Measuring Unemployment and the Informal Sector
Some Conceptual and Statistical Problems

Manfred Bienefeld and Martin Godfrey

When research workers collect statistics they tend to be taking photographs. Statisticians are in the business of making moving pictures and for this reason their work has infinitely greater potential but is also of much greater difficulty. To add to the inequality of the situation, the researcher is generally aiming to throw light on a particular and well defined set of problems, while the statistician has to be all things to all people.

What follows may appear to be simply another futile exercise by 'unrealistic' research workers, giving advice without any consideration or understanding of the real limitations placed on statisticians in their work. To some extent this charge is warranted, but, given that practical contraints are different at different times and in different places and that as research workers we can hardly define what these effective limits are, it is inevitable that this should be the case.

Furthermore, discussion of the kind presented in this paper can be justified by the need to obtain clarity about the type of information that ought ideally to be collected. In other words, it is essential to be clear about the theoretical framework within which one is identifying the problems to be studied and from which one derives the statistical measures to illuminate those problems.

Ultimately the question which we want to address is: What do we want to know about the problems of unemployment or about the informal sector? But, beforehand, we need to attempt to answer the question: *Why* do we want to know? Because the answer to this question largely determines what it is we want to know. In this paper we propose, first, to look historically at this connection between why people were interested in the issue under review and what information they collected as a result.

The essential issue is that of 'unutilized labour power' and historically this issue has always been a focus of concern for those interested in underdevelopment. But at different periods of time this interest has derived from very different perspectives.

There are three main reasons why one might be interested in the problem of unutilized labour power. One might see it (i) as a deplorable waste in the context of a concern with **growth** and economic development generally; (ii) as the critical issue which identifies, i.e. quantifies and locates. the **suffering** inflicted by the polarizing tendencies of capitalist development—suffering which results, in varying combinations, both from poverty due to low income and from the demoralization and indignity of having no productive role; and (iii) as a potential source of political **instability** or social unrest. Clearly, each of these concerns has its own statistical requirements and. even though the statistician is too often required to illuminate the problem from all perspectives, it is nevertheless useful to consider each of these issues separately[1] and to discuss them in conjunction with the shifts in emphasis which have taken place in both the analytical and the policy-oriented discussions of this subject.

Since economists have played an inordinately large role in the development debate, it is hardly surprising that the problem in question has frequently been seen from what we have termed a growth perspective. Initially many of the people who saw the problem from this perspective regarded the 'labour reserve' in a very optimistic way. They saw unutilized rural labour as a major idle resource the mobilization of which would and could overcome the problems of chronic poverty. The missing ingredient in this recipe was capital, and it followed that if only enough capital could be made available then development would follow.

The statistical implication of this approach was an intense concern with estimating the size of this potential productive resource. The question which was most often asked in this connection was: How many people could be removed from the land without reducing production of food. or sometimes, without reducing consumption of food?[2]

The measurement of the size of the labour reserve proceeded on the basis either of man-hours worked or of measures of productivity per head.

1 We will not, however, go into the statistical implications of the third reason for concern with unutilized labour, i.e. fear of political instability, since we do not share the view (which usually underlies it) that political stability at all costs and under all circumstances is a condition of development. It cannot be denied that, *ceteris paribus*, political stability is preferable to political instability, but, equally, there are situations where a period of political instability is a prerequisite of development.

2 See Robinson (1969) for a survey of this literature and bibliography.

Both procedures were sometimes used in ways which bordered on the irresponsible. When the measure was based on man-hours it was usual to estimate the total number of man-hours expended, and to calculate how many full working years that would make (incredibly, often without allowing for seasonality, while allowing for days off for weekends and public holidays). The figure so derived was then assumed to represent the number of fully employed rural workers who would be required to produce the current level of output. The difference between this number and the total labour force was therefore deemed to be the measure of the labour surplus available for 'development'.[3]

When the measurement was done on the basis of productivity, credibility fared little better. Here it was usual to assume the possibility of achieving some level of productivity per acre (often on the basis of some supposedly plausible comparison with arbitrarily selected countries), then to assume some plausible man-land ratio, and thence to determine how many people would be required to cultivate the acreage necessary to produce the current food supply at the assumed 'fictitious' level of productivity per acre.

