Living Standards
Measurement Study
Working Paper No. 23

LSM - 23
JAN. 1986

Collecting Panel Data in Developing Countries: Does It Make Sense?

LSMS Working Paper Series

No. 1. *Living Standards Surveys in Developing Countries.*

No. 2. *Poverty and Living Standards in Asia: An Overview of the Main Results and Lessons of Selected Household Surveys.*

No. 3. *Measuring Levels of Living in Latin America: An Overview of Main Problems.*

No. 4. *Towards More Effective Measurement of Levels of Living,* and *Review of Work of the United Nations Statistical Office (UNSO) Related to Statistics of Levels of Living.*

No. 5. *Conducting Surveys in Developing Countries: Practical Problems and Experience in Brazil, Malaysia, and the Philippines.*

No. 6. *Household Survey Experience in Africa.*

No. 7. *Measurement of Welfare: Theory and Practical Guidelines.*

No. 8. *Employment Data for the Measurement of Living Standards.*

No. 9. *Income and Expenditure Surveys in Developing Countries: Sample Design and Execution.*

No. 10. *Reflections on the LSMS Group Meeting.*

No. 11. *Three Essays on a Sri Lanka Household Survey.*

No. 12. *The ECIEL Study of Household Income and Consumption in Urban Latin America: An Analytical History.*

No. 13. *Nutrition and Health Status Indicators: Suggestions for Surveys of the Standard of Living in Developing Countries.*

No. 14. *Child Schooling and the Measurement of Living Standards.*

No. 15. *Measuring Health as a Component of Living Standards.*

No. 16. *Procedures for Collecting and Analyzing Mortality Data in LSMS.*

No. 17. *The Labor Market and Social Accounting: A Framework of Data Presentation.*

No. 18. *Time Use Data and the Living Standards Measurement Study.*

No. 19. *The Conceptual Basis of Measures of Household Welfare and Their Implied Survey Data Requirements.*

No. 20. *Statistical Experimentation for Household Surveys: Two Case Studies of Hong Kong.*

No. 21. *The Collection of Price Data for the Measurement of Living Standards.*

No. 22. *Household Expenditure Surveys: Some Methodological Issues.*

Collecting Panel Data in Developing Countries:

Does It Make Sense?

The Living Standards Measurement Study

The Living Standards Measurement Study (LSMS) was established by the World Bank in 1980 to explore ways of improving the type and quality of household data collected by Third World statistical offices. Its goal is to foster increased use of household data as a basis for policy decision making. Specifically, the LSMS is working to develop new methods to monitor progress in raising levels of living, to identify the consequences for households of past and proposed government policies, and to improve communications between survey statisticians, analysts, and policy makers.

The LSMS Working Paper series was started to disseminate intermediate products from the LSMS. Publications in the series include critical surveys covering different aspects of the LSMS data collection program and reports on improved methodologies for using Living Standards Survey (LSS) data. Future publications will recommend specific survey, questionnaire and data processing designs, and demonstrate the breadth of policy analysis that can be carried out using LSS data.

LSMS Working Papers
Number 23

Collecting Panel Data in Developing Countries: Does It Make Sense?

Does It Make Sense?

Orley Ashenfelter
Angus Deaton
Gary Solon

The World Bank
Washington, D.C., U.S.A.

Manufactured in the United States of America
First printing January 1986

This is a working document published informally by the World Bank. To present the results of research with the least possible delay, the typescript has not been prepared in accordance with the procedures appropriate to formal printed texts, and the World Bank accepts no responsibility for errors. The publication is supplied at a token charge to defray part of the cost of manufacture and distribution.

The World Bank does not accept responsibility for the views expressed herein, which are those of the authors and should not be attributed to the World Bank or to its affiliated organizations. The findings, interpretations, and conclusions are the results of research supported by the Bank; they do not necessarily represent official policy of the Bank. The designations employed, the presentation of material, and any maps used in this document are solely for the convenience of the reader and do not imply the expression of any opinion whatsoever on the part of the World Bank or its affiliates concerning the legal status of any country, territory, city, area, or of its authorities, or concerning the delimitation of its boundaries, or national affiliation.

The most recent World Bank publications are described in the annual spring and fall lists; the continuing research program is described in the annual *Abstracts of Current Studies.* The latest edition of each is available free of charge from the Publications Sales Unit, Department T, The World Bank, 1818 H Street, N.W., Washington, D.C. 20433, U.S.A., or from the European Office of the Bank, 66 avenue d'Iéna, 75116 Paris, France.

Orley Ashenfelter is professor of economics and director of the Industrial Relations Section at Princeton University, Angus Deaton is professor of economics and international affairs at Princeton University, and Gary Solon is assistant professor of economics at the University of Michigan at Ann Arbor; all are consultants to the World Bank.

Library of Congress Cataloging-in-Publication Data

Ashenfelter, Orley, 1942-
Collecting panel data in developing countries.

(LSMS working paper, ISSN 0253-4517 ; no. 23)
Bibliography: p.
1. Cost and standard of living. 2. Household surveys. 3. Panel analysis. I. Deaton, Angus. II. Solon, Gary, 1954- . III. Title. IV. Series.
HD6978.A69 1985 339.4'2 85-26405
ISBN 0-8213-0673-1

ABSTRACT

This Working Paper reviews a number of aspects of the collection and use of panel data from households in developing countries. Sampling issues are discussed in Section 1. The authors conclude that there are likely to be real, if modest, benefits from incorporating some panel element into household survey data collection in developing countries. The recognition that panel data are likely to be subject to substantial errors of measurement does not invalidate this conclusion. Section 2 discusses the measurement of income dynamics, an issue that cannot be addressed without panel data. Recent research using U.S. data is reviewed to show that comparable work for developing countries would add an important dimension to discussions of poverty, inequality, and development. It is in the third area of review, that of econometric analysis, that the real benefits of panel data appear most fragile. These issues are discussed in Section 3. While it is true that panel data offer the unique ability to deal with the contamination of econometric relationships by unobservable fixed effects, the presence of measurement error can compromise the quality of the estimates to the point where it is unclear whether cross-section or panel estimators are superior. This situation is in sharp contrast to that for sampling where errors of measurement typically cannot reverse the superiority of panel over cross-section estimators.

The authors conclude by arguing that panel data should be collected in both developing and developed countries. Benefits of well-designed data collection efforts are likely to outweigh the costs. However, it is easy to overstate the likely benefits of panel data. Their existence will not solve all outstanding problems of understanding poverty and household behavior in developing countries. While they will undoubtedly bring new and important insights, they will also bring new problems of interpretation and analysis.

ACKNOWLEDGMENTS

The views expressed in this paper are those of the authors and should not be attributed to The World Bank. The authors are grateful to Dennis de Tray and Jacques van der Gaag of the Development Research Department, World Bank for helpful comments. They would also particularly like to thank Greg J. Duncan for numerous suggestions and for detailed comments on an earlier draft.

TABLE OF CONTENTS

1. INTRODUCTION

There is no difference in principle between developing and developed countries in deciding whether or not it is desirable to collect panel data. In both cases, a balance has to be struck between statistical issues, likely errors of measurement, and the purposes to which the data are to be put. However, there are many relevant circumstantial differences in developing countries. The costs of data collection are different, response errors are perhaps higher and certainly so for some important items like income, and the mechanics of sampling are different. For example, American panels make heavy use of telephone interviews in a way that would not be feasible in most developing countries. Even so, most of the aims of panel data collection are shared between the United States and poor countries, particularly the documentation of levels of and changes in poverty, employment, unemployment and occupational structure. In consequence, much can be inferred from the American experience for the likely costs and benefits of collecting panel data in different circumstances. In this paper, we attempt to summarize both the conceptual issues and the relevant past experience.

It is not always possible or desirable to draw a sharp distinction between panel and cross-section data. The choice between repeated independent cross-sections on the one hand and a continuing panel on the other is by no means the only possibility. Rotating panels replace some fraction of households at each wave, so that such a design, with a variable proportion retained, has both traditional cross-sections and panels as special cases. The question then becomes whether it is desirable to retain any households, and if so, how many. The next section takes up the sampling aspects of this question. Note that, if there is a very long gap between the successive waves of a panel, attrition and the inability to track down previous participants will make it necessary to replace a large proportion of households. Hence, over such time periods, there is probably little effective difference between panels and independent cross-sections. At the other end of the spectrum, short-run panel elements are already frequently present in cross-sectional household surveys in developing countries, where repeat visits are used to control for seasonality, to observe time use, or to minimize the fatigue of long interviews. Such panel elements are really a matter of design for cross-sections, since the time between interviews is typically too short for the change that has taken place to be of very great interest in its own right. In this paper, we shall mostly contrast, on the one hand, repeated independent cross-sections a year or more apart, with, on the other hand, the option of retaining some or all households from one survey to the next.

We believe that, while there are genuine difficulties, there are good arguments for collecting panel data. We do not believe that there is any real justification for the mindless enthusiasm that is sometimes expressed by econometricians, nor do we believe it to be correct that the advantages of panel over cross-section data are destroyed by the recognition of substantial errors of measurement. On balance, we believe that collecting panel data in developing countries _is_ a sensible enterprise, though the net benefits should not be overstated. It is therefore wise to have a balanced view of both likely benefits and costs so that the exercise is not undertaken with unrealistic expectations. We also hope that the discussion below will deal with some of the negative arguments that are sometimes put against collecting panel data.

The paper is arranged as follows. Section 2 is concerned with the sampling and measurement issues in the context of measuring living standards and their changes over time. The traditional statistical arguments for panels are reviewed, and we discuss the all important issue of measurement error. Section 3 discusses those phenomena that can only be analysed with the aid of panel data, for example whether the poor always consist of the same individuals or whether individuals tend to move fairly rapidly in and out of poverty. Again, the issue of measurement error is discussed, and the American experience is reviewed. Section 4 is concerned with econometric attempts to use panel data to go beyond description towards analysis of processes and causation. While it is true that panel data hold great potential for econometric analysis, it is less than clear that this potential has so far been realized in hard results, rather than in work and renown for econometricians. Once again, the reason for this is almost certainly linked to the presence of errors of measurement in panel data, and these tend to be more of a problem for econometric analysis than for the statistical estimation of means and changes. Section 4 illustrates some of the problems with an analysis of the relationship between changes in labor supply and changes in wage rates using the data from the Michigan Panel Study of Income Dynamics (PSID). The difficulties encountered there would almost certainly be as or more severe with data from LDC's. Panel information can clearly contribute to our understanding of the world, but its analysis presents its own special problems, and its possession is not a magic key that will unlock the secrets of economic development.

