Reprinted from the
International Labour Review
Vol. 103, No. 6, June 1971

On Measuring the Social Opportunity Cost of Labour©

A. C. HARBERGER [1]

IN THIS PAPER I attempt to examine in some detail the commonly held notion that the opportunity cost of labour is represented by the product that is forgone from other activities as a consequence of its being labour for a given activity. The first variant of this notion that will be treated is the idea that in some poor countries the pool of labour in the agricultural sector is so abundant, and its marginal product so low (effectively zero, according to this idea), that other sectors can expand their demand for labour without entailing any significant loss in production elsewhere. The second variant to be considered is the idea, less restrictive than the first, that the product forgone in other sectors (in this case not necessarily zero or insignificant) is the appropriate measure of the social opportunity cost of labour. I shall argue that the data seem to contradict the idea that great masses of labour can be withdrawn from the agrarian sector without a palpable loss in product. I shall also contend that the use of forgone product as a measure of opportunity cost is an oversimplification, which can lead an analyst to wrong conclusions in a number of different ways. Far preferable is the concept that the supply price of marginal units of labour of given skill characteristics for given jobs in given labour-market areas is the relevant measure of social opportunity cost. But even this concept has some deficiencies, and I shall present analyses indicating that the true measure of social opportunity cost lies somewhere between the measurable supply price of labour and the market price actually paid in a given activity. In the course of this analysis, I shall discuss how to deal with the phenomenon of unemployment (of the chronic disguised, and the cyclical varieties) in evaluating labour's social opportunity cost. Finally, I shall comment briefly on the use of macro-economic approaches to derive measures of this opportunity cost.

[1] Professor of Economics, University of Chicago. Professor Harberger originally submitted this paper to the ILO Meeting of Experts on Fiscal Policies for Employment Promotion, Geneva, 4-8 January 1971.

The idea of zero marginal product

The idea that labour has an effectively zero marginal productivity in the agrarian sector of densely populated underdeveloped countries is widespread, and it is reflected in a number of different branches of the literature of economics. Cost-benefit analyses that take the benefits of a project to be the value of its output, and its costs to be the capital and materials outlays involved, implicitly (by failing to deduct anything in the way of labour costs) assume that labour's opportunity cost is zero. Likewise, models of economic growth (of which there are many, at varying levels of mathematical sophistication) which explain that growth solely as a function of capital accumulation fail to assign any role to increases in the labour force and hence implicitly assume that labour, in the relevant economic sense, is in oversupply. Similarly, many authors who use the term " disguised unemployment " in describing the labour-market situation of less developed countries have in mind the idea that workers can be drawn into productive employment from the pool of disguisedly unemployed with no loss in output in the sectors where they were formerly occupied.

I shall defer for the time being the detailed discussion of the phenomenon of abnormal open unemployment, and confine myself to the case where open unemployment is within the limits set by normal job turnover, migration, seasonal variations in demand, strikes, etc. The terms " frictional " and " seasonal ", as distinct from " chronic ", " structural ", and " cyclical " are commonly used to describe the types of unemployment that can be thought of as prevailing in the economic situations that will be discussed at this point.[1] Cases where unemployment is significantly in excess of the limits dictated by frictional and seasonal considerations will be discussed later.

My discussion will be based on the general economic principle that employers do not wittingly pay workers more than they (the employers) believe the incremental contribution of each worker to the value of output to be. This principle has an automatic implication: in a market economy in which most labour consists of hired workers, as distinct from family members, the marginal product of the labour actually employed will equal or exceed the wage it is paid. The wage may be set—by law, by collective bargaining, or even perhaps by custom and tradition—at a level higher than that which would cause employers to employ the full labour force, but this would lead to open unemployment above and beyond the normal frictional and seasonal amount, and that in turn would present us with a case whose discussion has explicitly been deferred to a later point in this paper. Hence if we have a market economy, with

[1] Seasonal and frictional unemployment are generally considered to average about 4 per cent of the labour force in the United States and may amount to as much as 6 or 7 per cent of the labour force in the urban labour markets of some less developed countries.

labour consisting mainly of hired workers, and without abnormal open unemployment, we can conclude that the marginal product of labour equals or exceeds its wage.

We may postulate, then, that cases where the marginal product of labour actually employed is below the going wage will be those where the marginal workers are family members, not hired hands. It is not surprising, therefore, that many of the writings based on the assumption of zero marginal product of labour in agriculture have referred to countries like India and its neighbours, where the extended-family system is a deep-rooted part of the culture. The extended-family system in effect requires a householder to take in and care for as a member of his own family any of a broad class of relatives who may appear on the scene. They, in turn, are required to behave as family members, contributing to the operation of the household either by finding an outside job or—what is most relevant in this case—by working on the family farm. If they cannot find an outside job and do end up working on the farm, they may, of course, simply displace some hired labour that the landowner was already using. In this case their marginal product would still equal or exceed the wage. However, if their landowner relative was originally hiring no labour, or if the number of extended-family members he was called upon to take in were to exceed the number of workers he otherwise would have hired, then it is possible that the marginal product of labour on his farm might fall below the wage level, and perhaps even come close to zero.

That this sort of phenomenon should be widespread in a country like India is certainly, on the surface, plausible. After all, well over 300 million people live in the rural sector of India, and the absorption into productive employment of the able-bodied labourers in such a large mass of people certainly appears to be a staggering task. Yet, initial appearances to the contrary, the Indian rural sector seems to have been quite successful in accomplishing that feat. Not only that, but the marginal productivity of labour in rural India appears to be closely approximated by the market wage.