Certainly in retrospect both of these approaches appear quite arbitrary but, more serious, the concern itself appears extremely sterile and uninteresting. This is the more surprising since these discussions occurred at a time when colonial governments were still confronting their primary problem with respect to the labour market, which was a perpetual labour shortage.

This unreality was significantly reduced when a new group of analysts led by Lewis (1954) and Nurkse (1955) addressed themselves to an analysis of the process by which this hypothetical surplus labour would be transferred to productive activities.

While Nurkse focused his attention on the difficulties of transferring the labour in question to new activities and especially on the related difficulties of maintaining and distributing agricultural production in the face of such transfers, Lewis concentrated his analysis on the way in which the market mechanism could be expected to respond to this situation. It was partly because of Lewis's faith in the market mechanisms and partly because Lewis seemed to provide a 'solution' that his model received a disproportionate amount of attention.

The Lewis model was basically optimistic in that it specified the mechanism, provided by the market, through which labour would be released from subsistence production in a way which would just meet the needs of employers. It was argued that this would happen because the wage is the supply price of labour and because under conditions of 'over supply' the wage would be fixed at a level which was just high enough to bring the necessary quantities of labour into capitalist production. If nothing else, this was a reaffirmation of faith in the existence of economic man. It was a dramatic departure from the backward-sloping supply curve of labour which had for so long been the rationale given by successive colonial administrations for forcing labour out of the subsistence sector by any means except providing them with higher wages and better conditions in their new employment! Unfortunately, for all its optimism, the 'fine tuning marginalism' that this position implied has not been justified in practice. Subsistence sectors have not remained passively waiting for their labour to be drained off through marginally attractive wage opportunities, and in the capitalist sectors the process of wage determination has proved to be a good deal more complex than the model suggested.

One should be careful not to accuse Lewis of all the sins committeed by those who used his model in ways which are basically departures from it. Fei and Ranis (1964) for instance have used the framework to derive conclusions about movement between the food and non-food sectors and by implication between rural and urban areas, whereas Lewis's distinction was between subsistence and capitalist production. An implication of the Fei/Ranis framework is that, in the absence of a 'dualistic landlord' on the lines of nineteenth-century Japan, the state must squeeze surplus out of rural areas by holding down rural wages and encouraging increases in productivity; action to restrain urban wage increases might also be necessary.[4] Indeed it is reasonable to presume that one of the reasons for the great popularity of this framework has been its emphasis on low wages. The model argues that low wages will lead to relatively high rates of growth and thus of employment creation. Unfortunately the beneficial utilization of the surplus generated in this way cannot be guaranteed.

Statistically this theoretical model has led to a concern with rural-urban income differentials. But the problems of measurement have proved to

3 See Godfrey (1967) for a more detailed description and critique of this procedure.

4 As Weeks (1971) points out, the model is "a marginalist translation of the Marxian model of capitalist development through 'primitive accumulation' ".

be very great indeed. The difficulties of measuring rural incomes, of measuring urban non-modern-sector incomes and, more especially of measuring both of these and modern-sector urban incomes on a comparable basis, are well known. Apart from the fundamental question of variations in the relationship between income and welfare, there are problems with the valuation of amenities and services provided in different areas, the collection of information on income apart from official wage payment, and finally the definition of the income earning unit (and the migrating unit).[5] In addition to such problems of measurement there were problems in utilizing the information thus gained analytically. Thus even if impeccable information were available, any attempt to relate labour supply to levels of differentials would require extensive disaggregation, since the relevant differentials for any particular group with any particular set of employment prospects would have to be defined and used as the operative differential.

For instance, the relevant differential for the poorest smallholders and landless labourers might be that between themselves and the lowest reaches of the urban informal sector; the more prosperous smallholders, on the other hand, might be responding, in their decisions to equip their children for migration by schooling, to the very high salaries available to the most educated in urban areas. Furthermore, the relationship between any particular differential and labour supply will inevitably change over time, will be different in the face of different levels of income both at the rural and at the urban end and, more awkwardly still, will be different given different patterns and rates of change of income levels at either the urban or rural end.