2. SAMPLING AND MEASUREMENT ISSUES

In this section we consider the question of using panel data to measure means and changes in means. In practice, there are many other important descriptive statistics that are routinely computed from panel data, for example, average changes for individuals over several time periods, or measures of the instability of individual experience over time. However, many of the same issues arise in these contexts and we begin with the simplest cases. The variable under consideration will be taken to be some measure of living standard (household per capita consumption, or household income) for some well-defined group, frequently the whole population. It matters for sample design what one is trying to measure and one of the problems with the design of household surveys has always been their multi-purpose nature and the consequent lack of a clear objective from which an optimal design can be derived. Nevertheless, average per capita income and consumption and changes in them over time are clearly quantities that are of interest to any survey-based attempt to measure living standards, so that it is sensible to ask what design will measure such things most accurately, or most accurately for a given cost. There is a standard answer to this question in terms of panels versus cross-sections, see for example, Hansen, Hurwitz, and Madow (1953), Kish (1965), or Cochran (1977). The algebra is summarized in the Appendix, but the argument runs as follows. Imagine that we are collecting data on two occasions and have the choice between independent cross-sections, a completely overlapping panel, or some compromise in which a fraction of households, p, say, is retained from one occasion to the next. We want to measure the mean in each period and the change from one period to the next. In general, there are trade-offs between the accuracy of measurement of the three quantities, so that it is necessary to specify our preferences for accuracy in one over another. Conceivably, we may only want to measure the change, but more likely we would not want to sacrifice all accuracy on means for accuracy on changes. It turns out that whatever balance is struck, it is always desirable to use means from both surveys to estimate the means of either one, as well as (more obviously) the change from one to the other. However, the choice of p, the proportion of overlapping households, depends on whether it is means or changes that must be measured most accurately. The optimal choice also depends crucially on the covariance between successive observations on the variable being measured. This is typically positive; households with high consumption in one period usually have high consumption in the next, though there are counterexamples, for example, an individual that has just bought a bicycle is unlikely to do so again for some time. Consider then estimating the change in per capita consumption. The variance of the estimate of a difference depends negatively on the covariance between the estimated means. Hence change is measured most precisely when the covariance is as large as possible. Independent cross-sections have a zero expected covariance, while a completely overlapping panel will exploit the positive correlation in the raw data to maximize the covariance between the estimated means. In consequence, the most accurate estimates of change from a given sample size are obtained from pure panel data. But we also want to get accurate estimates of means, and the higher is the positive correlation between periods, the more precision is sacrificed by not having the additional new information provided by the independent cross-sections. In consequence, paying attention to the precision of both means and changes will imply a rotating panel design in which the fraction of households retained is somewhere between zero and one.

Obtaining exactly the right balance between overlaps and new households requires more detail on preferences than most investigators could be expected to provide for a multi-purpose survey, but the point remains that a rotating design of some sort will generally be better than either a pure panel or independent cross-sections. It is also the case that getting the proportion exactly right is unlikely to be very important, since the efficiency gains are not very large unless the correlation over time in the raw data is very high, see Tables A.1 and A.2 in the Appendix. Besides, the correlation will be different for different variables in the survey, and only a single design can be chosen. Note too, that the further apart are the successive waves of the panel, the lower will tend to be the correlation between successive observations and thus the smaller the net advantage of the overlapping observations. Since changes over very short periods are typically less interesting than changes over a period such as a year or longer, this is not a prescription for bringing the successive waves close to one another. In the United States, the month-to-month change in employment and unemployment is considered to be an important statistic, and the Current Population Survey (CPS) is designed with that in mind. But most developing countries are concerned more with longer term changes in poverty and living standards than with short-run macroeconomic policy.

If survey responses were obtained, coded, and processed without error, that would be the end of the story, and any country that gets as far as contemplating regular cross-sectional household surveys would be well-advised to have some overlap from survey to survey. However, data collection is subject to all sorts of errors, and no serious analysis of survey procedures can ignore them. Of course, there are errors in <u>both</u> panel and cross-section data, but there are grounds for believing that (a) the same errors have different consequences for the two types of data, and (b) the different data collection modes produce different types and magnitudes of errors. We analyse each of (a) and (b) in turn.

For the reasons discussed above, panel data are often collected in order to analyse change over time. Since the change for each observation is often the difference between two large numbers, each of which is measured with error, the true change may be very small relative to the error with which it is measured. Panels may therefore be quite unsuccessful in measuring the very thing that they are designed for. All this is correct, but it is not an argument against panel data. If panel data are not collected, people will still attempt to measure changes over time, for example by comparing the means of successive independent cross-sections, and in most reasonable circumstances the presence of errors will only increase, not decrease, the relative precision of panel data. In the Appendix, a simple model of measurement error is developed to illustrate the point, but the issues are conceptually straightforward. For an individual respondent, or for the pairing of a respondent and an interviewer, two kinds of errors can usefully be distinguished: those that are essentially random, and will be independent over time, and those that are persistent, for example, when a given individual always understates his income. The total response error will then be the sum of both types. The random errors will have identical implications for cross-section and panel estimates of change. The estimates will be rendered less precise, but in the same way for both measures of the change. However, the panel data do better with the persistent errors since differencing for individuals will tend to remove at least part of the error and thus give a more precise estimate of the true change in means. It is only in the unlikely event of negative serial correlation in reporting errors that the estimate of change from the independent cross-sections would be more precise than that from the pure panel. For the same reason, rotating panels will tend to handle errors better than cross-sections, and even panel attrition cannot affect the direction of the result, since households that are lost can always be replaced with "look-alikes" or even random draws, and the remaining degree of overlap will always give the panel an edge over completely non-overlapping observations. If change is one of the objects of measurement, measurement error only enhances the case for a panel element in the survey.

Unfortunately, different sample designs introduce their own kinds of errors, and some assessment must be made of the likely effects of using a panel rather than an independent design. It is hard to come to firm conclusions in this difficult area, and we can do no more than rehearse some of the issues that are discussed in the literature. To preview, there are arguments going in both directions, and in our view there is, if anything, again a slight advantage in favor of the panel design.

Errors arise both at the individual level, from respondents, from interviewers, or from coding, and also at the aggregate level if the design fails to ensure continuing representativeness. Errors can be minimized in all designs, but usually at some cost, so it makes sense to consider relative costs at the same time as relative errors. Starting at the individual level, the literature discusses a number of frequently occurring errors in survey work, see for example, Sudman and Bradburn (1974). A common problem is "telescoping" or "border-bias," which in its basic form occurs in both panels and cross-sections. Respondents who are asked to recall purchases, or more generally, events of all sorts, will tend to bring distant events closer, so that, for example, purchases will be overstated by the inclusion of purchases that actually took place before the survey period took place. The border-bias reference is to Mahalanobis (1946), who discovered that reported yields on small plots were relatively overstated compared with yields on larger plots. The Bengali farmers who were unsure about the precise boundaries of their plots always tended to err on the side of including their neighbors' plants rather than run the risk of excluding their own. There is good evidence that cross-section household surveys suffer from this problem, see for example, Kemsley, Redpath and Holmes (1980) for evidence on the British Family Expenditure Survey, and there is no reason to suppose that panels would be any different. However, under some circumstances, repeat visits to the same household can help, as proposed in the method of "bounded recall" due to Neter and Waksberg (1964), since, after the first visit, new items can be checked to ensure that they have not already been recorded. This would not help for expenditures over a week or a month if the panel waves were three months or a year apart, but some major items of expenditure might be caught, and other major events (migration, job changes, demographic changes) might well be checkable over the longer period. Other questions involving memory recall are also important since one of the ways of inducing panel-type characteristics into cross-sections is to ask respondents about past data on the items currently being measured. A standard experimental result is that individuals tend to "forget" unpleasant events, and to forget them more frequently the further they are in the past. Hospital visits are a frequently quoted example; abnormally low incomes, financial losses, and still-births are others. Hence, although some questions, such as family life histories can be well investigated using recall, see Butz (1981) for evidence on the Malaysian Family Life Survey, it will always be difficult to obtain all of the desired information accurately. And, as pointed out by Duncan, Juster, and Morgan (1984), family life histories are less interesting if they cannot be linked to the economic histories of the household. It is therefore almost certainly the case that, in general, recall questions in cross-sections cannot fulfill the functions of genuine panels.

There is also an extensive literature on the question of "interviewer bias," whether the interviewer, by his or her manner of interviewing or method of recording, can affect responses. If so, there are all sorts of statistical consequences, including the non-independence of responses from different households handled by the same interviewer. For current purposes, most of the errors would be common between cross-section and panel designs. However, it is usual in panels for the same household to be handled by the same interviewer on repeat visits. If so, then the interviewer effects, if any, are likely to have a common component across visits, so that differencing to measure change will reduce the error variance. Once again, this would impart an advantage to the panel estimates. However, it may also be the case that repeated contact with the interviewer, or indeed the continuing presence of the household in the panel, may itself affect behavior. Such evidence as exists, as well as common sense, suggests that such "panel contamination" is likely to be small for behavior that is important to the respondents. The main problems seem to arise with attitude surveys, or with voting panels. It seems unlikely that contamination would be a serious problem in panels measuring living standards, consumption behavior, labor supply, or fertility. Indeed, experiments carried out by the Michigan PSID, in which replicate random samples were assigned to different interviewers, found essentially no variance between interviewers in the collection of economic data issues connected with labor supply.