I have two pieces of evidence to offer in support of this contention. The first and most important is the fact that successive surveys of the rural labour force in India reveal that some 25 per cent are landless workers who make their living as hired workers. This is precisely the group that would tend to be most disadvantaged under the extended-family system: they are the hired hands who would be displaced as relatives were absorbed on the family farm. And only after all hired workers of given skill qualifications were displaced on a given farm would the marginal productivity of that type of labour possibly be pressed below the going wage (in the case that the number of relatives who had to be absorbed exceeded the number of hired hands who were let go). The whole process by which the extended-family system leads to relatives being cared for would work to the detriment of hired hands. Thus it is

indeed surprising to find that the average landless agricultural labourer works some 250 full-time equivalent man-days per year, and that his annual earnings are approximately equal to the per capita income of the country.[1] Of course one must realise that annual earnings equal to a country's per capita income are likely to be well below the average level of earnings per worker in that country. None the less, though low, these earnings levels are far from negligible. In fact, the average earnings from all sources of members of the labour force in United States agriculture are also approximately equal to that country's per capita income. No one would dream of asserting that the marginal product of labour in United States agriculture was zero—and the above evidence indicates that such an assertion is equally inapplicable to India.

A second piece of evidence negating the idea of a zero marginal productivity of labour in Indian agriculture comes from Professor K. N. Raj's study of the great Bhakra Nangal Dam project.[2] In discussing the recruitment of labour for this project, Professor Raj states—

> It is a common assumption to make, on theoretical analyses, that the supply of unskilled labour in underdeveloped countries is almost infinitely elastic. This, obviously, is not always true. For instance, the additional demand for unskilled labour created by the Bhakra Nangal project, even at the peak level of activity during the construction of the canals, cannot be regarded as very large, when considered with reference to the investment undertaken or the area (and population) over which the construction work was spread. Yet the supply of unskilled labour from the areas adjoining the work proved hardly adequate. In March 1954, when the total number employed on the Bhakra canals was around 100,000, it would appear that as much as 60 per cent of the labour required had to be imported from other states.[3]

This result was observed in spite of the fact that the wages paid were comparatively good (typically, Rs. 2 per day plus free housing). It is obviously hardly consistent with the notion of a large pool of labour having zero marginal productivity, which can readily be drawn upon to meet additional labour requirements.

If Indian agriculture is not characterised by zero (or negligible) marginal productivity of labour, it seems hardly likely that the phenomenon is widespread in the rest of the underdeveloped world, in which incomes tend to be substantially higher and population densities substantially less than India's. Indeed, one might say that the zero marginal product hypothesis was a " straw man ", and that my efforts to refute it were unnecessary, if it were not for the fact, cited at the outset of this section, that the assumption is so widely made, explicitly or implicitly, in growth studies, project analyses and other economic writings.

---

[1] See Government of India, Cabinet Secretariat: *Wages, employment, income and indebtedness of agricultural labour households in rural areas*, National Sample Survey No. 33 (Delhi, 1960), pp. 18, 23 and 61.

[2] K. N. Raj: *Economic aspects of the Bhakra Nangal project* (Bombay, Asia Publishing House, 1960).

[3] Ibid., pp. 77-78.

I conclude, then, that the best approximation to the marginal product of rural labour is the going market wage applicable to each skill category, and that the social opportunity cost of such labour, in the absence of substantial open unemployment, is also best measured by the going wage.

Forgone product as a measure of social opportunity cost

The tendency for migration to flow in a more or less steady stream from rural to urban areas appears to be almost universal, and it has led some observers to conclude that, when new jobs are created in urban areas, they are filled, directly or indirectly, by migration from rural areas. The process is direct when a recent migrant himself takes the new job, and indirect when someone else takes the new job, leaving a vacancy which in turn is filled by another person, and so on until the last vacancy in the chain is occupied by a recent migrant. The conclusion that is normally drawn is that, at least for unskilled urban jobs, the relevant measure of social opportunity cost is the marginal product of agricultural labour in rural areas. This, it is said, is what is forgone when a worker moves from the country to the city. Implicit in the reasoning, of course, is the idea that the creation of a new job has the effect of inducing additional migration.

Plausible though it sounds, I believe the above argument contains a basic flaw. This arises from the fact that urban wages are in practically all countries substantially above those for similar labour in the rural areas from which migration flows. In discussing urban wages I want to distinguish those that are held up above the market-clearing level from those that are determined freely in the market. I call the former " protected-sector wages ", and the latter " unprotected-sector wages ". Wages can be held above the market-clearing level by minimum wage laws, by collective bargaining agreements in industries with strong unions or (as is often the case with large international concerns operating in less developed countries) by the policy of the hiring company itself. Protected-sector jobs can readily be identified because so many people want them. Companies paying wages higher than market levels for equivalent skills and working conditions tend to have very low labour turnover and long lists of applicants waiting for an opening to arise. Unprotected-sector companies tend to have more normal (i.e. higher) labour turnover and shorter waiting lists of applicants, and they may frequently have to engage in active recruiting efforts in order to obtain additional labour when they need it.

My hypothesis is that the unprotected-sector wage, for a given skill classification of labour, and in a given labour market area, is the best available measure of the social opportunity cost of that type of labour in that area. The reasoning is essentially the same as that of the forgone

agricultural product theory, except that it stops one step short (i.e. it takes the urban wage, which migrants receive after they migrate, as the measure of opportunity cost, rather than the rural wage they received before migrating).

Before elaborating on the reasons for my preference, I shall briefly recapitulate the process involved. A large international company increases its workforce in the capital city. Some of those it hires may have previously been working in unprotected-sector jobs, and some in other parts of the protected sector. But these latter leave behind the vacancies which in turn have to be filled. Ultimately they are filled from the unprotected sector, notably from recent migrants whose alternative city employment would have been in the unprotected sector.