In the midst of these struggles to come to grips with the statistical implications of the Lewis model there occurred a fundamental change in the perspectives through which the problem was viewed, a change in which all three of the reasons for concern mentioned above—growth, welfare and political stability—played a part. Instead of a chronic labour shortage and a concern with ways of prying labour out of the 'traditional' sectors, it suddenly appeared in the early 1960s that labour was leaving these sectors too rapidly. Urban areas were growing more rapidly than employment opportunities and it appeared that the rural underemployment which had been the focus of attention for so long was being transformed into urban unemployment. It was thought that this represented a net loss in terms both of growth and of welfare. And governments became aware of the potential political dangers of growing urban unemployment.

Although in this new situation attention continued to focus on the process by which labour was drawn from the traditional subsistence/rural sectors, there was now in addition a growing concern with the unemployment which was appearing in these economies.[6] Pessimistic arithmetic was ready to hand. With existing capital stocks even the most optimistic assumptions about capital-output and capital-labour ratios could be shown to be incapable of generating quantities of investible surplus and of employment which would absorb more than a fraction of the labour becoming available. As a result, concern with unemployment increased, and estimates of urban unemployment escalated to 25 per cent or more (Singer 1970) until Weeks (1974) showed that, using the procedures of the estimators, 100 per cent of people in the Third World would soon be unemployed.

Statistically this new perspective quite naturally led to an emphasis on the collection of unemployment statistics. Initially these statistics were collected mostly by labour-force surveys using the traditional and internationally accepted definitions of unemployment which are intended to identify workers not receiving an income who are both able and willing to work at any one point in time. To the surprise of some, the figures which emerged showed relatively modest levels of unemployment in many Third World urban areas. An explanation was urgently required.

Explanations centred quite correctly on the definition of unemployment. In many parts of the Third World unemployment is a luxury few can afford. In the absence of unemployment benefits and most other forms of social security there are powerful incentives to induce anyone without a regular source of income to earn some income even of an irregular kind. Nevertheless, it was generally agreed that the finding of relatively low levels of unemployment did not mean that a problem of unutilized labour did not exist. It was necessary under these circumstances to redefine the problem being examined.

Some analysts (Frank, 1968) simply argue that all those without regular wage incomes derived from the modern sector might as well be classified as unemployed. Others (ILO Colombia Report, 1970, Turnham and Jaeger, 1971) suggested the use of a 'minimum income' standard below which a person should be designated unemployed. This

5 See Knight (1971) for a discussion of these problems.

6 See particularly Todaro (1968) and Harris and Todaro (1970).

somewhat misleading use of language never found general acceptance and most analysts reverted to the idea of disguised unemployment or underemployment.

Unfortunately the usage of these concepts under these circumstances was not warranted, nor was it analytically useful. In fact it was a reversion to the rather sterile debates about the size of the underemployed labour pool, but this time in an urban rather than a rural setting. The analytical difficulties with the concept stem from the fact that the very concept of underemployment implies a concept of full employment. Under the circumstances the definition of full employment was arbitrary and any conclusions useless.

The term 'disguised unemployment' was first used by Joan Robinson (1937) in a paper dealing with the economic problems raised by the 1930's depression in the UK. Under those circumstances the concept was both valid and usable because the norm against which it compared the situation under discussion was clear and definable. The hours, or the income, or the efficiency of depression employment were merely being compared to the former employment of the same individual. When transposed to the underdeveloped world and used in an open-ended manner this clarity and usefulness disappeared.

The difficulty lies in the choice of a norm. Everyone everywhere could be defined as underemployed if the standard of comparison were not specified, since under any circumstances some improvement through technical change or the reorganisation of work is always possible. Yet whatever standard is chosen is inevitably arbitrary unless a precisely and comprehensively defined programme of labour mobilization is contemplated. In the absence of such a point of comparison the concept turned out to have little meaning for economic analysis. Basically the economist is prepared to deal only with existing labour supply or labour supply that may be generated through increased wages. The psychological preferences as well as the other conditions which underlie this relationship are generally deemed to lie outside the limits of his discipline.

So the problem was that just because people worked less than some arbitrary number of hours per day, or per week, or per month, or earned less than some arbitrarily determined level of income, this did not indicate the existence of a pool of surplus labour. Obviously it was possible that they did not wish to work longer or more intensively than they were doing at present. The economist's fall-back position was to say that the only person who could truly be said to be underemployed was one whose contribution to production was nil (sometimes wrongly referred to as a situation where marginal product was zero).[7] Unfortunately this 'escape' from the dilemma was more apparent than real. Quite apart from the fact that this position was not plausible, it also required the equally unwarranted assumption that the withdrawal of some people's labour would be costlessly compensated by the greater efforts of those remaining.