A well-documented problem with panel data, which has been widely discussed in the context of the U.S. Census Bureau's CPS, is the phenomenon known as rotation group bias, see in particular Bailar (1975). In the CPS, households are interviewed for four successive months, they then drop out of the panel for eight months, and finally return for four final months before being discarded. The sample is designed to be representative of the U.S. population at all stages, so that this is true for, say, all households in the sample for the third time, as well as those in the second month of their second spell. But it turns out that, in any given calendar month, the means are significantly different for the groups at different points in the rotation pattern. In particular, there tends to be a major difference between responses at the first interview and at all subsequent ones. For example, the estimate of unemployment from households in the sample for the first month is about 10 percent larger than the average over all eight rotation groups, U.S. Bureau of the Census (1978). There are other patterns for other reported items, in most of which the first month rotation group is anomalous. It is not known what causes these phenomena, though the Census Bureau lists a number of possible factors, including conditioning, interviewer bias, the effects of telephone interviewing, changes in the questionnaire with rotation group, and changes in the individuals within the household who are deputed to answer the questions. Note that it is not at all clear that the first month observations are incorrect; it is quite possible that it is the other way around. Even so, one plausible explanation of these panel conditioning effects is that respondent motivation improves with successive waves. Many respondents have little idea of survey purpose or importance in one-off surveys, but repeated contact generates improvements in cooperation and in response quality. If so, the discrepancies observed over time are a reflection of the superiority over cross-sections of panel data, at least in the later waves.

Even with accuracy of individual responses, surveys will provide inconsistent estimates of population means unless they are genuinely representative. This is one area where independent cross-sections have a clear advantage. Provided that the sample frame is adequate, and the sampling done correctly, each new cross-section is guaranteed to be representative. Indeed, in LDC's refusals to co-operate tend to be much rarer than in developed countries, at least for one-time surveys (figures of 99 percent versus 70 percent are typical for the Indian National Sample Survey and the British Family Expenditure Survey respectively), so that representativeness is almost certainly not a problem for traditional household surveys in poor countries. Panels, by contrast, even if originally representative, do not remain so without careful (and possibly costly) tending. One problem is attrition; even if households agree to participate at first, they frequently change their minds after the first interview, though there is usually less of a problem subsequently. As with cross-sections, this may well be much less of a problem in developing countries. But there is also attrition through households that have moved between waves and cannot be relocated. The CPS does not try to find such households, but replaces them with whoever happens to be living in the original address at subsequent rounds. This is obviously a low cost solution, and if measuring changes in means is all that is at issue, may not matter very much since there is likely to be a positive correlation between the behavior of different households in similar (in this case, identical) accommodations. The Michigan PSID spends considerable time, money and energy tracking down movers, see again Duncan, Juster, and Morgan (1985). It is hard to imagine that this source of attrition would not be a serious problem in most developing countries. It may be that the squatter camps of Delhi, or the slums of Cairo, have more order to them than appears to a Western eye, but it cannot be easy to relocate individual households in such areas a year or two years after the original interviews, even if they are still there. And those that move tend not to leave forwarding addresses. Rural areas present less of a problem, and at least one successful village panel has been maintained over a considerable period, see Binswanger (1979). Similarly, Bliss and Stern (1982) also had no difficulty in locating individuals in Palanpur in India some years after an earlier survey.

Even without attrition, panels become unrepresentative over long periods as households break up or are joined by new members. Within the population, individuals tend to form new households with other members of the population, something that is <u>not</u> replicated by individuals in the panel. Panels such as the Michigan PSID have a mechanism whereby newly-born individuals and families are absorbed into the panel with selection probabilities chosen in such a way that representativeness is maintained. This is generally cheaper than selecting new sample units, and it has the very great advantage of ultimately allowing study of the links between related individuals and families, for example between ex-spouses, or between split-off children and their parents. Note that these possibilities arise only in panels that run for many years; they are not available for short-term exercises or for "rolling" panels.

There are also some more general cost advantages for panel data. Constructing new samples is expensive, and panels do not require it. Second wave interviews are typically shorter and therefore cheaper than first-time interviews, and shorter interviews with individuals reasonably familiar with the survey's operating procedures may also be more accurate. Duncan, Juster, and Morgan claim substantial cost savings for panels over repeated cross sections, but in the United States considerable use is made of telephone interviews for rounds other than the first, and such a cost saving exercise would obviously be unavailable on a world-wide basis. It is also possible that the regular collection of panels makes it easier to maintain a standing survey organization, with an attendant increase in professionalism and precision and decrease in costs per observation. However, such advantages are shared by the regular collection of cross-sections, as for example by the National Sample Survey Office in India. In total, it would not be surprising if reasonably representative rotation panels could be maintained at a cost no greater than the cost of repeated cross-sections yielding similar or less precision of estimates. Longer running panels may also be a possibility in some areas. It is certainly hard to find, amongst all these arguments, any one that would militate strongly against collecting panel data at all.

3. DATA UNIQUE TO PANELS: INCOME DYNAMICS

All of the discussion so far has concerned cases where panel data and cross-sections are both available for estimating the same magnitudes, typically means or changes in means. However, there is an important class of issues that can only be illuminated by panel data. Single one-off surveys can tell us what fraction of the population is living in poverty, or what fraction is employed. Repeated individual cross-sections reveal how these proportions change over time. However, they tell us nothing about the "gross" flows, about whether or not it is the same individuals who are poor or unemployed, or what are the typical paths of individuals' living standards over time. Such questions are thought to be of great importance in the United States, and one would have thought that they were equally or more important in developing countries. How we feel about inequality depends a great deal on whether the poor are always poor and the rich always rich, or whether almost everyone encounters poverty and riches at some point in their life. The question of "who benefits from development" requires some sort of repeated observations (at least on groups) for an adequate answer, and the very name of the Michigan panel, the Panel Study on Income Dynamics, suggests the central role played by these topics in the rationale for collecting such data.

There is therefore no doubt that it would be desirable to collect such information for developing countries were it feasible and cost-effective to do so. We note three likely problems, none of which is necessarily insuperable. First, measurement error again causes problems. If individuals report random numbers for their living standards, then even if the mean is unbiased, the estimates of the degree of change from one period to the next will be overstated. It is easy to imagine a Peruvian smallholder trying his best to tell the interviewer what his income is, but only having a weak idea of the concepts involved and faced with a questionnaire designed by an MIT Ph.D. who has never set foot outside the United States. Even if the person's income had not changed since the previous interview, the answers might easily show considerable movement within the recorded income distribution. However, once again, the presence of error persistence will tend to offset such effects. The Appendix presents a simple model in which the random components of the errors tend to lead to overstatement of income mobility, while the fixed effects lead to understatement. One can only be aware of such problems; without an independent source of data on income dynamics, there is no way of assessing their importance. The second problem concerns attrition. If those who leave the panel, or who cannot be found, have changes in income or in living standards that are not randomly distributed in the population, then the changes for those that do remain will give a misleading picture of dynamics as a whole. Thirdly, there is the problem of the time between successive waves of the panel. If this is too short, most of the recorded movements will be of little interest, reflecting seasonal or shorter frequencies in living standards, and telling us very little about movements in and out of poverty. A longer interval or a longer running panel is better from this point of view, but increases the difficulties over attrition. The typical "rolling panel," with each individual present only for a year or two (or less), can tell us little or nothing about who benefits or loses from economic development.

Some of the advantages as well as limitations of this type of use of panel data can be illustrated from the American experience in measuring income dynamics. Most of the recent work on income mobility has been based on the Michigan PSID. This survey, administered by the University of Michigan, Institute for Social Research, has annually reinterviewed some 5,000 families since 1968. The results indicate a significant degree of both persistence and mobility in income levels.

One way to summarize the data is to describe movements among quintiles in the family income distribution. Duncan and Morgan (1984) report that 56 percent of the individuals whose family incomes were placed in the bottom quintile in 1971 remained in the bottom quintile in 1978. Forty-eight percent of those in the top quintile in 1971 remained in the top quintile in 1978. Overall, 40 percent of the 1971 sample stayed in the same quintile in 1978, 37 percent moved one quintile, and 23 percent moved at least two quintiles. Although Duncan and Morgan note a "substantial and perhaps surprising degree of income mobility" in these statistics, there is also quite a substantial degree of serial correlation.

More detailed analyses typically have focused on the low end of the income distribution. Hill (1981), for example, examines the persistence of poverty status as defined in official poverty statistics. Roughly speaking, the official poverty threshold for a family of given size and composition starts with an estimated minimal annual food budget, multiplies it by three, and makes some additional adjustments for family size and farm/nonfarm status. Applying this poverty standard to PSID data for 1969-78, Hill finds that the cross-sectional poverty rate averaged 7.7 percent over this period and that about 60 percent of those poor each year remained poor the next year. Seven-tenths of one percent of the sampled individuals were poor in all ten years, and 2.6 percent were poor in at least eight years. Again, the data reveal considerable turnover in the poverty population but also a significant degree of chronic poverty. Many longitudinal analyses of poverty have also investigated the characteristics of the persistently and temporarily poor. Some of the most striking findings pertain to racial composition. Hill, for instance, reports that 42 percent of the poor in the 1978 cross-section were black (as compared to 12 percent black in the U.S. population). Furthermore, 61 percent of those poor in all of the 1969-78 years and 62 percent of those poor in at least eight years were black. Thus, the concentration of poverty among blacks is even greater when the focus is on chronic poverty.

In principle, some aspects of income dynamics could be measured also from the Current Population Survey because the rotating panel design of the CPS permits year-to-year matching of a portion of the sample. Unlike the Michigan PSID, however, the CPS makes no effort to follow movers. If moves are correlated with income changes, as seems likely, the matched subsample would then be unrepresentative of the population. Indeed, Kelly (1973) notes substantial discrepancies between poverty estimates from the full CPS and the matched subsample. This suggests that, if panel surveys in developing countries are intended to generate accurate information on income dynamics, they must attempt to follow movers or else devise methods to correct for the bias.

Much of the United States research on income dynamics has examined individual earnings as opposed to family income. An outstanding example is Lillard and Willis' (1978) study of PSID data on working male heads of households. One part of their analysis involves a model that views the natural logarithm of an individual's earnings as the sum of a permanent component and a transitory (but possibly serially correlated) component. Their estimates of this model, based on 1967-73 data, measure the degree of persistence in individual earnings status. Their full-sample results, for example, indicate that the serial correlation in log earnings is .84 from one year to the next, declines to .78 at a two-year interval and .75 at a three-year interval, and falls asymptotically to .73. When they stratify their sample by race, they find less earnings mobility among blacks than whites. For whites, they estimate a one-year serial correlation of .83 that declines asymptotically to .71. For blacks, they estimate a one-year serial correlation of .89 that declines asymptotically to .81.