My argument turns on the fact that, practically everywhere, wages in the unprotected sector in urban areas are substantially higher than the wages for comparable workers in the rural regions from which migration flows. The annual earnings of casual construction workers, household sweepers and rickshaw drivers in major Indian cities are about double those of landless agricultural workers in the rural hinterland. The wages of unskilled or low-skilled workers in the highly competitive textile industry of Santiago de Chile are also about double those of workers of comparable skill levels in rural areas. This type of wage differential (though not always so large) seems to be replicated in country after country.

Let us suppose, taking India as an example, that unprotected-sector earnings of unskilled labour in Delhi amount to Rs. 1,000 a year, while those of corresponding labour in the rural areas are Rs. 500 per year. When a new job appears in the unprotected sector, and is filled by a migrant, is there a true social benefit equal to the Rs. 500 excess of his urban over his rural wage? This is what the forgone agricultural product approach implies, yet can it really be true? If rural workers truly can double their standard of living by migrating, why do they not come in hordes to the cities? Of course, if they did, they would drive down the wage in the unprotected sector, perhaps, in our example, all the way to Rs. 500. Why does this not occur? Why have the differentials between urban unprotected-sector wages and rural wages been so persistent over time rather than being eliminated as a consequence of more rapid migration flows? My answer is that the differential in question is very largely a self-equalising one. In rural areas of India most of the houses (or huts) of rural labourers are built by themselves (with help from family and friends), using local materials; they are very inexpensive, if not virtually free. By contrast, even the most rudimentary urban quarters have rents amounting to a third or more of the unprotected urban wage. Food in the cities is also considerably more expensive than in the villages, where largely local produce is consumed. Transportation to and from work also typically claims a considerable portion of the urban labourer's budget,

while in the villages the trip to the nearby fields is usually made on foot. In my view Rs. 1,000 in the cities (in my example) is attractive enough to induce a steady stream of migration; the average migrant probably finds it a bit better than Rs. 500 in the countryside. But it could not be substantially better, else the stream of migration would become a flood! It is far closer to the truth, in my opinion, to say that the migrants are approximately indifferent to earning Rs. 500 in the country or Rs. 1,000 in the city than it is to say that they double their real income on making the move.

I conclude, then, that the social opportunity cost of labour to be used on a project in a rural area is the going wage in that area (Rs. 500 in my example), while that of labour to be used in an urban project is the prevailing wage in the unprotected sector of the labour market in that area (Rs. 1,000 in my example). Some readers may feel that we have here lost the entire meaning of social opportunity cost as distinct from market price, but such is not the case. Many projects in the urban area will be in the protected sector, and will thus be paying wages far higher than the Rs. 1,000 of my example. And some projects in the rural area may be for factories, electrical generating plants, etc., and they may pay wages well above the going rural rate for farm labour (perhaps because they are subject to minimum wage laws that do not apply to farm labour). In all such cases the social opportunity cost of labour is lower than the wage actually paid and it, rather than the wage actually paid by the project, is the relevant price of labour to use in social decision-making.

The concept of supply price

One useful way of interpreting the example just discussed is that Rs. 500 is the amount for which members of the unskilled labour force are willing voluntarily to present themselves for work at a newly created job in the rural area, and Rs. 1,000 is the corresponding amount for which they are willing to present themselves for work at a newly created job in the city. In this section I shall elaborate upon the concept of voluntary supply price, and shall show that it provides a far more technically rigorous and subtle basis for approaching the measurement of social opportunity cost than does the concept of forgone product.

The issue turns, I believe, on the quite familiar but often forgotten deficiencies of the national accounts as measures of welfare. In the frequently cited example of the man who marries his housekeeper, welfare has, we presume, been enhanced, but national income as reflected in the national accounts has gone down, because the money the man gives his new wife does not appear as wages. Likewise, when a miner suffering from silicosis voluntarily gives up a $7-an-hour job in a coal mine in order to take a $2-an-hour job as a grocery clerk, we must presume, because he acted voluntarily, that he conceives his welfare to have

improved (or at worst stayed the same), in spite of the fact that the national income and product aggregates have fallen as a consequence of his move. On the other hand, the contribution to national income of the farm-to-city migrant of our earlier example was doubled as a consequence of his move, but we have seen that there are good reasons to believe that his welfare was probably only marginally affected.

Let us now ask what is the true purpose behind the use of social opportunity costs (shadow wages) in the evaluation of investment projects. I believe the answer is that, where there is an excess of wages actually paid over social opportunity costs, this excess should be counted as part of the benefits of the project. Put another way, instead of counting against the project all of its wage costs, we charge it only with that part which represents the true social opportunity cost. The notion of voluntary supply price fits quite naturally and elegantly into this framework. Suppose three workers on a project are hired for identical jobs at $10 a day. One of them might have been willing to present himself for that job for as little as $6, and the second for as little as $8, while the third might have been just barely willing to come for $10. While the wages bill actually paid by the project is $30 a day, the social opportunity cost of the labour used is only $24. When labour is valued at social opportunity cost, the project is given credit for the $4 " profit " gained by the first worker and for the $2 " profit " gained by the second worker.

This is quite as it should be. If only we knew the voluntary supply prices of the workers to be hired by a project, we would surely want to use them as our measures of the social opportunity cost of each man's labour. The reason why we resort to market information is not that it gives the ideal measure, but rather because we cannot get inside the head of each worker in order to find out his voluntary supply price for a given job (nor indeed can we predict, when we analyse a prospective investment project, who will be employed on what job). Unprotected-sector wages serve in nearly all cases as a good proxy for supply price, because they represent, in effect, what the workers who leave the unprotected sector to take jobs in the protected sector were previously willing to offer their services for. If a worker was willing to work in the unprotected sector for $10 and takes a job in the protected sector for $15, we presume that he has made a " profit " of $5, which should be attributed to the project in evaluating it with the use of social criteria. Similarly, if the $15 job is filled by worker A, who was previously earning $12 in another part of the protected sector, and if A is replaced by B, who was previously earning $10 in the unprotected sector, we presume that A has made a gain of $3 and B a gain of $2, which should be attributed to the project. Of course, we cannot in practice identify A's and B's individual gains, but we know that the total presumed gain will be the difference between the newly created $15 wage and the $10 wage corresponding to workers of equivalent skill and capacity in the unprotected sector. And this applies

no matter how many links there may be in the vacancy replacement chain, so we do not have to worry about how many there are or are likely to be.