Statistically this new-found concern with the size of the urban labour surplus had two major effects. On the one hand those who insisted on attempting to quantify the amount of underemployment returned to the basically arbitrary and sterile calculations which were familiar from the early days when such calculations were made for the rural labour force (Ray, 1966). On the other hand the more rigorous economists concerned themselves with the collection of data to be used in connection with (generally linear) production functions relating output to capital, labour and other variables such as land and fertilizer use. These exercises generally cast grave doubt on the assumption of marginal labourers with zero product, though a few of them suggested that under conditions of family labour such a possibility might exist (Desai and Mazumdar, 1970).

The shift from this sterile and involuted debate to the discussion of the 'informal sector' was a welcome relief and a considerable advance. This new formulation entailed a renunciation of the morbid fascination with the measurement of the exact size of the unutilized labour pool, and focused attention on the real issue, which was the role played by the small-scale, less-regulated producers in the process of accumulation and distribution, as well as an analysis of that process itself.

Unfortunately the definition of the 'sector' is very loose and inexact with the consequence that the statistical implications are not at all clear. The reason for this is that the sector is defined to include heterogeneous sets of activities and of people, whose definition is imprecise and who have no identifiable, analytically useful common characteristics.

Indeed, attempts to define the sector *a priori* may well be futile. It may be preferable to follow Gerry's example (1974), and to use a "framework in which the productive ensemble is viewed as an ensemble" focusing on "the relationships between the different elements of the ensemble,

7 See Sen (1962) for a clarification of this distinction.

rather than stress the mutually exclusive characteristics of one element *vis-à-vis* another". In which case an important purpose of collecting statistics would be to throw light on the nature of these relationships.

There are two views about this. Those who recommend 'increasing links' (e.g. the ILO Kenya Report, 1972) implicitly assume that the relationship between large and small firms is a benign one. Others (e.g. Leys, 1973, Bose, 1974) take a more pessimistic view of the relationship as essentially exploitative. Small producers, it is suggested, (a) produce at very low prices, under pressure of extreme competition, inputs for large producers; (b) produce, again at low prices and under pressure of extreme competition, wage-goods and services for the employees of the large producers, enabling wages to be lower than they otherwise would need to be, thus in another way transferring surplus to large producers; (c) possibly perform a reserve-army role holding down the general level of wages; (d) serve only 'residual' markets in which large firms are not currently interested—as soon as small producers have developed a market to the extent that it is of interest, then large firms take it over.

Moreover, whatever one's primary focus of interest, whether the productive potential of producers, the income distribution implications of the present situation, or the nature of the linkages between the smaller scale producers and the modern sector, the so-called 'informal sector' includes a large variety of people and activities situated very differently in relation to each of these issues. It is therefore essential that the sector under discussion be substantially disaggregated in such a way that its components become analytically significant and that each can be defined in a way which is statistically useful.

In disaggregating this concept a number of principles should be observed. First, it is necessary to be clear whether one is classifying activities or people. Since at present we are discussing issues from a growth perspective it seems advisable to begin with the disaggregation which is based on separating different sets of activities. Here it is almost certainly advisable to begin by making three major distinctions: between activities which (i) produce tradeable commodities; (ii) involve the production of services connected with distribution and finance, and (iii) involve personal services.

In spite of the fact that orthodox economists are generally averse to such distinctions it seems important to isolate in this way activities which are necessarily derivative from production elsewhere, namely the distribution and financing of activities carried on elsewhere or the provision of personal services to those employed elsewhere. Of course, many of the activities producing saleable commodities will also be derivative in the sense that they are closely tied to various kinds of production in the modern sector, but it seems essential to make this distinction in order to seek out those parts of the small-scale productive sector which have the opportunity to expand in a more independent manner. This clearly means that in practice the three divisions suggested here need to be even further disaggregated, but the precise nature of that disaggregation is not likely to be generalizable for all situations.[8]

Ultimately this study of a productive process focuses on the accumulation of capital and of productive resources, and on technical change and the related generation of employment opportunities. If such an approach were to be taken seriously it would imply the abandonment of the all too prevalent mailed questionnaires, and the establishment of groups within the statistical offices who would monitor, on a continuous and permanent basis, the growth and development of particular sectors of the economy.