Like many of the studies of family income, Lillard and Willis' earnings analysis devotes special attention to the low end of the distribution. They define low earnings in a particular year as earnings below half the median observed for males in that year's CPS. By this standard, 3 percent of the whites and 13 percent of the blacks in the 1967 PSID sample had low earnings. Lillard and Willis estimate that, of the low-earning men in a given year, about 45 percent of the whites and 65 percent of the blacks will have low earnings the next year, and about 25 percent of the whites and 50 percent of the blacks will have low earnings in both of the next two years.

An important question concerning all of the above analyses of income and earnings dynamics is whether the measurement of mobility is substantially distorted by measurement error. A long history of studies, including Herriot and Spiers (1975) and Miller and Paley (1958), has compared matched survey and administrative data on income and found widespread discrepancies. Similar discrepancies have been observed in comparisons of matched data from different surveys, see Mellow and Sider (1983), Miller (1964), and Pritzker and Sands (1958). More recently, Michigan's Institute for Social Research conducted a PSID-like survey of employees whose employer already had accurate records on earnings and other employment variables. Duncan and Hill (1985) report that the survey's ratio of error variance to true variance was .15 for annual earnings and .32 for their logarithm.

Where mobility analyses have mentioned the problem of measurement error at all, they typically have suggested that errors tend to exaggerate the extent of mobility. As demonstrated in the Appendix, however, this need not be the case. If the serial correlation in measurement error is large enough relative to the true serial correlation in income, mobility might be understated. In any case, it is somewhat reassuring that earnings mobility studies based on administrative data from the U.S. social security system have yielded results qualitatively similar to those from survey-based studies. McCall (1973), for example, tabulates 1962-63 transition rates across a $1500 threshold, which was approximately half the median covered earnings at that time. Of the white men, aged 25-34, with 1962 earnings below $1500, 42 percent remained below $1500 in 1963. In both the 35-44 and 45-54 age categories, the persistence rate was 45 percent. Also, like Lillard and Willis' PSID study, McCall's study found higher persistence rates for blacks.

4. ECONOMETRIC ANALYSIS OF PANEL DATA

Ultimately, the greatest payoff to collecting panel data will come if it is possible to go beyond description and discover something new about the processes governing living standards. The availability of panel data in the United States over the last decade or so has generated a great deal of often very sophisticated econometric analysis, in many cases dealing with phenomena that could not have been investigated without such data. Some of this work is the analytical counterpart to the descriptive material discussed in the previous section. For example, flows of individuals into and out of unemployment have been much studied, with particular reference to the effects of unemployment benefits and the effects, if any, of unemployment itself on the reemployability of the individual. Both areas are ones of much activity, not to say acrimony. Although many would disagree with us, we feel that this work has been more successful in illuminating the special problems associated with the econometric analysis of panel data, rather than in providing convincing answers to the original problems. It would not therefore be sensible to expect that increased availability of panel data from poor countries would lead to a great increase in knowledge about the processes of development. (It would, however, almost certainly increase the flow of good students into the subject, and that would do a great deal for the current state of development economics).

A. Studies of Food Consumption

We illustrate first with perhaps the most obvious use of panel data, which is to control for individual fixed effects. It is also an example that is important for the current use of household expenditure survey data in developing countries. Most countries that collect household budget data on consumption use it to fit Engel curves that summarize the relationship between purchases of various commodities and household income or total expenditure. One of the major uses of such Engel curves is for projection. Even without some form of central planning, most poor countries attempt to forecast the evolving structure of food demand to compare with their assessment of the growth of supplies. The problem is that there are nearly always major inconsistencies between the results from the surveys and the development of consumption patterns over time. For example, food as a whole always declines as a share of the budget as total expenditure rises across a budget survey; it typically declines much more slowly with increases in real income over time. There are similar problems with other categories of the budget. In consequence, forecasting experience with Engel curves estimated from budget data has been very poor. The essential problem is that the variables that are being implicitly held constant in the budget survey are not the same as those being held constant over time. To illustrate, consider the food equation

$$q_h = a + b x_h + f_h + u_h \qquad (1)$$

where q_h is expenditure on food by household h, x_h is household total expenditure, and a and b are parameters of the Engel curve - the linear form is used only for convenience. The quantity f_h is the household fixed effect; it represents that household's idiosyncratic demand for food and is constant over time. The error u_h is random, and is also part of the unpredictable part of h's demand; unlike f_h however, the u_h's in different time periods are independent of each other. Now, if the distribution of the f's in the population remains constant over time, the regression of aggregate food demand on aggregate expenditure will have a slope of b, the true marginal propensity to consume. However, the same regression in a cross-section will not estimate b except in the unlikely event that the fixed effects f are uncorrelated with total expenditure x. Even if other variables are included in the regression in an attempt to control for the fixed effects (household demographics, education, occupation, and so on), it is always implausible that <u>all</u> such effects can be accounted for.

Panel data, even for two successive observations can in principle deal with this problem. For if equation (1) is first-differenced using two observations from the same household, the fixed effects vanish, and the regression of the change in food consumption on the change in total expenditure yields an unbiased and consistent estimate of the Engel slope b. Unfortunately, errors of measurement once again interfere. Assume in (1) that x is measured with error. There will then be two sources of bias in the cross-section regression, one from the omitted fixed effect, and one from the error of measurement. The panel estimator suffers only from measurement error. However, it is a standard result that the bias from measurement error depends on the ratio of the measurement error variance to the total variance. In the cross-section, even poorly measured total expenditure will have a relatively small ratio of error to total variance, while in the panel differenced regression, the ratio is likely to be much larger. Once again, the difference between two large numbers is likely to be subject to large relative errors. It is very easy in these circumstances to construct examples where the inconsistency of the panel first-differenced estimator is worse than that of the original estimate from the cross-section. It may be, however, that given sufficient panel data, methods can be developed that will yield consistent estimates of the desired parameters, and this is very much a topic of current research, see, for example, Griliches and Hausman (1984).

There is similar evidence of the influence of measurement error in American work on labor supply using the PSID data. Indeed, it is in the area of labor supply analysis that panel data have been most widely used, and it would hardly be an exaggeration to claim that the existence of panel data has transformed the nature of applied research in the area. We therefore conclude with a review of the uses of panel data in labor supply analysis where the focus once again ends up being on the consequences of measurement error. As before, we are essentially confined to reviewing the American work.

B. Life-Cycle Labor Supply Studies Using Longitudinal Data

Econometric analyses of labor supply have focused on the life-cycle because of the many policy issues raised by the difference between the effects of permanent versus transitory changes in incentives. The only way these effects can be sorted out is if both permanent and transitory wage movements can be observed. In a single cross-section of data this is not usually possible. Hence, questions of whether a short term tax change will have the same effect on labor supply as a permanent tax change simply cannot be addressed without some kind of longitudinal data analysis.

The study of life-cycle labor supply with longitudinal data is an especially fine example where well designed economic models have been confronted and analyzed with the conceptually appropriate data. This does not mean these models have already generated compelling empirical estimates of key theoretical parameters. Indeed, the difficulties of squaring up the data with the models are constructed. Examining the course of the difficulties in matching the models to the longitudinal data is instructive for the light it sheds on how such data can structure and alter our economic analyses. It is especially instructive with regard to how measurement error in the longitudinal data can cause severe problems with the subsequent econometric analyses.

The simple theoretical models of life-cycle labor supply serve two purposes. First, they show the clear connection between the life-cycle labor supply model and the permanent income theory of consumption. Indeed, the latter is simply the consumption plan derived from the former, and permanent income is nothing more than the appropriately discounted present value of future wage earnings. Second, to be tractable, the empirical analyses are going to require some form of linearity <u>and</u> some simple method for summarizing a consumer-worker's future prospects. The specialized assumptions of the available models rationalize the assumptions necessary to produce such implementable models.

(i) <u>Simple Models with Perfect Foresight</u>

One simple theoretically based model that generates a linear earnings function is the Stone-Geary utility function that has been used by Ashenfelter and Ham (1979). It leads to an equation that explains the choice of a worker's desired labor earnings as consisting of two parts. One part is a constant throughout an individual's life. This constant part is proportional to an index of the discounted present value of the individual's lifetime discretionary income. Discretionary income in this model is the difference between maximum feasible earnings and minimum feasible consumption, where "feasible" is defined by constant taste parameters. Variations in permanent incomes across individuals will, of course, result in variation in hours worked across individuals.

The second part of desired labor earnings in this model varies over the life-cycle according to how the wage rate varies over the life-cycle. This part is greatest in periods when wage rates are highest and is smallest when wage rates are lowest. Moreover, hours worked are also greatest in this model when wage rates are highest. To be consistent with the simplest utility maximization models the elasticity of hours worked with respect to within life-cycle movements in the wage rate, holding constant permanent income, must be positive. Efforts to test this proposition and to estimate this intertemporal labor supply elasticity make up much of the empirical literature analyzing panel data on labor supply.

The Stone-Geary utility function also generates a lifetime plan for consumption. Desired consumption also contains two parts. One part is a constant and is proportional to permanent discretionary income while the other part varies with life-cycle variation in the real price of consumption. This model would therefore rationalize the Engel curve given by equation (1) above. Considerably less effort has been devoted to the study of panel consumption data using such models in the United States, however, perhaps because so much less consumption information is available.

In the Stone-Geary model intertemporal movements in labor earnings are solely a result of life-cycle or time-series movements in the wage, w_t. These movements are governed entirely by a parameter γ_h in such a way that the change in desired earnings, $\Delta w_t h_t$, moves according to the simple rule

$$\Delta w_t h_t = \gamma_h \Delta w_t \qquad (2)$$

The key point is that, as before, longitudinal data allow us to "difference out" the fixed effect that represents permanent discretionary income and avoid bias. The proportional change in earnings is

$$\Delta(w_t h_t)/w_t h_t = (\gamma_h/h_t)[\Delta w_t/w_t] \qquad (3)$$

In this model the intertemporal elasticity of labor supply is $(\gamma_h/h_t)-1$. Since the intertemporal elasticity of labor supply must be non-negative, this establishes that in the regression (2), the coefficient γ_h must be larger than h_t for the model to be consistent with the utility maximizing principle.

