to whether $a_i > 0$ or < 0 . Thus C_i is a true Gini, if $a_i > 0$ and $-C_i$ is a true Gini, if $a_i < 0$.

Now
$$\sum w_i C_i = G \sum a_i$$

Hence C_i represents an exact decomposition, if $\sum a_i = 1$. This can be proved as follows:

$$a_i = \frac{\sum_{r=1}^n (\hat{x}_{ir} - \bar{x}_i)(y_r - \bar{y})}{\sum (y_r - \bar{y})^2}$$

$$\sum a_i = \frac{\sum_{r=1}^n (y_r - \bar{y}) \sum_1^m (\hat{x}_{ir} - \bar{x}_i)}{\sum (y_r - \bar{y})^2}$$

$$= \frac{\sum (y_r - \bar{y}) (y_r - \bar{y})}{\sum (y_r - \bar{y})^2} = 1$$

The estimated factor incomes \hat{x}_{ir} thus represent an exact decomposition of y_r .^{1/} Furthermore, the Gini coefficients of these estimated factor incomes (g_i) are exactly related to the Gini coefficient of total income as follows:

1/

We may note, parenthetically, that

$$\bar{y} = \sum_{i=1}^m \bar{x}_{ir} = \sum_{i=1}^m b_i + \bar{y} \sum_{i=1}^m a_i = \sum_{i=1}^m b_i + \bar{y}$$

Therefore $\sum b_i = 0$

Hence
$$\sum_{i=1}^m \hat{x}_{ir} = y_r \sum a_i = y_r$$

$$\sum w_i g_i k_i = G$$

where $k_i = 1$, when the regression slope $a_i > 0$

-1 , when the regression slope $a_i < 0$

We refer in the text to $g_i k_i$ as the Estimated Factor Income Gini (EFI).

Annex 2

Answers to Standard Questionnaire

A. Objectives, Strategy and Results

Q1. To what extent did the study fulfill its objectives?

A1. To a limited extent.

Q2. Is the general problem to which the research is addressed of relevance to the Bank?

A2. Yes.

Q3. Were the objectives of the study clearly formulated?

Did they change as the study was undertaken?

A3. The objectives of the study were not clearly formulated.

They continued to evolve as the study was undertaken.

Q4. How do the results correspond with what was originally expected?

To what can be ascribed any differences between original and actual objectives?

A4. Not applicable.

Q5. Who are the intended beneficiaries? (Bank staff; planning authorities and decision makers in developing countries; other researchers).

A5. Not clear who the beneficiaries are.

Q6. Has the project assisted in developing research or other analytical capacity in the countries under study?

A6. Yes.

Q7. Were efforts made to coordinate work with other studies underway in the Bank or outside, to enhance the comparability of results or avoid duplication?

A7. Not applicable.

B. Design

- Q1. Does the research improve in specific ways upon a well-established methodology, or is the analytical framework relatively innovative?
- Q2. Were the theoretical approaches and the methodology employed in the study appropriate to its stated purposes? Were difficulties encountered in applying the methodology? If so, how were they overcome?
- Q3. How reliable were the data? Does their reliability depend on the design and coverage of sample surveys conducted as part of the project? If so, were such surveys properly designed and carried out, and their results adequately incorporated in the research?

A1, A2, A3

The research did not use a well established framework. It was innovative but not very useful. Not enough information is given on the data so no evaluation can be given. DRC judged the data to be of average quality.

C. Organization

- Q1. Did the research tasks follow a logical sequence? Were there opportunities to review progress at intermediate stages?
- Q2. How effective were the consultants or consulting firms employed? How open were the channels of communication between Bank staff and consultants?
- Q3. Was the extent of Bank staff involvement in design, implementation, and supervision adequate to meet the study's objectives?

A1, A2, A3

There was continuous correspondence and communication between the consultants and Bank staff but the relationship was somewhat strained for reasons given in the report.

Q4. What was the nature and extent of awareness, support, or participation among:

- Bank operating departments?
- Local research institutes?
- Government agencies?

A4. Support was given by local research institutes.

D. Dissemination

Q1. Are the research outputs written and presented in a manner which makes them accessible to the intended audience(s)?

A1. The FRK book is the main output disseminated. Not easily accessible since it is written in an opaque style.

Q2. By what means have findings been communicated to the intended beneficiaries?

A2. Not applicable.

E. Cost

Q1. How does the overall cost and efficiency of the study compare with the initial estimates? Did it take longer than expected? Were there significant cost overruns?

Q2. What appreciation, however broad, can be given to cost-effectiveness?

A1, A2

The study took longer than expected. It kept within the original budget but only because out of three countries, only one was done.

F. Lessons for the Conduct of Future Research

Q1, Q2, Q3

See paragraph 27 of covering memo.

OFFICE MEMORANDUM

SEP 17 1980

TO: Evaluation Panel
FROM: Benjamin B. King, DEDDR *BBK*
SUBJECT: Evaluation of RPO 670-80 ← file

DATE: September 17, 1980

1. Thank you for your comments. I attach two draft insertions: the introduction; and a paragraph to follow the present para. 24 in response to a suggestion by El Serafy. Other changes will be mainly editorial except:

- (i) Dropping paragraphs 16 and 17 (Baird and Hawkins).
- (ii) Dropping Annex 2. On reflection it seems unnecessary; I might add a sentence or two to the text instead.

Please give me any reactions by c.o.b. Friday.

2. You might be interested in the attached table, which I put together from the book and from the Pyatt-Chen-Fei article. The two sets of data are not quite comparable, but the differences are not so great as to invalidate the point. As you can see, at the decile level, the true factor Gini, the EFI Gini and the concentration ratio are in the same ball-park, though not near enough to establish the kind of identity to justify causal relationships between factor Ginis and total Ginis. The difference between the factor Gini at the decile level and at the individual level makes a mockery of the FRK argument.

AttachmentsDistribution

Messrs. Acharya
El Serafy
Mohan
Baird
Hawkins
Nankani
Duloy
Pyatt

Introduction

1. In current dollars this project probably cost something of the order of \$300,000.^{1/} It cannot be judged a success. It was carried out by consultants, who appear to have thought at the outset that they had received a grant for a proposal which they could thenceforward carry out with a relatively free hand. From the very beginning and throughout the project, the department chosen to sponsor the project, the DRC, were critical of its conduct - in the panel's view, rightly. The result was an acrimonious relationship between the Bank and the consultants. The final product was a book, which is so flawed and so lacking in integrity that, in the panel's view, its publication under the Bank's imprimatur does the Bank no credit.

2. Since the ultimate product illustrates much of what was wrong with the project, we shall start our report by describing its flaws (Section 2). Then we shall describe the course of events (Section 3). We believe that, despite the flaws in the book, the project has not been wholly without merit (Section 4). Finally, we draw our conclusions (Section 5).

^{1/} Including staff-time of about 1 man-year which may well be an underestimate. The dollars are approximately 1975 vintage.

Insert (para. 24a)

24a. The panel felt that, in its conduct as sponsors of the project, the DRC did its best in very difficult circumstances. They pursued the task of maintaining the quality of the research conscientiously and efficiently; it is evident that they put in much more staff time than is shown in the Completion Report. The fact that the consultants took their criticisms with an ill grace and that the final product is not one to be proud of, cannot be laid at the DRC's door.

Factor Income Indicators, 1968

	<u>Decile level</u> Gini (1)	<u>Decile level</u> EFI Gini (2)	<u>Decile level</u> Conc. Ratio (3)	<u>Individual</u> <u>level</u> Gini (4)
<u>Household</u>				
<u>Groups</u>				
<u>All</u>				
Wage	.293	.280	.260	.519
Property	.460	.483	.466	.699
Agric.	.182	.150	.167	.806
<u>Urban</u>				
Wage	.273	.263	.240	.461
Property	.425	.431	.531	.751
Agric.	n.a.	n.a.	-.194	.916
<u>Rural</u>				
Wage	.188	.161	.249	.591
Property	.278	.283	.241	.431
Agric.	.337	.344	.277	.442

- Sources: (1) FRK : p. 92
(2) FRK : calculated from slope (p. 94), share and total Gini (p. 92)
(3) PCF : Table II, col. 2
(4) PCF : Table III, col. 3

OFFICE MEMORANDUM

File

TO: Evaluation Panel

DATE: September 9, 1980

FROM: Benjamin B. King

B.B.K.

SUBJECT: Evaluation of RPO 670-80

As I had a rather unexpected and pressing chore last week, I am a bit behind. However, here is a draft (except for the introduction) along the lines of my outline. It needs tightening, pruning and much other modification, but I thought it best to get even this crude version to you. I would be grateful for your general reaction by phone and specific ones on paper (marked-up copy, for example). Neither Annex 2 nor Annex 3 is attached; the former because it has not been written; the latter, because it has already been sent to you as the answers to the questionnaire.

Attachment

Distribution

Messrs. Acharya
El Serafy
Mohan
Baird
Hawkins
Nankani
Duloy
Pyatt

Evaluation Report

1. Introduction

[still to be written]

2. What is Wrong with the Book?

1. The purpose of the book can be briefly stated in the authors' own words. They note that previous writers, in their work, 'discern an inverse U-shaped relation between growth and equity. They conclude that..... the distribution of income must first worsen before it can improve.....

Taiwan is one exception..... Taiwan's family distribution of income in the 1950s was not very different from the unfavorable levels most LDCs seem to be prey to in the early years of their transition effort. But that distribution has substantially improved during two decades of rapid growth. This "deviant" record should therefore be of interest to academicians and policy makers.' They wanted to explore how, and if possible, why it was that Taiwan achieved this unusual result. There can be little question that the subject is an interesting one.

2. Unfortunately, the authors saddled themselves, if we can so put it, with an albatross. This was the methodology they used and which they clung to, despite observations of various commentators, at different stages of the work, which cast doubt on its adequacy.

The methodology also has the disadvantage, which we will come to later, that the conclusions to be drawn from it are obscure.

3. To do justice to the debate on this methodology, would require something much longer than this report. Nevertheless, we feel that a

minimum description of it and its alleged faults is necessary. We will not go through all the details of the arguments, but will attempt to justify our assertions by reference to other documents and to annexed notes of our own, where necessary.

4. The central unit of observation used by the authors is family income distribution (FID). The Gini coefficient is used as the index of inequality. The heart of the methodology is to decompose this Gini coefficient into component parts as follows:

$$G = \sum G_1 \phi_1$$

The Gini coefficient (G) in this equation is equal to the sum of the Gini coefficient (G_1) of different factor incomes (agricultural, wage, property) multiplied by the weights (ϕ_1) of these factor incomes in total income. The authors then go on to explain changes in the Gini for total income by changes in its six component parts (the three factor Ginis and the three factor shares): or, to be more precise, they combine these changes in a particular way and describe the three combinations as three "effects".

5. The heart of the difficulty lies in the fact that no such precise decomposition exists or can exist. Two similar ones do. In one case the Gini coefficient for the factor income is replaced by something known as the "concentration ratio" or pseudo-Gini, as the authors call it. In the other the true Gini is replaced by what we shall call the estimated factor income (EFI) Gini, which is what the authors use. A note on these is given in Annex 1. It is the authors' contention that the error introduced by using the EFI Gini is sufficiently small that it can be neglected. Their case rests on the validity of this contention.

6. The principal evidence that the authors bring to bear to support their contention is on pages 92 and 93. Here, they give figures for the total error which is indeed quite low. However, it is somewhat difficult to reconcile this with other figures given on pages 102 and 103, which suggest that the change between 1964 and 1968 for the true total Gini is positive, while the change based on the estimated factor income Ginis is negative.^{1/}

7. It is not clear how the overall error was calculated, since there are certain residual incomes for which no information is given. In any event, the overall error is not the only error to be considered. This overall error is the sum of three component errors for each factor income. There is solid evidence that these individual errors are much larger, but offset each other. This casts considerable doubt on the validity of some of the assertions.

8. We would, however, regard these approximations as not the principal threat to the validity of the argument, despite their seriousness. It is entirely possible that if they were properly documented and if the consequent limitations to the conclusions were properly stated, they would not necessarily destroy the major part of the analysis.

^{1/} "Finding 3.4" on p. 102 states that one "highly favorable" effect overwhelmed (sic) a "highly unfavorable" effect to cause the slight worsening of income distribution.

9. A more serious deficiency in the analysis arises because the authors use grouped data, arranged by deciles of the population according to total income. A Gini coefficient constructed from decile averages is different from a Gini coefficient constructed from incomes of the total population. But the direction of the error introduced is well known and it is bounded. This in itself is then not too serious a problem.

10. The same is by no means true for the individual factor incomes. Because individual factor incomes are ordered by total income, group or decile averages may conceal extreme variations; for example, some of the factor incomes in each decile are intrinsically likely to be zero and the number of such zero incomes is likely to change over the period considered. The possibilities for error are thus very great. Indeed, calculations from ungrouped or computerized data carried out after the main work on the project was completed, but well before the publication of the book, demonstrate that the errors indeed are great.^{1/} It is hardly putting it too strongly that the story that is told is about a statistical artifact rather than about the real world.

11. What is perhaps most disturbing is that few of these points are immediately obvious to even a well qualified reader. Most readers would not have the patience to undertake the analysis required to uncover the possible sources of error, much less to assess their magnitude. Indeed in many cases, it is impossible from the material presented to attempt an assessment of this magnitude.

^{1/} The authors refer to the possibility of this kind of error in the last two paragraphs of the book (p. 409), pointing out that they had no choice but to use grouped data.

12. We now turn to questions of data. Most of the statistics on which the authors rely originate from household surveys conducted by one organization, the Directorate-General of Budget, Accounting and Statistics (DGBAS) between 1964 and 1976. The fact that one organization was responsible for the majority of the data is a distinct plus. Nevertheless, household surveys such as these are subject to well-known sources of error, to which the authors do indeed refer.^{1/} They in fact state that the inferences they derive "should not be sensitive to small variations" (page 13).

13. We are inclined to doubt whether the authors have followed their own prescription. This is especially the case because DGBAS data were not available for the city of Taipei from 1968 onwards. So surveys by the administration of that city had to be substituted. This inevitably raises questions of continuity. They would not be so serious, were it not for the fact that the authors place a great deal of emphasis on the year 1968 as a "turning point" when surplus labor ceased to become available. There is such a sharp break in 1968 in some of the data, that inevitably one is led to the suspicion that the abrupt changes ascribed to economic factors may, in fact, have been due to the simple fact of discontinuity in the data sources. Moreover, series constructed from computerized ungrouped data, show no evidence of a particular break in 1968.

^{1/} On page 97 there is a table comparing the shares of the three factor incomes according to (a) the household surveys used and (b) the national accounts. This phenomenon was observed in extenso in Latin America by Oscar Altimir.

14. The Bank employed a highly qualified reader (Professor A. B. Atkinson) to review the book before publication. While he recognized the interest in the subject and the possible contribution that the authors were making to it, he had criticisms, both in general and in detail. He made a series of comments on two successive drafts. Most, if not all, of our criticisms can be found in his reviews. The comments which are of particular interest to us are those in which he expresses his doubt whether the "data are sufficiently robust to stand" the use to which they were put. In his second round of comments, he noted that inadequate notice had been taken of his more general criticisms on the first round. It is our opinion that this was true on the second round as well. In fact, there is at least one instance where he pointed out that a numerical error had been made, both on the first round and on the second round. It has not been corrected in the final edition (page 69, line 3, third column). [Other cases are given in Annex 2].

15. The style of the book is opaque and often confusing to the reader. As a simple example, it is often unclear whether the authors are referring to the Gini coefficient of a particular factor income in the true sense (using grouped data) or the Gini coefficient of the estimated factor income, that we have described as the EFI Gini; there is a good example in the table on page 76. The authors complicate their description of the process of decomposition by introducing transfer income, which has or is likely to have special characteristics different from other factor incomes. This is unnecessary, because they never use transfer

income subsequently. The discussion could as well have been put in an appendix. The mathematical treatment is in many cases unnecessarily labored. We doubt whether anyone other than specialists will be tempted to read very far.

16. Much is made of the three "effects" referred to earlier which are as follows:

(i) The reallocation effect. This is the effect that follows from the decline of the share of agriculture in income. It is favorable if the agricultural Gini is higher than that for the rest of the economy and unfavorable if it is lower (on the assumption that the decomposition "works").

(ii) The functional distribution effect. This depends on the relationship between wage and property incomes. If the share of income with the lower Gini (presumably wages) increases, the effect is favorable (on the same assumption as above).

(iii) The factor Gini effect. This is the weighted sum of the changes in the three factor Ginis. If it is negative, it is favorable.

17. We doubt whether this prior determination of the interesting effects is helpful. It might have been more helpful to find out what had happened and then to isolate the main causes.

The principal findings of the authors are that:

- (1) Through 1968, income distribution in agriculture improved very much and largely offset the deterioration in income distribution in other factor incomes; the net result appears to have been ambiguous (see pp. 102/3).

(ii) After 1968, the improvement in distribution was general.

(iii) The two other "effects" (Reallocation and Functional Distribution) had some importance but considerably less than the Factor Gini effect.

It is, in fact, rather difficult to extract these points from the text.

18. In several respects the authors were criticized by Prof.

Atkinson for not taking sufficient account of previous work in particular fields, for example, on taxation with reference to Chapter 6 of the book.

From our own point of view, we note that, throughout, the analysis was

in terms of family income. It is by now well established--in large part,

through work conducted by the DRC--that family income distribution is

not the same as individual income distribution; the correspondence between

low-income families and low-income individuals is far from being one-to-one,

for the obvious reason that families differ in size and composition. This

fact might at least have been mentioned.

19. The objective of the book was to relate the particular path of growth that Taiwan took which caused or influenced the unusual phenomenon of greater equality. The investigation of this connection was more or less neglected after Chapter 2. Chapter 2 itself is full of assertions rather than analysis. A richer treatment would have related the trends in the disaggregation of income inequality with real economic phenomena.

Such phenomena would include a description of the particular activities or occupations which resulted from Taiwan's particular growth path; the

causes of an effective land reform; the tracing of the movement of population; the mix and concentration of industry. As it stands, however, the book is rather sterile and little insight is gained on how or why Taiwan was able to achieve growth with equity.

3. How did this come about?

20. The Bank's formal research program started in 1972 and this project was one of the early ones. Understandably, at the beginning, there was some anxiety to launch the program and little experience to go on. A proposal was submitted by the Economic Growth Center at Yale for a much more ambitious project, covering seven countries, than finally emerged. It was criticized strongly by the Development Research Center (DRC). Their criticism, which was intended as an internal document, was passed on to the Yale Growth Center. Considerable pressure was brought to bear on the DRC to moderate their criticism in view of the high reputation of the Center and the researchers involved. The present formal system of review did not at that time exist; the review was undertaken hastily by two members of the Research Committee who had several other review commitments at the same time.

21. The Economic Growth Center, on its part, had previously received grants from USAID to support its research. There is evidence to suggest that they regarded the grant from the Bank as a similar sort of core support, rather than a source of funds for carrying out a particular

project, in the form and content of which the Bank had a considerable interest. Indeed, the bureaucratic procedures of the Center, as it turned out, were not well adapted to coping with grants that depended on the delivery of a specific product. This misunderstanding was one source of antagonism; another was the fact that they had been shown the DRC's critical memorandum, which would probably have been worded differently had the authors known where it was going. There seems to have been an unfortunate atmosphere in which the authors felt they were being "got at" by the DRC and the latter in its turn, felt that the authors were not taking seriously reasonable and sincere criticism. The authors at least gave the appearance of feeling that the way to deal with this situation was to appeal to higher authority. The atmosphere was not exactly ideal for carrying on a cooperative project.

22. In retrospect, it appears that the authors must have set their minds on using this particular project to try out the particular methodology they had chosen. Little else would explain their persistence in carrying on with it in the face of various kinds of adverse criticism. It is not so much the methodology per se that is wrong, but the weight and the faith that they place on it. The Bank's interest lay, quite differently, in exposing, by whatever eclectic means was available, the root causes of the success story of Taiwan. Where any particular methodology would help, that methodology could be used. But the purpose was not to use one particular methodology or another, but to get at the truth as far as possible. Thus, there was another fundamental source of conflict.

23. The process of publication was in some ways a repetition of the process of acceptance of the proposal. The DRC, by now, had so accepted its adversary role that it agreed with the view that it should not act as judge and jury. This deprived the Editorial Sub-Committee, which was supposed to pass on the merits of the book, of valuable advice. To be sure, the Bank's reviewer (Professor Atkinson) did an excellent and conscientious job. As we have pointed out, he made many serious criticisms; but he recommended that in view of the importance of the subject and of the novelty of the authors' approach, the book should be published, once his criticisms were taken into account.

24. Again, there was pressure from higher authorities to hasten the process of publication. And, again, his identity was revealed to the authors without his knowledge. In the event the necessary follow-up to determine whether Atkinson's criticisms had been taken into account fell between two stools. The Editorial Sub-Committee probably did not have the expertise or the mechanism to see that this was done. The DRC clearly felt that, once the decision in principle had been made to publish, the responsibility for oversight was no longer theirs.

4. The Saving Grace

25. Despite all we have said, we feel that this project has paid off. While the proposal as such had serious faults in the way it was conducted from start to finish, it is not merely the project itself that we should be thinking about. The project was itself part of a process.

Much further enquiry has been stimulated into the nature of the decomposition of indices of income distribution and into the magnitude of the difference between the analysis of grouped and ungrouped data. The search continues for a better description and explanation of the evolution of income distribution in Taiwan. One fitting outcome of the controversy is an article, written jointly by one of the authors of the book and a staff member of the DRC who criticized it. We should recognize that false steps can have their virtues.

5. The lessons

There is a formal list of questions which evaluation panels are asked to answer. We append ours as Annex 3. However, we prefer to treat these questions as a kind of checklist. We will also put our own conclusions in our own way. The principal ones are as follows:

- (1) The review process. This should not be subjected to hierarchical pressure and should be conducted by an independent panel. In fact, this is now the case, as far as we can see.
- (2) Clarification of the Bank's role as a source of funds. The Bank should always make it clear, when collaborative research is undertaken, what its role is and what the conditions for the release of funds are. We believe that its performance in this respect is now much superior to what it was. However, to the extent that the Bank moves into special arrangements to support research activity in LDCs, the need for clarity may become more acute.

- (3) The responsibility of sponsors. Under the present system, all research proposals have to be sponsored by a particular department in the Bank. Unless this system is changed in some way, we feel that the responsibility of the sponsoring department must be respected. Part of that responsibility lies in making the judgment as to when a difference of opinion with outside consultants on research ceases to be merely a difference of opinion and becomes a question of quality control. This responsibility should not be diluted without cause.
- (4) Anonymity of reviewers. This should be respected, whether the review is form or informal. Otherwise, the Bank will be unable to command honest reviews.
- (5) Follow-up procedure by the Editorial Sub-Committee
When a decision to publish is made subject to the authors "taking into account reviewers' comments", we feel that a specific procedure should be set up to ensure that this is done.

Annex 1

The Gini coefficient

There are several formulae for the Gini coefficient. A convenient one for present purposes is as follows:

$$G = \frac{2}{nS} \sum_{r=1}^n (y_r - \bar{y}) r$$

where y_r is a set of incomes (in this case) ordered so that they increase monotonically, i.e., for any $r = i, j$ and $i > j$, then $y_i \geq y_j$.

In addition $S = \sum_{r=1}^n y_r$

The concentration ratio 1/

Suppose there are, as in the present case, two sets of related variables; total income y_r ; and a factor income x_{ir} forming part of each total income y_r . Then the concentration ratio for x_{ir} , in relation to total income, is defined as:

$$C_i = \frac{2}{ns_i} \sum_{r=1}^n (x_{ir} - \bar{x}_i) r$$

where $s_i = \sum_{r=1}^n x_{ir}$

It should be noted that, here, the ordering depends on y_r , not x_{ir} . Clearly, if x_{ir} are large for small r , the concentration ratio can be negative.

1/ Or Pseudo-Gini,; see FRK pp. 352 ff.

If there are m such factor incomes equal to total income, such that $\sum_{i=1}^m x_{ir} = y_r$, it follows that

$$\begin{aligned} \sum_{i=1}^m C_i s_i &= \frac{2}{n} \sum_{r=1}^n r \sum_{i=1}^m (x_{ir} - \bar{x}_i) \\ &= \frac{2}{n} \sum_r (y_r - \bar{y}) \\ &= GS \end{aligned}$$

If we denote the weight of each factor income in total income as $w_i = s_i/S$.

$$\sum w_i C_i = G$$

The Estimated Factor Gini

Suppose, as before, there is a set of factor incomes x_{ir} equal to total income y_r . Suppose, further, that each factor income is regressed on total income with the result that the estimated factor income, denoted by \hat{x}_{ir} , is related to y_r as follows:

$$\hat{x}_{ir} = b_i + a_i y_r$$

Hence $\hat{x}_{ir} - \bar{x}_i = a_i (y_r - \bar{y})$

If we form the concentration ratio for x_{ir} :

$$\begin{aligned} C_i &= \frac{2}{n s_i} \sum_{r=1}^n (x_{ir} - \bar{x}_i) r \\ &= \frac{2 a_i}{n s_i} \sum_{r=1}^n (y_r - \bar{y}) r \\ &= \frac{a_i S}{s_i} G \\ &= \frac{a_i G}{w_i} \end{aligned}$$

Note that \hat{x}_{ir} increases or decreases monotonically with r , according to whether $a_i > 0$ or < 0 . Thus C_i is a true Gini, if $a_i > 0$ and $-C_i$ is a true Gini, if $a_i < 0$.