How carefully and how precisely this can be done clearly depends on the size of the economy and the number of establishments to be covered, as well as the size and resources of the statistical office in question. In this way, and in this way alone, is it possible to come to terms with the need to understand the processes of change which are generating the problems in which we are interested.

In practice the kinds of statistics that such monitoring would generate would include, in order of rough priority, statistics on levels of employment (regular, casual, unpaid), length of time in business, value of output, some information on capital equipment in use, estimates of profit and of value added, prices paid for standard inputs, prices received for standard outputs (where possible), information on labour force structure (sex, age, wages, seniority in job and firm), as well as information on the types of markets served (production to order, production for sale to own retail outlet, production for sale to commercial buyers). While this is not intended to be a comprehensive list it is suggested that these variables would be amongst the most important which would need to be collected on a regular basis.

In most cases it is advisable that any such system

[8] See Bienefeld (1975) for a discussion of these issues in relation to Tanzania.

be established only after a comprehensive household survey has sought to establish the dimensions and scale of various types of activities not normally included in establishment surveys. Once such procedures were operating it is almost certain that in most areas of production, teams which were constantly in touch with the producers themselves could maintain relatively accurate records indicating the disappearance of old firms and the appearance of new ones.

Statistics of this type would present a picture of the structure of the economy and would in time reveal the structural shifts occurring in the economy and allow one to assess their employment implications, especially the impact of new entrants to the market, as well as possible externalities arising out of certain types of investment.

The resource implications of such a change in statistical policy cannot, of course, be ignored. We realise that statistical officers in many countries are already overworked. Yet the problem should not be exaggerated. For one thing we are recommending the new approach as a **substitute** for, not an addition to, time-consuming existing practice. For another, serial sampling of varying intensity could considerably reduce the time needed for such investigations. After all, the most illuminating of our insights into the operation of the informal sector have emerged from case-histories of a relatively small number of operators (e.g. Gerry, 1974; Bose, 1974; King, 1974).

If our interest is in growth, then we need to try to understand the process of accumulation and technical change if we are to understand the process of change which leads to the non-intensive utilization of much of the available labour.

On the other hand, if we are primarily interested in welfare, the object of statistical investigation must be the household, since this is the primary income sharing mechanism, and the operative distinction has to be one based on poverty levels or on income levels. This means of course dropping the formal/informal sector distinction since individuals within the same household may work in different sectors of the economy. The concept of the informal sector. disaggregated in the manner described previously, must be reserved for productive activities.

Practically, focusing on households must begin with the definition of the household. This is no easy matter and even the most commonly used criterion, i.e. eating together, brings with it difficulties in situations where membership of households is often relatively fluid and where the degree of economic involvement within the household varies. Once defined, the information to be obtained from each household includes composition and changes in composition over time, income-earning activities, other economic activities, and consumption patterns and levels. In this context also, some attempt can be made to measure and classify various types of labour which might reasonably be considered as an existing labour surplus. It is clear from this description that periodic household surveys, including household budget surveys, are essential.

Although we are now discussing the statistical implications of a concern with welfare we should not lose sight of the need to make a connection between activities and households. For instance, if those monitoring the productive activities of particular sectors of the economy could also be involved in laying the basis for small-scale sample surveys of the households of those engaged in these activities, this would help to make the connection between sectoral changes in production and changes in welfare. (In the case of most small-scale agriculture this question would not arise since the household is not only the income-sharing mechanism but also the unit of production).

The need to separate the interest in productive potential from the interest in welfare and to focus on units of production in the first case and on households in the second becomes very clear when one considers some recent interesting studies concerning the poor in various Third World urban areas. These suggest a need in general to avoid premature typologies of the informal sector and particularly to avoid jumping to conclusions about the causes of poverty. It is possible, for instance, that in some cases employment in the informal sector may at one and the same time be very low-paid, very unstable, and generally an inadequate source of income, but make a positive contribution to the welfare of the poor. This would occur when the majority of participants in informal activities were young people, old people or people who otherwise would for one reason or another be excluded from full economic activity.[9] At the same time the majority of formal-sector income-earners might be heads of households. In these circumstances an automatic characterization of the poverty problem as one

9 Indeed, there is evidence that those males of prime working age who are involved in informal sector activities earn comparable or higher incomes than their counterparts (standardized for age, education) in the 'formal' sector. This is the strong conclusion of a study of Brazil appearing in the **Journal of Developing Areas**, and it is suggested in some Tanzanian data (e.g. Bienefeld 1974).

of low rewards in the informal sector—with rewards in the formal sector, being regarded as if anything 'too high'—would clearly be fallacious.