This model has been fit to data from the Michigan PSID by Ashenfelter and Ham (1979), but before reporting those results it is useful to turn to some data that provide a very simple method for estimating γ_h. This scheme is based on the observation that equation (2) is a regression without a constant term. One consistent estimator for γ_h is therefore the ratio of the mean across workers of the change in earnings to the mean of wage change across workers in a longitudinal data set. The advantage of this estimator is that it remains consistent even when zero-mean measurement errors are appended to Δw_t and $\Delta w_t h_t$ in equation (2), see Wald (1940).

To get a feeling for the estimates obtained in this way consider the mean changes in real earnings and real wage rates reported from the PSID in Table 1. The third column reports the estimates of γ_h. First, they are very unstable. Second, in three of the years considered they are lower than the actual mean of hours worked. Discarding as extreme outliers the results for 1969-70 and 1973-74 leads to an average of the ratios $\gamma_h/h_t = 1.04$, which implies an intertemporal labor supply elasticity of .04. Obviously, with these data virtually any estimate of γ_h may be obtained depending on what the empirical analyst wants to see.

There are clearly difficulties in using this simple estimator to calibrate the size and stability of the intertemporal labor supply elasticity. Perhaps most disturbing is the possibility that aggregate supply shocks or their determinants will obscure movements along the supply schedule (2). More generally, anything that might successfully and correctly be removed from the panel data by the addition of year dummy variables will produce a specification bias in these results. It is important to emphasize, however, that the consistency of many of the estimates of the intertemporal labor supply elasticity reported below are dependent on the same assumption necessary to ensure the consistency of the estimates of the labor supply elasticity in Table 1 <u>and</u> typically on further assumptions.

TABLE 1: **Changes in Real Earnings and Real Wage Rates: Panel Survey of Income Dynamics (White Males, 25–50 years old in 1967)**

Year	Change in Real Earnings	Change in Real Wage	γ_h	Mean Hours Worked
1967–68	486	.19	2,558	2,416
1968–69	313	.20	1,565	2,403
1969–70	-101	.01	-10,000	2,370
1970–71	206	.18	1,144	2,352
1971–72	561	.12	4,675	2,367
1972–73	396	.16	2,475	2,370
1973–74	-371	-.02	18,550	2,328

Source: Appendix of Ashenfelter and Ham (1979); earnings and wage rate deflated by consumer price index, 1967 = 1.0.

Equation (2) has also been fitted to the PSID micro data directly by Ashenfelter and Ham (1979). These results are even more disappointing. The estimates of γ_h based on the pooled covariances in the data are around 1,900 hours. This result implies a negative intertemporal elasticity of labor supply.

One possible explanation for the poor performance of this kind of model is based on the idea that desired hours may differ systematically from actual hours worked for a variety of reasons. Ham (1985) has used panel data to follow up the possibility that unemployment is a measure of the gap between actual and desired hours at work. In his study he finds considerable evidence for this hypothesis, and it does appear that allowing for this fact permits the generation of positive, but small, intertemporal labor supply elasticities, although not without some allowance for measurement errors.

A different model that leads to a log linear labor supply function has been suggested by Heckman and MaCurdy (1980) and MaCurdy (1981). As Abowd and Card (1983) observe, under certain conditions this model leads to precisely the log linear labor supply function initially proposed by Lucas and Rapping (1970). In this case, the change in desired hours at work will depend only on the change in wage rate. Holding permanent income constant the proportionate change in labor supply over the life-cycle is governed by

$$\Delta \ln h_t = \eta \Delta \ln w_t, \tag{4}$$

where $\Delta \ln h_t$ and $\Delta \ln w_t$ are the (approximate) proportionate hours and wage changes, and η is the intertemporal labor supply elasticity. Again, there are some straightforward estimates of the intertemporal labor supply elasticity available from the ratios of the means of the hours changes to the means of the wage change in a panel of data. To provide a feeling for the size of these estimates Table 2 contains the data on the change in log hours and log wages from the PSID computed by Abowd and Card (1983).

The estimates of η, the labor supply elasticity, in Table 2 are qualitatively consistent with the results in Table 1 except for the year 1970-71. In general, however, the data in Table 2 are far more congenial to an estimate of the intertemporal labor supply elasticity that is positive and large in magnitude. Only two of the ten estimates of the elasticity are negative, and the simple average of all the estimates is 1.14. Deleting the two extreme outliers leads to an estimate of the intertemporal labor supply elasticity of .89. (This is equivalent to deleting the two estimates with the denominators closest to zero in absolute value.) As before, however, these estimates are very unstable and this instability casts serious doubt on the credibility of this model.

TABLE 2: **Changes in Log Real Earnings, Log Hours, and Log Real Wages: Panel Survey of Income Dynamics (Male Heads of Households, 21-64)**

Year	Change in Log Earnings	Change in Log Hours	Change in Log Wage	η a/	Change in Unemployment Proportion b/	Change in Log Manhours c/
1969-1970	.032	-.011	.043	-.26	.009	-.02
1970-1971	.030	.003	.027	.11	.007	-.01
1971-1972	.072	.021	.051	.41	-.004	.04
1972-1973	.048	.021	.027	1.29	-.007	.04
1973-1974	-.051	-.042	-.009	4.67	.006	.00
1974-1975	-.041	-.027	-.014	1.93	.023	-.05
1975-1976	.046	.012	.034	.35	-.008	.04
1976-1977	.024	.002	.022	.09	-.006	.04
1977-1978	.005	-.003	.008	-.38	-.007	.05
1978-1979	-.055	-.042	-.013	3.23	.001	.03

a/ Calculated as the ratio of the mean change in log hours to the mean change in log wages.

b/ Change in unemployment proportion for males aged 35-44.

c/ Change in the logarithm of the payroll series data on the aggregate weekly hours index.

Source: Abowd and Card (1983), Table 2.
Employment and Training Report of the President, Table A-30, C-13, 1982.

There are several ways to use covariances in the data to estimate the labor supply elasticity. The simplest method is simply to compute the regression coefficient of the proportionate changes in hours on the proportionate changes in wages. In the absence of measurement error the estimate should, of course, be positive. Abowd and Card (1983) report all of the necessary data to do this from the PSID and from the National Longitudinal Survey of Older Men (NLS). This regression coefficient is always negative and significantly different from zero at quite small probability levels. In the PSID it is -.36 and in the NLS it is -.28 for the data reported by Abowd and Card (1983). This is, of course, essentially the same "fact" discovered by Ashenfelter and Ham (1979), since it is implied by their finding that γ_h is typically less than hours worked in the sample.

It is possible to use the covariances in the data to find a regression coefficient with the sign implied by the utility maximization hypothesis. For example, MaCurdy (1981) observes that adding $\eta \Delta \ln h_t$ to both sides of (4) produces the relationship

$$\Delta \ln h_t = \{\eta/(1 + \eta)\}\Delta \ln w_t h_t \qquad (5)$$

This suggests computing the regression coefficient of the change in the log of hours on the change in the log of earnings. In the PSID and NLS data these regression coefficients imply estimates of the labor supply elasticity of around .78 and .61 respectively.

It seems clear, however, that simple applications of either the linear earnings equation (2) or the log linear hours equation (4) will require some manipulation of the data before they will produce credible estimates of the intertemporal labor supply elasticity. As a result, these estimates are likely to be sensitive to the model specification, although preliminary indications are that they are not likely to be larger than .7 or .8.

(ii) Models of Life-Cycle Labor Supply with Measurement Error

The presence of measurement error has been suggested as one important reason for modifying the longitudinal analyses of labor supply thus far reviewed. One suggestion is to recognize the presence of measurement error in both wage changes and hours changes at the micro level. As we have observed, the simple ratio of means estimates in Tables 1 and 2 need not suffer from bias induced by measurement error. On the other hand, the usefulness of this simple procedure depends critically on the assumption that unmeasured economy wide shocks to real interest rates or other aggregate variables can be safely ignored. Using the covariances in the panel data with time means subtracted out does not run this risk.

MaCurdy (1981) circumvents these issues by using an instrumental variables estimator. With time means subtracted out of the data his estimates of the regression coefficient of the change in log hours on the change in log wages are .10 and .15 with standard errors of about the same magnitude. These are not large elasticities and the imprecision of their estimation is disturbing. The imprecision no doubt results from the inevitably poor quality of what are essentially time-invariant instrumental variables. (It is well known that wage rates in a cross-section are roughly a semilogarithmic function of schooling, experience, and experience squared. The first-difference in the log wage is therefore approximately a linear function of experience, the main instrument available.)

In an empirical study Altonji (1985) reports several efforts to account for measurement error in an attempt to estimate (4). He reports three alternative sets of results from the PSID data. The first set uses an instrumental variables scheme designed to reproduce MaCurdy's results. The estimated intertemporal labor supply elasticity falls in the range of .08 and .50 depending on whether time means are subtracted out of the data and whether age is included in the labor supply equation. Estimated sampling errors fall in the range .12 to .4, however, so that elasticities are still imprecisely estimated. A second procedure uses an alternative (but contemporaneously measured) wage variable as an instrument for the wage in a classical instrumental variables setup for handling measurement error. With this procedure the estimates of the intertemporal substitution elasticity are around .04 to .07, depending on specification. Estimated sampling errors are very small also, at around .07, so that substitution elasticities larger than .25 may be ruled out. In a third procedure Altonji recognizes that a contemporaneously measured alternative wage variable may be contaminated by common measurement errors or, in a model with uncertainty, correlated with labor supply function errors. Using a lagged alternative wage variable Altonji estimates intertemporal labor supply elasticities around .05, but estimated sampling errors are now around .45. All of these estimates rule out substitution elasticities greater than unity. Altonji concludes that these estimates suggest an intertemporal labor supply elasticity in the range 0 to .35, although we prefer to state all these results and their limitations so that they speak for themselves.