The use of unprotected-sector wages as the best approximation to supply price and therefore to social opportunity costs has the great added advantage that data can be obtained on these wages, classified into a wide range of occupational and skill groups. We all recognise in daily life the enormous heterogeneity of the labour factor, but we too often forget this in conducting economic analyses. I have had a long-standing complaint against most labour ministries (and the United States' own Department of Labour), in that I believe they have neglected the important task of gathering good information on the wages and salaries of workers at many different skill levels, classified by labour market area, industry, size of firm, etc. Such information can be of enormous usefulness for many purposes, not the least of which is the measurement of social opportunity costs of various types of labour. Fortunately, one does not have to wait for massive studies of the entire labour market in order to obtain the information needed for this latter purpose, since normally one can easily distinguish the protected from the unprotected sector. Once the unprotected sector has been identified, a sampling from the personnel records of a number of firms drawn from within it should provide the information needed to ascertain what wages are typically being paid to workers of different skill and capacity characteristics. That this is feasible has already been demonstrated by Professor Sergio de Castro, who for a study that is as yet unpublished has gathered detailed data on over 10,000 workers from some 600 farms in Santiago de Chile.[1]

I have argued above that unprotected-sector wages are the best available approximation to the supply price of labour. Let us now investigate the cases in which some adjustment might be called for. First, there is the obvious case of non-comparable working conditions. If our project provides better working conditions than typically prevail in the unprotected sector, the social opportunity cost of labour to it should presumably be lower than the corresponding unprotected-sector wage, and conversely if our project entails working conditions more unpleasant than normal. Second, and perhaps less obvious, is the case in which some of the workers who are employed on a project, or who are hired by other parts of the protected sector somewhere along the vacancy replacement chain, may be newly drawn into the labour force

[1] One minor clarification is perhaps in order. It should be borne in mind that some firms may be in the protected sector with respect to certain skill categories of labour and in the unprotected sector with respect to others. A firm without a strong union might be in the protected sector with respect to its lowest-skill workers (because the minimum wage is effective for them) but not for other workers, whose market price is above the minimum in any case. Similarly, a firm might be in the protected sector for its blue-collar workers (who have a strong union) but not with respect to its white-collar workers (who are not unionised). In this case, data on the salaries of white-collar workers would be useful for estimating social opportunity cost, even though the information on blue-collar wages would not.

as a consequence of the project. For example, a newly created job at $15 a day might be filled by someone who was unwilling to work in the unprotected sector at $10 a day. His supply price could be anywhere between, say, $11 and $15. The fact that a project in the protected sector can draw some of its labour from the ranks of the voluntarily unemployed (voluntarily as against working in the unprotected sector), or from outside the labour force, means that the unprotected-sector wage tends to be an underestimate of the effective average supply price of the labour hired. However, the degree of underestimation is likely to be negligible except in clearly identifiable cases where the bulk of the labour supply is obtained from the voluntarily unemployed or from outside the labour force.[1]

The concept of supply price, as applied to entry into the labour force or exit from voluntary unemployment, points up sharply another important fact, to which we shall return later. This is that supply price (and therefore social opportunity cost) is not zero even for involuntarily unemployed labour. Such labour invariably has a reservation price, below which it is not willing to work, and the " profit " which an unemployed worker makes upon getting a job is accordingly not the full wage, but the excess of that wage over his reservation price (supply price).

Chronic urban unemployment: the case of Panama

I shall begin the discussion of chronic urban unemployment with the case of Panama, which exhibits in extreme form characteristics that are probably found in lesser degree in a number of other less developed countries. Panama had been a booming country during the Second World War, when the passage of ships through the Canal reached an all-time peak. The end of the war, however, witnessed a sharp slackening-off of Canal activity, and left the country with a substantial pool of unemployed in the so-called Metropolitan Area (principally Panama City and Colón) adjacent to the Canal Zone. What has puzzled many observers is that this pool of unemployed, though fluctuating in size over time, has remained large in spite of substantial—even spectacular—economic growth. Panama's real gross national product grew at a compound annual rate of 5.5 per cent in the decade of the 1950s and of about 8 per cent in the decade of the 1960s, in spite of which the urban unemployment problem remains severe.

How can this paradox be explained? I believe the answer lies in the wage regulations prevailing in the Metropolitan Area, together with the phenomenon of migration. In brief, minimum wage legislation is applicable to nearly all employments in the Metropolitan Area. What we called

[1] A characteristic case of this kind is the establishment of a canning factory in an agricultural region. These are often staffed by married women or students on summer vacation who enter the labour force only for the brief canning season.

in earlier sections the unprotected sector hardly exists for unskilled workers in this area. For clarity of exposition, I shall here assume that there is no unprotected sector for the unskilled—i.e. all jobs are covered by the minimum wage.

Now let us recall what, in the analysis of the forgone product as a measure of social opportunity cost, was the relationship between developments in the unprotected sector and the flow of migration. It should be clear from that earlier discussion that the wage level in the unprotected sector operated as a regulator of the flow of migrants. If initially the wage levels were so high as to attract floods of migrants, the forces of supply and demand would soon bring wages down to an equilibrium position in which the steady stream of migrants was more or less in balance with the growth of urban demand for such workers.