Only by separating initially the analysis of activities from that of households and later carefully making the connection between them can we hope to understand the process by which poverty is generated.

References

Bienefeld, Manfred (1974) The Self Employed of Urban Tanzania, **IDS Discussion Paper** No. 54, IDS Sussex.

Bienefeld, Manfred (1975) The Informal Sector and Peripheral Capitalism: the Case of Tanzania, **IDS Bulletin,** Vol. 6 No. 3, February, 1975.

Bose, A. N. (1974) **The Informal Sector in the Calcutta Metropolitan Economy,** ILO, World Employment Programme, Geneva.

Desai, M. and Mazumdar, D. (1970) A Test of the Hypothesis of Disguised Unemployment, **Economica,** Vol. XXXVII.

Fei, John C. H. and Ranie, Gustav (1964) **Development of the Labour Surplus Economy,** Irwin, Homewood, Illinois.

Frank, C. R. (1968) Urban Unemployment and Economic Growth in Africa, **Oxford Economic Papers,** Vol. 20 No. 2.

Gerry, Chris (1974) **Petty Producers and the Urban Economy: a Case Study of Dakar.** ILO World Employment Programme, Dakar.

Godfrey, E. M. (1967) Measuring the Removeable Surplus of Agricultural Labour in Low-Income Countries, **Journal of Economic Studies,** Vol. 1, No. 2.

Harris, J. R. and Todaro, M.P. (1970) Migration, Unemployment and Development: a Two-Sector Analysis, **American Economic Review,** March.

International Labour Office (1970) **Towards Full Employment: A Programme for Colombia,** Geneva.

International Labour Office (1972) **Employment, Incomes and Equality: a Strategy for Increasing Productive Employment in Kenya,** Geneva.

King, Kenneth (1974) Kenya's Informal Machine Makers: a Study of Small-Scale Industry in Kenya's Emergent Artisan Society, **World Development,** Vol. 2, No. 4/5.

Knight, J. B. (1971) Measuring Rural-Urban Income Differentials, paper presented to conference on urban employment in Africa, IDS Sussex.

Lewis, W. Arthur (1954) Economic Development with Unlimited Supplies of Labour, **Manchester School,** Vol. 26, pp. 1-32.

Leys, Colin (1973) Interpreting African Underdevelopment: Reflections on the ILO Report on Employment, Incomes and Equality in Kenya, **African Affairs,** Vol. 72 No. 289.

Nurkse, Ragnar (1955) **Problems of Capital Formation in Lesser-Developed Areas,** Blackwell, Oxford.

Ray, Robert S. (1966) **Labour Force Survey of Tanzania 1965,** Ministry of Economic Affairs and Development Planning, Dar es Salaam.

Robinson, Joan (1937) **Essays in the Theory of Employment,** Macmillan, London.

Robinson, W. C. (1969) Types of Disguised Rural Unemployment and Some Policy Implications, **Oxford Economic Papers,** Vol. 21 No. 3.

Sen, A. K. (1962) **Choice of Techniques: an Aspect of the Theory of Planned Economic Development,** Blackwell, Oxford.

Singer, H. W. (1970) Brief Note on Unemployment Rates in Developing Countries, **Manpower and Unemployment Research in Africa,** Vol. 3, No. 1.

Todaro, Michael P. (1968) The Urban Unemployment Problem in less Developed Countries: an Analysis of Demand and Supply, **Yale Economic Essays.**

Turnham, David and Jaeger, Ingelies (1971) **The Employment Problem in Less Developed Countries,** OECD Development Centre Studies, Employment Series No. 1, Paris.

Weeks, John (1971) The Political Economy of Labour Transfer, **Science and Society.**

Weeks, John (1974) A Brief Note on the Unemployment Crisis in Poor Countries, **Manpower and Unemployment Research in Africa,** Vol. 7, No. 1.