Abowd and Card (1983) have presented some persuasive evidence that much of the variation in both hours and earnings in the available longitudinal data may be a result of measurement error. To see the nature of this evidence suppose the level of any variable is measured with error that is serially uncorrelated. The successive changes in such a variable will have a serial correlation of -1/2 if the true variable has changes that are serially uncorrelated. The reason is that the covariance of successive changes in the imperfectly measured variable will, if the covariance in the true variable is zero, be just one-half the negative of the total variance. The serial correlation coefficients at all higher lags will be zero.

Abowd and Card present data that imply first-order autocorrelation coefficients for hours and earnings in the NLS and in the PSID data of about -.35. Neither the second nor third order autocorrelation coefficients in either data set is as large (in absolute value) as .04. Although this is hardly conclusive, it suggests the possibility that much of the panel data movement in hours and earnings may be composed of measurement error. Indeed, Abowd and Card take the null hypothesis against which they test their labor supply equations to be simple models of measurement error. Although they reject the measurement error models as a complete explanation for the data, a major message of their paper is the importance of dealing with this problem.

Altonji (1985) provides further evidence of measurement error in the main wage series used in a typical panel data study of labor supply. The change in the conventional wage measure in these studies is the change in the ratio of labor earnings to annual hours at work. For hourly workers in the PSID the change in the hourly wage rate is also recorded. It follows that the correlation between these two different measures of the wage rate provides an indication of the fraction of the variance in them which is correctly measured. If there were no measurement error at all, the correlation between the two wage measures would be perfect. If only independent measurement error were contained in the two wage measures, however, the correlation between them would be negligible. Altonji reports that these tests indicate that between 70 and 90 percent of the variance in the two wage measures may be measurement error. Such estimates may seem horrendously large, but they are entirely consistent with the PSID validation study discussed in Section 3 where it was estimated that the ratio of error to total variance in measured hourly earnings was around 75%.

These results do not imply that the wage change data contain _only_ measurement error. Still, it seems clear that the longitudinal data series available for the estimation of labor supply models are very noisy. It is possible that a more carefully measured data set would remedy this problem, although it is also possible that a different theoretical model may more appropriately describe the offered combinations of wages and hours we observe. In either case, it is clear that while the existence of panel data on hours and wages holds out the promise of "clean" and theoretically appropriate estimates of relationships that are central to the understanding of behavior and welfare, that promise has not yet been fulfilled in the United States. We believe that the villain in the piece is the presence of measurement error and the consequences of that measurement error for the particular types of questions being asked. Greg J. Duncan has emphasized to us that these labor supply studies probably place panel data in their worst possible light. Wage rates are usually constructed by dividing annual earnings by annual hours, and quite small errors in the original data can be magnified by such a procedure. The consequences of measurement error clearly depend on what is done with the data and on the type of model that we are trying to estimate. Nevertheless, even if labor supply analysis is the worst case, it is an extremely important one, and we have included the largely negative results of this section because the ability to model labor supply is frequently cited as one of the reasons for collecting panel data.

5. SUMMARY AND CONCLUSIONS

This paper has reviewed a number of aspects of the collection and use of panel data from households in developing countries. In Section 1, the sampling issues were discussed, and we concluded that there are likely to be real, if modest, benefits from incorporating some panel element into household survey data collection in developing countries. The recognition that panel data are likely to be subject to substantial errors of measurement does not, in our view, invalidate this conclusion. Section 2 discussed the measurement of income dynamics, an issue that cannot be addressed without panel data. Some of the recent American research was reviewed; the results are clearly of considerable interest and importance, and comparable work for developing countries would add an important dimension to discussions of poverty, inequality and development. Errors of measurement must also be recognized in this context. While the likely biases can be assessed at a theoretical level, it is hard to assess their empirical importance without independent sources of evidence on individual income changes. Even so, once again we feel that the recognition of the difficulties does not invalidate the genuine insights that have been obtained. It is in the third area of review, that of econometric analysis, that we are perhaps most sceptical of the real benefits of panel data. The issues are discussed in Section 3. While it is true that panel data offer the unique ability to deal with the contamination of econometric relationships by unobservable fixed effects, the presence of measurement error can compromise the quality of the estimates to the point where it is unclear whether cross-section or panel estimators are superior. This situation is in sharp contrast to that for sampling, where errors of measurement typically cannot reverse the superiority of panel over cross-section estimators. Our review of American labor supply studies using panel data illustrates both the advantages and disadvantages, though this may be an area where panel data are seen in the worst possible light.

We believe that panel data _should_ be collected in both developing and developed countries. Benefits of well-designed data collection efforts are likely to outweigh the costs. However, it is very easy to overstate the likely benefits of panel data. We have no doubt that they have contributed to our understanding of economic behavior in the past, and see no reason to suppose that new data will not do so in the future. But their existence will not solve all the outstanding problems of understanding poverty and household behavior in developing countries. They will undoubtedly bring new and important insights, but they will also bring new problems of interpretation and analysis. While as economists we are anxious to be able to work on these problems, we should not be wise to promise too much too quickly.

REFERENCES

Abowd, J.M. and D. Card. "Intertemporal Substitution in the Presence of Long Term Contracts," Working Paper No. 166, Industrial Relations Section, Princeton University, September, 1983.

Altonji, J. "Intertemporal Substitution in Labor Supply: Evidence from Micro Data," Journal of Political Economy, forthcoming, 1985.

Ashenfelter, O. "Macroeconomic Analyses and Microeconomic Analyses of Labor Supply," Carnegie-Rochester Conference Series on Public Policy, 21, pp. 117-56, 1984.

Ashenfelter, O. and J. Ham. "Education, Unemployment, and Earnings," Journal of Political Economy, Vol. 87, No. 5, Part 2, pp. S99-S116, 1979.

Bailar, B.A. "The Effects of Rotation Group Bias on Estimates from Panel Surveys," Journal of the American Statistical Association, Vol. 70, pp. 23-30, 1975.

Binswanger, H.P. "Attitudes Towards Risk: Theoretical Implications of an Experiment in Rural India," Economic Journal, Vol. 91, pp. 867-890, 1981.

Bliss, C.J. and N.H. Stern. Palanpur: The Economy of an Indian Village, Oxford, Clarendon Press, 1982.

Butz, W.P. "Why Collect Retrospective Data?" Rand Corporation, Santa Monica, California, 1981.

Cochran, W.G. Sampling Techniques (3rd Edition). New York, Wiley, 1977.

Duncan, G.J. and D.H. Hill. "An Investigation of the Extent and Consequences of Measurement Error in Labor Economic Survey Data," Journal of Labor Economics, forthcoming, 1985.

Duncan, G.J., F.T. Juster and J.N. Morgan. "The Role of Panel Studies in a World of Scarce Research Resources," in S. Sudman and M.A. Spaeth (eds.), The Collection and Analysis of Economic and Consumer Behavior Data, Champaign, Illinois, Bureau of Economic and Business Research, 1984.

Duncan, G.J. and J.N. Morgan. "An Overview of Family Economic Mobility," Chapter 1 in G.J. Duncan, Years of Poverty, Years of Plenty, Ann Arbor: University of Michigan, Institute for Social Research, 1984.

Griliches, Z. and J.A. Hausman. "Errors in Variables in Panel Data," Technical Working Paper No. 37, NBER, Cambridge, MA, 1984.

Ham, J. "Testing Whether Unemployment Represents Intertemporal Labor Supply Behavior," Review of Economic Studies, forthcoming, 1985.

Hansen, M.H., W.N. Hurwitz and W.G. Madow. Sample Survey Methods and Theory, Volume II, New York, Wiley, 1953.

Heckman, J. and T.E. MaCurdy. "A Life Cycle Model of Female Labor Supply," Review of Economic Studies, Vol. 47, pp. 47-74, 1980.

Herriott, R.A. and E.F. Spiers. "Measuring the Impact on Income Statistics of Reporting Differences between the Current Population Survey and Administrative Sources," Proceedings of the Social Statistics Section, American Statistical Association. pp. 147-158, 1975.

Hill, M.S. "Some Dynamic Aspects of Poverty," Chapter 3 in M.S. Hill, D.H. Hill, and J.N. Morgan (eds.), Five Thousand American Families--Patterns of Economic Progress, Volume 9, University of Michigan, Institute for Social Research, Ann Arbor, 1981.

Kelly, T.F. "The Creation of Longitudinal Data from Cross-Section Surveys: An Illustration from the Current Population Survey," Annals of Economic and Social Measurement, Vol. 2, pp. 209-214, 1973.

Kemsley, W.F.F., R.V. Redpath and M. Holmes. Family Expenditure Survey Handbook, Her Majesty's Stationery Office, London, 1980.

Kish, L. Survey Sampling, New York, Wiley, 1965.

Lillard, L.A., and R.J. Willis. "Dynamic Aspects of Earning Mobility," Econometrica, Vol. 46, pp. 985-1012, 1978.

Lucas, R.E. and L.A. Rapping. "Real Wages, Employment, and Inflation," in E. Phelps et al., Microeconomic Foundations of Employment and Inflation Theory, W.W. Norton, 1970.

McCall, J.J. Income Mobility, Racial Discrimination, and Economic Growth, Lexington, MA: Lexington Books, 1973.

MaCurdy, T.E. "An Empirical Model of Labor Supply in a Life Cycle Setting," Journal of Political Economy, Vol. 89, pp. 1059-86, 1981.

Mahalanobis, P.C. "Recent Experiments in Statistical Sampling in the Indian Statistical Institute," Journal of the Royal Statistical Society, Vol. 109, pp. 3236-70, 1946.

Mellow, W. and H. Sider. "Accuracy of Response in Labor Market Surveys: Evidence and Implications," Journal of Labor Economics, Vol. 1, pp. 331-44, 1983.

Miller, H.P. Rich Man, Poor Man, New York, Thomas Y. Crowell Company, 1964.

Miller, H.P. and L.R. Paley. "Income Reported in the 1950 Census and on Income Tax Returns," in An Appraisal of the 1950 Census Income Data, Studies in Income and Wealth, Volume 23, Princeton, Princeton University Press, 1958.

Neter, J. and J. Waksberg. "A Study of Response Errors in Expenditures Data from Household Interviews," Journal of the American Statistical Association, Vol. 59, pp. 18-55, 1964.

Pritzker, L. and A. Sands. "The 1950 Census and the Post-Enumeration Survey," in An Appraisal of the 1950 Census Income Data, Studies in Income and Wealth, Vol. 23, Princeton, Princeton University Press, 1958.