Now
$$\sum w_i C_i = G \sum a_i$$

Hence C_i represents an exact decomposition, if $\sum a_i = 1$. This can be proved as follows:

$$a_i = \frac{\sum_{r=1}^n (\hat{x}_{ir} - \bar{x}) (y_r - \bar{y})}{\sum (y_r - \bar{y})^2}$$

$$\sum a_i = \frac{\sum_{r=1}^n (y_r - \bar{y}) \sum_1^m (\hat{x}_{ir} - \bar{x})}{\sum (y_r - \bar{y})^2}$$

$$= \frac{\sum (y_r - \bar{y}) (y_r - \bar{y})}{\sum (y_r - \bar{y})^2} = 1$$

We may note, parenthetically, that

$$\bar{y} = \sum_{i=1}^m \bar{x}_{ir} = \sum_{i=1}^m b_i + \bar{y} \sum_{i=1}^m a_i = \sum_{i=1}^m b_i + \bar{y}$$

Therefore $\sum b_i = 0$

Hence
$$\sum_{i=1}^m \hat{x}_{ir} = y_r \sum a_i = y_r$$

The estimated factor incomes \hat{x}_{ir} thus represent an exact decomposition of y_r . Furthermore, the Gini coefficients of these estimated factor incomes (g_i) are exactly related to the Gini coefficient of total income as follows:

$$\sum w_i g_i k_i = G$$

where $k_i = 1$, when the regression slope $a_i > 0$

-1 , when the regression slope $a_i < 0$

We refer in the text to $g_i k_i$ as the estimated factor Gini.

SEP 15 1980

File → RPO 670-84

Mr. Jack Duloy, DRC

September 5, 1980

Benjamin B. King, DED

FRK (pp. 92-95)

1. I have been intrigued by the point made by various people that, whatever FRK's total "non-linearity" error may be, it could be the aggregate of larger but offsetting errors. I have not seen any numerical evidence and so did an exercise myself.
2. The results are in the attached table. The first line in each group is the product of (i) the relevant regression slope (.292) and the total Gini (.3208). The second line is the product of (iii) the factor share (.4324) and (iv) the factor Gini (.2365). Numbers refer to the 1964 wages Gini as an example; except for the slope (p. 94) they are all from p. 92. As expected, the individual errors are much larger and offsetting.
3. But now comes the puzzle. Given that there is a residual share and a residual slope (both deducible from the tables) and a Gini for the residual (not available), how did FRK arrive at their error? I have done my best by assuming that the residual Gini is the same as the total Gini. It helps in two cases, but not in the third.
4. Do you know the answer? Have you done the same calculations or similar ones for more cases? In short, any more light, as and when you are not preoccupied by Other Matters?

Attachment

cc: Messrs. Mohan
Acharya ✓

BBKing:gm

All Households

	<u>1964</u>	<u>1966</u>	<u>1968</u>
<u>Wages Gini</u>			
Weighted Est.	.0937	.1210	.1418
" True	<u>.1023</u>	<u>.1284</u>	<u>.1485</u>
Difference	-.0086	-.0074	-.0067
 <u>Prop. Gini</u>			
Weighted Est.	.1168	.1110	.1340
" True	<u>.1077</u>	<u>.1049</u>	<u>.1277</u>
Difference	.0091	.0061	.0063
 <u>Agr. Gini</u>			
Weighted Est.	.0950	.0694	.0228
" True	<u>.0976</u>	<u>.0722</u>	<u>.0277</u>
Difference	-.0026	-.0028	-.0049
Total difference	-.0021	-.0041	-.0053
 <u>Residual Gini</u>			
Weighted Est.	.0154	.0213	.0274
" True	<u>.0167</u>	<u>.0182</u>	<u>.0207</u>
Difference	-.0013	.0031	.0067
Total (incl. residual)	-.0034	-.0010	.0014
FRK "error"	.0009	.0004	.0019

AUG 22 1980

OFFICE MEMORANDUM

TO: Distribution Below

FROM: Benjamin B. King, DEDDR *BANK*

SUBJECT: Evaluation (RPO 670-80) ← *File*

DATE: August 21, 1980

1. The silence since we met does not imply inaction. Rakesh Mohan and I have obtained a copy of the Editorial Subcommittee's records on the RFK book, which is as thick as the document we all have. We (mostly Rakesh) have been summarizing this, trying to sift the wheat from the chaff on methodological problems and making a preliminary effort at identifying issues and lessons. It isn't easy. Moreover, next week I will be on leave. In view of the need to do a good job on this, Shankar Acharya and I have agreed that our report should be delivered on or before Sept. 15 (rather than August 30th). I hope I can send you a draft by September 8th.
2. Rakesh prepared some notes on the meeting, which are attached. If you think they need modification, please let him know (61274). We do not intend to make this part of the final report, so minor questions of wording aren't important. Possible exception: the questionnaire answers on page 3 and the graph.
3. I attach a quick-and-dirty outline which will form the basis for the draft report, unless we have second thoughts. Please let us have your comments: in writing to both Rakesh and me; orally to Rakesh.

Attachment.

cc: Messrs. Acharya (VPD) ✓
El Serafy (OED)
Mohan (DED)
Baird (EAL)
Hawkins (AEA)
Nankani (VPD)
Duloy (DRC)
Pyatt (DRC)

OUTLINE

1. Introduction

Some sort of brief summary to introduce what follows.

2. What is wrong with the book?

- (i) Methodology. Brief and confined to assertions. In short, there are two related to decomposition: first, approximation not adequately qualified--the lesser of the two; aggregation by deciles--in our mind, much the more serious. An annex will support the assertions. We will also refer to the series break in 1968, which they consider a crucial date. Final point, even a well-qualified reader could be deceived, if he didn't know.
- (ii) Data. Inadequate statement of sources and qualification.
- (iii) Casual attitude to reviewer's comments. There are plenty of cases where Atkinson's comments were ignored, even a numerical error. Bald statement, also documented in an annex.
- (iv) Presentation. Opacity, excessive laboring of point, difficulty in finding out what they are up to, etc. (panel's comments on this are welcome).
- (v) Failure to relate discussion to historical events. General impression that it does not tell a story, but not yet adequately sorted out. Treatment of population movements, land reform possible instances. Duloy/Pyatt's thoughts would be welcome.

3. How this came about?

- (i) Initial conditions. Early project. Mistaken view of the Bank as a foundation. Antagonism from the outset with (probably) a view by the authors that they could override DRC.
- (ii) Preoccupation with methodology. And a faulty one at that.
- (iii) Publication. In some ways a repetition of the initial conditions. Unfortunate situation with both the DRC (for reasons Jack Duloy gave) and the Editorial Subcommittee both unwilling to say no. Failure on part of ES to oversee incorporation of Atkinson's comments.


Outline (continued)

4. Saving grace. Whatever the quality of the product, something good has come out of the project and some lines of enquiry have been pursued as a result, e.g., Pyatt-Chen. (Essentially El Serafy's point).

5. Lessons.

- (i) Proper research committee review process, not subject to pressure (now in place, hopefully).
- (ii) Bank should either operate as a foundation or not and make it clear which it is doing (no longer a problem, but may become one again if and when we do "institution building" in LDC's.)
- (iii) Sponsoring departments should be given full responsibility for deciding at what point differences of opinion cease to be that and become a question of quality control. They should not be overridden without cause.
- (iv) The anonymity of reviewers (inside or outside) should be protected.
- (v) Follow-up procedure by Ed. Subcommittee needs strengthening. Should feel freer to tell authors to go elsewhere.

OFFICE MEMORANDUM

TO: Mr. Benjamin B. King, DED
FROM: Rakesh Mohan, DEDRB 
SUBJECT: Evaluation of "Growth, Employment and Size Distribution of Income"
RPO 670-84

DATE: August 20, 1980

Minutes of Panel Meeting: Friday, August 15, 1980

1. The evaluation panel for the above project consisting of Messrs. M. Baird, J. Duloy, S. El Serafy, E. Hawkins, B.B. King and R. Mohan met on Friday, August 15, 1980. Messrs. S. Acharya, C. Bell, G. Nankani and G. Pyatt were unable to attend.
2. Mr. Duloy gave a short introduction on matters not covered in the documentation of the project provided to the panel. He reminded the panel that this project was initiated in 1973 when the Bank's research activities were just starting up. There were not yet any set procedures guiding the evaluation of research project proposals as they exist now. The original project proposal sent by Yale faculty members was a much larger one, of the order of about \$1.5 million (1973 dollars). They expected to study 7-8 countries and to involve a large number of faculty members. A memo from DRC dated February 26, 1973 to VPD argued against the funding of the project on the grounds that neither the objectives nor the methodology of the project were well thought out. VPD felt that this was an area where research needed to be done and that since the Yale Economic Growth Centre was an institution of high reputation and quality ways should be found to encourage such development research there. As a result, after much discussion, the project was funded in mid 1973 at a much scaled down version of about \$200,000 for a 3 year period.
3. Mr. Duloy felt that two problems had soured relations between the DRC and the Yale Growth Centre from the beginning of the project. One misunderstanding arose because the Yale Growth Centre viewed the Bank much as a foundation which gave funds for research but was not in a supervisory position. There was no set policy in the Bank on outside research funding and the DRC regarded the Yale Growth Centre as consultants who had to be supervised. This divergence of understanding naturally soured relations when the researchers perceived the supervision as interference. The second problem was that some adverse comments written as internal memos by DRC staff found their way to Yale so that the researchers perceived the DRC as adversaries from the start.

4. Nevertheless, the DRC was successful to some extent in their supervision activities in scaling down the research project; in persuading FRK (Fei, Ranis, Kuo) finally to make adjustments for the inclusion of Taipeh in the data and in making sure changes connected with the Gini decomposition controversy.

5. Mr. Duloy provided a succinct summary of the Gini decomposition controversy. The crux of the matter is that what FRK call Gini coefficients in the decompositions are really concentration ratios. Further, decomposition problems are caused because all calculations are based on decile grouped data. In illustration he provided the attached graph which shows how the so called pseudo Gini coefficients ($C(w,y)$ and $C(a,y)$ (or concentration ratios) are quite different in concept as well as empirically from the true Ginis ($G(w)$, $G(a)$) of the factor income component. These graphs were derived from calculations done by Pyatt and Chen from the raw data. These results will appear in the QJE with Pyatt, Fei and Chen as authors. Mr. Duloy noted that these results were available to FRK as early as November 1978 but no correction or caveat appears in the book.

6. FRK and Pyatt had finally reached agreement on the Gini decomposition theoretical problems. The remaining controversy rested on the importance of the errors caused by the FRK methodology in the final empirical results. The DRC view (as set out in the Bell/Alluwallia memorandum) was that the errors in estimation made their conclusions unwarranted while FRK believe that they are unimportant for their conclusions.

7. Mr. El Sarafy was heartened by the scrutiny and diligence with which the DRC had subjected this project and felt that there were important negative conclusions to be derived from it: both procedurally as well as substantially. The substantive negative conclusion is perhaps that the Gini decomposition method does not greatly help in illuminating how growth can be achieved with equity.

8. Mr. Mohan inquired after the positive publication decision when the DRC has consistently had such reservations about the project. Mr. Duloy said that, in agreement with Mr. Chenery, the DRC felt that it could not act simultaneously as prosecutor and judge and had therefore decided that they would not object to the publication of the book. Mr. Mohan remarked that consistently high publication criteria were important for Bank publications so that they would not be devalued for other contributors.

9. The panel quickly reviewed the guidelines for evaluation and agreed to the following answers.

A. Objectives, Strategy and Results

1. No
2. Yes
3. The objectives of the study were not clearly formulated. They continued to evolve as the study was undertaken.
4. Not applicable.
5. Not clear who the beneficiaries are.
6. Yes
7. Not applicable.

B. Design

- 1, 2, 3. The research did not use a well established framework. It was innovative but not very useful. Not enough information is given on the data so no evaluation can be given. DRC judged the data to be of average quality.

C. Organization

- 1, 2, 3. There was continuous correspondence and communication between the consultants and Bank staff but the relationship was somewhat rocky for reasons given by Mr. Duloy earlier.
4. Support was given by local research institutes.

D. Dissemination

1. The FRK book is the main output disseminated. Not easily accessible since it is written in an opaque style.
2. N.A.

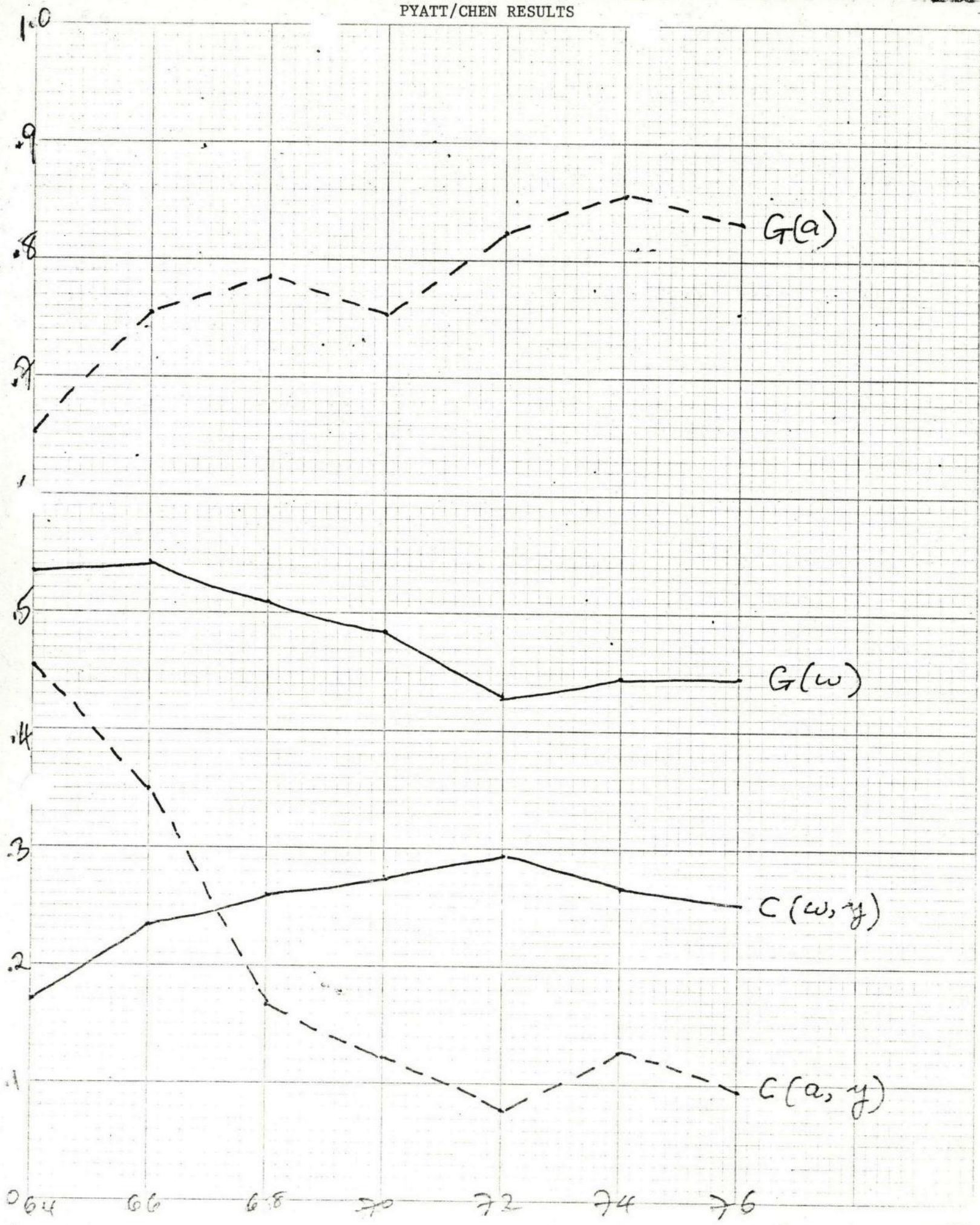
E. Cost

- 1, 2. The study took longer than expected. It kept within the original budget but only because out of 3 countries, only one was done.

F. Lessons for the conduct of Future Research

1. Bank should be clear whether it acts as a foundation or as research supervisor.
2. Care should be taken in communications with consultants that internal memoranda remain internal memoranda.
3. The methodology employed did not in fact contribute to an understanding of "Growth with Equity." The work became merely one of decomposition of Gini Coefficients.

PYATT/CHEN RESULTS



RPO 670-84

OFFICE MEMORANDUM

TO: Distribution

DATE: July 24, 1980

FROM: Shankar Acharya, VPD

SUBJECT: Evaluation of "Growth, Employment and
Size Distribution of Income (RPO 670-84)

1. An evaluation panel comprising Messrs. B.B. King (DED), Chairman, S. El Serafy (OED), R. Mohan (DED), M. Baird (EAL) and E. Hawkins (AEA)* has been formed to evaluate the above proposal. It will meet on August 15 at 2:30 p.m. in Room I8-218.

<u>Research Project</u>	<u>Supervisor</u>	<u>Total Amount Spent</u>
Growth, Employment and Size Distribution of Income	J. Duloy	\$209.5

2. The panel is expected to evaluate the output of the research project in the light of its objectives in the original research proposal. The attached 'Guidelines for the Evaluation of Research Projects' suggests a set of questions and a format that the panel is urged to follow.

3. In order to aid the evaluation process, the following documents are attached:

- (i) Completion Report (September 21, 1978);
- (ii) The original research proposal (May 21, 1973);
- (iii) The review panel's memorandum (May 24, 1973);
- (iv) The Research Adviser's memorandum (July 17, 1974);
- (v) A note from Professor G. Ranis (May 6, 1975);
- (vi) The Research Adviser's memorandum (May 23, 1975);
- (vii) Professor Ranis' notes (June 9 and 13, 1975);
- (viii) Mr. Pyatt's note (July 15, 1975);
- (ix) The Research Advisor's letter and memorandum (July 9 and 16, 1975 respectively); and
- (x) Messrs. Bell and Pyatt's memorandum (November 29, 1976).

4. The final product of the project was a World Bank publication, "Growth with Equity: The Taiwan Case" by J. Fei, G. Ranis, S. Kuo. Panel members who do not already possess a copy may call Ms. Thampy (x69013) to obtain one.

5. The evaluation memorandum should be received by August 31, 1980.

* Subject to confirmation.

Distribution: Panel Members, Messrs. G. Nankani, C. Bell, J. Duloy, G. Pyatt

GN:lt

Reviewer's Comments

"Growth and the Family Distribution of Income by Factor Components: The Case of Taiwan" by J.C.H. Fei, G. Ranis and S. W. Kuo

1. Probably the most illuminating general comment on the paper is that if I were refereeing it for a journal, I would accept it -- subject to a number of reservations discussed below. This is, I suppose, equivalent to saying that I find the research competent and useful, if not outstanding. Aside from questions of substance, I think it is also worth raising a point of style: the paper is written in very grandiose fashion, and claims a lot of "findings" and growth theory "explanations" thereof, which simply aren't there. In a revised version for the IBRD or publication, this mannerism should be ruthlessly suppressed.

2. Ahluwalia and Duloy point out in their comments that the sufficient condition for a decomposition of the Gini coefficient is that factor incomes should be ranked across individuals in the same way as total incomes -- this is in fact clear in the proofs in the appendix of the Fei-Ranis-Kuo (FRK) paper, which all depend on ranking methods. Pyatt gives a formula for the decomposition of the aggregate Gini which depends on covariances between income components across individuals and the overall income ranking. These observations suggest several ways in which the FRK analysis can be fruitfully extended:

(a) Is the ranking condition at least approximately satisfied for factor income components of interest? The high correlation coefficients FRK report suggest that it is, but they aren't really looking in the right place. The main exceptions to the ranking

rule are likely to be relatively high incomes from assets among some low income people (the famous "widows and orphans"; the aged in general, particularly in societies where extended families are breaking down) and low or non-existent wage incomes among the rich. FRK should be able to check these exceptions fairly easily in their data. They won't affect the aggregate Gini, but then nothing does, very much. On the other hand, people at the extremes of the distribution are ripe targets for redistributionist policies. Another major problem with FRK is that they ignore such possibilities completely, in the obsession with the Gini coefficient.

(b) If one wants to work with Ginis, then the obvious generalization of the FRK approach is in the direction of sampling theory. Their linear regressions satisfy "adding-up" conditions on incomes, but then so do many other families of Engel curves, some of which can capture U-shaped relationships between income components and total incomes (such as the relationship between asset incomes and total income, if "widows and orphans" are economically significant). Perhaps Pyatt's formula and Kakwani's work in the Development Research Center can provide a groundwork for this type of analysis. I might add that I personally am rather sympathetic to an explicit statistical decomposition analysis of large bodies of individual data, and think that procedures such as those used by FRK can teach us a great deal about what is occurring in the economy, even if they can't "explain" it.

3. FRK put a lot of effort into trying to relate their decomposition to growth theory in a dual economy, going so far as to presuppose at times that the reader is intimately familiar with the Development of the Surplus Labor Economy (e.g. the non-definition of terms such as J on p. 17). In fact, the shifts of Gini coefficients within the income groupings adopted in the paper dominate the "explanation" of shifts in the aggregate Gini. This doesn't mean the growth theory classifications are wrong, but only inadequate. In further work, FRK should seek more fruitful and policy-related classification schemes, perhaps stressing specific target groups. Fishlow provides a good example and the authors' attempt to differentiate their decomposition procedure from his (on p. 53) is unworthy.

4. I cannot comment on the significance of the "separation of Taipei City" from the data (p. 24 footnote) but it surely deserves more than burial at the bottom of the page. Also, the allocation of proprietors' incomes to wages and profits should be discussed -- this will amount to 20 or 30 percent of national income and its method of assignment to the two traditional categories could affect the aggregate results. The following specific comments on the empirical results are in order:

P. 26 (footnote): high correlations between factor incomes and total income do imply that the ranking criterion discussed above is approximately satisfied. Of course, they say nothing at all about linearity of the relationship, or lack of it. The obvious test is to stick some quadratic terms into the regression.

P. 29: Discovering a historical "turning point" in eight years' worth of data would be a major achievement. The authors have no real statistical reason to assert that they have done any such thing, and should not say so.

P. 30: Although the authors talk a lot about the Kuznets hypothesis, it is only stated in something like its original form near the middle of this page. It cannot be tested by the type of aggregate functional decomposition of incomes the authors adopt, and they should not advert otherwise.

P. 31: Lots of talk here and elsewhere about the narrowing of the "wage gap", but numbers to show that this has occurred are not produced.

P. 33: Is spatially dispersed industry "rather unique"? It seems that we have at least some historical experience with putting-out systems.

P. 35: The growth theory discussion in the middle of the page is so much mumbo-jumbo.

P. 36: One doesn't just "assume" that initial exports came from rural-based industry, nor can direct statements about bias in technical change be made from the type of data analyzed here, even if all of neoclassical production theory is maintained hypothesis. In fact, strictly speaking, neoclassical theory is inconsistent with the authors' linear regressions of factor incomes on output, since only a Cobb-Douglas technology will generate this type of relationship.

P. 39: Rural wage incomes are sometimes distributed more unequally than their urban counterparts, sometimes not. The assertion should be tested.

P. 42: Again, assertions about dualism can't be tested from this type of data.

5. Despite these major reservations, the story FRK tell about Taiwan is an interesting one, and can in fact be made even better if the authors take these and other criticisms into account. However, it remains true that

(a) Decomposition procedures of the type adopted here are mainly diagnostic tools to be used imaginatively; by their nature they cannot indicate causality and should not be interpreted as doing so.

(b) For policy purposes, the decompositions are most interesting when used with a disaggregation scheme tied to "interesting" social groups, e.g. those at the extremes of the distribution or which might be affected by politically feasible policy initiatives. There probably remains much work to be done in Taiwan (as in most countries) along these lines.

(c) Analytically, combination of Gini coefficients with Engel curves has some interest, particularly when carried out in connection with sampling theory. Kakwani in the IBRD is working on this, and FRK should be encouraged to do so, as well.

WORLD BANK RESEARCH PROGRAM

COMPLETION REPORT

Date of Submission

SEP 21 1978

I. PROJECT IDENTIFICATION AND APPROVAL OF REPORT

Title: Growth, Employment and Size Distribution of Income 2. Project No. 670-84

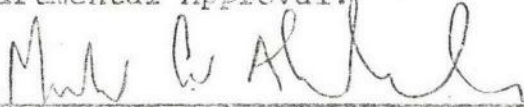
Department: DRC

4. Division: IDD

Staff Participation
a. Principal Supervisor: M.S. Ahluwalia
b. Others Responsible:

6. Date Approved by
Research
Committee: June '73

Departmental Approval:

x 
Division Chief (signature)
M.S. Ahluwalia


Department Director (signature)
J. Duloy

II. IMPLEMENTATION

Date of First Contract: Sept. 1973 2. Date of Final Disbursement: 3. Date Final Report Completed.

Parts

a. Final Report "Equity with Growth: The Taiwan Case", Fei, Ranis, Kuo, ✓

b. Other Publications/Reports

"Growth and the Family Distribution of Income by Factor Components, Ranis, Fei, Kuo, ✓
QJE, February 1978.

ES: Item II, 3: Date when final report available for dissemination.
II, 4: Journal articles, Staff Working Papers, Departmental reports, technical memoranda, and other reports prepared under the project. List authors and dates of completion or publication.