In the case of Panama, with no urban unprotected sector, quite clearly the unprotected-sector wage cannot perform the function of regulating the flow of migration. Yet somehow this flow will in fact be contained. If the minimum wage is so high as to attract floods of migrants under conditions of high employment, the way in which migration will be held down to more normal levels is obviously through increased unemployment.[1] Given the rate of growth of demand for unskilled labour at the minimum wage, there will be a certain equilibrium rate of unemployment which will stem the potential flood of migration and hold it down to a stream which is consonant with the given rate of growth of demand. But it is important to realise that in a steady-state unemployment equilibrium situation the number of migrants will regularly exceed the number of jobs created.[2]

The following table illustrates the way in which the equilibrating process works. In the initial situation we have an urban unskilled labour force of 1,000, of whom 800 are employed at the minimum wage and 200 are unemployed. The equilibrium rate of unemployment is 20 per cent. Now 100 extra jobs are created. Initially, these are filled from the ranks of the urban unemployed, and as a consequence the unemployment rate is reduced to 10 per cent. But this makes movement to the city extremely

[1] This conclusion was independently arrived at on the basis of an analysis quite similar to my own in an article by Michael P. Todaro: "A model of labor migration and urban unemployment in less developed countries", in *American Economic Review* (Menasha (Wisconsin)), Mar. 1969.

[2] Actually, the growth of the urban labour force stems both from the natural increase of the urban population and the flow of migration from outside the urban area, but for simplicity I am here disregarding the natural demographic growth of the urban labour force. This enables me to compare new job creation in the urban area directly with the flow of migration. No generality is lost in this simplification, however. When I speak of an equilibrium relationship between new job creation and migration, readers can interpret this as meaning an equilibrium relationship between the number of new jobs and the net increase of the urban labour force from both sources. Since the natural growth from within the area is overwhelmingly governed by demographic considerations, the element which responds significantly to changes in the unemployment rate will in any case be the flow of migration.

attractive and sets in motion a process of migration which only terminates when once again the equilibrium unemployment rate of 20 per cent is restored. Paradoxically, the creation of 100 new jobs in the end actually increases the absolute number of unemployed (though naturally not the unemployment rate).[1]

Urban labour force	Initial equilibrium situation	After creation of 100 new jobs	
		Immediate response	Ultimate equilibrium situation
Total	**1 000**	**1 000**	**1 125**
Employed	800	900	900
Unemployed	200	100	225
Percentage unemployed	20	10	20

One should not infer that the over-all unemployment rate in the urban area measures the actual average experience of new migrants. The newly arrived migrant will almost certainly be unemployed for a time. The longer he remains in the area, the greater will be his chances of having a job at any given moment. The point is that his chance of having a job after, say, a month in the city might be 50 per cent when the over-all unemployment rate is 20 per cent, with this chance jumping to 75 per cent when the over-all unemployment rate is 10 per cent. We are thus using the over-all unemployment rate (among unskilled workers) as a proxy for the whole spectrum of job probabilities that the typical migrant would face after different lengths of stay in the urban employment market. The important point is that, when the unemployment rate is below equilibrium, the typical migrant evaluates life in the city as being very much better than life in the country (which tends to produce a flood of migration), while when the unemployment rate is at its equilibrium level, the typical migrant is more or less on the margin of indifference between the prospects of life in the two places (which explains why the rate of migration is constrained to an amount consistent with the rate of creation of new jobs while maintaining the equilibrium rate of unemployment).

This fact—of migrants being more or less on the margin of indifference when the equilibrium rate of unemployment prevails—enables us to

[1] This, of course, has to happen if the unemployment rate remains constant. Rarely, however, is it possible to check the process in detail, because the rate of growth of new jobs is relatively small (compared to the labour force) and steady, as is the rate of immigration. However, Professor Frederick Harbison reports on a case where the number of jobs in the urban sector of Nairobi was increased rapidly by nearly 15 per cent in 1964. " This acted like a magnet attracting new workers into the urban labour markets. . . . In the end, the volume of unemployment, as a consequence of the expansion of the modern labour force . . . was probably increased rather than decreased " (cited in Todaro, op. cit., pp. 140-141).

reintroduce the notion of supply price, though in a somewhat more complicated context than before. When he makes his move, the migrant is in effect buying a series of lottery tickets—a 40 per cent chance of a job after two months, a 60 per cent chance after four, and perhaps an 80 per cent chance after six months. Let us suppose that these are in fact the probabilities of employment, and let us trace in the following table the employment experience of 100 typical unskilled migrants into, say, the Panama Metropolitan Area at a time when the minimum wage is $5 a day. Column (4) gives the expected average earnings after two, four and six months, for which the migrants were willing to come to the city. It represents, in a sense, the supply price of migrants after different durations of stay. But it says nothing whatsoever about the supply price to be attached to new jobs created by an investment project. This is given in column (5) and, quite surprisingly, it turns out to be precisely equal to the minimum wage.

Period	Percentage of migrants		Total wages earned ($)	Wages earned per migrant ($)	Wages earned per employed migrant ($)
	Employed (1)	Unemployed (2)	(3)	(4)	(5)
After two months . .	40	60	200	2	5
„ four „ . .	60	40	300	3	5
„ six „ . .	80	20	400	4	5

The paradox is quite easily explained. When the equilibrium rate of unemployment is 20 per cent (and still abstracting from local demographic increase in the city), the creation of eighty new jobs will induce the migration of 100 workers. Ultimately, an average of eighty of these will be employed at the minimum wage at any one time (although of course natural labour turnover will mean that the particular individuals employed will not always be the same). Once that point (represented by the bottom row of the table) is reached, the average migrant will be earning $4 a day. But only eighty new jobs were created, so the cost is $5 per man-day actually worked. This outcome of eighty out of 100 migrants being employed at $5 per day is what is necessary to satisfy the expectations that induced them to migrate in the first place.