Sudman, S. and N.M. Bradburn. Response Effects in Surveys, Chicago, Aldine, 1974.

U.S. Bureau of the Census. The Current Population Survey: Design and Methodology, Technical Paper 40, Washington, D.C., U.S. Government Printing Office, 1978.

Wald, A. "The Fitting of Straight Lines If Both Variables Are Subject to Error," Annals of Mathematical Statistics, Vol. 11, pp. 284-300, 1940.

APPENDIX
SAMPLING RESULTS AND THE EFFECTS OF MEASUREMENT ERROR

A. Sampling with Rotating Panels

The results in this section are adapted from Hansen, Hurwitz and Madow (1953, pp. 268-72).

Assume that sampling takes place on two occasions, 1 and 2, and that in both surveys there are n households, a proportion p of which are retained from period 1 through period 2. In standard situations, the means from the various parts of the sample will be sufficient for the population means that we want to estimate.

Let x_1 be the mean in period 1 of the np overlapping households

Let x_2 be the mean in period 2 of the np overlapping households

Let y_1 be the mean in period 1 of the $n(1-p)$ -"old" households

Let y_2 be the mean in period 2 of the $n(1-p)$ "new" households

Let S_1 and S_2 be the population standard deviations of the variable sampled in 1 and 2, and let r be the population correlation between the two periods. We wish to estimate μ_1, μ_2, the two population means, and $\mu_2 - \mu_1$, the change from 1 to 2. Implicit or explicit is a weighting function on the sampling errors of the three estimates. We leave it implicit for the moment.

Consider linear estimates of the form

$$\hat{\mu}_1 = a_{11}x_1 + a_{12}x_2 + b_{11}y_1 + b_{12}y_2 \tag{1}$$

$$\hat{\mu}_2 = a_{21}x_1 + a_{22}x_2 + b_{21}y_1 + b_{22}y_2 \tag{2}$$

$$\hat{\Delta\mu} = a_{31}x_1 + a_{32}x_2 + b_{31}y_1 + b_{32}y_2 \tag{3}$$

It is required (a) that the estimators be unbiased, and (b) that the appropriately weighted average of sampling errors be as small as possible. One can also impose $a_{12} = b_{12} = 0$ if the estimate of the first mean has to be computed before the second survey has been taken. It turns out that, given p, the optimal choice of the a's and b's is independent of the weighting function. The formulae for $\hat{\mu}_2$ and $\hat{\Delta\mu}$ turn out to be

$$\hat{\mu}_2 = \frac{rp(1-p)}{1-(1-p)^2r^2} \frac{S_2}{S_1} (y_1 - x_1) + \frac{p}{1-(1-p)^2r^2} x_2$$

$$+ \frac{(1-p)\{1-(1-p)r^2\}}{1-(1-p)^2r^2} y_2 \tag{4}$$

$$\hat{\Delta\mu} = \frac{(1-p)\{1-(1-p)r^2\}}{1-(1-p)^2r^2} (y_2 - y_1) + \frac{p}{1-(1-p)^2r^2} (x_2 - x_1)$$

$$+ \frac{p(1-p)r}{1-(1-p)^2r^2} \left[(y_1 - x_1) \frac{S_2}{S_1} - (y_2 - x_2) \frac{S_1}{S_2}\right] \tag{5}$$

and these are the formulae that would be used in practice given p. Note, that although it is natural to try to optimize on p too, we cannot have different p's for different items in the survey. Since different items will have different correlations and different S_2/S_1 ratios, what is an optimal design for one item will not be so for another.

Rather than specify an explicit objective function, consider the estimation standard errors corresponding to (4) and (5). For the change, $\hat{\Delta\mu}$, in the simple case where $S_1 = S_2$, the variance is

$$\sigma_\Delta^2 = \frac{2(1-r)s^2}{n\{1 - (1-p)r\}} \qquad (6)$$

Hence, as we would expect, if $r > 0$, the optimal design has $p = 1$, that is, complete overlap, whereupon σ_Δ^2 takes its minimum value of $2(1-r)s^2/n$. It is interesting to see how much loss there is in <u>not</u> choosing complete overlap. Table A.1 gives the ratios of optimal standard errors to actual standard errors for various combinations of r and p. Clearly, completely independent surveys sacrifice a great deal of efficiency, particularly when the correlation is high. But even with 50 percent retention, the loss is never more than 30 percent and is much less for what we would judge to be reasonable values of the correlation.

TABLE A.1: **Fractions of Optimal Precision Attained by Various Designs: Difference Estimator**

r	p = 0	0.25	0.50	0.75	1.00
0.5	0.71	0.79	0.87	0.94	1
0.6	0.63	0.74	0.84	0.92	1
0.7	0.55	0.69	0.81	0.91	1
0.8	0.45	0.63	0.77	0.89	1
0.9	0.32	0.57	0.74	0.88	1
0.95	0.22	0.54	0.72	0.87	1
0.99	0.10	0.51	0.71	0.87	1
1.0	0	0.50	0.71	0.87	1

TABLE A.2: **Fractions of Optimal Precision Attained by Various Designs: Second Period Estimator**

r	optimal p	p = 0	0.25	0.50	0.75	1.00
0.5	0.46	0.97	0.99	1.00	0.99	0.97
0.6	0.44	0.95	0.99	1.00	0.98	0.95
0.7	0.41	0.93	0.99	1.00	0.97	0.93
0.8	0.38	0.89	0.99	0.99	0.96	0.89
0.9	0.30	0.85	1.00	0.98	0.92	0.85
0.95	0.24	0.81	1.00	0.96	0.89	0.81
0.99	0.12	0.76	0.98	0.92	0.84	0.76
1.0	0.00	1.00	0.88	0.75	0.63	0.50

Table A.2 turns to the precision associated with the second period $\hat{\mu}_2$. The variance in this case is given by

$$\sigma^2_{\mu_2} = \frac{S_2^{\ 2}}{n} \ \frac{1 - r^2(1-p)}{1 - r^2(1-p)^2} \tag{7}$$

a formula that holds whether or not $S_1 = S_2$. The optimal design is now given by

$$p = 1 - \frac{1 - \sqrt{(1-r^2)}}{r^2} \tag{8}$$

and these values are shown in the second column of Table A.2. Note that for $\hat{\mu}_2$, the higher the correlation, the lower the degree of desirable overlap, because repeated observations contain less information than new ones. There is therefore a potential conflict between designs for estimating the mean and for estimating the change. However, the figures in the rest of the table show that the loss of precision to an inappropriate design is typically small unless the correlation is very close to unity. In consequence, and comparing the two tables, it is clear that, unless r is near one, a rotating design can obtain a good deal of the possible gains from a full panel in estimating changes without sacrificing much precision in estimating means.

B. Panel versus Cross-Sections with Measurement Errors

It is not possible to come to general conclusions about the influence of measurement error, since results depend on the precise form taken by both the errors and the data. However, the purpose of this section is to try to argue that the advantage of panels in estimating change is unlikely to be subverted by the presence of measurement error.

Start with a cross-section in which respondent i reports a value y_i made up of an error ε_i and a "true" value x_i, that is

$$y_i = x_i + \varepsilon_i \tag{9}$$

Since the respondents in one survey are different from those in the next, the values of x_i and ε_i will be independently drawn from their respective distributions in successive surveys. Hence, write

$$\Delta\hat{y} = \bar{y}_2 - \bar{y}_1 \tag{10}$$

for the cross-section changes estimator. We then have

$$E(\Delta\hat{y}) = \mu_2 + \mu_1 \tag{11}$$

$$\text{s.e.}\ (\Delta\hat{y}) = \sqrt{\{\sigma^2_{x_1} + \sigma^2_{x_2} + 2\sigma_\varepsilon^{\ 2}\}/n} \tag{12}$$

In the panel, we recognize the time subscripts explicitly and decompose the error ε_i in (9) into a fixed effect v_i and a random effect w_{it}; hence

$$y_{it} = x_{it} + v_i + w_{it} \tag{13}$$

Note that v_i does not have a time subscript; it represents the misreporting for individual i that is common across time periods. This is not the only way of modelling the error process, but it is a reasonable one. Differencing

(13) and taking means gives the panel estimator of change $\Delta\tilde{y}$ and we have immediately

$$E(\Delta\tilde{y}) = \mu_2 - \mu_1 \tag{14}$$

$$\text{s.e. } (\tilde{\Delta}y) = \sqrt{\{\sigma^2_{\Delta x} + 2\sigma^2_w\}/n} \tag{15}$$

where $\sigma^2_{\Delta x}$ is the variance over the population of the change in x. To compare (12) and (15), note that

$$\sigma^2_\varepsilon = \sigma^2_v + \sigma^2_w \tag{16}$$

on the assumption that fixed and random errors are independent. We also have

$$\sigma^2_{\Delta x} = \sigma^2_{x_1} + \sigma^2_{x_2} - 2\text{cov}(x_1, x_2) \tag{17}$$

In consequence, a positive correlation between x_1 and x_2 is a sufficient but not necessary condition for the panel estimator to be more efficient than the cross-section estimator. Clearly, the presence of autocorrelated measurement error, like the presence of genuine correlation, enhances the attractiveness of the panel design.

C. Measuring Income Dynamics with Measurement Error

In descriptive work on income dynamics, individuals are usually traced over time as they pass through various positions in the income distribution. To show how such tracking can be affected by measurement error, consider the simplest possible Markovian model, that is,

$$x_{it} = \rho x_{it-1} + u_{it} \tag{18}$$

where x_{it} is income in period t, u_{it} is an error, and ρ a parameter to be estimated. Note that the <u>lower</u> is ρ, the <u>higher</u> is the degree of income mobility over time. Assume that the recorded observations are not x_{it} but y_{it} with

$$y_{it} = x_{it} + v_i + w_{it} \tag{19}$$

where, as before v_i and w_{it} are fixed and random effect errors of measurement.