III. SUMMARY OF RESOURCES

Expenses through FY75: \$136.7	2. Research Committee Authorizations as from FY76: \$72.8	3. Total Research Committee Financing: \$209.5
Expenses as from FY76: FY 76 : 72.8 FY ___ : FY ___ : ___ : Total: 72.8	5. Other Financing: a. Bank Departments FY ___ : FY ___ : FY ___ : FY ___ : Total	b. Outside Sources FY ___ : FY ___ : FY ___ : FY ___ : Total

Staff Time (weeks):	Professional	Assistant	Total
FY 74 :	2.0		2.0
FY 75 :	4.0		4.0
FY 76 :	6.0		6.0
FY ___ :			
Total	12.0		12.0

- TES: Item III, 1: The sum of actual expenses, if any, from Research Committee authorizations in FY75 and earlier years.
- III, 2: The "total authorization" of the project, i.e., the authorization for FY76 (if any) and all subsequent years.
- III, 3: Equals V, 1 + V, 2.
- III, 4: Actual expenditures of funds authorized by the Research Committee, FY76 through FY of final disbursement.
- III, 5a: Departmental discretionary funds spent primarily or exclusively on project (if any).
- III, 5b: Breakdown of contributions by donor. Include estimates, to extent possible, of local contributions.

V. DESIGN AND ORGANIZATION

- methodology or analytical framework employed; difficulties encountered in application of methodology; how these difficulties were or were not overcome
 - sequencing of research tasks
 - reliability of data
 - performance of consultants or consulting firms
 - extent of Bank staff involvement in design, implementation, supervision
 - extent of awareness, support, or participation among Bank operating departments, local research institutes, government agencies
 - main reasons for overruns or savings in cost and time
1. The methodology employed consists of two parts. *First*, the researchers have documented trends in inequality over time. The observed levels of inequality in each year are decomposed into component parts using two different decomposition methods and changes in overall inequality are thus reduced into changes in the individual components of this inequality. The decomposition techniques used are in part an application of existing techniques (which had been developed theoretically but had not been empirically used) and in part they represent development of new decomposition techniques. The application of this method is definitely innovative. *Second*, the researchers attempt to relate the observed changes in components of inequality to the nature of the underlying growth process in Taiwan thus attempting to explain the factors which generated the overall trends in inequality. This second part of the study is necessarily more impressionistic than the first.
 2. The sequencing of the research tasks originally considered was as follows. First it was planned to develop a methodology and apply it to Taiwan and in a second and third stage it was to be applied to two other countries. In the event the DRC was not satisfied with the early definitions of the methodology at the end of the first year, and again with the first application at the end of the second year. The project was therefore narrowed down by agreement to the completion of one country study with a more thorough treatment of the issues involved.
 3. The reliability of data is a major problem in the area of income distribution. In my view the data used by the study are of "average quality" in terms of data generally available in LDCs. The authors should have devoted greater effort towards documenting the problems with the data. The monograph does not contain sufficient discussion of data quality or comparability.
 4. The performance of the consultants on this project has been broadly satisfactory. We have had many discussions and disagreements, including a few acrimonious exchanges, arising out of our dissatisfaction with the rigour with which the analysis was being conducted and the consultants' feeling that we were applying unduly strict standards. Much of this was an inevitable consequence of the lack of a well defined methodology for decomposing inequality into component parts and interpreting changes in these component parts. The major thrust of our comments and criticisms was to force the researchers to concentrate on improving the study of Taiwan instead of broadening the study to include two additional countries. The consultants have been fully cooperative in responding to these pressures although they have felt that we were insufficiently appreciative of their work.

Narratives in the following sections summarize the results of the project in relation to its objectives and describe the design, organization, and dissemination strategy adopted. The categories and topics are intended to correspond closely to those of the "Guidelines for Evaluation of Completed Research Projects." Using additional space as necessary, the principal supervisor should give particular attention to the points listed.

IV. OBJECTIVES AND STRATEGY

- objectives of the research, as originally formulated and with later modifications
- intended beneficiaries (Bank staff; planning authorities and decision makers in developing countries; other researchers)
- contribution to research or other analytical capacity in the countries under study
- efforts to coordinate work with other research in the Bank and outside

- (1) The original objectives of the research were to document trends in inequality in three countries and attempt to explain these trends in terms of the nature of the underlying growth process. The methodology for examining the causal factors was intended to be heuristic rather than formal, although the documentation of trends itself was expected to be fairly rigorously based on available survey data. In subsequent discussions it became clear that attempting to cover three countries would lead to a superficial treatment of each and the research contribution of the project would be greatly enhanced if the researchers were to concentrate on doing a thorough job for one country.
- (2) The intended beneficiaries included the development community in general. The project was not narrowly focussed on Bank operational or policy concerns but it did attempt to fill a crucial gap in our knowledge viz. the extent of change in inequality over time in LDCs and factors affecting these changes.
- (3) A significant involvement of local institutions was planned and was achieved. One of the researchers is based in Taiwan.
- (4) The question of coordination with others which did not arise since there was almost no ongoing research effort in this particular area in other institutions at that time.

V. Design and Organization (cont'd)

5. Substantial amounts of Bank time were involved. These were not inordinate but were more than we originally budgeted. This reflects our own inexperience in estimating the time required to get consultants to produce a satisfactory product. I have no doubt that without such intensive inputs the study would have fallen considerably short of its present standard.
6. Bank operating departments were not involved at all, mainly because of the absence of Bank operational involvement in Taiwan. Local institutes were involved in the project.
7. There were no cost overruns on the total budget but, of course, this is because the project budget was fully exhausted for one country study. The consultants almost certainly underestimated the size of the research task in terms of the effort needed to produce a product of sufficient quality. On the whole the total cost \$210,000 is probably somewhat high for the final product. In retrospect, the study should have cost around \$150,000. However, this overrun is probably not greater than the "full cost" overrun on most research projects (i.e. where allowance is made for staff time spent over the initially budgeted amounts).

VI. RESULTS

- nature of project findings; correspondence with what was originally intended; reasons for differences between intended and actual objectives
- suggestions for follow-up, including other research topics

Except for the reduction in country coverage from three countries to one the project results correspond with original intentions. The monograph provides a substantial description of change in income inequality over time in Taiwan and a discussion of the relationship between these changes and some aspect of Taiwanese growth.

Project Findings

The major empirical findings of the research project can be summarised as follows:

- (i) Inequality in Taiwan was more or less constant upto 1968 and then declined. The authors argue that this shows that inequality and high growth are in principle compatible.
- (ii) For *all households* as well as for *urban households*, the pattern is one of constant inequality before 1968 and a decline thereafter, while for *rural households* there is declining inequality before 1968 and constancy thereafter.
- (iii) For *urban households* the decline in inequality is due to the increase in wage share after 1968. This is argued as demonstrating the value of a labour intensive growth strategy with its resultant effects upon employment and real wages. This is the central factor accounting for the reduction in overall inequality in Taiwan.
- (iv) For *rural households* the authors find that the income from agriculture is more unequally distributed than the non-agricultural or "new-income" from rural based industrialisation. The rural industries are very labour intensive and over time rural households have benefited increasingly from these industries and the share of non-agricultural rural income in rural household income has increased.
- (v) In examining wage inequality the authors conclude that "warranted differentials" are far more important in explaining inequality than unwarranted ones. In other words institutional discrimination in the labour market is not a problem. A related conclusion is that "the varying family ownership of high grade labour must be regarded as the most important cause of overall family wage inequality" ("high grade" in this context means one or a combination of three things "prime age", "male" and "highly educated").

Some Reservations

In monitoring this project we have expressed technical reservations about the decomposition method actually used and its empirical application. We have achieved considerable narrowing down of the basis for disagreement but some issues remained unsettled. These were subjected for consideration to a distinguished external referee, Professor A.B. Atkinson. Professor Atkinson's conclusion is that while the issues raised by the DRC (Graham Pyatt) are valid, and should ideally be taken into account, the monograph as it stands could be regarded as a publishable finished product which makes an important contributions.

VII. DISSEMINATION

- presentation of results to facilitate access to intended audiences
- methods (published reports, seminars, conferences, etc.) for dissemination of findings to these audiences

The monograph will be published as a book by the World Bank. Furthermore, an article has been accepted by the Quarterly Journal of Economics and will be included in the Bank Reprint Series. No further special efforts at dissemination are needed.

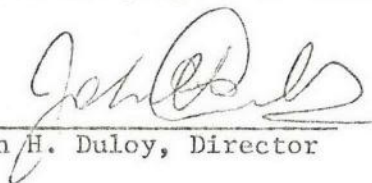
VI. Results (cont'd)

Further Work

The study of changes in inequality over time and its relationship with growth is a major area for continued work. This project represents a particular methodological approach which has been on the whole fruitful. Further work in other countries would be extremely valuable and in my view should have very high priority. In future application, however, more attention should be paid to tracing the history of socio-economic groupings and also documenting with special detail what happens to poverty groups.

SUPPLEMENTARY EVALUATION
by John H. Duloy

I am fully in agreement with the thoughtful and balanced evaluation of this project prepared by Mr. Ahluwalia. In my own view, this project was successful in terms of its original objectives, for which the credit must accrue to the authors. It has also required a substantial staff input (far more than originally envisaged). This need is often underestimated (by DRC as well as by others) and is one of the main lessons to be drawn from the project in terms of the management of research.


John H. Duloy, Director

File (with only)
① with completion report in completion report
② In project file
③ Had to R.C.

September 20, 1978

Professor Gustav Ranis
Economic Growth Center
Yale University
New Haven, Connecticut 06520

Dear Gus:

The completion of the Taiwan project has been a long haul, and I am very pleased that we are now at an advanced stage in the process of publication. Because of the sometimes stormy exchanges which have occurred in the past, I thought it appropriate to take the unusual step of providing you with a copy of the Completion Report which we are required to write as our input into the Bank's evaluation process. You will appreciate that this is a purely internal document prepared for purposes internal to the Bank. My motivation in sending you the document is that you have sometimes expressed a concern that the record should be kept straight.

I hope that you will agree that the report represents a successful attempt at balance. The panel (independent of DRC) which will review the project will have available also the article and the manuscript. If you were interested and willing, I think that your participation at the review meeting would be valuable and could be arranged. Please let me know if you are interested and if you have travel plans, so that we can suggest some dates to the Research Committee.

Finally, I hope that you and your co-authors will accept my congratulations and those of my colleagues on the outcome of the project.

With kind regards,

Yours sincerely,

John H. Duloy
Director
Development Research Center

Enclosure

- cc: H. Chenery (VPD)
- 1) B. Balassa (DRC) ✓
- 2) S. Bery (VPD) ✓
- M. Ahluwalia (DRC)
- G. Pyatt (DRC)

VRes - 670.84
cc 670.85
cc. 670.94
cc. 671.41

Mr. Montek Ahluwalia, DRC

August 18, 1978

Suman Bery, VPD

Completion Reports

On February 23, 1978 and on prior occasions completion reports for the following projects have been requested from you. On the understanding that these projects are now complete no narrative on them has been included in the forthcoming "Abstracts of Current Studies". Since these projects have been completed, please let me have completion reports on them by c.o.b. August 30, 1978 in order that they may be forwarded to the Research Committee at its next meeting.

<u>Project No.</u>	<u>Title</u>
670-84	Growth, Employment and Size Distribution of Income
670-85	Urban Income Distribution in Latin America
670-94	Employment and Income Distribution in Malaysia
671-41	Indirect Estimation of the Size Distribution of Income

July 1978

Growth with Equity

To begin with, I would repeat what I said in the earlier report - the material is very valuable, there are important innovations in methodology, and the analysis of Taiwanese experience appears extremely interesting to a non-expert. There can be little doubt that the manuscript is the basis for an outstanding book. On the other hand, in its present form there are two problems :

a) I was certainly expecting a more extensive revision than has been carried out (the report referred, for example, to a systematic re-writing), and that this would go further to meet the points made (for example, the robustness of the conclusions, with respect to errors in the data - e.g. from grouping),

b) there is the debate between Graham Pyatt and the authors.

As far as the former is concerned, I have again gone through the manuscript and made detailed comments. Some of these (such as mistakes in tables) are editorial points; others are more serious, questioning the strength of the conclusions which can be drawn. In my view, too much weight is placed on the results and insufficient prominence given to the qualifications. Do the authors really regard a change of 0.0002 in the Gini coefficient as worth discussing? This is reinforced when we turn to the FRK/Pyatt debate, on which I have written a separate set of comments. Here there seems to be no disagreement of principle but one of the empirical importance of the error in ungrouped data (the authors have demonstrated that the error is small for the grouped, decile data). This cannot be resolved without detailed analysis of the computer tapes, and until this has been carried out the validity of the FRK conclusions cannot be assessed.

There seem to be three possible courses of action :

- i) publish the book as it stands (with editorial revisions),
- ii) ask the authors to revise the statement of the findings so as to bring out the necessary qualifications (which would require more than simply adding a few footnotes),
- iii) wait until the computerised data are available, so that the sensitivity of the results can be checked.

The first of these would in my view mean that the authors were open to criticism even within the terms of reference of the data they actually have; the third would mean that the appearance of an important - book would be considerably delayed. On balance, I would favour the second course, although it has to be recognised that the findings may turn out to need substantial modification when richer data are available. If this is the course followed, then the attention of the authors should be particularly drawn to the comments marked with an * in the detailed comments.

Ginis, Concentration Ratios and Decomposition

As I read the various letters and notes, there is no essential disagreement about the theory; what is disputed is the empirical importance of different factors. This is largely a matter of judgment, based on experience of calculations with actual and hypothetical data, and I have no particular standing. What I have done therefore is to make some comments on the note from Professor Ranis (22 May 1978) in response to the Pyatt/Chen calculations (4 May 1978), and then try to give a brief assessment of the position.

(a) Grouping Error

As both sides have agreed, the problem arises because of the conjunction of two factors (cf. Ch.VII, p.45) - the use of grouped (consolidated) data and the use of concentration rather than Gini ratios. The latter arises only with the decomposition, so that I begin for simplicity with the case of total income. The use of grouped data coupled with a particular method of calculation (of the coefficient) means that the Gini is under-stated. As I pointed out in my earlier report, there has been a substantial literature on this and I am surprised that no real use has been made of the article by Gastwirth (REStat 1972). As far as Ranis' comments on p.4 are concerned, I feel that his criticism of the Chen/Pyatt Gini is over-stated. Obviously any calculation is conditional on the fineness of the grouping but it is possible to calculate bounds which are not arbitrary (and do not involve 'guessing'). Moreover one can make use of other authors' results. For example, Gastwirth refers to the results from the US CPS sample 1968 (approx. 60,000 incomes) where the exact Gini was 0.4014. If the data were grouped, the following bounds could be calculated

No. of groups	Lower bound	Upper bound
10	0.3883	0.4083
28	0.4001	0.4020

This suggests that the gain in accuracy from increasing the number of groups from 10 to 28+ may be quite worthwhile. Moreover, the lower bound corresponds to that used in the book, whereas the true figure is typically closer to the upper bound.

(b) Grouping and Decomposition

The numerical example given by Ranis and the associated diagram are helpful. At the same time, they would have been closer to the case considered in Ch.III if the grouped wage patterns had been monotonically related to total family income. Thus, as he notes, the wage Gini and wage concentration ratio are identical for the grouped published data. (The numerical example would be closer to the actual case if it were changed to $40 = W_3$ and $60 = W_6$.)

I prefer therefore to think about it in terms of the algebra. From (6.8)

$$\begin{aligned}
 G_y &= \sum_i \phi_i R_i G_i \\
 &= \underbrace{\sum_i \phi_i R_i^* G_i}_{\text{approx-}} + \underbrace{\sum_i \phi_i G_i (R_i - R_i^*)}_{\text{error}} \\
 &\quad \text{imation}
 \end{aligned}$$

where R_i^* is an assumed value of the 'correlation' coefficient (R_i^* is assumed to be 1 for all classes of income except transfers where it is -1). The issue is now how close R_i is to the assumed value. For the decile grouped data the authors demonstrate that the error is small, but for the ungrouped data the question is left open. Crucial to this is how much variation in wage (or other income) there is within a given total income grouping. This is an empirical question, as explained on p.77 of Ch.III, and the value of the Pyatt/Chen analysis is that it throws light on the empirical

validity of the assumptions made. Here I find Ranis' response rather puzzling. When he says that one cannot substitute the Pyatt/Chen Ginis, I assume that he means (taking approximate values for 1966) that

	ϕ_i	R_i	G_i (from Pyatt's charts)
Wage	0.48	1	0.53)
Property	0.25	1	0.67)
Agriculture	0.21	1	0.75)
Transfer	0.06	-1	0.80)

) imply approximation
of 0.531

(since I have not seen the original Chen data, I may have misinterpreted the figures). This is indeed 'ridiculously' high, but this means that the error term for the computerised data is much larger than for the grouped data. The fact that this is so is not a matter of design but of testing the hypothesis on which the FRK factor component analysis is based.

(c) Assessment

The question of the grouping error is the more straightforward. As I agreed in my earlier report, and have suggested again in the detailed comments on the revised manuscript, the error from grouping plus sampling error may be sufficiently large to cast doubt on the significance of some of the conclusions drawn. For this reason I have urged caution.

The error which may arise from the differences between R_i and R_i^* is much harder to assess. The Pyatt/Chen calculations suggest that the divergence may be substantial but one cannot say from these that the FRK conclusions are necessarily wrong. What is needed is a detailed study of the micro (computerised) data tapes. I would not like to hazard even a guess as to the outcome of this, and if I were the authors I would not be happy about going into print without some idea of what it would show (although of course the QJE article is already out).

Detailed Comments

Chapter I reads very well and gives a clear introduction to the contents. I have only minor comments.

(1) The footnote acknowledging Cannan is welcome, but his name should be spelled correctly !

(2) Page 17. As noted elsewhere, the grouping (consolidation) issue is of considerable importance and more use should be made of the earlier literature. The passage here needs some re-writing in the light of this.

(3) Page 24. The use of the expression 'deterministic theory' is slightly strange. Usually there is a contrast between deterministic and stochastic theories but this is obviously not what is meant. What is meant is presumably a fully specified theory.

Chapter II is in general a well-structured account of the development of the Taiwanese economy over the past 25 years, which is very helpful to the reader (like myself) who knows little of the background.

(4) A brief chronology of recent Taiwanese history would be useful (for example, is there particular significance to the vertical lines from 1961 and 1968 in the diagrams ?).

(5) Pages 12-13. I am not very happy with this initial discussion of the income distribution evidence, which treats it in a casual way quite out of keeping with the rest of the book. First, international comparisons are difficult to make for well-known reasons. One has to know the basis for the figures, making allowance for household v. family data, etc. Just turning the pages of Jain, one can find examples of Ginis as low as those for the 1960s. They may not be comparable, and I am sure that the statement is correct, but it needs substantiation or qualification.

*(6) Going on from this, there is the question of the standards by which one judges changes. What constitutes a 'substantial'

reduction ? I am not wanting to suggest an answer but the authors should either address the issue or be more guarded in their use of words. In the same way, what does 'virtual constancy' mean ? Here issues such as the standard errors of the Ginis (referred to in my earlier comments) become relevant.

* (7) Page 43. In response to my earlier comment, the authors have added a footnote but I am still uneasy about the approach adopted. It is the same issue as involved in the concentration Gini debate, and as there it is an empirical question whether the use of rankings by farm size makes much difference. Without any support the assumption made does seem heroic. Both the level of concentration and the year-to-year changes may be considerably in error.

* (8) Page 50. The authors have again modified the text, and the points made seem very reasonable. They do however seem to have touching faith in the accuracy of the data. Many readers may I fear be sceptical about statistics quoted to four decimal places based on a farm income survey carried out in 1952.

(9) Page 60. There is still a mistake in this table, or else I still do not understand it. In 1953 agricultural income of 6096 out of 19542 equals 0.3119 as a proportion.

(10) Page 63. Given what we know of the quality of the data, it seems to me unwise to say that the evidence is conclusive.

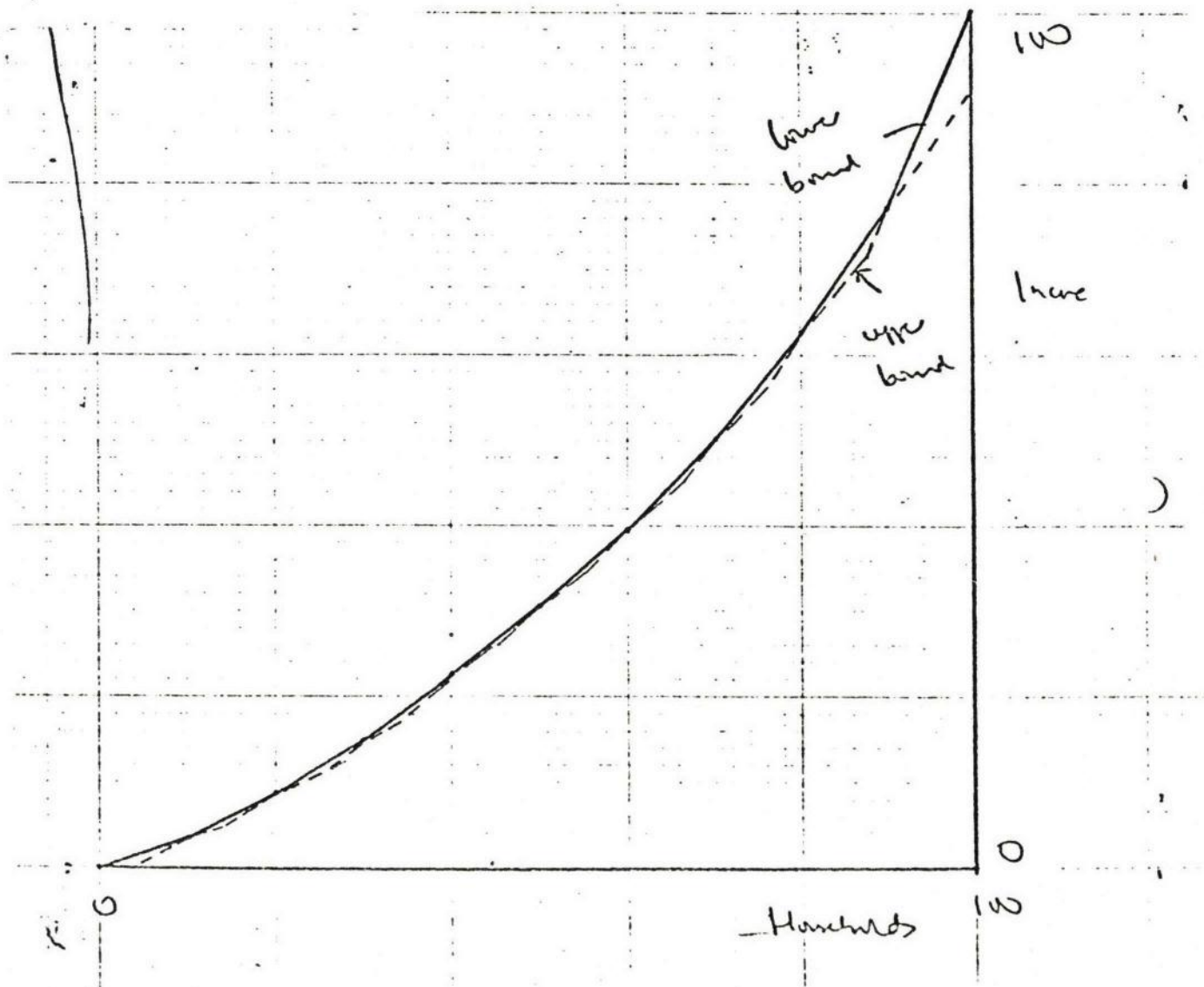
Chapter III raises the issue of concentration and Gini coefficients which I have discussed separately. However, apart from that I am rather disappointed by the relatively small changes which appear to have taken place in this chapter (mainly the addition of footnotes). A clearer account of what the authors are doing would help reduce misunderstanding (particularly since the QJE article is of necessity a rather compressed account of the methodology).

*(11) This criticism may be illustrated by the passage inserted on p.36a which refers to the point made in my earlier report about statistical significance. I am afraid that this paragraph does not in my view meet the objection. Indeed it actually makes things worse, since the earlier conclusions remain unqualified, yet the reader is led to believe that even stronger statements might be justified. In my view the authors should either (a) calculate standard errors or (b) set a consistent quantitative standard for the use of terms such as 'significant' or 'moderate' (e.g. ± 0.02 change in Gini is deemed significant). In the latter case, this would imply an expansion of what is contained in the paragraph at the top of p.32.

(12) Pages 1-10 seem to me to need re-writing with the needs of the reader in mind. For example, the passage on pp.5-6 discussing what would have happened if a negative sign had not been attached to G_N is a digression (what the reader wants to know is why a negative sign is correct). It may indeed be better to leave out Type I income in the initial exposition.

*(13) Page 19 et seq. As I mentioned in my earlier report, there is no adequate treatment of grouping error, and I am surprised that the authors give no calculations of the possible error. As shown by Gastwirth (REStat 72) upper and lower bounds can be calculated on a range of different assumptions (some of which, the gross bounds, are quite weak assumptions). The point is also discussed (in the context of wealth data) in A.B. Atkinson and A.J. Harrison The Distribution of Personal Wealth in Britain (Cambridge University Press, 1978), pp.132-137, where we try to give a more intuitive account. Not having access to the data used by the authors, I have drawn the Lorenz curve attached from the Jain data for Taiwan 1972 (NL, HH) using the decile shares. Since the deciles were not

available, this means that the upper bound is less 'tight' (it means in effect taking the range for the i -th decile of \bar{Y}_{i-1} to \bar{Y}_{i+1} , where these denote the means). The solid line is the lower bound using group means as in FRK; the dashed line is the loose upper bound. Even ignoring the top decile, the difference is of the order of 0.005. This can be sharpened using the deciles; on the other hand, this may not be enough to offset the contribution from the top decile.