The same kind of analysis applies in the intervening period of adjustment. After two months, forty are employed and sixty unemployed, so the average earnings per migrant are only $2 a day. This is what was necessary to satisfy the over-all expectations of the group after two months; but since only forty migrants are employed at this point in time, the cost to the employers which just satisfies the expectations of the migrants is $5 per man-day in any case.

I must apologise here for the need to introduce such a degree of precision into this discussion. It is quite clear that no individual migrant actually has in mind a precise picture of how the probability of his being employed will vary with his length of stay. Nor is it very easy to picture the whole group of 100 as being in some sense just satisfied when after two months forty of them are employed at $5 a day and the remaining sixty unemployed. (Clearly, the forty will probably be quite happy, and the sixty quite unhappy.) But none the less, I feel that this degree of abstraction is necessary to cope rigorously with the problem at hand. The essence of the notion of an equilibrium relation between the unemployment rate and migration is that typical migrants are more or less near the borderline of indifference. For if the typical migrant were exceedingly happy with the way things turned out, there would be a flood of friends and relatives into the city, and if the typical migrant were very unhappy, there would be a mass reflux to the farm.

One way of putting the problem in more down-to-earth terms is to recognise that those migrants who are employed within a month or two are very happy—they, in a sense, have gained a benefit that they did not fully expect. They will probably be writing home telling their friends and relatives how good life in the city is, and some additional migration will be induced. On the other hand, some of those who fail to find employment in three, four, or five months will probably give up in despair and go home, thus counterbalancing the inflow of the friends and relatives of the lucky ones. Those who stay in the city may very well receive benefits beyond what they expected, but those who return home disappointed have suffered costs which in principle should be set off against the extra benefits of the lucky ones. The abstraction that the typical migrant is more or less indifferent is really a shorthand way of recognising this offsetting effect.

Chronic urban unemployment: the general case

In the analysis contained in the sections on forgone product and supply price we concluded that, in the absence of abnormal open unemployment, the relevant social opportunity cost of labour was best measured by the wage rates paid in the unprotected sector. In the section on chronic urban unemployment in Panama we concluded that, in the absence of an unprotected sector, migration to the cities would be held in check by the equilibrium level of unemployment, and that, at that level, the social opportunity cost of unskilled labour was best measured by the minimum wage, which was assumed to be effective for nearly all unskilled jobs. We now consider the possible case where, even in the presence of an unprotected sector of reasonable size, chronic unemployment in excess of normal seasonal and frictional levels persists. I will continue to assume, because the phenomenon is so widespread, that migration into the urban

employment market continues in spite of the existence of relatively high unemployment. The easiest way to depict this case is as a sort of combination of the previous two, with the introduction of a new element into the argument—namely the costs of searching for a job. It is also well to think of the protected sector of the employment market as having several different levels—very high wage rates in international companies, quite high ones in local companies with strong unions, middling high wage rates in government employment, and minimum wage rates which for unskilled workers are somewhat higher (in the areas covered by the minimum wage laws) than they can obtain in the unprotected market.

Consider now the case of an unskilled migrant entering this market. His options may include taking a job in the unprotected sector, but this will interfere with his looking for a job at higher wages in the protected sector. If he chooses to spend his time searching for a high-paying job rather than working at a low-paying job, this means that his supply price lies above the unprotected-sector wage. And I am sure that cases are quite common in which this happens. The migrant may have brought with him some accumulated savings, which may last for a few months of search, or he may have friends who are willing to put him up for a certain period of time, but not indefinitely. If he uses this time to search for a high-paying job, we can say that during this period his supply price lies above the unprotected-sector wage; when his money (or his welcome) runs out, he takes a job in the unprotected sector, at which point his effective supply price to the protected sector falls to the level of the unprotected-sector wage.

The existence of search costs can thus lead to a situation of chronic unemployment even in the presence of an ample unprotected market. But now the rate of migration is a function of two things—the rate of unemployment and the wage structure (including the unprotected-sector wage). The higher the fraction of jobs located in the protected sector, and the higher their wage rates, the higher will be the fraction of people willing to bear the costs of search and the higher will be the equilibrium level of unemployment. And while we cannot be quite as specific about the social opportunity cost of labour in this case as in the Panama case, two important statements can be made—

— first, the supply price that is relevant for measuring the social opportunity cost of labour for jobs in the protected sector is at least equal to the wage of corresponding labour in the unprotected sector;

— and second, the greater the fraction of the labour force that chooses to be unemployed while engaging in search, the greater will be the excess of this supply price over the unprotected-sector wage.

This is in close correspondence to the Panama case, treated earlier. If, with a minimum wage of $5, the equilibrium level of unemployment was 20 per cent, we can be sure that this equilibrium level would be lower

if the minimum wage were $4.50 and higher if it were $5.50. And since in the Panama case the minimum wage itself turned out to be a relevant measure of social opportunity cost, we have in that case the same relationship as in the case presently being discussed: higher equilibrium levels of unemployment are associated with higher social opportunity costs of labour. The emphasis here, of course, must be on the word equilibrium. The more attractive are opportunities in the protected sector, the higher will be the rates of unemployment that migrants will be willing to suffer, and still come to the city. The social opportunity cost of labour in a sense includes the cost of this unemployment, and hence must go up if more attractive wages or an increased fraction of jobs in the protected sector induce a rise in the equilibrium level of unemployment.