If we use (19) to substitute for x_{it} in (18), we can calculate the effects of the errors of measurement on the estimation of ρ. For example, if ρ is estimated by $\hat{\rho}$ the OLS estimator, then it is straightforward to show that

$$\text{plim}\ \hat{\rho} = \rho\{1-\frac{\sigma_v^2 + \sigma_w^2}{\sigma_y^2}\} + \frac{\sigma_v}{\sigma_w^2 + \sigma_v^2}\ (\frac{\sigma_v^2 + \sigma_w^2}{\sigma_y^2}) \tag{20}$$

so that plim $\hat{\rho}$ is a weighted average of ρ, the true autocorrelation in income, and $\sigma_v^2/(\sigma_w^2 + \sigma_v^2)$, which is the autocorrelation of the measurement error. The weights are the proportions of the measured income variance that are attributable respectively to true income variance and measurement error. Clearly, then, mobility will be overstated or understated as ρ is greater or less than the autocorrelation of the measurement error.

DISTRIBUTORS OF WORLD BANK PUBLICATIONS

ALGERIA
Office des Publications Universitaires
1, place centrale de Ben-Aknoun
Algiers
Tel: 60-44-60, 60-54-60, 53 534

ARGENTINA
Carlos Hirsch, SRL
Galeria Guemes
Florida 165, 4th Floor-Ofc. 453/465
1333 Buenos Aires
Tel: 30-7122, 33-1787

AUSTRALIA, PAPUA NEW GUINEA, FIJI, SOLOMON ISLANDS, AND VANUATU
Info-Line
Overseas Document Delivery
Box 506, GPO
Sydney, NSW 2001
Australia
Tel: (02) 282-1614

AUSTRIA
Gerold and Co.
A-1011 Wien
Graben 31
Tel: (0222) 52 22 35

BAHRAIN
MEMRB
P.O. Box 22103
Manama Town 317
Tel: 232934

BANGLADESH
Micro Industries Development Assistance Society (MIDAS)
G.P.O. Box 800
Dhaka
Tel: 326427

BELGIUM
Publications des Nations Unies
Av. du Roi 202
1060 Brussels
Tel: 02/538.51.69

BRAZIL
Publicacoes Tecnicas Internacionais Ltda.
Rua Peixoto Gomide, 209
01409 Sao Paulo, SP
Tel: (011) 257-1640, 258-8167, 258-8442

CANADA
Le Diffuseur
C.P. 85, 1501 rue Ampère
Boucherville, Quebec
Canada J4B 5E6
Tel: (514) 641-1334

CHILE
Editorial Renacimiento
Miraflores 354
Santiago
Tel: 35244

COLOMBIA
Enlace Ltda.
Apartado Aero 34270
Bogota D.E.
Tel: 245-0645

COSTA RICA
Libreria Trejos
Calle 11-13
Av. Fernandez Guell
San Jose
Tel: 303-2858 USA

CYPRUS
MEMRDB
P.O. Box 2098
Nicosia
Tel: 21-99333

DENMARK
SamfundsLitteratur
Rosenoerns Alle 11
DK-1970 Copenhagen V
Tel: 01-35 19 42

EGYPT, ARAB REPUBLIC OF
Al Ahram
Galaa Street
Cairo
Tel: 758333

FINLAND
Akateeminen Kirjakauppa
P.O. Box 128
SF-00101
Helsinki 10
Tel: 651-1221121

FRANCE
World Bank Publications
66 Avenue d'Iéna
75116 Paris
Tel: 723-54-21

GERMANY, FEDERAL REPUBLIC OF
UNO-Verlag
D-5300 Bonn 1
Simrockstrasse 23
Tel: (0228) 21 29 40

GREECE
MEMRB
24, Ippodamou Street
Athens-11635
Tel: 7018082, 7016965

HONG KONG, MACAU
Asia 2000 Ltd.
6 Fl., 146 Prince Edward Road, W.
Kowloon
Hong Kong
Tel: 3-805294/5

INDIA
UBS Publishers' Distributors Ltd.
Post Box 7015
Ansari Road
New Delhi 110002
Tel: 273601-3
10 First Main Road
Gandhi Nagar
Bangalore 560009
Tel: 28044, 29206
Apeejay Chambers, P.O. Box 736
5 Wallace Street
Bombay 400001
Tel: 260562, 260827, 266971
8/1-B, Chowringhee Lane
Calcutta 700016
Tel: 212451, 249473
7/188, 1(CA), Swarup Nagar
Kampu 208002
Tel: 81300
Sivaganga Road
Nungambakkam
Madras 600034
Tel: 476355

INDONESIA
Pt. Indira Limited
Jl. Sam Ratulangi 37
Jakarta Pusat
P.O. Box 181
Tel: 881-018

IRELAND
TDC Publishers
12 North Frederick Street
Dublin 1
Tel: 744835

ITALY
Licosa Commissionaria Sansoni SPA
Via Lamarmora 45
Casella Postale 552
50121 Florence
Tel: 579.751/2/3

IVORY COAST
Centre d'Edition et de Diffusion Africaines (CEDA)
04 B.P. 541
Abidjan 04 Plateau
Tel: 22-20-55 and 32-60-02

JAPAN
Eastern Book Service (EBS)
37-3, Hongo 3-Chome, Bunkyo-ku 113
Tokyo
Tel: (03) 438-2611

JORDAN
MEMRB
P.O. Box 3143
Jabal
Amman

KENYA
Africa Book Service (E.A.) Ltd.
P.O. Box 45245
Nairobi
Tel: 23641/330272

KOREA, REPUBLIC OF
Pan Korea Book Corporation
P.O. Box 101, Kwangwhamun
Seoul
Tel: 723-1421

KUWAIT
MEMRB
P.O. Box 5465
Kuwait
Tel: 421843

MALAYSIA
University of Malaya Cooperative Bookshop, Limited
P.O. Box 1127, Jalan Pantai Baru
Kuala Lumpur 22-11
Tel: 565-000

MEXICO
INFOTEC
San Lorenzo 153-11, Col. del Valle Deleg. Benito Juarez
03100 Mexico City D.F.
Tel: 559-5211

MOROCCO
MEMRB
2, rue Molière Racine
Casablanca
Tel: 259461

THE NETHERLANDS
Medical Books Europe, BV (MBE)
Noorderwal 38,
7241 BL Lochem
Tel: 05730-4866

NEW ZEALAND
R. Hill and Son Ltd.
Private Bag
New Market
Auckland
Tel: 54-8119

NIGERIA
University Press Limited
Three Crowns Building Jericho
Private Mail Bag 5095
Ibadan
Tel: 0-22 411356

NORWAY
Tanum-Karl Johan, A.S.
P.O. Box 1177 Sentrum
Oslo 1
Tel: (02) 80 12-60

PAKISTAN
Mirza Book Agency
65, Shahrah-e-Quaid-e-Azam
P.O. Box No. 729
Lahore 3
Tel: 66839

PANAMA
Ediciones Libreria Cultural
Panamena, S.A.
Av. 7, Espana 16
Panama Zone 1
Tel: 23-5628

PERU
Editorial Desarrollo SA
Apartado 3824
Ica 242, OE. 106
Lima 1
Tel: 28-53-80

THE PHILIPPINES
National Book Store
P.O. Box 1934
Metro Manila
Tel: 40-82-12

PORTUGAL
Livraria Portugal
Rua Do Carmo 70-74
1200 Lisbon
Tel: 32 82 20

SAUDI ARABIA
Jarir Book Store
P.O. Box 3196
Riyadh—11471
Saudi Arabia
Tel: 4773140

SINGAPORE, TAIWAN, BURMA
Information Publications Private, Ltd.
02-06 1st Fl., Pei-Fu Industrial Bldg.
24 New Industrial Road
Singapore 1953
Tel: 283-1798

SPAIN
Mundi-Prensa Libros, S.A.
Castello 37
28001 Madrid
Tel: 276-997

SRI LANKA AND THE MALDIVES
Lake House Bookshop
P.O. Box 244
100, Sir Chittampalam A. Gardiner Mawatha
Colombo 2
Tel: 32104

SWEDEN
For single titles:
ABCE Fritzes Kungl. Hovbokhandel
Regeringsgatan 12, Box 16356
S-103 27 Stockholm
Tel: 08 23 89 00
For subscription orders:
Wennergren-Williams AB
Box 30004
S-104 25 Stockholm
Tel: 08-54 1200

SWITZERLAND
Librairie Payot
6 rue Grenus
Case postal 381
CH 1211 Geneva 11
Tel: 31-89-50

TANZANIA
Oxford University Press
P.O. Box 5299
Dar es Salaam
Tel: 29209

THAILAND
Central Department Store
306 Silom Road
Bangkok
Tel: 233-6930-9

TRINIDAD & TOBAGO
Systematics Studies Unit
55 Eastern Main Road
Curepe
Trinidad, West Indies
Tel: 1-(809) 662-5654

TUNISIA
Société Tunisienne de Diffusion
5 Avenue de Carthage
Tunis
Tel: 255-000

TURKEY
Haset Kitapevi A.S.
469, Istiklal Caddesi
Beyoglu-Istanbul
Tel: 144 94 70/71

UGANDA
Uganda Bookshop
P.O. Box 7145
Kampala
Tel: 43756

UNITED ARAB EMIRATES
MEMRB
P.O. Box 6097
Sharjah
Tel: 592233

UNITED KINGDOM AND NORTHERN IRELAND
Microinfo Ltd.
P.O. Box 3
Alton, Hampshire GU 34 2PG
United Kingdom
Tel: Alton 86848

VENEZUELA
Libreria del Este
Aptdo. 60.337
Caracas 1060-A
Tel: 32 23 01/33 26 04

WESTERN SAMOA
Wesley Bookshop
P.O. Box 207
Apia
Tel: 22 822

The World Bank

Headquarters
1818 H Street, N.W.
Washington, D.C. 20433, U.S.A.

Telephone: (202) 477-1234
Telex: WUI 64145 WORLDBANK
RCA 248423 WORLDBK
Cable Address: INTBAFRAD
WASHINGTONDC

European Office
66, avenue d'Iéna
75116 Paris, France

Telephone: (1) 47.23.54.21
Telex: 842-620628

Tokyo Office
Kokusai Building
1-1 Marunouchi 3-chome
Chiyoda-ku, Tokyo 100, Japan

Telephone: (03) 214-5001
Telex: 781-26838

Ashenfelter and others