Another way of assessing the possible grouping error is from studies on other data. For example, J.L. Gastwirth and M. Glauberan 'On the Interpretation of the Lorenz Curve and Gini Index from Grouped Data' give the following for US IRS data :

Year	No. of groups	Bound : lower	upper	Δ
1963	29	.4423	.4443	.0020
1964	18	.4440	.4492	.0052
1969	21	.4597	.4669	.0072

This lends some support to the view that the difference could be 0.005 or more with 10 groups.

*(14) Page 20. The kind of calculation of grouping error helps put in perspective the non-linearity error shown in Table 1. In the same way it would be interesting to have calculations of the sampling error. Again this was referred to in my earlier report, and the authors' response on p.36a is not quite fair. There has been some discussion of this issue, see Glasser JASA 1962, and Kendall and Stuart, Vol.I, ch.10.

*(15) Taking the grouping and sampling error together suggests that it would be better to delete all references to changes less than 0.005 : e.g. to avoid such statements as 'virtually no deterioration (+0.0002)' on p.30 or 'very modest deterioration' on p.34.

(16) Page 42. As noted earlier, I would prefer to see a more systematic comparison with the cross-country evidence on the Kuznets effect.

*(17) Page 57. My comment on this stands - the conclusion is a non sequitor.

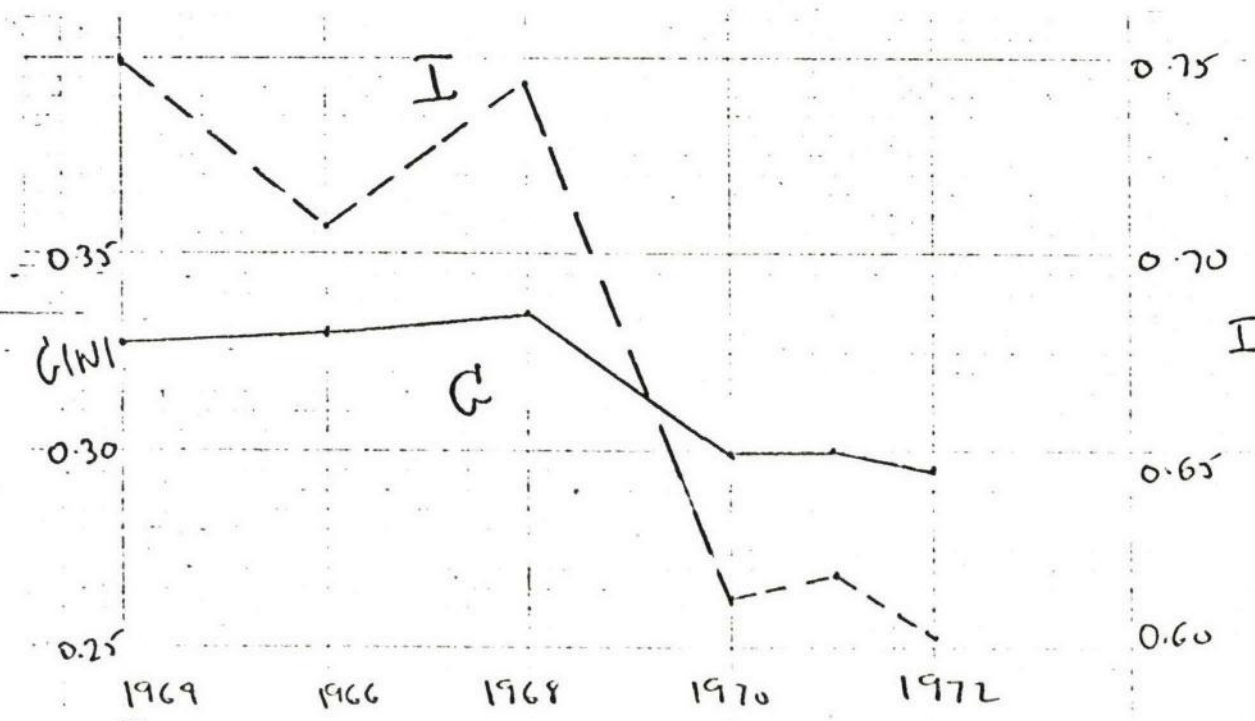
*Chapter IV. As I noted in my earlier comments, little account seems to have been taken of recent work on earnings, and I do not find the brief discussion of this very helpful. I am afraid that labour

economists will find this analysis primitive. To take three examples, there is (i) the difference between age and experience referred to in the earlier report, (ii) the problem of sample selection for married women, and (iii) the problems introduced by using family income as a right hand side variable.

Chapter V. The basic theory of decomposition is now much clearer.

(18) Page 12a. The reference to eqn.5 is a mistake. I am also unclear on two points which are probably obvious but have escaped me. Why is there an error of estimation? Secondly, from the table on p.14, $Z_1 + Z_2 = 2$, which implies from the definitions that $h_1 = h_2 = \frac{1}{2}$, so it looks as though something is wrong.

*(19) Page 13. My comment on the fragile nature of the conclusions still stands, since the amendment does not come to grips with the basic problem. Do the authors really believe that the difference between 0.2907 (farm) and 0.2876 (non farm) is significant? Also the behaviour of I from year to year gives rise to concern. As the diagram below shows, the year-to-year movements are much more pronounced than for G. This reflects no doubt the greater sensitivity to extreme values, but some comment seems called for.



Chapter VI contains a lot of interesting material but its treatment of tax incidence falls short of present 'best practice'. I referred before to the Pechman/Okner work and more account should certainly be taken of the incidence issues.

(20) Page 21. The discussion of rank reversal is reassuring in one respect but not in another. In particular, it brings out the dangers in using grouped data. At the level of individual households the tax system will certainly lead to considerable rank changes : for example because of differences in consumption patterns, differences in portfolio choices, etc.

Chapter VII - see separate comments.

*Chapter VIII. The question to be asked about this chapter is whether it adequately summarises the earlier results, with all the necessary qualifications. In my view it draws together the threads very well but does not give adequate attention to the qualifications. Results, such as those on the contribution of 'institutional discrimination', are stated quite baldly without the reader being made aware of the objections which can be raised to the method of analysis. The same applies to the findings on taxation. Thus the authors state that the tax burden is neutral with respect to FID and draw the conclusion that fiscal policy will probably have a small role to play. This may be true but it is certainly not conclusively demonstrated.

OFFICE MEMORANDUM

TO: Mr. M. Ahluwalia, DRCID
FROM: Armeane M. Choksi, VPD
SUBJECT: Completion Reports *Am*

DATE: February 23, 1978

On October 31, 1977 completion reports for the research projects listed below were requested from you. So far, I have not received them.

If these projects have been completed, would you please send these reports to me by c.o.b. Wednesday, March 15. If for some reason, you cannot send me these reports by the above date, please let me know as soon as possible (Ext. 76003).

Project No.Title

✓670-84	Growth, Employment and Size Distribution of Income
670-85	Urban Income Distribution in Latin America
670-94	Employment and Income Distribution in Malaysia
671-39	Price Intervention in Agriculture
671-41	Indirect Estimation of the Size Distribution of Income

Mr. Montek S. Ahluwalia, DRC

October 31, 1977

Orville F. Grimes, Jr., VPD

Completed Research Projects

1. In discussions with you and from the receipt (and non-receipt) of recent status reports, we have ascertained that a number of the research projects under your supervision are completed. You and I have agreed on the need to prepare completion reports for each of the projects in question. They are listed below:

- 670-84 ("Growth, Employment and Size Distribution of Income")
- 670-85 ("Urban Income Distribution in Latin America")
- 670-94 ("Employment and Income Distribution in Malaysia")
- 671-39 ("Price Intervention in Agriculture")
- 671-41 ("Indirect Estimation of the Size Distribution of Income")

2. Completion reports were originally requested for two of these projects (670-84 and 670-85) last June. Since then, as you know, we have issued a revised form, copies of which are attached.

3. We would appreciate receiving the filled-in completion reports no later than Wednesday, November 30. Thank you.

Attachments

cc(w/o att): Messrs. DuLoy,

cc(w/o att): Mr. DuLoy, Ms. Stout

OFGGrimes:tgr

Growth, Employment, and Size Distribution of Income

Ref. No. 670-84

Evidence from specific countries, such as the Republic of China (Taiwan), leads to some questioning of the hypothesis that the size distribution of income appears to worsen as development proceeds. Even if a negative historical relationship exists between growth rates of Gross National Product (GNP) and trends in the distribution of income, the issue still remains whether this relationship is inevitable or whether an effective development policy can reduce the conflict between growth and the size distribution of income.

This study analyzes the relationship between economic growth, governmental policies and income distribution in Taiwan. The underlying assumption is that in a mixed economy, the long-run trend in the distribution of income is determined by forces reflecting factor endowments, production conditions, and technology, which are modified by government intervention. The study attempts to isolate the factors responsible for the observed changes in income inequality in Taiwan from 1964-72. The technique adopted for this purpose is a decomposition of the Gini coefficient into the contribution of factor income shares and the degree of concentration of factor incomes. Changes in the Gini coefficient

are, therefore, seen as attributable to changes in these components. The impact of government policy on the different sectors of the economy is examined in order to draw generalizations about the relationship between government policy and the processes of growth and distribution.

Responsibility: Development Research Center - Montek S. Ahluwalia. The researchers are John C. H. Fei, Gustav Ranis, and Gary S. Fields of the Economic Growth Center, Yale University, in collaboration with Wan-Yong (Shirley) Kuo of the Economic Planning Council in Taiwan.

Completion date: December 1977.

Reports

Fei, John C.H., and Fields, Gary S. The Indexability of Ordinal Measures of Inequality. Center Discussion Paper No. 205. New Haven, Conn: Economic Growth Center, Yale University, May 1974.


Fei, John C. H., and Ranis, Gustav. Income Inequality by Additive Factor Components. Center Discussion Paper No. 207. New Haven, Conn.: Economic Growth Center, Yale University, June 1974.

Fei, John C.H.; Ranis, Gustav; and Kuo, Shirley. Growth and the Family Distribution of Income by Factor Components: The Case of Taiwan. Center Discussion Paper No. 223. New Haven, Conn.: Economic Growth Center, Yale University, March 1975.

OFFICE MEMORANDUM

TO: Mr. Goddard Winterbottom, IPA

DATE: June 23, 1977

FROM: Graham Pyatt, DRC SUBJECT: "Equity with Growth: The Taiwan Case"

I have had further conversations and written exchanges with Ranis and, more especially, Fei in an attempt to extract some agreed conclusions concerning our earlier differences over the manuscript, Equity with Growth: The Taiwan Case, which you are now considering for publication.

Firstly, I enclose a note which I wrote recently and have made available to both Ranis and Fei. This has elicited a response from Fei in the attached letter and amended page 51 of Chapter 7. I would be grateful if you would insert the amendment as indicated.

The view in the DRC is that an amendment to Chapter 8 does not entirely satisfy our difficulties since the technical point now agreed has implications for other parts of the volume, notably Chapter 3. However, I am advised by Fei that much of this Chapter 3 material is now accepted for publication in article form by the Quarterly Journal of Economics. In this sense the material has clearly passed the important test of editorial review by a respected journal. We would therefore be reluctant to push our objections further beyond the present point. Your reviewers may however choose to agree with us but at this point must obviously be allowed to form an independent judgement.

For the future, I have agreed with Fei to write a joint article embodying the results which are obtained when exact decomposition methods are employed. It is only by doing this that we can hope to produce a final evaluation of the relevance of the approximations employed in the volume you are considering. Meanwhile, our recommendation as a Department in transmitting the volume to you is that it should be seriously considered for publication and that it would be helpful to have reviewers competent to form an independent judgement of the technical issues and their implication for the analysis. Depending on how much weight you wish to put on our judgement, the volume is important and controversial. We have made arrangements to proceed with a dialogue aimed at resolving a major controversy. However, this will inevitably take time and it would be wrong to delay a decision on publication until this future work is completed.

Attachments

Cleared with & cc: Messrs. Duloy, Ahluwalia, C. Bell.

cc: Messrs. Chenery - for information

B.B. King/Grimes

DRC Senior Staff

Mr. M. Ahluwalia, DRCID

June 23, 1977

Orville F. Grimes, Jr., VPD

Completion Report for Research Project No. 670-84
("Growth, Employment, and Size Distribution of Income")

Our records show that this project has been completed. I would therefore appreciate your filling out the attached completion report and returning it to me by Friday, July 8. Thank you.

Attachment

cc: Mr. Duloy
Ms. Stout

OFGrimes:gm

0-84

OFFICE MEMORANDUM

TO: Mr. Montek Ahluwalia
FROM: Graham Pyatt, *GP*
SUBJECT: Ranis-Fei Manuscript

DATE: May 6, 1977

A by-product of the recent Bellagio Conference was the opportunity to re-open dialogue with Gus Ranis, following the somewhat impossible situation that had been reached via recent written communications.

I explained to Gus that his manuscript had already been forwarded to the Publications Committee but that a substantive memorandum of transmission had yet to be sent by us to the Committee. I expressed the view that, for the most part, our earlier concerns with the manuscript were no longer of immediate concern -- the Publications Committee would be appointing readers and it would be for them to adjudicate the merits or otherwise of the text. Meanwhile, however, we retain some worries about the methodology and would feel it incumbent on us to urge the Committee to appoint a referee with particular qualifications to adjudicate in this area if the primary difficulty could not be resolved between us. Gus accepted all this as being fair and in further conversation he conceded that we might be right in stating that many of his Gini coefficients were in fact concentration ratios and therefore would not bear the interpretation which was given to them in the text. In other words, the stylized facts of Taiwan are changed if we are correct on this methodological issue. There would therefore be some need to consider redrafting.

By agreement with Gus, I have contacted John Fei since returning to Washington and enclose a copy of the memorandum which I have sent to him. On the telephone at least he also seemed open to persuasion that there was a methodological problem and potentially asks if he might incorporate my memorandum in their text, if he agrees with it. He also advised that he has drafted an article on their methodology for submission to a journal and will be sending a copy for my comments. I explained that the methodological point had significant implications for the text, so that a few extra pages on methodology would not quite meet the case. I also indicated that Gus had proposed a joint paper to me, if in fact they came to see things my way. I indicated to John that such an arrangement would be entirely acceptable to me, assuming, of course, that we can see eye to eye on the basis of the attached note.

Attachment

cc: Messrs. Chenery
B.B. King
Duloy
C. Bell

May 6, 1977

Dr. John Fei
Economic Growth Center
Yale University
New Haven, Conn. 06520

Dear John:

As promised, I am sending some notes on Gini coefficients versus concentration ratios. I fear that these fall short of being immediately ready for publication but they do, I hope, make clear my concern. Having read them, you might perhaps like to let me know your views and these can be forwarded to me in Malaysia if you were to write in reply to the office here. Moreover, if you feel that there is a need to spend some time together on these questions, I could perhaps extend my proposed stopover in Taiwan. However, this would require some bureaucratic action at this end, so it would be necessary for you to write to me via my secretary, if you feel the case arises.

For my part, I am glad that we are back in correspondence on these issues and trust we can stay with it to the point of resolving outstanding difficulties. Meanwhile, Montek has copies of all correspondence and is aware that you would hope to clarify this technical problem with me before the DRC comments on your manuscript to the Publications Committee. Perhaps, therefore, you should copy any future letters for me to him during my absence.

Yours sincerely,

Graham Pyatt
Senior Adviser
Development Research Center

Attachment

cc: Mr. Ahluwalia

May 6, 1977

A NOTE ON DECOMPOSITION OF GINI COEFFICIENTS

Graham Pyatt

Inequality in total income, y , among a group of individuals or households can be decomposed from two points of view. Both are covered in my paper, "On the Interpretation and Disaggregation of Gini Coefficients", DRC (mimeo) February 1975. The first approach is relevant when the total population is grouped into sub-populations or classes. This leads to a decomposition

$$\begin{aligned} G(y) &= \text{Within group effect} \\ &\quad + \\ &\quad \text{Between group effect} \\ &\quad + \\ &\quad \text{Overlap effect} \end{aligned}$$

and the precise formula and derivation of the result has been set out in my EJ article of June 1976 which gives a more polished version of the relevant proofs than was contained in the earlier paper. It has been noted that if income level is used as the grouping criterion (e.g. by defining classes as decile groups with respect to income), then there will be no overlaps of the income distributions of different groups so that the overlap effect will be zero. Moreover, if all individuals in a group are assumed to be concentrated at some income level (e.g. the mean for the group), then there will be no within group effect. Hence $G(y)$ is captured simply by the between group effect.

If there is variation within groups but this variation is ignored, then the between group effect is an approximation to $G(y)$ which is biased downwards, i.e. $G(y)$ is underestimated by the between group effect. This underestimation may not be serious, however. It decreases as

the number of groups increases and corresponds exactly to drawing the Lorenz curve as a series of connected line segments rather than as a smooth curve.

In their main (Chapter 3, Table 1) analysis, Fei-Ranis-Kuo (FRK) estimate the Gini coefficient for income using decile groups for defining income classes, and ignoring within-group variation. As noted, this will lead to underestimations of the Gini but the error is not likely to be large. In these terms their Gini coefficients for income, y , are conventional and reasonable approximations.

The second approach to decomposition is relevant when income is disaggregated into its component parts, i.e.

$$y = \sum_{i=1}^k x_i$$

Here two exact results are available. The first, due to Rao, is

$$G(y) = \sum_{i=1}^k \phi_i C(x_i)$$

where ϕ is the share of x_i in total income, y . These shares are calculated over the total population so that if subscript j refers to individual j for $j = 1, \dots, n$, then

$$\phi_i = \frac{\sum_{j=1}^n x_{ij}}{\sum_{j=1}^n y_j}$$

It follows that

$$\sum_i \phi_i = 1$$

so that $G(y)$ is simply a weighted average of the variables $C(x_i)$.

The variables $C(x_i)$ are Rao's concentration ratios. They are not Gini coefficients except in a special case. This can be seen from the second exact decomposition which is given in my earlier paper but is not discussed in the published version. The result is

$$G(y) = \sum_{i=1}^k \phi_i \frac{R(x_i, r(y))}{R(x_i, r(x_i))} G(x_i)$$

where G 's are Gini coefficients and ϕ_i 's are defined as before. The new variables which enter are correlation coefficients: $R(u, v)$ is simply the first-order coefficient of correlation between u and v . In the above expression, u always corresponds to x_i i.e. to income of the i th type. The variables $r(y)$ and $r(x_i)$ are rankings of individuals with respect to the respective variables, i.e. they are integers running from 1 to n . Hence, if individuals rank the same with respect to x_i as they do with respect to y then $G(y)$ is simply

$$G(y) = \sum_i \phi_i G(x_i)$$

Conversely, if the rankings differ, then

$$G(y) = \sum_i \phi_i C(x_i) < \sum_i \phi_i G(x_i)$$

More specifically, since

$$C(x_i) = \frac{R(x_i, r(y))}{R(x_i, r(x_i))} G(x_i)$$

$C(x_i)$ is less than $G(x_i)$ if the ranking of individuals with respect to x_i is different from their ranking with respect to y . (This result is easy to prove formally.)

The above discussion shows that the difference between a concentration ratio and a Gini coefficient depends on how individuals are ranked. If individuals are ranked by income, then a graph of cumulative type i income, i.e. x_i , against cumulative numbers of people - both taken in the order dictated by their total income ranking - is a concentration curve. This will differ from a Lorenz curve only if the ordering of people according to x_i is different from their ordering according to y . The concentration coefficient is then calculated as the proportion of the area below the 45° diagonal which is above the concentration curve, i.e. in a strictly analogous manner to the calculation of the Gini coefficient from the Lorenz curve. However, it can be noted that while a Lorenz curve must be convex, a concentration curve need not be. Indeed the concentration curve will be convex only if the ordering of people is the same with respect to x_i as it is with respect to y i.e. only when the concentration curve is identical to the Lorenz curve.

Now the Gini coefficients for income components, x_i , reported by FRK appear to be some compromise between concentration ratios and true Ginis. Specifically they work with ten groups of households (decile groups)

defined with respect to income, and then re-order groups so that the average value of x_i between groups is monotonic increasing. Thus, if there were only ten households they would have true Ginis for each x_i . And if there was no re-ordering of groups then they would have true concentration ratios. But as it is, the numbers must lie somewhere within the interval $C(x_i)$ to $G(x_i)$: but it is hard to discern where.

The problem can be seen alternatively in terms of the first disaggregation discussed above. To calculate $G(x_i)$ households are put into ten groups, i.e. deciles of the distribution of income, y . $G(x_i)$ is then approximated by the between-group effect. Thus, within-group and overlap effects are ignored. This is as before in calculating $G(y)$. But now note that the approximation is much cruder. In calculating $G(y)$ within-group effects will be small because the grouping is with respect to y . But in calculating $G(x_i)$ the within-group effect may be large since the grouping criterion remains the size of y when it should be the size of x_i . Moreover, in calculating $G(y)$, overlap effects are zero by definition since grouping is with respect to y . But in calculating $G(x_i)$ with groups defined with respect to y , there must be overlaps unless x_i and y have the same ranking for all individuals. On both counts $G(x_i)$ will be underestimated more seriously than $G(y)$ using the FRK methodology.

Does this matter? I think it does because trends in $G(x_i)$ as calculated by FRK are not trends in true Ginis. One marked example is income from agriculture, $x_i = a$. FRK show $G(a)$ declining from 30 percent to 10 percent in round terms. This must be false. If 50 percent of the population have zero income from agriculture, then $G(a)$ must be at least 50 percent: and as the proportion of the population with agricultural

income falls, so $G(a)$ will rise. Thus, the FRK result with respect to $G(a)$ seems wrong both with respect to level and trend. Accordingly, the stylised facts they infer are potentially misleading. In truth what seems to be happening is that the correlation between agricultural income and total income is declining - being relatively poor and having agricultural income are becoming less closely associated attributes. Hence, while $G(a)$ may be rising (and almost certainly is) $C(a)$ is falling. To the extent that the FRK Ginis for income components lie closer to the latter, their interpretation must be quite different since changes in $C(x_i)$ depend on changes in $R(x_i, r(y))$ as well as on $G(x_i)$. Unfortunately, there seems to be no way of ascertaining from the FRK data as it now stands just how important the two different contributions might be. What may be possible, however, is to calculate approximate $C(x_i)$'s, i.e. to unscramble any re-ordering of deciles, and to tell the story in terms of trends in ϕ_i 's and $C(x_i)$'s. Unfortunately, however, trends in $C(x_i)$'s compound correlation and Gini effects. Since these cannot be separately identified, the stylised facts in terms of trends in $C(x_i)$'s are not likely to contain so much meat.

WORLD BANK GROUP

ROUTING SLIP		DATE December 27, 1976	
NAME		ROOM NO.	
Messrs. Duloy (o/r)			
Pyatt			
Ahluwalia (o/r)			
	APPROPRIATE DISPOSITION	NOTE AND RETURN	
	APPROVAL	NOTE AND SEND ON	
	COMMENT	PER OUR CONVERSATION	
	FOR ACTION	PER YOUR REQUEST	
X	INFORMATION	PREPARE REPLY	
	INITIAL	RECOMMENDATION	
	NOTE AND FILE	SIGNATURE	
REMARKS			
FROM		ROOM NO.	EXTENSION
B. King		F1233	6001

0-84
12/27/76

"Extract from Yale Growth Center
75/76 Report"

Income Distribution

John Fei and Gustav Ranis completed a draft of a book-length study of the interrelationship between income distribution and economic growth for the case of Taiwan. They show that income distribution, as measured by the Gini coefficient, is very much affected by the particular forces of growth that a country is experiencing. To do this they developed a new method of analysis in which the family distribution of income is decomposed into its factor components. Changes in the distribution of income are then linked to changes in the nature of the growth path itself. Using the Taiwanese case, they show that the beginnings of rapid growth in a developing country need not necessarily be associated with a worsening of the distribution of income.

Gary Fields has conducted research in several areas concerning income distribution. His research on income distribution in Colombia yielded a paper on the determinants of intersectoral wage structure and another on the relationship between education and economic mobility and its effects on income distribution. In addition, he prepared a guide to the use of Colombian microeconomic data. He wrote two papers demonstrating that the rapid alleviation of absolute poverty in Brazil took place at the same time as relative income inequality was increasing. His theoretical research developed a welfare economic approach relating the changing pattern of income distribution to various strategies of economic development over time.

DEC 17 1976

B. King

OFFICE MEMORANDUM

TO: Mr. Hollis B. Chenery, VPD
FROM: Graham Pyatt, DRC *GP*
SUBJECT: Fei, Ranis, Kuo Volume on Taiwan

DATE: December 16, 1976

0-84

Since sending our comments on the Taiwan report to its authors, there have been two developments. The first is a most welcome letter from Shirley Kuo: a copy is attached.

The second new piece of information is that the team organized by Walter Galenson to write a report on Taiwan Post-War Economic Development now has several chapters in draft. It seems very likely that they will produce an excellent report and certainly they have brought together an amazing wealth of facts and data. Galenson is not convinced that 1968 was a turning point, although the initiative in summing up lies with Ian Little. Meanwhile, Erik Thorbecke supports the '68 turning point hypothesis in terms of an end to the slack rural labor market.