Disguised unemployment

Much has been written about disguised unemployment, especially with reference to less developed countries. It is often contended that this phenomenon gives rise to a situation where the social opportunity cost of labour will lie below the market wage, and may even be zero. There are some real elements of truth in this contention, but most expositions of the argument fail to put it in proper perspective. Some treatments associate disguised unemployment simply with all low-wage, low-skill jobs—sweepers, gardeners, casual construction workers, rickshaw drivers, truck loaders, etc., in the case of India. This variant simply says that the disguisedly unemployed earn wages substantially below the average wage of unskilled workers in a given labour market. Obviously, there is virtually no distinction between this definition of disguised unemployment and what I would call the employment of labour in the unprotected sector. And if one considers the case where new jobs are created in the protected sector, my analysis, which would take as the first approximation to social opportunity cost the wage prevailing in the unprotected sector, would, like that of the proponents of this variant of the disguised unemployment argument, apply to a protected-sector project a shadow wage below that which the project actually paid.

The second variant associates disguised unemployment not just with low wages but with situations in which the marginal productivity of labour lies below the actual wages earned. This is clearly a quite different concept, which could among other things apply to high-wage as well as to low-wage workers. There are a variety of activities to which this argument applies. A classic example is that of fishermen on a lake. The addition of more fishermen increases the total catch, but not proportionately, yet the last fisherman has an equal chance of making a given catch as the first. The expected catch is the same for all, and is equal to their average productivity. But, owing to the fact that the total catch does not increase in proportion to the number of fishermen, the marginal productivity of

a fisherman is less than what he earns. Other cases, more frequently cited in the economic development literature, are the shoeshine boys in a given square, where the presence of the last boy does not proportionately increase the number of shines, or the hawkers and vendors found in the streets of less developed countries, where the addition of another man selling a given product does not proportionately increase the amount sold but has the effect of somewhat reducing the average amount received by each.[1]

Several points are relevant to the connection between this type of disguised unemployment and the measurement of the social opportunity cost of unskilled labour. The first is that activities with the characteristic of marginal product less than earnings generally represent only a small part of the unprotected sector of the market for unskilled labour. The necessary conditions do not, in general, apply to hired labourers—be they sweepers, construction workers, runners, loaders or whatever. Here the presumption remains that their marginal product is at least equal to the wage. The second point is that when new jobs are created in the protected sector, in the absence of migration, the workers to fill them would normally come from all parts of the unprotected sector, with only a relatively small fraction coming from activities in which labour's marginal product lies below its earnings. (In this case, a small downward adjustment of the social opportunity cost of labour below the unprotected-sector wage would be justified, reflecting the fact that product has not gone down by the full amount that, say, the coat-hanger hawker who took a job in the protected sector had previously been earning. But the adjustment would be small because only a small fraction of the jobs created would be filled by hawkers, vendors, shoeshine boys, etc.)

The third point is that just as the creation of new jobs in the protected sector can result in an increase in the number of openly unemployed, so also can it lead to an increase in the number of disguisedly unemployed in the sense now under discussion. If the hawkers and vendors taking protected-sector jobs are not fully replaced, there would be, as just indicated, a downward adjustment of the social opportunity cost of unskilled labour below the figure it would otherwise have; if migration just managed to replace those that left, no adjustment would be warranted, and if migration more than fully replaced the vendors and hawkers who left for protected-sector jobs, this would warrant an upward adjustment in the social opportunity cost of unskilled labour, over and above the level that it otherwise would have.

In my judgment, the number of people who would abandon jobs with marginal product less than earnings would be so small, for every 100 new jobs created in the protected sector, and the likelihood of their being substantially replaced by migration so high, that any adjustment on account of disguised-unemployment effects would probably be of negligible magnitude.

[1] Even in these cases the discrepancy between earnings and marginal productivity arises only when, as is sometimes the case, the market price (of shoeshines or, say, of coathangers sold on streetcorners) is a "conventional" price which does not adjust downward to clear the market. With full price adjustment the "last" worker would in fact earn his marginal product, which consists in part of the value of the net increment to output due to his presence and in part of the alternative earnings of the labor displaced by the downward price adjustment that his presence brought about.

Cyclical unemployment

In a way, the key question with respect to cyclical unemployment has already been answered, in the sense that the relevant social opportunity cost of a given unemployed worker is his voluntary supply price. In this section I shall further develop this idea, and indicate conditions under which further adjustments would be indicated.

In the first place, it is important to note that cyclical unemployment rates tend to differ substantially for different classes and grades of labour, the unskilled tending to suffer far wider swings in employment than the skilled, and skilled blue-collar workers in turn tending to have somewhat sharper variations in employment than white-collar workers. This being the case, the setting of a shadow wage below the market wage in cases of cyclical variation would very likely apply only to certain types of labour.

In the second place, it is important to note that it is quite possible to have substantial amounts of cyclical unemployment, even when a well-developed unprotected sector exists. This is because, even in the best of labour markets, the adjustment of wages to balance supply and demand is generally sluggish, particularly when the pressure on wages is in the downward direction. This observation carries with it the consequence that, for categories of labour suffering from substantial cyclical unemployment, the social opportunity cost of labour can for the period of such unemployment lie significantly below even the unprotected-sector wage. (It could not be zero, however, being limited to the reservation price below which such labour would not accept employment.)

In the third place, while for most cases of cyclical unemployment the estimated reservation price of the unemployed would be an appropriate measure of their opportunity cost, special further downward adjustment from the market wage is warranted in cases (like the Great Depression) of massive cyclical unemployment extending to all or nearly all categories of workers. The reason for this is the multiplier effect that applies in such cases. If, when a group of unemployed workers are given jobs, a quarter or so of their increased spending has the effect of absorbing still more unemployed, this effect, plus subsequent multiplier rounds, should be taken as an external benefit associated with their employment and should therefore reduce still further the social cost charged against the project hiring them. The magnitude of the adjustment should not be exaggerated, however, for several reasons, viz.—

(a) The multiplier effect comes into being from the increase in total (first round) spending generated by the workers' gaining employment. Since he was already spending something even while unemployed, the increase in spending is not likely to equal the wage he is paid even if he spends it all.