It is apparent that the report of the team will overlap the work of Fei, Ranis and Kuo at many points. Indeed Ranis and Kuo are involved in both studies.

Attachment

cc: Messrs. Duloy
Ahluwalia
C. Bell

ECONOMIC PLANNING COUNCIL
EXECUTIVE YUAN

118, HWAI-NING STREET, TAIPEI, TAIWAN
REPUBLIC OF CHINA

REFERENCE NO.

CABLE ADDRESS: EPCEY TAIPEI
TELEX: 11385 TAIPEI

December 10, 1976

Dr. Graham Pyatt
Senior Adviser
Development Research Center
The World Bank
1818 H Street, N.W., Washington D.C. 20433
U. S. A.

Dear Dr. Pyatt:

Thank you very much for your kind letter commenting on my method of approximate Gini disaggregation and for sending me your splendid paper.

Your excellent comment gives me the idea that I should rewrite a part of my chapter giving a clearer explanation of the abstract methodology aspect of my paper with the aid of a numerical example. For this will make it easier for readers.

I will send you the revised version of my chapter as soon as I finish it.

With best regards,

Sincerely,

Shirley W. Y. Kuo

Shirley W. Y. Kuo

cc: Messrs. Chenery)
Duloy o/r) 12/16/76
Ahluwalia o/r)

0-684

December 3, 1976

Prof. Gustav Ranis
Universidad de los Andes
Facultad de Economía
Apartado Aereo 4976
Bogota, Colombia

Dear Gus:

Your letter of October 18 awaited my return from Pakistan-- sorry for the delay in responding. You are correct that there is an accommodation constraint at Bellagio so that we are limited to either you or John, but not both. Montek will put out a detailed agenda when he returns early in January. However, we are interested more in the policy implications of the Taiwan study than in the algebraic intricacies-- which suggests you rather than John.

I regret to be a day or so over the end-November date for comments. However, here is a set prepared by Clive Bell and Graham Pyatt. As you will note, they cover what is by now familiar ground, viz. decomposition procedures and data. In the revised manuscript, I would urge the use of an exact decomposition as well as a careful statement of how the discontinuity in the data series was handled. Presumably, it will be possible to produce a much slimmer version for consideration for publication than the rather voluminous report on the project. This might be achieved, for example, by publishing the core of the methodology in a journal and confining the monograph to the analysis and implications thereof, with only references to the methodological literature.

For your information, the Bank procedure in this matter is for a submitted manuscript to be evaluated by its Editorial Committee, which seeks guidance from invited readers. While the transmitting department (in this case the DRC) can make recommendations, the decision is the sole responsibility of the Committee.

Yours sincerely,

John H. Duloy, Director
Development Research Center

Enclosure

cc: M/s. John Fei (Taiwan), Chenery, B. King, Pyatt, Weaving (w/incoming) JED/vec

OFFICE MEMORANDUM

TO: Mr. M. Ahluwalia

DATE: November 29, 1976

FROM: Clive Bell and Graham Pyatt

SUBJECT: 'Equity with Growth: The Taiwan Case' by Fei, Ranis and Kuo (Draft)

1. The authors of this volume are refreshingly frank in expressing their disappointment that a theory of growth and distribution has not emerged from their analysis. It remains the case, however, that it was worth a try, while the interest of their report would be enhanced if their own reflections on why such a theory is elusive were set out.

2. Our views on this question fall under two heads. First, the approach adopted concentrates on what has happened to inequality (measured by the Gini coefficient). This is partial and it is difficult to see how a general theory can be developed without simultaneous consideration of growth. Yet the volume says relatively little about this and the disaggregation of growth as opposed to inequality is seriously underplayed. Moreover, the methodology with respect to inequality relies heavily on statistical decompositions without an explicit attempt to drive ex post observations back to structural and policy foundations. Accordingly, there is limited scope for a general theory of growth and redistribution to emerge.

3. Within these limitations (which we can see more clearly with hindsight), the authors have been working under further handicaps which are, in part at least, self-imposed. Thus, with respect to statistical decomposition, they have chosen to work with Gini rather than Theil. This creates difficulties which are well known. Also, within the Gini framework, the formal problems of statistical analysis have been solved by Mangahas, Mehran, Rao and Pyatt's work based on Rao and Bhattacharya/Mahalanobis. None of these contributions appears to have been deployed so that instead the statistical methods adopted are rife with approximations and perversions of results available in the literature. It can be argued, of course, that this doesn't matter insofar as the approximations are close. However, there is the obvious point that the literature should be taken on board by an academic study so that at least a new chapter/appendix is needed to set out the literature and relate the methods used to it. If this were done (much of our time has been taken up trying to disentangle the approximations) then some other points would emerge. One is that the approximate methods adopted are not invariably close and can involve offsetting errors. Hence the decomposition can be more misleading than the aggregate error suggests, both at a point of time and over time.

4. As it turns out, Table 1 (III:20) provides the wherewithal to examine the importance of this last point with respect to time. The following first differences have been obtained from Table 1 using equation (1.8) of the text:

0-84

Mr. Hollis B. Chenery

November 29, 1976

John H. Duloy

CONFIDENTIAL
DECLASSIFIED

Fai-Ranis-Kuo on Taiwan


DEC 19 2019

WBG ARCHIVES

I would appreciate your views on the attached draft letter to Gus Ranis, before it is sent. To refresh your mind on the past history of the project, I attach also a selected set of previous correspondence.

Attachments

cc: Mr. B.B. King
JHb/vec



Dear Gus:

We promised to let you have comments on your draft, "Equity with Growth: The Taiwan Case," by the end of this month and I apologize for the fact that we are doing so at the last minute. However, we have had some real problems with your report and it has taken time to clarify the issues and to determine our position at this stage.

Essentially our view is that the concerns we have expressed over the course of the project about (a) the use of Gini approximation formulae, and (b) quality of basic data, continue to obtain. The attached note by Clive Bell and Graham Pyatt sets out our view of these issues in some detail, and I hope you will agree that their arguments are both considered and weighty.

There is no point in my rehearsing here the reasons why we cannot regard your Gini approximations as satisfactory. At the very least we think you should take the published literature in this field much more seriously and in the final analysis you should make use of the various exact decompositions which are available.

We realize that it would be a major task to re-do your analysis using exact decompositions. This would have been relatively straightforward had the data base been set up in machine-readable form. However, your chosen strategy has been to work manually with the data so that considerable costs (which we are not in a position to meet) would now be involved in any attempt to use exact methods. Yet it is our view that exact decompositions are necessary as a basis for a satisfactory publishable product.

It is, of course, entirely your decision whether to embark on a new analysis involving exact decompositions. Even if you did so, however, there would remain doubts about whether the data base is good enough to justify the effort and required resources. These doubts center around the discontinuity

in the coverage of the series between 1968 and 1970. Your draft report does not provide an adequate basis to judge whether the observed discontinuity is wholly or partially attributable to the exclusion of Taipei in the basic source after 1968 and your subsequent adjustment of the figures to correct for this. Our view is that this question should not be left in the air and ought to be resolved explicitly and in depth in any potential publication.

As you know, we have completed disbursements to you under the agreed research contract. The fact that we have major difficulties with your report, as outlined above, implies that substantial revisions would be necessary if the report were to be given any further consideration as a potential publication by the Bank. In these circumstances you may well wish to consider alternatives. One possibility would be for you to proceed to publication independently of the Bank, and thereby to bypass the comments and suggestions we have made. In this case, I would have to request that you do not include any reference or acknowledgment of Bank support and participation in the study.

I very much regret that this project has not worked out better from both our points of view. We feel that we have tried to make constructive criticisms in the past, and that they have focussed on issues which are important. The fact that you have not incorporated our suggestions we see as a matter for regret. My concern, then, is that we now be able to agree to disagree, since it is presumably too late in the day for you to meet our objections.

Yours sincerely,

John H. Duloy, Director
Development Research Center

Attachment: note by C. Bell and G. Pyatt

OFFICE MEMORANDUM

TO: Mr. M. Ahluwalia

DATE: November 29, 1976

FROM: Clive Bell and Graham Pyatt

SUBJECT: 'Equity with Growth: The Taiwan Case' by Fei, Ranis and Kuo (Draft)

1. The authors of this volume are refreshingly frank in expressing their disappointment that a theory of growth and distribution has not emerged from their analysis. It remains the case, however, that it was worth a try, while the interest of their report would be enhanced if their own reflections on why such a theory is elusive were set out.
2. Our views on this question fall under two heads. First, the approach adopted concentrates on what has happened to inequality (measured by the Gini coefficient). This is partial and it is difficult to see how a general theory can be developed without simultaneous consideration of growth. Yet the volume says relatively little about this and the disaggregation of growth as opposed to inequality is seriously underplayed. Moreover, the methodology with respect to inequality relies heavily on statistical decompositions without an explicit attempt to drive ex post observations back to structural and policy foundations. Accordingly, there is limited scope for a general theory of growth and redistribution to emerge.
3. Within these limitations (which we can see more clearly with hindsight), the authors have been working under further handicaps which are, in part at least, self-imposed. Thus, with respect to statistical decomposition, they have chosen to work with Gini rather than Theil. This creates difficulties which are well known. Also, within the Gini framework, the formal problems of statistical analysis have been solved by Mangahas, Mehran, Rao and Pyatt's work based on Rao and Bhattacharya/Mahalanobis. None of these contributions appears to have been deployed so that instead the statistical methods adopted are rife with approximations and perversions of results available in the literature. It can be argued, of course, that this doesn't matter insofar as the approximations are close. However, there is the obvious point that the literature should be taken on board by an academic study so that at least a new chapter/appendix is needed to set out the literature and relate the methods used to it. If this were done (much of our time has been taken up trying to disentangle the approximations) then some other points would emerge. One is that the approximate methods adopted are not invariably close and can involve offsetting errors. Hence the decomposition can be more misleading than the aggregate error suggests, both at a point of time and over time.
4. As it turns out, Table 1 (III:20) provides the wherewithal to examine the importance of this last point with respect to time. The following first differences have been obtained from Table 1 using equation (1.8) of the text:

	<u>1966/64</u>	<u>1968/66</u>	<u>1970/68</u>	<u>1971/1970</u>	<u>1972/71</u>
All Households					
ΔG_y	0.0018	0.0024	-0.0332	0.0022	-0.0053
$\Delta \theta$	-0.0005	0.0015	-0.0010	0	-0.0001
Urban Households					
ΔG_y	--	0.0060	-0.0502	0	0.0019
$\Delta \theta$	--	0.0001	0	0.0062	-0.0052
Rural Households					
ΔG_y	--	-0.0358	0.0765	0.0167	-0.0297
$\Delta \theta$	--	-0.0001	-0.0838	0.0039	0.0227

where ΔG_y is the change in Gini with respect to income and $\Delta \theta$ is the change over time in the error of approximation due to 'non-linearity error.' (It can be noted that these figures are based on our own calculations of θ and $\Delta \theta$. These differ from the data of Table 1 which erroneously shows low values of θ in the four instances in which we detect calculation errors.) With respect to our calculations, in some cases, to be sure, $|\Delta \theta|$ is indeed "small" compared with $|\Delta G_y|$; but then again, in others it is obviously not. Moreover, two of the "acceptable" instances occur in the period 1970/68, when there was a pronounced "step" in each of the observed G_y 's, a point to which we return below. Meanwhile, it appears that changes in the estimates obtained from the "correct approximation equation" are matched, or even dominated, by those in the approximation error. Accordingly we cannot subscribe to the authors' view that "from now on we can safely ignore the non-linearity term in our empirical analysis" (III:26).

5. In parallel with formal statistical methods, there is a need for much more detail on basic data. This is vital because the main time series to be explained is (Table 1, Chapter 3)

Year	1964	1966	1968	1970	1971	1972
Gini	.32	.32	.33	.29	.29(5)	.29

The discontinuity between 1968 and 1970 is obvious. And this happens to correspond to a discontinuity in the data. It is inadequate for the authors to refer the reader to Kuo "Income Distribution by Size in Taiwan Area - Changes and Causes" for two reasons. One is that their main thesis is that 1968 was a turning point, so the reconciliation of inconsistent data for 1964 to 1968 and 1970 to 1972 is crucial and should be reported at length as part of this volume. A second objection is that the Kuo paper referred to is wholly inadequate on the subject -- there is just one paragraph which reads:

" (1) Adjustment of DGBAS Data to include Taipei City

Since the coverage of DGBAS data for the period 1964-1968 is different from the coverage of the period 1970-1972, i.e., the former including Taipei City data and the latter excluding it, a necessary adjustment is required in this regard. The DGBAS data of 1970, 1971 and 1972 without Taipei City are adjusted to include Taipei City firstly. "

The importance of this issue is highlighted by the fact that "F.4B: the significant improvement of G (after 1968) is contributed mostly by the now favourable non-agricultural Gini effect ..." Clearly, then, documentation on the data is needed to dispel the possibility that nothing more interesting has been happening in Taiwan than a change in data coverage.

6. All the above points could be taken on board by a revised draft: meanwhile, we think they are serious criticisms of the present manuscript.

7. Further limitations of the study derive from the fact that the data (except for 1966) have not been computerized. This makes "reruns" a virtual impossibility and has, perhaps, led to the use of numerical methods which compound the problems with approximation methods of decomposition. Thus in Chapter 6 there is a decomposition of income after tax into consumption and savings. The method used is exact. (It is the Rao result.) It gives

$$G = \phi_c R_c G_c + \phi_s R_s G_s$$

The ϕ 's are correctly calculated ($\phi_c = apc$; $\phi_s = aps$). However, the data are handled in such a way that R_c and R_s are not the correlation coefficients between consumption or savings and income; and the G_c and G_s are not the Gini coefficients for consumption and savings either. This is obvious from the fact that R_s is reported as 1.000 for 1972 -- an amazing result if it were true. A little further reflection indicates that this must be curious: some individuals have negative savings so it is logically impossible to calculate a Gini coefficient for savings.

8. So the R 's are not correlation coefficients using individual data; and the G 's are not Gini coefficients.

9. This conundrum seems to be explained by the fact that all the analysis is based on data grouped by household income level. These groupings never change, i.e. there is no regrouping according to consumption level in order to calculate G_c . Rather, G_c is obtained from the Lorenz curve for consumption built up by ordering household groups according to average consumption level when the groups themselves are defined by income level (percentile groups). This seems to be the implication of the new appendix which John Fei has sent in response to an inquiry (see attached). It implies that every Gini coefficient in the volume, except the Gini coefficients for total family income, are under-estimates of true Ginis since the wrong grouping criterion is being used. At no point is this acknowledged in the main text which talks throughout as if

Gini coefficients were what they purport to be. This can be extremely misleading. Three examples may suffice. (i) The tax system is shown to have no influence on redistributing income in 1973 (see Chapter 6, Table 10). Yet if income before tax is the grouping criterion then the Gini coefficient for income after tax must be understated so that the tax system must be regressive, i.e., the regressive indirect taxes more than offset the progressive direct taxes. (ii) A second example (Chapter III, Table 1) shows the Gini coefficient for agricultural income declining 1964 to 1973 from 0.35 to 0.11. This is spurious -- as the proportion of the population in agriculture declines, so the inequality in income from agriculture across all households must rise (and if half the population have no agricultural income then the Gini must be at least 0.5). The decline from 0.35 to 0.11 therefore has little to do with the Gini coefficient for agricultural incomes. What it reflects is that the grouping criterion (total family income) is increasingly irrelevant as a proxy for income from agriculture. (iii) As a third example, Finding F.3b reports "virtually no deterioration [in rural income inequality] for 1968-72." Yet inspection of Table 1 (III.20) reveals that G_y^r rose from 0.2842 to 0.3477 over that period, a deterioration of more than 20 percent. We are therefore left wondering whether the authors have misinterpreted their data or we have misinterpreted what they mean by a Gini coefficient.

10. We do not exclude the possibility of having misunderstood what is going on. The explanations we have are by no means clear on the precise methods used or the interpretations which should be given to particular statistics. But if as a result of earlier exchanges with the authors we now have the correct diagnosis, then it follows that much of the text and argument is seriously misleading. The three examples given suffice to make the point that the findings reported are simply not well founded.

11. If we are right, what can be done? One solution is to spell out what is happening within the data and explain to the reader that all Gini's (other than for total income) are not Gini's and so should not be read as such. This would be very messy. A second solution would be to cut out all Gini's other than G_y and try to restate the history of Taiwan with the data that remains. This may be possible and could result in a nice (albeit slim) volume. It would, of course, require a major effort.

12. Perhaps enough has been said to express concern and disquiet. However, we must draw attention to two further problem areas.

13. First, Shirley Kuo's disaggregation in Chapter 5 is wrong. This (and much of the data) comes from her paper referred to above. The error is pointed out in the appended correspondence. It can be seen most clearly by noting that if a population is divided into groups so that there is only one person in each group, then the Kuo formula gives a Gini of zero irrespective of what the true Gini is. Her formula works well when the criteria for disaggregation result in groups which in fact have similar income distributions, i.e., the criteria don't discriminate. This is in fact the case throughout

much of the chapter. Its argument is therefore that the criteria she looks at don't matter much. This is interesting. But the Gini disaggregation is unnecessary to making the point.

14. Finally, the 1966 data was computerized and some regressions emerge. The earnings function fitted takes little cognisance of the vast literature on such functions and is badly specified. One in particular is the inclusion of family income as an explanatory variable for individual wage income as a potential test of nepotism. We just can't let this sort of thing pass. Moreover, to go from regression to Gini disaggregation of predicted values is somewhat gratuitous -- there are well established techniques within the regression framework for analyzing how much each variable contributes to total explanation. (And the link between Gini's and regression is set out by Rao.)

15. So we are left at the end of the day with a number of mixed feelings. Essentially we are very disappointed, perhaps because of wanting to see some glimpses of solutions to the exciting challenge that this project took on. And perhaps they are there. But in general we find ourselves in a pervasive fog. A lot of work has gone into this study, so we hesitate to write it off and have enough respect for the authors to retain some unease that we may be seeing difficulties where none exists. But we must conclude by saying that for the most part we have got nothing but misgivings out of reading the manuscript.

16. Graham Pyatt's annotated copy of the draft is available for return to the authors.

cc: Messrs. B.B. King
Duloy

September 20, 1976

Professor John Fei
Economic Growth Center
Yale University
New Haven, Conn. 06520

Dear John:

I am reading your study with Gus and Shirley Kuo on "The Taiwan Case" and so far have a number of comments on the early chapters. However, I have got stuck at Table 1 in Chapter 3 and must ask for some clarification at this point in order to proceed.

My problem concerns the way in which the statistics G_w , G_π and G_A have been calculated. Footnote 1, p. 3.19, states that the decile group with respect to total income was determined for each household. Does this mean that G_w , say, is obtained by ranking households (to the nearest decile) according to y and then obtaining G_w from cumulative wages versus cumulative numbers of households? My reading of the figures is that this must have been the procedure. If it is correct, then I think that there are a number of points which follow:

- (i) G_w , G_π and G_A are not Gini coefficients but rather concentration coefficients in Rao's terminology, or pseudo-Gini's (\bar{G} 's) in yours.
- (ii) If they are pseudo-Gini's, then the disaggregation

$$G = \sum \phi_i \bar{G}_i$$

is exact. Thus the errors θ in Table 1 are due to statistical problems, e.g. the fact that $\sum \phi_i \neq 1$ in your data, rather than to 'Gini error' as you imply.

If my conjecture above is wrong, then you must have ranked households separately by (deciles of) w , π and A . But if this is so, then the numbers seem to be rather odd. Thus, if 50 percent of the population have no agricultural income, then G_A cannot be less than 0.5. Yet for 1972 you have 0.11. Similarly, G_A for all households cannot be less than G_A for rural households (since adding new population members with zero income can only increase a Gini coefficient). Yet the ordering is always the other way round in Table 1.

You will gather that I have a rather important problem which obviously has implications for your analysis. Hopefully you will be able to resolve the issue for me -- perhaps because there is some third alternative which I have missed. In any event, an early reply would be most helpful since the above needs to be resolved before I can proceed with my assimilation of your analysis.

My reading beyond p. 3.20 has been more cursory so far. However, in Chapter 5, I have a difficulty which I can take this opportunity of also raising with you. It concerns equation (1). I follow the appendix proof down to mid p.53, the result there being a version of my own Gini disaggregation formula. However, you then make the simplifying assumption $f_1(x) = f_2(x)$. I read this as implying that the distribution of Y in sector 1 differs from that in sector 2 by a scalar (due to mean differences between sectors). However, this would imply $G_1 = G_2$. So if in fact $G_1 \neq G_2$, then formula (7) in the appendix can only be an approximation. Is this correct? If so, it might be interesting to explore the nature of the approximation by reference to my exact decomposition.

I am sorry to be bothering you with all this, but anticipate that you will share my concern to have a clear understanding of these issues. Since Gus is abroad, I understand, I am not copying this letter to him. No doubt you will be keeping in touch with him meanwhile and will let me know if I should write him direct.

With best wishes,

Yours sincerely,

Graham Pyatt
Senior Adviser
Development Research Center

ECONOMIC PLANNING COUNCIL
EXECUTIVE YUAN

Oct. 4, 1976

Dr. Graham Pyatt
The World Bank
1818, H Street
N.W., Washington, D.C., 20433
U.S.A.

Dear Graham:

Ms. Shirlry Kuo gave me her copy of our manuscript on Income Distribution, I am now ready to answer the question which you raised on the first page of your letter of Sept.20. The intricate nature of the issue can be best analyzed with aid of numerical example which is attached.

In respect to the question you raised on the second paper, I have given your comment to Shirley Kuo, who will no doubt write you about it. Shirley is the main author of Chapter 5 of our manuscript.

I am very happy that you can give our manuscript a very careful reading. In order to make the point clearer to the future readers, the few pages attached should be included as a separate appendix of Chapter 3 .

Best regards,

Sincerely yours

John C.H. Fei
John C.H. Fei

cc : Messrs Dulay
Ahlwalia

Appendix: Impact on Factor Gini Coefficient due to the Grouping Data

For the computation of Gini Coefficient and Factor Gini Coefficients of this Chapter, we made use of statistical data which is published in a "grouped form" according to the stratification of families into different income classes. This may be illustrated by Table A1. In this Table there are 4 "income classes" with the average income shown in col(1), while the number of families shown in col(2). Their products (ie. total class income) are shown in col(3), while col(4) and (5) show total wage income and total agricultural income for each income class respectively. (col(3) is the sum of col(4) and (5)) This is the way in which the original data are published.

Table A1: Original Source Data

Income class	Average income (1)	No. of families (2)	Total class income (3)	Total wage income (4)	Total agricultural (5)
1	4	1	4	4	0
2	5	3	15	12	3
3	12	3	36	33	3
4	15	3	45	30	15
		10	100	79	21

In computing the Gini Coefficient, this Table is first processed into Table A2 showing the income pattern for a typical member within each income class. Col(3) of Table A2 is the same as col(1) of Table A1, while col(4) and (5) of this Table are obtained from col(4) and (5) of Table A1, divided by the number of families within the each class. Thus for the 10 families in this Table, the total family, wage income and agricultural income patterns are represented by the following vectors.

Table A2: Typical Earning Pattern

Income class (1)	No. of families (2)	Average family income (3)	Average wage income (4)	Average agricultural income (5)
1	1	4	4	0
2	3	5	4	1
3	3	12	11	1
4	3	15	10	5

ECONOMIC PLANNING COUNCIL
EXECUTIVE YUAN

- 1.a) $Y = (4, 5, 5, 5, 12, 12, 12, 15, 15, 15)$ $G_Y = .24 = \bar{G}_Y$
 b) $W = (4, 4, 4, 4, 11, 11, 11, 10, 10, 10)$ $\bar{G}_W = .19$
 c) $A = (0, 1, 1, 1, 1, 1, 1, 5, 5, 5)$ $\bar{G}_A = .44 = G_A$

With the indicated Pseudo Gini Coefficient \bar{G}_Y , \bar{G}_W , \bar{G}_A , The underlying assumption when Table A1 is processed into Table A2, is that every person belonging the same income class has exactly the same income pattern. Whenever one makes use of data published in the form of Table A1, this assumption must obviously be made.