(b) Since one of the effects of unemployment is to deplete the savings of the unemployed and their relatives, some considerable effort is likely

to be made, when they become employed, to replenish the depleted family coffers. This means that they will not be spending their entire wage, but possibly saving a significant fraction of it.

(c) Not all the extra spending will result in the employment of unemployed labour. Very likely a substantial fraction of it will accrue to capital or categories of labour that are fully employed, and in some countries a good part of it would go for imports.

(d) The extra benefit to be assigned to the employment of extra workers is not equal to the full amount of wages paid to the labour newly employed as a multiplier effect of the first group's increase in spending. Once again, it is the excess of the wages paid to the second group over and above their individual reservation prices.

All these reasons suggest that even when employment is extremely widespread—say, 25 or 30 per cent of the labour force—the downward adjustment of the social opportunity cost of labour in order to take into account its multiplier effect would probably not exceed 10 or at most 20 per cent of its wage. The same reasons indicate that for moderate recessions, when only a limited number of categories of labour are seriously affected by unemployment, the adjustments of the social opportunity cost of labour to account for multiplier effects are small enough to be neglected.

Finally, a caveat concerning the making of adjustments for cyclical unemployment when evaluations of investment projects are prepared. If cyclical unemployment is present in the economy and the project is to be initiated soon, it is by all means correct to adjust for cyclical unemployment in estimating the social opportunity cost of labour during the first year or so of the project's life. But it would be a serious mistake to use the same adjustment for the labour cost that will apply over the whole ten or twenty or fifty years of the project's life. Beyond the immediately foreseeable future the best assumption that we can make is that any given year will be one neither of cyclical boom nor cyclical recession, and we should therefore project the social opportunity cost of labour on the basis of the normal employment situation that we expect to prevail. Such an employment situation can quite easily include substantial amounts of chronic unemployment of the migration-regulating type, but not cyclical unemployment as such. And, as we have seen, the adjustment of the social opportunity cost of labour for the influence of migration-regulating unemployment is actually in the opposite direction from the adjustment that is indicated in the cyclical case.

Concluding comments

In this paper, I have attempted to present the case for using prevailing wage levels in what I have called the unprotected sector as the point

of departure for estimating the social opportunity cost of labour in a given labour market area.

In the section on forgone product, this was shown to be the appropriate value for the simple case of a full-employment economy, disregarding the presence of people not currently employed, but with reservation prices above the unprotected-sector wage.

In the next section, the concept of supply price was introduced, and it was shown that in the presence of people with higher reservation prices the unprotected-sector wage somewhat underestimates the true social opportunity cost. In the section on chronic urban unemployment in Panama a rigorous model was presented showing how, in the absence of an unprotected sector and in the presence of a relatively high minimum wage, the phenomenon of migration can lead to an equilibrium level of chronic unemployment in the urban sector. In that model, the relevant social opportunity cost of unskilled labour was shown to be the minimum wage itself. In the next section, the model was adapted to a more general situation where both the phenomenon of migration and an unprotected sector were present. Here there once again turned out to be an equilibrium level of chronic unemployment, and the relevant social opportunity cost of labour turned out to exceed the unprotected-sector wage by an amount which was greater, the greater the level of chronic unemployment. The concept of disguised unemployment was then discussed, and it was shown that one variant of that concept coincided closely with my own line of analysis—especially in the form presented in the study of the forgone product. The second variant—quite different—was shown to give rise to modest downward adjustments of the social opportunity cost of labour below the unprotected-sector wage in the absence of migration, but in its presence the adjustment was shown to be indeterminate in sign and probably negligible in magnitude. I then turned to the phenomenon of cyclical unemployment; it was shown that, even with the existence of an ample unprotected market, cyclical unemployment could cause the social opportunity cost of labour to fall below the unprotected-sector wage—the relevant measure of opportunity cost here being the reservation price of the affected labour. But it was emphasised that, particularly for the social evaluation projects, adjustment of social opportunity costs on the basis of cyclical unemployment was strictly a near-term phenomenon. Looking ahead for three or more years, one can make no better assumption than that the economy will be at a normal level of activity, giving rise to no adjustment for cyclical unemployment for the great bulk of most prospective projects' lives.

All this means that, with modest qualifications and occasional adjustments (usually upward), the unprotected-sector wage stands as the basic measure of social opportunity cost. As against alternative measures, most of which are based on macro-economic analyses of one form or other, it has the great advantage of being readily capable of reflecting the com-

plexity and subtlety of labour market phenomena. This measure can show that the social opportunity cost of the same type of labour might be Rs. 200 in the countryside, Rs. 750 in Hyderabad and Rs. 1,000 in Calcutta. It can easily branch out to take separate account of an indefinite number of skill and capacity classifications of labour. It subtly reflects these differences of location and skill, as well as the tastes of workers (even, with an adjustment, the tastes that lead to chronic unemployment).

By contrast, the estimation of the social opportunity cost of labour by macro-economic approaches seems crude indeed. Rarely do programming approaches distinguish more than two skill categories of labour, and equally rarely do they distinguish more than one labour market area (the whole economy or a single part of it). And it is difficult if not impossible to imagine more than a modest additional amount of disaggregation under a programming approach. The data requirements are far too severe, and the computational requirements astronomical. Similar comments can be made about global econometric models of the economy. These models can serve many useful purposes, but among them the task of estimating the social opportunity costs of twenty classes of labour in six major markets of a country simply cannot be counted. The approach here advocated takes the infinitely complex machinery of the economy itself as its computer and finds in the data generated by that machinery—in the form of unprotected-sector wages—the best approach to measuring the social opportunity cost of labour—by type, skill and location.