In order to compute the Factor Gini Coefficient, notice that the wage vector in 1.b) is not arranged in a monotonic order. When it is rearranged in a monotonic order, it becomes:

2) $(4, 4, 4, 4, 10, 10, 10, 11, 11, 11)$ $G_W = .21$

with the indicated wage Gini Coefficient. While the Pseudo Factor Gini Coefficients satisfy the condition

3) $G_Y = \phi_W \bar{G}_W + \phi_A \bar{G}_A = (.79) (.19) + (.21) (.44) = .24$

exactly, the estimated value of Gini Coefficient is:

4) $\hat{G}_Y = \phi_W G_W + \phi_A G_A = (.79) (.21) + (.21) (.44) = .26$

Thus there is a nonlinear error:

5) $\theta = \hat{G}_Y - G_Y = .26 - .24 = .02$

There is another type error due to the fact that the published data (i.e. Table A1) is different from the true data. Since, the original data are obtained from sample questionnaires in household survey, the returns appear in the form of Table A3

Table A3: True Income Patterns

Family No.	True Income Pattern		
	Total income	Wage income	Agricultural income
1	4	4	0
2	5	5	0
3	5	5	0
4	5	2	3
5	12	12	0
6	12	12	0
7	12	9	3
8	15	15	0
9	15	15	0
10	15	0	15
Total	100	79	21
Share	1.00	$\phi_w = .79$	$\phi_A = .21$

in which the true income patterns of the 10 families are shown by the 10 rows of this table. Thus the true total family income, wage income, agricultural income patterns are:

$$\begin{aligned}
 6.a) \ Y^T &= (4, 5, 5, 5, 12, 12, 12, 15, 15, 15) & G_y^T &= .24 \\
 b) \ W^T &= (4, 5, 5, 2, 12, 12, 9, 15, 15, 0) & G_w^T &= .37 \\
 c) \ A^T &= (0, 0, 0, 3, 0, 0, 3, 0, 0, 15) & G_A^T &= .81
 \end{aligned}$$

Notice that when the questionnaire returns ie. Table A3 was processed into Table A1 and then processed into Table A2, there is a fouling of data because of the grouping error. This error can be seen by comparing

the true patterns in 6.a.b.c) and the foul patterns in 1.a.b.c). In other words, while in this simple case there is no fouling of the total income pattern. " $Y^T = Y$ " there is a fouling of both wage income pattern " $W^T \neq W$ " and agricultural income pattern " $A^T \neq A$ ". Generally, there is underestimation of the true degree of inequality for each factor component:

$$\begin{aligned} 7.a) \quad G_W^T &> G_W && (.37 > .21) \\ b) \quad G_A^T &> G_A && (.81 > .44) \end{aligned}$$

as stated in footnote 1 of section 3. Such a fouling of data is ordinarily encountered whenever the raw data are grouped into frequency distribution Table. However in the calculation of Factor Gini Coefficient such a fouling can be a serious one. For example in the true income pattern of Table A3 more than 50% of the families receive no agricultural income at all. Hence, the true Gini Coefficient is greater than .5 (ie. $G_A^T = .81$). However the agricultural Gini based on the grouped data is much smaller (ie $G_A = .44$). This error is strictly due to the consolidation of data, and has nothing to do with the nonlinear error .

As long as one makes use of published data in the form of Table A1, the consolidation error can not be avoided. The only way they can be avoided is through the use of the original questionnaire. Thus in this Chapter, we make use of the published data, while in Chapter 4 and 6 we make use of the data from the original questionnaire (see appendix of chapter 4). Thus the factor Gini Coefficient of this Chapter differ (and are generally smaller than) the Factor Gini Coefficient of Chapter 4 and 6, because of the consolidation error.

ECONOMIC PLANNING COUNCIL
EXECUTIVE YUAN

118, HWAI-NING STREET, TAIPEI, TAIWAN
REPUBLIC OF CHINA

CABLE ADDRESS: EPCEY TAIPEI
TELEX: 11385 TAIPEI

REFERENCE NO.

October 2, 1976

Dr. Graham Pyatt
Senior Advisor
Development Research Center
The World Bank
1818 H Street N. W.
Washington D.C. 20433
U. S. A.

Dear Dr. Pyatt:

Professor John Fei who is currently with us at the Economic Planning Council, referred your letter of September 20 to him to me.

On the second page of your letter, you raised the question concerning the decomposition formula of chapter 5 in our manuscript. Since chapter 5 is essentially my work, I made use of a decomposition equation which I developed some two years ago. (See pages 105-109 in the paper I attached). I am happy to know that we have been working along parallel lines.

The issue which you mentioned in your letter to Dr. Fei certainly deserves further investigation. Will you please send me a reprint of your work on this issue. In case I make any further progress in studying this issue, I will certainly write to you immediately.

Thank you for reading our manuscript.

With best wishes,

Sincerely,

Shirley W. Y. Kuo
Shirley W. Y. Kuo

November 18, 1976

Dr. Shirley W.Y. Kuo
Economic Planning Council
Executive Yuan
118, Hwai-ning Street
Taipei, Taiwan
Republic of China

Dear Dr. Kuo:

Many thanks for your letter of October 2 and the enclosed copy of your paper on income distribution in Taiwan. In reply I enclose a copy of my recent paper on Gini coefficient disaggregation on which your comments would be welcome. Please note that I am sending a copy of this letter to John Fei and no doubt you will wish to discuss its contents with him.

I have been unable to resurrect your approximate Gini disaggregation formula (3) (p.109) from my own (exact) result and think that the trouble lies with your formula (1) (p.108). This states that

$$f_w(x) = h_a \left[\frac{y_w}{y_a} \right] f_a(x) + h_n \left[\frac{y_w}{y_n} \right] f_n(x) \quad (1)$$

where w indicates whole population
 a " agricultural households
 n " non-agricultural households
 h_a, h_n are population proportions such that

$$h_a + h_n = 1$$

and y_a, y_n and y_w are mean incomes such that

$$y_w = h_a y_a + h_n y_n$$

This leaves f_a, f_n and f_w to be defined as density functions of the variable x where x is income normalized by mean income for the (sub) population in question.

Consider drawing an income at random from the whole population. The probability that this is less than xy_w , i.e., x times the average y_w , can be denoted $\Pr(y < xy_w)$ and your first equation at the top of p.108 can be written

$$\Pr (y < xy_w) = \Pr (y < xy_w/y \text{ is an agricultural income}). h_a \\ + \Pr (y < xy_w/y \text{ is a non-ag. income}). h_n$$

This can be written equivalently as

$$\Pr \left(\frac{y}{y_w} < x \right) = \Pr \left(\frac{y}{y_a} < x \frac{y_w}{y_a} / y \text{ is an ag. income} \right). h_a \\ + \Pr \left(\frac{y}{y_n} < x \frac{y_w}{y_n} / y \text{ is non-agricultural} \right). h_n$$

Now differentiating both sides with respect to x gives

$$f_w(x) = h_a \frac{y_w}{y_a} f_a \left(x \frac{y_w}{y_a} \right) + h_n \frac{y_w}{y_n} f_n \left(x \frac{y_w}{y_n} \right) \quad (1)'$$

It is apparent that this is not the same as your equation (1), but the two will be close if $y_a = y_w = y_n$. This condition holds for the data you present in Table 13 but can't be relied on in general. If you combine it with the other assumption on which your decomposition is based, viz. $f_a(x) = f_n(x)$, then the total effect is to assume that income distributions are similar in the two sub-sectors.

The usefulness of your decomposition formula (2) (p.109) is obviously limited if it does not take full account of differences in means between sub-groups. I would therefore very much welcome your comment on the above and, meanwhile, will look further into the question of more robust approximations based on $f_a(x) = f_n(x)$. Hopefully you will be able to spare me this effort by pointing out some error in my thinking. But the fact is that I think (1)' is right and the double use of (1) in the 'proof' of (2) is relatively serious.

Yours sincerely,

Graham Pyatt
Senior Adviser
Development Research Center

cc: Dr. John Fel
Dr. G. Ranis, Yale
Mr. M. Ahluwalia, DRC

	<u>1966/64</u>	<u>1968/66</u>	<u>1970/68</u>	<u>1971/1970</u>	<u>1972/71</u>
All Households					
ΔG_y	0.0018	0.0024	-0.0332	0.0022	-0.0053
$\Delta \theta$	-0.0005	0.0015	-0.0010	0	-0.0001
Urban Households					
ΔG_y	--	0.0060	-0.0502	0	0.0019
$\Delta \theta$	--	0.0001	0	0.0062	-0.0052
Rural Households					
ΔG_y	--	-0.0358	0.0765	0.0167	-0.0297
$\Delta \theta$	--	-0.0001	-0.0838	0.0039	0.0227

where ΔG_y is the change in Gini with respect to income and $\Delta \theta$ is the change over time in the error of approximation due to 'non-linearity error.' (It can be noted that these figures are based on our own calculations of θ and $\Delta \theta$. These differ from the data of Table 1 which erroneously shows low values of θ in the four instances in which we detect calculation errors.) With respect to our calculations, in some cases, to be sure, $|\Delta \theta|$ is indeed "small" compared with $|\Delta G_y|$; but then again, in others it is obviously not. Moreover, two of the "acceptable" instances occur in the period 1970/68, when there was a pronounced "step" in each of the observed G_y 's, a point to which we return below. Meanwhile, it appears that changes in the estimates obtained from the "correct approximation equation" are matched, or even dominated, by those in the approximation error. Accordingly we cannot subscribe to the authors' view that "from now on we can safely ignore the non-linearity term in our empirical analysis" (III:26).

5. In parallel with formal statistical methods, there is a need for much more detail on basic data. This is vital because the main time series to be explained is (Table 1, Chapter 3)

Year	1964	1966	1968	1970	1971	1972
Gini	.32	.32	.33	.29	.29(5)	.29

The discontinuity between 1968 and 1970 is obvious. And this happens to correspond to a discontinuity in the data. It is inadequate for the authors to refer the reader to Kuo "Income Distribution by Size in Taiwan Area - Changes and Causes" for two reasons. One is that their main thesis is that 1968 was a turning point, so the reconciliation of inconsistent data for 1964 to 1968 and 1970 to 1972 is crucial and should be reported at length as part of this volume. A second objection is that the Kuo paper referred to is wholly inadequate on the subject -- there is just one paragraph which reads:

" (1) Adjustment of DGBAS Data to include Taipei City

Since the coverage of DGBAS data for the period 1964-1968 is different from the coverage of the period 1970-1972, i.e., the former including Taipei City data and the latter excluding it, a necessary adjustment is required in this regard. The DGBAS data of 1970, 1971 and 1972 without Taipei City are adjusted to include Taipei City firstly. "

The importance of this issue is highlighted by the fact that "F.4B: the significant improvement of G (after 1968) is contributed mostly by the now favourable non-agricultural Gini effect ..." Clearly, then, documentation on the data is needed to dispel the possibility that nothing more interesting has been happening in Taiwan than a change in data coverage.

6. All the above points could be taken on board by a revised draft: meanwhile, we think they are serious criticisms of the present manuscript.

7. Further limitations of the study derive from the fact that the data (except for 1966) have not been computerized. This makes "reruns" a virtual impossibility and has, perhaps, led to the use of numerical methods which compound the problems with approximation methods of decomposition. Thus in Chapter 6 there is a decomposition of income after tax into consumption and savings. The method used is exact. (It is the Rao result.) It gives

$$G = \phi_c R_c G_c + \phi_s R_s G_s$$

The ϕ 's are correctly calculated ($\phi_c = \text{apc}$; $\phi_s = \text{aps}$). However, the data are handled in such a way that R_c and R_s are not the correlation coefficients between consumption or savings and income; and the G_c and G_s are not the Gini coefficients for consumption and savings either. This is obvious from the fact that R_s is reported as 1.000 for 1972 -- an amazing result if it were true. A little further reflection indicates that this must be curious: some individuals have negative savings so it is logically impossible to calculate a Gini coefficient for savings.

8. So the R's are not correlation coefficients using individual data; and the G's are not Gini coefficients.

9. This conundrum seems to be explained by the fact that all the analysis is based on data grouped by household income level. These groupings never change, i.e. there is no regrouping according to consumption level in order to calculate G_c . Rather, G_c is obtained from the Lorenz curve for consumption built up by ordering household groups according to average consumption level when the groups themselves are defined by income level (percentile groups). This seems to be the implication of the new appendix which John Fei has sent in response to an inquiry (see attached). It implies that every Gini coefficient in the volume, except the Gini coefficients for total family income, are underestimates of true Ginis since the wrong grouping criterion is being used. At no point is this acknowledged in the main text which talks throughout as if

117

Gini coefficients were what they purport to be. This can be extremely misleading. Three examples may suffice. (i) The tax system is shown to have no influence on redistributing income in 1973 (see Chapter 6, Table 10). Yet if income before tax is the grouping criterion then the Gini coefficient for income after tax must be understated so that the tax system must be regressive, i.e., the regressive indirect taxes more than offset the progressive direct taxes. (ii) A second example (Chapter III, Table 1) shows the Gini coefficient for agricultural income declining 1964 to 1973 from 0.35 to 0.11. This is spurious -- as the proportion of the population in agriculture declines, so the inequality in income from agriculture across all households must rise (and if half the population have no agricultural income then the Gini must be at least 0.5). The decline from 0.35 to 0.11 therefore has little to do with the Gini coefficient for agricultural incomes. What it reflects is that the grouping criterion (total family income) is increasingly irrelevant as a proxy for income from agriculture. (iii) As a third example, Finding F.3b reports "virtually no deterioration [in rural income inequality] for 1968-72." Yet inspection of Table 1 (III.20) reveals that G_y^r rose from 0.2842 to 0.3477 over that period, a deterioration of more than 20 percent. We are therefore left wondering whether the authors have misinterpreted their data or we have misinterpreted what they mean by a Gini coefficient.

10. We do not exclude the possibility of having misunderstood what is going on. The explanations we have are by no means clear on the precise methods used or the interpretations which should be given to particular statistics. But if as a result of earlier exchanges with the authors we now have the correct diagnosis, then it follows that much of the text and argument is seriously misleading. The three examples given suffice to make the point that the findings reported are simply not well founded.

11. If we are right, what can be done? One solution is to spell out what is happening within the data and explain to the reader that all Gini's (other than for total income) are not Gini's and so should not be read as such. This would be very messy. A second solution would be to cut out all Gini's other than G_y and try to restate the history of Taiwan with the data that remains. This may be possible and could result in a nice (albeit slim) volume. It would, of course, require a major effort.

12. Perhaps enough has been said to express concern and disquiet. However, we must draw attention to two further problem areas.

13. First, Shirley Kuo's disaggregation in Chapter 5 is wrong. This (and much of the data) comes from her paper referred to above. The error is pointed out in the appended correspondence. It can be seen most clearly by noting that if a population is divided into groups so that there is only one person in each group, then the Kuo formula gives a Gini of zero irrespective of what the true Gini is. Her formula works well when the criteria for disaggregation result in groups which in fact have similar income distributions, i.e., the criteria don't discriminate. This is in fact the case throughout

much of the chapter. Its argument is therefore that the criteria she looks at don't matter much. This is interesting. But the Gini disaggregation is unnecessary to making the point.

14. Finally, the 1966 data was computerized and some regressions emerge. The earnings function fitted takes little cognisance of the vast literature on such functions and is badly specified. One in particular is the inclusion of family income as an explanatory variable for individual wage income as a potential test of nepotism. We just can't let this sort of thing pass. Moreover, to go from regression to Gini disaggregation of predicted values is somewhat gratuitous -- there are well established techniques within the regression framework for analyzing how much each variable contributes to total explanation. (And the link between Gini's and regression is set out by Rao.)

15. So we are left at the end of the day with a number of mixed feelings. Essentially we are very disappointed, perhaps because of wanting to see some glimpses of solutions to the exciting challenge that this project took on. And perhaps they are there. But in general we find ourselves in a pervasive fog. A lot of work has gone into this study, so we hesitate to write it off and have enough respect for the authors to retain some unease that we may be seeing difficulties where none exists. But we must conclude by saying that for the most part we have got nothing but misgivings out of reading the manuscript.

16. Graham Pyatt's annotated copy of the draft is available for return to the authors.

cc: Messrs. B.B. King
Duloy

Mr. B.B. King
1) ~~10-84~~
2) 0-84

Mr. M. Ahluwalia, DRCID

November 22, 1976

Graham Pyatt, DRCDR

'Equity with Growth: The Taiwan Case' by Fei, Ranis and Kuo (Draft)

1. The authors of this volume are refreshingly frank in expressing their disappointment that a theory of growth and distribution has not emerged from their analysis. It remains the case, however, that it was worth a try, while the interest of their report would be enhanced if their own reflections on why such a theory is elusive were set out.
2. My personal views on this question fall under two heads. First, the approach adopted concentrates on what has happened to inequality (measured by the Gini coefficient). This is partial and I do not see how a general theory can be developed without simultaneous consideration of growth. Yet the volume says relatively little about this and the disaggregation of growth as opposed to inequality is seriously underplayed. Moreover, the methodology with respect to inequality relies heavily on statistical decompositions without an explicit attempt to drive ex post observations back to structural and policy foundations. Accordingly, there is limited scope for a general theory of growth and redistribution to emerge.
3. Within these limitations (which I see more clearly with hindsight), the authors have been working under further handicaps which are, in part at least, self-imposed. Thus, with respect to statistical decomposition, they have chosen to work with Gini rather than Theil. This creates difficulties which are well known. Also, within the Gini framework, the formal problems of statistical analysis have been solved by Mangahas, Mehran, Rao and my own work based on Rao and Bhattacharya/Mahalanobis. None of these contributions appears to have been deployed so that instead the statistical methods adopted are rife with approximations and perversions of results available in the literature. It can be argued, of course, that this doesn't matter insofar as the approximations are close. However, there is the obvious point that the literature should be taken on board by an academic study so that at least a new chapter/appendix is needed to set out the literature and relate the methods used to it. If this were done (much of my own time has been taken up trying to disentangle the approximations) then some other points would emerge. One is that the approximate methods adopted are not invariably close and can involve offsetting errors. Hence the decomposition can be more misleading than the aggregate error suggests, both at a point of time and over time.
4. In parallel with formal statistical methods, there is a need for much more detail on basic data. This is vital because the main time series to be explained is (Table 1, Chapter 3)

Year	1964	1966	1968	1970	1971	1972
Gini	.32	.32	.33	.29	.29(5)	.29

The discontinuity between 1968 and 1970 is obvious. And this happens to correspond to a discontinuity in the data. It is inadequate for the authors

to refer the reader to Kuo "Income Distribution by Size in Taiwan Area - Changes and Causes" for two reasons. One is that their main thesis is that 1968 was a turning point, so the reconciliation of inconsistent data for 1964 to 1968 and 1970 to 1972 is crucial and should be reported at length as part of this volume. A second objection is that the Kuo paper referred to is wholly inadequate on the subject -- there is just one para which reads:

" (1) Adjustment of DGBAS Data to include Taipei City

Since the coverage of DGBAS data for the period 1964-1968 is different from the coverage of the period 1970-1972, i.e., the former including Taipei City data and the latter excluding it, a necessary adjustment is required in this regard. The DGBAS data of 1970, 1971 and 1972 without Taipei City are adjusted to include Taipei City firstly. "

The importance of this issue is highlighted by the fact that "F.4B: the significant improvement of G (after 1968) is contributed mostly by the now favourable non-agricultural Gini effect ..." Clearly, then, documentation on the data is needed to dispel the possibility that nothing more interesting has been happening in Taiwan than a change in data coverage.

5. All the above points could be taken on board by a revised draft: meanwhile, I think they are serious criticisms of the present manuscript.

6. Further limitations of the study derive from the fact that the data (except for 1966) have not been computerized. This makes "reruns" a virtual impossibility and has, I think, led to the use of numerical methods which compound the problems with approximation methods of decomposition. Thus in Chapter 6 there is a decomposition of income after tax into consumption and savings. The method used is exact. (It is the Rao result.) It gives

$$G = \phi_c R_c G_c + \phi_s R_s G_s$$

The ϕ 's are correctly calculated ($\phi_c = \text{apc}$; $\phi_s = \text{aps}$). However, the data are handled in such a way that R_c and R_s are not the correlation coefficients between consumption or savings and income; and the G_c and G_s are not the Gini coefficients for consumption and savings either. This is obvious from the fact that R_s is reported as 1.000 for 1972 -- an amazing result if it were true. A little further reflection indicates that this must be curious: some individuals have negative savings so it is logically impossible to calculate a Gini coefficient for savings.

7. So the R's are not correlation coefficients using individual data; and the G's are not Gini coefficients.

8. This conundrum seems to be explained by the fact that all the analysis is based on data grouped by household income level. These groupings never

change, i.e. there is no regrouping according to consumption level in order to calculate G_c . Rather, G_c is obtained from the Lorenz curve for consumption built up by ordering household groups according to average consumption level when the groups themselves are defined by income level (percentile groups, I think). This seems to be the implication of the new appendix which John Fei has sent to me in response to an inquiry (see attached). It implies that every Gini coefficient in the volume, except the Gini coefficients for total family income, are underestimates of true Ginis since the wrong grouping criterion is being used. At no point is this acknowledged in the main text which talks throughout as if Gini coefficients were what they purport to be. This can be extremely misleading. Two examples may suffice. The tax system is shown to have no influence on redistributing income in 1973 (see Chapter 6, Table 10). Yet if income before tax is the grouping criterion then the Gini coefficient for income after tax must be understated so that the tax system must be regressive, i.e., the regressive indirect taxes more than offset the progressive direct taxes. A second example (Chapter III, Table 1) shows the Gini coefficient for agricultural income declining 1964 to 1973 from 0.35 to 0.11. This is spurious -- as the proportion of the population in agriculture declines, so the inequality in income from agriculture across all households must rise (and if half the population have no agricultural income then the Gini must be at least 0.5). The decline from 0.35 to 0.11 therefore has little to do with the Gini coefficient for agricultural incomes. What it reflects is that the grouping criterion (total family income) is increasingly irrelevant as a proxy for income from agriculture.

9. I do not exclude the possibility that I have misunderstood what is going on. But if as a result of earlier exchanges with the authors I now have the correct diagnosis, then it follows that much of the text and argument is seriously misleading. The two examples given suffice to make the point that the findings reported are simply not well founded.

10. If I am right, what can be done? One solution is to spell out what is happening within the data and explain to the reader that all Gini's (other than for total income) are not Gini's and so should not be read as such. This would be very messy. A second solution would be to cut out all Gini's other than G_y and try to restate the history of Taiwan with the data that remains. This may be possible and could result in a nice (albeit slim) volume. It would, of course, require a major effort.

11. Perhaps I have said enough to express my concern and disquiet. However, I must draw attention to two further hangups I have.

12. First, Shirley Kuo's disaggregation in Chapter 5 is wrong. This (and much of the data) comes from her paper referred to above. The error is pointed out in the appended correspondence. It can be seen most clearly by noting that if a population is divided into groups so that there is only one person in each group, then the Kuo formula gives a Gini of zero irrespective of what the true Gini is. Her formula works well when the criteria for disaggregation result in groups which in fact have similar income distributions,

i.e., the criteria don't discriminate. This is in fact the case throughout much of the chapter. Its argument is therefore that the criteria she looks at don't matter much. This is interesting. But the Gini disaggregation is unnecessary to making the point.

13. Finally, the 1966 data was computerized and some regressions emerge. The earnings function fitted takes little cognisance of the vast literature on such functions and is badly specified. One in particular is the inclusion of family income as an explanatory variable for individual wage income as a potential test of nepotism. We just can't let this sort of thing pass. Moreover, to go from regression to Gini disaggregation of predicted values is somewhat gratuitous -- there are well established techniques within the regression framework for analyzing how much each variable contributes to total explanation. (And the link between Gini's and regression is set out by Rao.)

14. So I am left at the end of the day (actually at the end of a weekend) with a number of mixed feelings. I frankly find it impossible to be more constructive in my comments than I have been here. Essentially I am very disappointed, perhaps because I have wanted to see some glimpses of solutions to the exciting challenge that this project took on. And perhaps they are there. But I personally find myself in a pervasive fog. A lot of work has gone into this study, so I hesitate to write it off and have enough respect for the authors to retain some unease that I may be seeing difficulties where none exists. But I must conclude by saying that for the most part I have got nothing but misgivings out of reading the manuscript.

15. My annotated copy of the draft is available for return to the authors.

Attachments

cc: Messrs. B.B. King ✓
Duloy
Bell

November 18, 1976

Dr. Shirley W.Y. Kuo
Economic Planning Council
Executive Yuan
118, Hwai-ning Street
Taipei, Taiwan
Republic of China

Dear Dr. Kuo:

Many thanks for your letter of October 2 and the enclosed copy of your paper on income distribution in Taiwan. In reply I enclose a copy of my recent paper on Gini coefficient disaggregation on which your comments would be welcome. Please note that I am sending a copy of this letter to John Fei and no doubt you will wish to discuss its contents with him.

I have been unable to resurrect your approximate Gini disaggregation formula (3) (p.109) from my own (exact) result and think that the trouble lies with your formula (1) (p.108). This states that

$$f_w(x) = h_a \left(\frac{y_w}{y_a} \right) f_a(x) + h_n \left(\frac{y_w}{y_n} \right) f_n(x) \quad (1)$$

where w indicates whole population
 a " agricultural households
 n " non-agricultural households
 h_a, h_n are population proportions such that

$$h_a + h_n = 1$$

and y_a, y_n and y_w are mean incomes such that

$$y_w = h_a y_a + h_n y_n$$

This leaves f_a, f_n and f_w to be defined as density functions of the variable x where x is income normalized by mean income for the (sub) population in question.

Consider drawing an income at random from the whole population. The probability that this is less than xy_w , i.e., x times the average y_w , can be denoted $\Pr(y < xy_w)$ and your first equation at the top of p.108 can be written

$$\Pr (y < xy_w) = \Pr (y < xy_w/y \text{ is an agricultural income}). h_a \\ + \Pr (y < xy_w/y \text{ is a non-ag. income}). h_n$$

This can be written equivalently as

$$\Pr \left(\frac{y}{y_w} < x \right) = \Pr \left(\frac{y}{y_a} < x \frac{y_w}{y_a} / y \text{ is an ag. income} \right). h_a \\ + \Pr \left(\frac{y}{y_n} < x \frac{y_w}{y_n} / y \text{ is non-agricultural} \right). h_n$$

Now differentiating both sides with respect to x gives

$$f_w(x) = h_a \frac{y_w}{y_a} f_a \left(x \frac{y_w}{y_a} \right) + h_n \frac{y_w}{y_n} f_n \left(x \frac{y_w}{y_n} \right) \quad (1)'$$

It is apparent that this is not the same as your equation (1), but the two will be close if $y_a \doteq y_w \doteq y_n$. This condition holds for the data you present in Table 13 but can't be relied on in general. If you combine it with the other assumption on which your decomposition is based, viz. $f_a(x) \doteq f_n(x)$, then the total effect is to assume that income distributions are similar in the two sub-sectors.

The usefulness of your decomposition formula (2) (p.109) is obviously limited if it does not take full account of differences in means between sub-groups. I would therefore very much welcome your comment on the above and, meanwhile, will look further into the question of more robust approximations based on $f_a(x) \doteq f_n(x)$. Hopefully you will be able to spare me this effort by pointing out some error in my thinking. But the fact is that I think (1)' is right and the double use of (1) in the 'proof' of (2) is relatively serious.

Yours sincerely,

Graham Pyatt
Senior Adviser
Development Research Center

cc: Dr. John Fei
Dr. G. Ranis, Yale
Mr. M. Ahluwalia, DRC



UNIVERSIDAD DE LOS ANDES
FACULTAD DE ECONOMIA
APARTADO AEREO 4976
BOGOTA - COLOMBIA

cc: Montek

October 18, 1976

Dear Jack:

It was good to see you and Montek last week.

As I told you orally, I shall be ~~xx~~ glad to attend the April conference. I do, however, have a question which did not occur to me ~~xxxx~~ during our meeting, i.e. are you planning to invite John as well, or, given the shortage of accommodations at Bellagio, only one partner in each team. If the latter should be the case, once we know ~~xxxxxx~~ more precisely what you intend to focus on in the more detailed agenda, we could decide perhaps which one of us would most beneficially represent us. Please let me know what the situation is so that, if need be, I can get in touch with John.

I am looking forward to receiving your comments on our manuscript before the end of November. If it is not too much trouble could ~~x~~ you simultaneously send a copy to John in Taipei. Thanks, and best regards.

Sincerely yours,

Sur

ECONOMIC PLANNING COUNCIL
EXECUTIVE YUAN

Oct. 4, 1976

Dr. Graham Pyatt
The World Bank
1818, H Street
N.W., Washington, D.C., 20433
U.S.A.

Dear Graham:

Ms. Shirley Kuo gave me her copy of our manuscript on Income Distribution, I am now ready to answer the question which you raised on the first page of your letter of Sept. 20. The intricate nature of the issue can be best analyzed with aid of numerical example which is attached.

In respect to the question you raised on the second ^{page} paper, I have given your comment to Shirley Kuo, who will no doubt write you about it. Shirley is the main author of Chapter 5 of our manuscript.

I am very happy that you can give our manuscript a very careful reading. In order to make the point clearer to the future readers, the few pages attached should be included as a separate appendix of Chapter 3.

Best regards,

Sincerely yours

John C.H. Fei
John C.H. Fei

cc : Messrs Dulay
Ahluwalia

ECONOMIC PLANNING COUNCIL
EXECUTIVE YUAN

Appendix: Impact on Factor Gini Coefficient due to the Grouping Data

For the computation of Gini Coefficient and Factor Gini Coefficients of this Chapter, we made use of statistical data which is published in a "grouped form" according to the stratification of families into different income classes. This may be illustrated by Table A1. In this Table there are 4 "income classes" with the average income shown in col(1), while the number of families shown in col(2). Their products (ie. total class income) are shown in col(3), while col(4) and (5) show total wage income and total agricultural income for each income class respectively. (col(3) is the sum of col(4) and (5)) This is the way in which the original data are published.

Table A1: Original Source Data

Income class	Average income (1)	No. of families (2)	Total class income (3)	Total wage income (4)	Total agricultural income (5)
1	4	1	4	4	0
2	5	3	15	12	3
3	12	3	36	33	3
4	15	3	45	30	15
		10	100	79	21

In computing the Gini Coefficient, this Table is first processed into Table A2 showing the income pattern for a typical member within each income class. Col(3) of Table A2 is the same as col(1) of Table A1, while col(4) and (5) of this Table are obtained from col(4) and (5) of Table A1, divided by the number of families within the each class. Thus for the 10 families in this Table, the total family, wage income and agricultural income patterns are represented by the following vectors.

Table A2: Typical Earning Pattern

Income class (1)	No. of families (2)	Average family income (3)	Average wage income (4)	Average agricultural income (5)
1	1	4	4	0
2	3	5	4	1
3	3	12	11	1
4	3	15	10	5

ECONOMIC PLANNING COUNCIL
EXECUTIVE YUAN

$$\begin{aligned}
 1.a) \ Y &= (4, 5, 5, 5, 12, 12, 12, 15, 15, 15) & G_y &= .24 = \bar{G}_y \\
 b) \ W &= (4, 4, 4, 4, 11, 11, 11, 10, 10, 10) & \bar{G}_w &= .19 \\
 c) \ A &= (0, 1, 1, 1, 1, 1, 1, 5, 5, 5) & \bar{G}_A &= .44 = G_A
 \end{aligned}$$

With the indicated Pseudo Gini Coefficient \bar{G}_y , \bar{G}_w , \bar{G}_A , The underlying assumption when Table A1 is processed into Table A2, is that every person belonging the same income class has exactly the same income pattern. Whenever one makes use of data published in the form of Table A1, this assumption must obviously be made.

In order to compute the Factor Gini Coefficient, notice that the wage vector in 1.b) is not arranged in a monotonic order. When it is rearranged in a monotonic order, it becomes:

$$2) \ (4, 4, 4, 4, 10, 10, 10, 11, 11, 11) \ G_w = .21$$

with the indicated wage Gini Coefficient. While the Pseudo Factor Gini Coefficients satisfy the condition

$$3) \ G_y = \phi_w \bar{G}_w + \phi_A \bar{G}_A = (.79) (.19) + (.21) (.44) = .24$$

exactly, the estimated value of Gini Coefficient is:

$$4) \ \hat{G}_y = \phi_w G_w + \phi_A G_A = (.79) (.21) + (.21) (.44) = .26$$

Thus there is a nonlinear error:

$$5) \ \theta = \hat{G}_y - G_y = .26 - .24 = .02$$

There is another type error due to the fact that the published data (i.e. Table A1) is different from the true data. Since, the original data are obtained from sample questionnaires in household survey, the returns appear in the form of Table A3

Table A3: True Income Patterns

Family No.	True Income Pattern		
	Total income	Wage income	Agricultural income
1	4	4	0
2	5	5	0
3	5	5	0
4	5	2	3
5	12	12	0
6	12	12	0
7	12	9	3
8	15	15	0
9	15	15	0
10	15	0	15
Total	100	79	21
Share	1.00	$\phi_w = .79$	$\phi_A = .21$

in which the true income patterns of the 10 families are shown by the 10 rows of this table. Thus the true total family income, wage income, agricultural income patterns are:

$$\begin{aligned}
 6.a) Y^T &= (4, 5, 5, 5, 12, 12, 12, 15, 15, 15) & G_y^T &= .24 \\
 b) W^T &= (4, 5, 5, 2, 12, 12, 9, 15, 15, 0) & G_w^T &= .37 \\
 c) A^T &= (0, 0, 0, 3, 0, 0, 3, 0, 0, 15) & G_A^T &= .81
 \end{aligned}$$

Notice that when the questionnaire returns ie. Table A3 was processed into Table A1 and then processed into Table A2, there is a fouling of data because of the grouping error. This error can be seen by comparing

ECONOMIC PLANNING COUNCIL
EXECUTIVE YUAN

the true patterns in 6.a.b.c) and the fowl patterns in 1.a.b.c). In other words, while in this simple case there is no fouling of the total income pattern. " $Y^T = Y$ " there is a fouling of both wage income pattern " $W^T \neq W$ " and agricultural income pattern " $A^T \neq A$ ". Generally, there is underestimation of the true degree of inequality for each factor component:

$$\begin{aligned} 7.a) \quad G_W^T &> G_W && (.37 > .21) \\ b) \quad G_A^T &> G_A && (.81 > .44) \end{aligned}$$

a stated in footnote 1 of section 3. Such a fouling of data is ordinarily encountered whenever the raw data are grouped into frequency distribution Table. However in the caculation of Factor Gini Coefficient such as a fouling can be a serious one. For example in the true income pattern of Table A3 more than 50% of the families receive no agricultural income at all. Hence, the true Gini Coefficient is greater than .5 (ie. $G_A^T = .81$). However the agricultural Gini based on the grouped data is much smaller (ie $G_A = .44$). This error is strictly due to the consolidation of data, and has nothing to do with the nonlinear error .

As long as one makes use of published data in the form of Table A1, the consolidation error can not be avoided. The only way they can be avoided is through the use of the original questionnaire. Thus in this Chapter, we make use of the published data, while in Chapter 4 and 6 we make use of the data from the original questionnaire (see appendix of chapter 4). Thus the factor Gini Coefficient of this Chapter differ (an are generally smaller than) the Factor Gini Coefficient of Chapter 4 and 6, because of the consolidation error.

ECONOMIC PLANNING COUNCIL
EXECUTIVE YUAN

118, HWAI-NING STREET, TAIPEI, TAIWAN
REPUBLIC OF CHINA

REFERENCE NO.

CABLE ADDRESS: EPCEY TAIPEI
TELEX: 11385 TAIPEI

October 2, 1976

Dr. Graham Pyatt
Senior Advisor
Development Research Center
The World Bank
1818 H Street N. W.
Washington D.C. 20433
U. S. A.

Dear Dr. Pyatt:

Professor John Fei who is currently with us at the Economic Planning Council, referred your letter of September 20 to him to me.

On the second page of your letter, you raised the question concerning the decomposition formula of chapter 5 in our manuscript. Since chapter 5 is essentially my work, I made use of a decomposition equation which I developed some two years ago. (See pages 105-109 in the paper I attached). I am happy to know that we have been working along parallel lines.

The issue which you mentioned in your letter to Dr. Fei certainly deserves further investigation. Will you please send me a reprint of your work on this issue. In case I make any further progress in studying this issue, I will certainly write to you immediately.

Thank you for reading our manuscript.

With best wishes,

Sincerely,

Shirley W. Y. Kuo
Shirley W. Y. Kuo

September 20, 1976

Professor John Fei
 Economic Growth Center
 Yale University
 New Haven, Conn. 06520

Dear John:

I am reading your study with Gus and Shirley Kuo on "The Taiwan Case" and so far have a number of comments on the early chapters. However, I have got stuck at Table 1 in Chapter 3 and must ask for some clarification at this point in order to proceed.

My problem concerns the way in which the statistics G_w , G_π and G_A have been calculated. Footnote 1, p. 3.19, states that the decile group with respect to total income was determined for each household. Does this mean that G_w , say, is obtained by ranking households (to the nearest decile) according to y and then obtaining G_w from cumulative wages versus cumulative numbers of households? My reading of the figures is that this must have been the procedure. If it is correct, then I think that there are a number of points which follow:

- (i) G_w , G_π and G_A are not Gini coefficients but rather concentration coefficients in Rao's terminology, or pseudo-Gini's (\bar{G} 's) in yours.
- (ii) If they are pseudo-Gini's, then the disaggregation

$$G = \sum \phi_i \bar{G}_i$$

- is exact. Thus the errors θ in Table 1 are due to statistical problems, e.g. the fact that $\sum \phi_i \neq 1$ in your data, rather than to 'Gini error' as you imply.

If my conjecture above is wrong, then you must have ranked households separately by (deciles of) w , π and A . But if this is so, then the numbers seem to be rather odd. Thus, if 50 percent of the population have no agricultural income, then G_A cannot be less than 0.5. Yet for 1972 you have 0.11. Similarly, G_A for all households cannot be less than G_A for rural households (since adding new population members with zero income can only increase a Gini coefficient). Yet the ordering is always the other way round in Table 1.

You will gather that I have a rather important problem which obviously has implications for your analysis. Hopefully you will be able to resolve the issue for me - perhaps because there is some third alternative which I have missed. In any event, an early reply would be most helpful since the above needs to be resolved before I can proceed with my assimilation of your analysis.

My reading beyond p. 3.20 has been more cursory so far. However, in Chapter 5, I have a difficulty which I can take this opportunity of also raising with you. It concerns equation (1). I follow the appendix proof down to mid p.53, the result there being a version of my own Gini disaggregation formula. However, you then make the simplifying assumption $f_1(x) = f_2(x)$. I read this as implying that the distribution of Y in sector 1 differs from that in sector 2 by a scalar (due to mean differences between sectors). However, this would imply $G_1 = G_2$. So if in fact $G_1 \neq G_2$, then formula (7) in the appendix can only be an approximation. Is this correct? If so, it might be interesting to explore the nature of the approximation by reference to my exact decomposition.

I am sorry to be bothering you with all this, but anticipate that you will share my concern to have a clear understanding of these issues. Since Gus is abroad, I understand, I am not copying this letter to him. No doubt you will be keeping in touch with him meanwhile and will let me know if I should write him direct.

With best wishes,

Yours sincerely,



Graham Pyatt
Senior Adviser
Development Research Center

670-84
 Aug 24, 1976

Summary: The Fei/Ranis/Kuo Report -- Equity with Growth: The Taiwan Case

The purpose of this report is to analyze the interaction of income distribution and rapid economic development in Taiwan. The standard view that there necessarily exists a simple trade-off between equity and growth is not confirmed by an empirical analysis.

METHODOLOGY

Decomposing the Gini coefficient: The first task of the research decomposes the Gini coefficient of total family income (G_Y) into three factor components measuring: wage (G_W), property (G_π), and transfer (G_N) income distribution. An estimated total Gini coefficient is defined as the weighted average of factor Gini coefficients:

$$(1) \quad G_Y = \phi_W G_W + \phi_\pi G_\pi - \phi_N G_N$$

where $\phi_i = \frac{\text{factor income across all families}}{\text{total income}}$

Transfer income does not apply for Taiwan and equation (1) becomes:

$$(2) \quad \begin{aligned} \text{a) } G_Y &= \phi_\pi G_\pi + \phi_W G_W \\ \text{b) } \phi_W + \phi_\pi &= 1 \end{aligned}$$

In order to differentiate between agricultural and nonagricultural activity, (2) can be modified as follows:

$$(3) \quad \begin{aligned} \text{a) } G_Y &= \phi_X G_X + \phi_A G_A \\ \text{b) } G_X &= \phi'_W G_W + \phi'_\pi G_\pi \end{aligned}$$

where G_A = Gini coefficient for agricultural income
 G_X = Gini coefficient for all nonagricultural income
 ϕ_A = share of agricultural income in total income
 ϕ_X = share of nonagricultural income in total income
 $\phi'_W = \phi_W / \phi_X$ = wage income share within nonagriculture
 $\phi'_\pi = \phi_\pi / \phi_X$ = property income share within nonagriculture
 so that $1 = \phi'_W + \phi'_\pi + \phi_A$
 $1 = \phi_W + \phi_\pi$

Equity and Growth: The next stage of the project decomposes the change through time in the total Gini coefficient as measured in (3) into three effects: the reallocation effect (R), which reflects the shift of the economy from an agricultural to nonagricultural concentration; the functional distribution effect (D), which describes the net effect on G_Y of changes in the relative shares of capital and labor; and the factor Gini effect (B), which traces the net effect on the

aggregate changes in the individual factor Ginis. These are obtained by differentiating (3) with respect to time:

$$(4) \quad a) \quad \frac{dG_Y}{dt} = R + D + B$$

$$b) \quad \text{Reallocation Effect: } R = (G_A - G_X) \frac{d\phi_A}{dt}$$

$$c) \quad \text{Functional Distribution Effect:} \\ D = (G_W - G_\pi) \left(\frac{d\phi'_W}{dt} \right) \phi_X$$

$$d) \quad \text{Factor Gini Effect:} \\ B = \left(\frac{dG_A}{dt} \right) \phi_A + \left(\frac{dG_W}{dt} \right) \phi_W + \left(\frac{dG_\pi}{dt} \right) \phi_\pi$$

RESULTS FOR TAIWAN

Based on the above decomposition, the following conclusions about income distribution in Taiwan were reached:

(i) The Time Pattern of G_Y :

For all households, G_Y increases slightly between 1964 and 1968 but declines consistently thereafter;
 For urban households, G_Y shows the same time pattern as all households but slightly more pronounced;
 For rural households, G_Y declines significantly between 1964 and 1968 but remains relatively constant thereafter.

The authors conclude from these results that the Kuznets effect -- things have to get worse before they get better (distribution-wise) -- is a complex phenomenon which is mainly relevant to the nonagricultural sector. Further, the more urban centered the growing nonagricultural activity, the more significant is the Kuznets effect. Thus where agricultural activity is important and industrialization is decentralized, as in Taiwan, there need not be a conflict between growth and equity, even before the so-called turning point (the exhaustion of surplus labor -- 1968 for Taiwan).

(ii) The Reallocation Effect:

For rural households, the reallocation of labor from agricultural activity to rural industries improved family income distribution (FID) equality throughout 1964-1972.
 For all households, the labor reallocation improved FID equity before the turning point (1968) and worsened equity thereafter.

(iii) The Functional Distribution Effect:

Before the turning point the strong labor using bias of technology change in the rural industries contributes to FID equality for rural families.

After the turning point, capital deepening and labor using bias both contribute to FID equality.

(iv) Factor Gini Effect:

The nonagricultural factor Gini effect is very unfavorable before 1968 and very favorable thereafter for all households; slightly favorable before and very unfavorable after 1968 for rural households; slightly favorable before and very favorable after 1968 for urban households.

The effect due to agricultural income (G_A) is always favorable.

The above results lead the authors to conclude that the most effective method of tackling the maldistribution of income is via a change in the nature of the growth path itself. Equity and growth are not inconsistent goals and governments need not rely on obvious tools such as transfers to achieve a more equitable distribution of incomes. The Kuznets effect can be overcome by other forces since it is felt only in the centralized (urban) nonagricultural sector.

COMMENTS AND FURTHER RESULTS

1. Perhaps the most interesting implication of the results described above is their applicability outside the Taiwanese context. The authors themselves agree that the experience of Taiwan is unique in that the country had a not too unfavorable income distribution at the start. The whole of Chapter II is devoted to a description of the country's "rare and remarkable" achievements by any LDC standards which are a function of its initial human resources endowment, its colonial experience, and heavy American support.
2. The reader must also be careful in equating wage income distribution equality with consumer welfare equality. First of all, family wage income equality deals with only the primary labor force and not with the poorest segments of the population. Secondly, public expenditures on health, education, etc. should be imputed to a welfare measure.
3. Chapter IV studies the causes of inequality underlying the distribution of family income. The impact of labor heterogeneity (by characteristic) on the wage rate is determined. A linear regression equation describing the wage rate is used that includes a proxy variable for "political pull" -- total family income. This variable may be highly collinear with the education variable. The authors conclude that the sex and age premiums are extremely high in the city.
The degree of inequality of wage income is decomposed into

the labor characteristic components in another part of the chapter. The education and age characteristics account for 2/3 of inequality in the wage rate.

4. The economic structure decomposition of Chapter V concludes that the fall of the internal inequality within the nonfarm sector was the most significant contribution to the change in overall income inequality.

5. Chapter VI on taxation concludes that the Taiwanese tax burden is highly regressive because of its heavy reliance on indirect taxes.

4670-84
August 19, 1976

Professor Gustav Ranis
Director, Economic Growth Center
Dept. of Economics
Yale University
P.O. Box 1987, Yale Station
New Haven, Conn. 06520

Dear Gus,

Many thanks for Equity With Growth: The Taiwan Case, which I have just received but have not yet read.

For your information, the Research Committee has recently instituted a formal evaluation process for research projects. To this end, it is necessary to write a Completion Report which goes to a Committee panel along with the output of the project itself. We will send you a set of comments on Equity With Growth within four weeks. You might want to revise your report in the light of these or, on the other hand, you may prefer to have it go forward for evaluation as it stands.

The above process should be distinguished from the more usual process of refereeing for publication. This occurs through our Editorial Committee and, of course, usually leads to some rewriting.

If I can receive your assurance that your team is willing to prepare the manuscript for publication at no additional cost to the Bank, I see no reason to hold up disbursing the remainder of the contract funds. Please let me know if this is OK.

Yours sincerely,

John H. Duloy, Director
Development Research Center

cc: Mr. H. B. Chenery
Mr. B. B. King
Mr. M. Ahluwalia
Mr. G. Pyatt
Ms. M. Gary
Ms. F. Stone

Growth, Employment, and Size Distribution of Income

Ref. No. 670-84

Extensive study of national data has given rise to the hypothesis that the size distribution of income appears to worsen as development proceeds. However, evidence from specific countries, such as the Republic of China (Taiwan), leads to some questioning of this hypothesis. Even if a negative historical relationship exists between growth rates of Gross National Product (GNP) and trends in the distribution of income, the issue still remains whether this relationship is inevitable or whether an effective development policy can reduce the conflict between growth and size distribution of income.

This study analyzes the relationship between economic growth, government policies, and income distribution in Taiwan. The underlying assumption is that in a mixed economy, the long-run trend in the distribution of income is determined by forces reflecting factor endowments, production conditions, and technology, which are affected by government intervention. The study attempts to isolate the factors responsible for the observed changes in income inequality in Taiwan from 1964-1972. The technique adopted for this purpose is a decomposition of the Gini coefficient into the contribution of factor income shares and the degree of concentration of factor incomes. Changes in the Gini coefficient are therefore attributable to changes in these components. The impact of government policy on the different sectors of the economy is examined in order to draw generalizations about the relationship between government policy and the processes of growth and distribution.

Responsibility: Development Research Center - Montek S. Ahluwalia.

The researchers are John C.H. Fei, Gustav Ranis, and Gary S. Fields of the Economic Growth Center, Yale University, in collaboration with Kuo Wan-Yong (Shirley) of the Economic Planning Council in Taiwan.

Completion date: August 1976.

Reports

Fei, John C.H., and Fields, Gary S. The Indexability of Ordinal Measures of Inequality. Center Discussion Paper No. 205. May 1974. New Haven, Conn.: Economic Growth Center, Yale University.

Fei, John C.H., and Ranis, Gustav. Income Inequality by Additive Factor Components. Center Discussion Paper No. 207. June 1974. New Haven, Conn.: Economic Growth Center, Yale University.

Fei, John C.H., Ranis, Gustav, and Kuo, Shirley. Growth and the Family Distribution of Income by Factor Components: The Case of Taiwan. Center Discussion Paper No. 223, March 1975. New Haven, Conn.: Economic Growth Center, Yale University.

July 30, 1976

Mr. Montek Ahluwalia
Chief, Income Distribution Division
International Bank for
Reconstruction and Development
1818 H Street, N. W.
Washington, D. C. 20433

Dear Montek:

I am today sending you under separate cover the draft volume entitled "Equity with Growth: The Taiwan Case" which has been completed under Contract No. RPO-284 with the World Bank.

Under the terms of the agreement of July 31, 1975, I hope you will now release the final installment of \$30,000 to Yale University.

We look forward to your comments in the preparation of the final draft.

With best regards,

Sincerely yours,

Gustav Ranis

GR:clh

cc: Benjamin B. King ✓
Joseph S. Warner
Monica Zucker

May 12, 1976

Mr. Gustav Ranis
Department of Economics
Yale University
Box 1987, Yale Station
New Haven, Connecticut

Dear Mr. Ranis:

Your letter of April 30, 1976 to Mr. King was referred to me for reply.

The time extension requested in your letter has been approved and we will issue an internal administrative extension of the contract through July 31, 1976. None of the other provisions of your contract letter dated July 31, 1975 are affected by this change.

Thank you for advising us of this change.

Sincerely,

Myrna L. Gary
Administrative Officer
Development Research Center

cc: Messrs. King ✓
Erimes
Ahluwalia
Miss Lenthe

MGary:mhg

Yale University *New Haven, Connecticut*

MAY 6

REC'D

File: 0-84

Orv.
I have already
spoken to Gus R.
and Jack D. who
this. Can you
ask N.L. who
should write him
(see 06 personal files)

DEPARTMENT OF ECONOMICS

Economic Growth Center

Box 1987, Yale Station

GUSTAV RANIS

Professor of Economics

April 30, 1976

Mr. Benjamin B. King
Research Advisor
International Bank for Reconstruction and Development
1818 H Street, N. W.
Washington, D. C. 20433

Dear Ben:

I am referring to Miss N. Lenthe's letter to us of July 31, 1975, concerning the agreement between Yale University and the Bank pursuant to which we have agreed to complete the work involved in the Taiwan Case Study. Said letter also states that it is expected that the completed monograph will be available by June 30, 1976 but that any slippages in completion time will not involve any additional cost to the Bank.

My purpose in writing today is to inform you that it now appears that the draft of the completed Taiwan manuscript will be ready by the end of July instead of the end of June. This modest anticipated slippage has been occasioned by unforeseen difficulties in logistics and communication among the authors, who are dispersed between Taipei and New Haven.

The second paragraph of Miss Lenthe's letter indicates that our present agreement will terminate on June 30, 1976 unless otherwise mutually agreed. I am therefore requesting at this time that the agreement be extended to July 31, 1976, with all other provisions of the July 31, 1975 letter left unaffected.

Thank you for your attention.

Sincerely yours,

Gustav Ranis

Gustav Ranis

Approved by:

Joseph S. Warner
Joseph S. Warner

Director,

Grant and Contract Administration
Yale University

GR:clh

cc: M. Ahluwalia
N. Lenthe
J. Warner

Faint header text at the top of the page.

TO: [Name]
FROM: [Name]
SUBJECT: [Subject]

Dear Sir,
I am writing to you regarding the matter of [Subject].
The information provided to me indicates that [Details].
I am sure you will find this information of interest.

I am sure you will find this information of interest.
I am sure you will find this information of interest.
I am sure you will find this information of interest.

I am sure you will find this information of interest.
I am sure you will find this information of interest.
I am sure you will find this information of interest.

Very truly yours,
[Signature]

INCOMING MAIL UNIT

1976 MAY - 5 PM 5:48

RECEIVED

Faint text at the bottom right of the page.

MAR 31 1976

Project 670 - 84

Supervisor M. Ahluwalia

Expenditure Estimates
((\$'000))

	<u>Present Estimate</u>	<u>New Estimate</u>
1. FY76/1	72.8	72.8
2. <u>FY77</u>	<u>0.0</u>	<u>-</u>
3. Subtotal FY76/77	72.8	72.8
4. <u>FY78 and later</u>	<u>0.0</u>	<u>-</u>
5. Total/1	72.8	72.8

/1 Present estimate for FY76 is FY76 authorization.
Present estimate for Total is total authorization.

Note: The important figure is the subtotal on line 3. Distribution between FY76 and FY77 is less important.

March 31, 1976

Comments:

M.S. Ahluwalia
J.H. Duloy
G. Pyatt

In reviewing your paper "Who Benefits from Economic Development - a Re-examination of Brazilian Growth in the 60s", we find we have some problems with the procedure used to make the 1970 distribution comparable with the 1960 distribution. We find that the results are materially affected by your particular choice of procedure, which suggests that some consideration be given to justify this choice. Our difficulties with your procedure are explained in Section I and the results of an alternative procedure reported in Section II. Some general comments on the data base used are offered in Section III.

I. The Fields Procedure

Your procedure as described in the footnote for Table 2 on page 7 is

- (a) to deflate the 1970 income classes by a general price deflator to convert the 1970 classes into equivalent income classes in 1960 N Cr\$, and
- (b) to recompute the population frequency in the deflated 1970 income intervals to obtain the frequency for income intervals corresponding to the 1960 distribution.

We have two problems with step (b) in this procedure.

- (i) You assume that the distribution of population percentages over each income interval of the 1970 distribution can be approximated by a linear interpolation. Thus you calculate that of the 31.7 percent of the population in the deflated income bracket 0-2.8 in 1970, approximately $(2.1 \div 2.8) \times 31.7 = 23.8$ were in the range 0-2.1, and the remaining 7.9 percent in the range 2.1 - 2.8. The assumption that the cumulative percentage distribution of population is piecewise linear in absolute incomes is clearly restrictive as you yourself note. It does not conform with prior expectations about the shape of the distribution function. The Pareto distribution for example would suggest that

that the relationship is log-linear. If this is so, your procedure overstates the number of people below the absolute cut off line.

(ii) You also assume that the share of income accruing to a subgroup within each deflated 1970 income range can also be obtained by linear interpolation. On page 11 in equation 4, you imply that the income share accruing to the population below an income level of 2.1 in 1970 as $(2.1 \div 2.8) \times 8.0 = 6.0$ where 8.0 is the percentage income share of the population in the income range 0-2.8. Such a result would be valid if all people in the income range 0-2.8 received the mean income of this group, but in that case the cumulative distribution of population percentages across absolute income is a step function and not the piecewise linear function assumed in (i) above. This inconsistency in your method of interpolating income shares and the method for interpolating population causes serious problems. If population is really distributed as described in (i) above then there are an equal number of people at each income level within an income bracket. In that case, the mean income of the population in the range 0-2.1 is $(2.1 \div 2) = 1.05$ and the mean income of population in the range 0-2.8 is 1.4. Given that 3/4 of the people in the income range 0-2.8 are in the first subgroup the ratio of income shares of this group to the total is $(3 \times 1.05) \div (4 \times 1.4) = .56$. Since we know that the population in the range 0-2.8 has an income share of 8 percent this gives an income share for the population in the range 0-2.1 of 4.5 percent and not 6.0 percent.

Alternatively, if you really believe that all people in an income bracket have the mean income, then the mean income of the 31.7 percent of the population who are in the range 0-2.8 in 1970 (and whose income share is 8%) is given by $(8 \div 31.7) \times 7.31 = 1.84$ (where 7.31 is the mean income of the whole population in 1970). In this case all the population in this income range are below the 2.1 poverty

cut off and together with 11.7 percent with zero incomes this gives a poverty group in 1970 constituting 43.4%. This result is quite different from that obtained in your exercise, where you first estimate the population below 2.1 and then estimate their mean incomes.

We suspect your method first overstates the percentage of the population in poverty but much more seriously it over-estimates the income share of the lower subgroup in a range. Since the poverty group is defined at an observed point for 1960 and an interpolated point for 1970 the comparison exaggerates income share in 1970.

II. Fitting a Lorenz Curve

The alternative to your method is to fit a Lorenz curve directly to the data on cumulative population and income percentages, and then use the fitted curve to estimate the relevant magnitudes. Applying one such procedure to your data, we found the results are substantially altered^{1/}. (We note in passing that the fitted curves have coefficients of determination of 0.9995 for 1960 and 0.998 for 1970.) Our results from this exercise can be summarised as follows.

(i) The percentage of the population in poverty is given by the point where the slope of the Lorenz curve is equal to the ratio of the poverty line to the mean income. This can be calculated directly from the Lorenz curve given the estimated parameter values. Taking the mean incomes of the two years in constant 1960 US\$ as 513 and 679 respectively, and using a poverty line of \$195 (which

^{1/} The procedure is taken from Kakwani and Podder and is described on pages xii-xiv of the Shail Jain monograph. Note that the Jain monograph reports the income shares for each decile using your data sources in Col. 3 on page 15 (1960) and Col. 1 on page 16 (1970).

corresponds to 2.1 1960 N Cr\$'000s), we find that the percentage population in poverty drops from 38 percent in 1960 to 34 percent in 1970.^{1/} This shows a slightly larger proportional decline than your estimate of a drop from 37 percent to 35.5 percent.

(ii) More seriously, our estimates of the income shares of the poverty groups in the two years (and therefore their mean incomes) are quite different. We find that the income share of the poverty group was 5.75 percent in 1960 but dropped to 4.14 percent in 1970. This compares with your estimate of 5.2 percent in 1960 and 6 percent in 1970.^{2/} These differences lead to radically different estimates of the growth of income of the poor. Our estimate implies that the mean income of the poor rose from 77.63 in 1960 to 82.67 in 1970 - an increase of 6.5 percent compared to an increase in mean income of 32 percent. Against this, your estimate implies an increase in mean income of the poor of 62 percent!

Our general results on the evaluation of Brazilian experience are summarised in Table 1, which reports cumulative income shares by cumulative population percentages, the mean income for each cumulative population percentage, and the 1960-1970 growth in mean incomes. The basic pattern that emerges is that there has been a substantial squeezing of the middle in favour of the rich. The lowest 10 percent have zero incomes but the next decile appear to have substantially improved their income share so that the lowest 20 percent appear to have experienced an accelerated growth of income. However, as shown

^{1/} Note that the actual percentage below 2.1 in 1960 is 37 percent whereas the fitted Lorenz curve gives 38 percent. This gives an indication of the error in the fitting procedure.

^{2/} The 6.0 percent estimate is obtained in your paper by using the linear interpolation $(2.1 \div 2.8) \times 8.0 = 6.0$, where 8.0 is the income share of the population in the range 0-2.8. Note that in our results (Table 1) a 6.0 percent income share is reached only at 40 percent of the population. This reflects the extent of overestimation of income shares in your method.

in Shail Jain, from the third decile onwards, each decile except the top experienced a squeezing of income shares. As a result, if we take the lowest 30 percent, their growth of income was more or less the same as of the population as a whole and from the fourth decile onwards, there is a substantial squeezing of cumulative income shares until we get to the top 10 percent (see Col. 2 of Table 1). As shown in Table 1 the growth of mean income for cumulated population percentages remains less than the mean right up to the lowest 90 percent.

We should emphasise that we do not make any claim that the Kakwani-Podder procedure used above is necessarily optimal.^{1/} For one thing, the use of ordinary least squares to fit the log form of the equation for the Lorenz curve (see page xiii) of Shail Jain) raises obvious questions about the likely behaviour of the error terms and the suitability of the OLS procedure in this context. However, the fits obtained are quite good, and we suspect that it is probably as good a procedure as any from this point of view. The point we would emphasise is that the particular procedure used must be chosen with care.

In the light of the difficulties pointed out in (ii) above, your procedure should be re-examined.

III. Other Aspects of the Data Set

Finally, there are two aspects of the Brazil data that limit its usefulness for the purpose of studying poverty.

(i) The data relate to earnings of the economically active labour force including individuals with zero income. From the point of view of poverty measurement we should be looking at the distribution of households or indi-

^{1/} Indeed we are quite suspicious of the estimated increase of 82 percent for the mean incomes of the lowest 20 percent in Table 1. This reflects an increase in the estimated income share from 0.8 to 1.1. Note that the absolute difference is quite small but it makes an enormous proportional difference at the lower end.

viduals by per capita household income (or consumption). This is the preferred indicator we are using in our country specific studies. Our experiments with available data suggest that low income earners provide at best a poor overlap with the poverty group.

(ii) There is some evidence that the income position of socio-economic groups may be more volatile than percentile groups and that offsetting changes in socio-economic groups may yield much more stable distribution by percentile groups. Do the Brazil data permit any investigation of this phenomenon? For example, could it be the case that the Northeast has suffered a marked relative decline, while other socio-economic groups at the lower end of the distribution have not, so that the overall income growth position of the lowest 35 percent reflects an average of these two disparate trends? It would be interesting to explore this possibility if you have access to the relevant data.

Table 1

	Cumulative Income Shares		Mean Income (1960 US\$)		Percentage Growth of Income
	1960	1970	1960	1970	1960-70
Lowest Decile	0	0	0	0	
Lowest 20 Percent	.8	1.1	20.52	37.35	82.0
Lowest 30 Percent	3.1	3.1	53.01	70.16	32.4
Poverty Group ^{1/}	5.75	4.14	77.63	82.67	6.5
Lowest 40 Percent	6.5	6.0	83.36	101.85	22.2
Lowest 50 Percent	11.3	10.2	115.94	138.5	19.4
Lowest 60 Percent	17.6	15.8	150.48	178.8	18.8
Lowest 70 Percent	26.1	23.3	191.28	226.01	18.2
Lowest 80 Percent	37.8	33.8	242.39	286.88	18.4
Lowest 90 Percent	55.0	49.7	313.50	374.96	19.6
Total Population	100	100	513.0	679.0	32.4

^{1/} The poverty group (defined as population below US\$195) are estimated by us as 38 percent in 1960 and 34 percent in 1970. Mean incomes of each group in the table are calculated as $[\text{Income Share} \div \text{Population Share}] \times \text{Mean Income of Total Population}$. The zero values for the lowest decile arise because the data are for income earners and include those in the labour force with zero incomes (presumably unemployed and unpaid family helpers). The legitimacy of using the distribution of income earners is discussed in Section III.

Table 2 - Changes in FID and Decomposition of ΔG_y

	Total Income Gini	E F F E C T S			Decomposition of (Net) Factor Gini Effects				
		Reallo- cation	Function- al Distri- bution	(Net) Factor Gini	Agri.	Non- Agri.	Non-Agri.		
							Property	Wage	
G _y	R	D	B	G _A	G _x	G _π	G _w		
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)		
I. All Households									
(1) B.T.P.	(A) Actual Change in FID ¹								
	1. Average Annual Δ	+0.0019				-0.0374	+0.0128	+0.0025	+0.0151
	2. % Δ over Period	+2.4%				+41.2%	+16.6%	+2.2%	+25.1%
	(B) Decomposition of ΔG ¹								
	1. Estimated Avg. Ann. Δ ²	-0.0050	-0.0017	-0.0001	-0.0032	-0.0103	+0.0071	+0.0006	+0.0065
2. Distribution of Change ³	-100%	-34%	-2%	-64%	-322%	+222%	+8%	+92%	
	Index of Findings	F.3a	F.5a	F.8a	F.3a	F.12	F.11a		
(2) A.T.P.	(A) Actual Change in FID ¹								
	1. Average Annual Δ	-0.0095				-0.0029	-0.0012	-0.0094	-0.0039
	2. % Δ over Period	-11.4%				-42.9%	-12.4%	-8.0%	-11.8%
	(B) Decomposition of ΔG ¹								
	1. Estimated Avg. Ann. Δ ²	-0.0104	+0.0018	-0.0016	-0.0106	-0.0035	-0.0071	-0.0026	-0.0045
2. Distribution of Change ³	-100%	+17%	-15%	-102%	-33%	-67%	-37%	-63%	
	Index of Findings	F.3a	F.5a	F.8b	F.3a	F.12	F.11a		
II. Rural Households									
(1) B.T.P.	(A) Actual Change in FID ¹								
	1. Average Annual Δ	-0.0182				-0.0082	-0.0132	-0.00246	-0.0014
	2. % Δ over Period	-11.2%				-4.6%	-10.4%	-14.4%	-1.3%
	(B) Decomposition of ΔG ¹								
	1. Estimated Avg. Ann. Δ ²	-0.0171	-0.0070	-0.0019	-0.0092	-0.0054	-0.0028	-0.0025	-0.0003
2. Distribution of Change ³	-100%	-41%	-11%	-48%	-66%	-33%	-89%	-11%	
	Index of Findings	F.3c	F.5b	F.8a	F.3b	F.12	F.11b		
(2) A.T.P.	(A) Actual Change in FID ¹								
	1. Average Annual Δ	+0.0008				-0.0087	+0.0098	+0.0163	+0.0090
	2. % Δ over Period	+1.1%				-10.1%	+17.2%	+22.2%	+17.4%
	(B) Decomposition of ΔG ¹								
	1. Estimated Avg. Ann. Δ ²	-0.0034	-0.0030	-0.0003	-0.0001	-0.0046	+0.0045	+0.0016	+0.0029
2. Distribution of Change ³	-100%	-88%	-9%	-3%	-4600%	-4500%	+36%	+64%	
	Index of Findings	F.3c	F.5b	F.8b	F.3b	F.12	F.11b		
III. Urban Households									
(1) B.T.P.	(A) Actual Change in FID ¹								
	1. Average Annual Δ	+0.0042					+0.0022	+0.0032	-0.0010
	2. % Δ over Period	+2.5%					+1.3%	+1.5%	+0.7%
	(B) Decomposition of ΔG ¹								
	1. Estimated Avg. Ann. Δ ²	+0.0018		+0.0014	+0.0004		+0.0004	+0.0010	-0.0006
2. Distribution of Change ³	+100%		+78%	+22%		+100%	+250%	-150%	
	Index of Findings	F.3b		F.8a	F.11c		F.11c		
(2) A.T.P.	(A) Actual Change in FID ¹								
	1. Average Annual Δ	-0.0125					-0.0124	-0.0100	-0.0104
	2. % Δ over Period	-14.8%					-14.5%	-9.2%	-14.8%
	(B) Decomposition of ΔG ¹								
	1. Estimated Avg. Ann. Δ ²	-0.0112		-0.0019	-0.0093		-0.0093	-0.0034	-0.0059
2. Distribution of Change ³	-100%		-17%	-83%		-100%	-37%	-63%	
	Index of Findings	F.3b		F.8b	F.11c		F.11c		

¹Part A of each model shows the actual changes in G_y, G_A, G_x, G_π and G_w, while Part B is based on the decomposition equations ((2.4) and (2.8)). Hence ΔG_y of A.1^π (based on actual change) and B.1 (estimated from the decomposition equation (2.4)) generally differ slightly.

²Columns 2+3+4=Column 1; columns 5+6=column 4; columns 7+8=column 6.

³Columns 2+3+4=100%; columns 5+6=100%; columns 7+8=100%.

(Handout HZID - Feb 76)

DIAGRAM labc Panel (A)
Factor Ginis, Total Gini, and Degree of Over-estimation

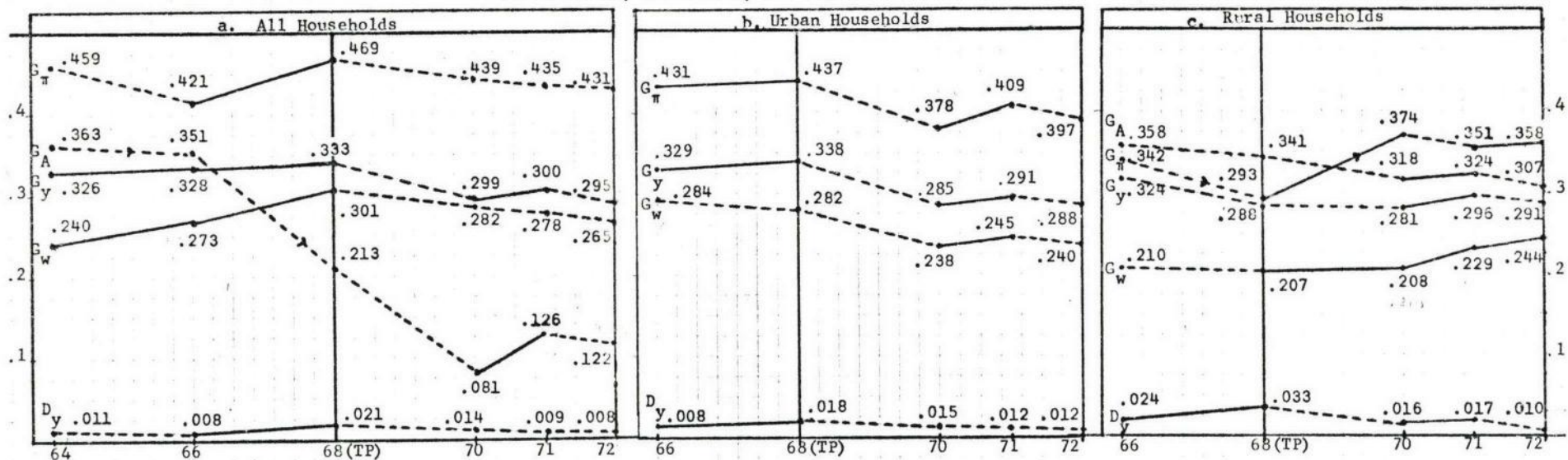
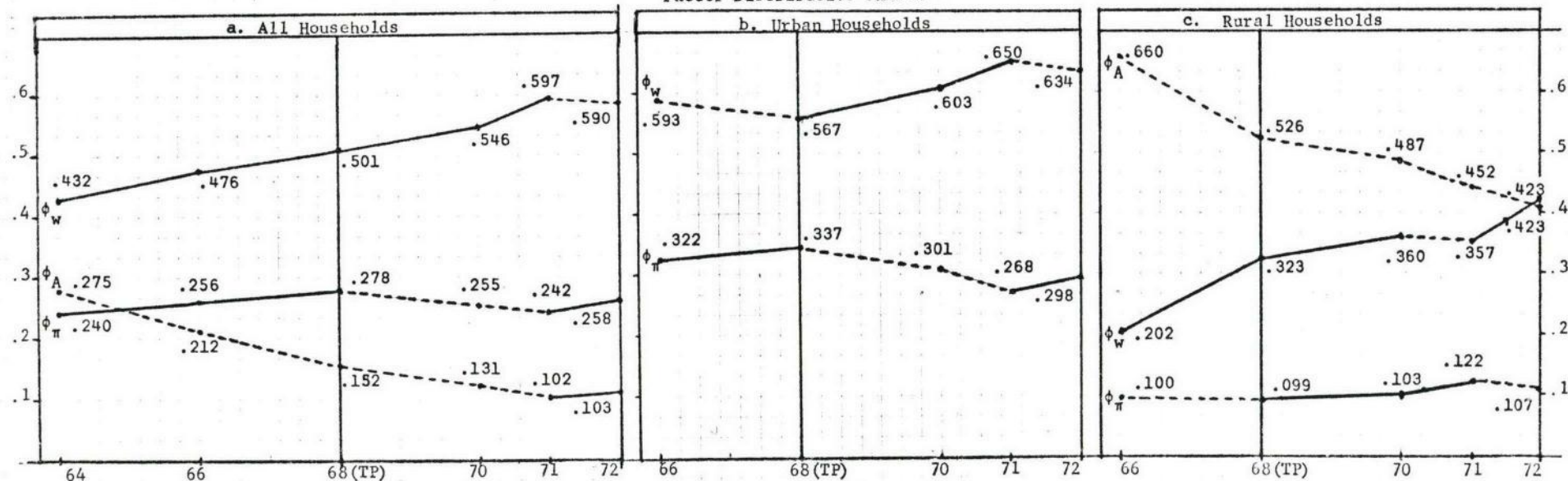


DIAGRAM labc Panel (B)
Factor Distributive Shares



Note: — increasing - - - - decreasing

0-84

March 30, 1976

Dr. Gary S. Fields
Asst. Prof. of Economics
Economic Growth Center
Yale University
P.O. Box 1987 Yale Station
New Haven, Conn. 06520

Dear Gary:

Because your paper, "Who Benefits from Economic Development-- a Re-examination of Brazilian Growth in the 60s," comes up with very different conclusions on the Brazilian experience from those of previous writers, we thought it useful to inquire into them rather carefully. The results of our attempts to reproduce your analysis with a different methodology are given in the attached note.

Basically, we have two problems. The first has to do with comparisons of two distributions over time. We have found in the past that the errors involved in linear interpolation of even a single distribution can be large; the results from a difference between two such distributions may then reflect mainly the errors (depending on whether or not they happen to be off-setting). We, therefore, have found it necessary to use the Kakwani-Podder procedure. Applying this to your data, we obtain a difference equal to an order of magnitude in what is, perhaps, the most important single numerical result. That is, we find an income increase for the "poor" of about 6% in contrast to your finding of 62%.

Second, your main innovation is to introduce a poverty concept. We have doubts about how this relates to a distribution of the economically active labor force. Relating it to average per capita family income would seem more appropriate.

I hope that these comments may prove useful. At the minimum, they do suggest the need for circumspection in drawing conclusions about the contentious issue of the beneficiaries of development in Brazil.

Dr. Gary S. Fields

- 2 -

March 30, 1976

Finally, I must apologize for the length of time it has taken us to react substantively to your paper--it only occurred to us recently that there appeared to be an incongruity between your results and our study on the Northeast. It was this that suggested the need to explore more carefully.

I enclose also a complimentary copy of Shail Jain's compilations.

Best wishes,

Yours sincerely,

John H. Duley, Director
Development Research Center

Enclosures

cc: M/s. Chenry
Karaosmanoglu
Pyatt
Ahluwalia
B. King

JHD/vec

Mr. M. Ahluwalia, DRCID

November 6, 1975

Benjamin B. King, VPD *Signed B. B. King*Yale Discussion Paper 235 (670-84)

Hollis suggests making this into a Working Paper. I see no objection as long as it's OK with Yale. Gary Fields raises interesting questions.

However, I do have one question about the paper. Perhaps the answer to it, if relevant, could be incorporated. Briefly, it can be put as: "How come?" Gini goes up but the poor are relatively better off.

Suppose we use the following notation:

G = Overall Gini
 G_p, G_r = Within-group poor and rich Gini's
 Q_p, Q_r = Proportions of poor and rich by number
 (sum = 1)
 Y_p, Y_r = Proportions of poor and rich by income
 (sum = 1)
 Then $G = G_p Q_p Y_p + G_r Q_r Y_r + Q_p - Y_p$

Now suppose proportions by number don't change and also within-group Gini's don't change, then differentiating:

$$dG = dY_p (G_p Q_p - G_r Q_r - 1)$$

If dY_p is positive, dG is negative.

Ergo, in Brazil, either the G's or the Q's changed. We can rule the latter out by definition (though there may be statistical quirks). Therefore, at least one within-group must have got worse.

In other words, if the within-the-rich Gini got worse, do we care? Isn't this the gist of the question on page 25?

Of course, the within-the-poor Gini might have got worse too (or it may be the main reason). In this case I suppose we care a bit more. But, if all the poor were equally poor to start with, should we cavil at some of them getting out of the rut?

But, if we aren't very concerned with the within-Gini's, aren't we back at square one with a simple before and after comparison of percentages of total income, such as:

	<u>Before</u>	<u>After</u>
Poor	10	12
Middle	35	34
Rich	55	54

With some indices of absolute income, doesn't that tell us enough, without having to "explain" why the Gini has a credibility gap?

cc: Messrs. Chenery
Grimes

BBKing:gm

1) *Ginn*
2) 670-84

OFFICE MEMORANDUM

TO: Mr. B. King, VPD

DATE: October 30, 1975

FROM: Hollis B. Chenery, *HBC* VP, Development Policy

SUBJECT: Proposed Discussion Paper by Gary Fields

1. The attached Discussion Paper on Income Distribution in Brazil was prepared by Fields on our RPO 284 with Yale. Montek and I feel that it is a very useful paper and could be circulated to advantage under the Bank's Working Paper series. I wish you would have a look at it and if you agree, Montek will get in touch with Fields to secure his concurrence.

2. Although there is not a large volume of such memoranda prepared outside the Bank under our research projects, I think it would be useful for you and Orville to look at the products that are prepared with a view to including them in the Working Paper series whenever appropriate.

Attachment - *orig returned to HBC 11/13/75*
cc: Mr. M. Ahluwalia

HBC:gss
HBC:gss