INTRODUCTION TO COST-BENEFIT ANALYSIS

P A R T   I

PROFILES, FOREIGN EXCHANGE AND CAPITAL

Arnold C. Harberger

University of California, Los Angeles

November 2008

This paper is not intended to give its readers the knowledge and skills needed to carry out economic evaluations of real-world projects. That task would require a major study program, ideally extending over several months or more of hard, full-time work, comprising classroom hours, extensive readings, and many exercises in which participants deal with a whole gamut of issues in a sequence of simulated real-world cases.

Instead, we try here to convey an understanding of what is involved in cost-benefit analysis: what are the analytical foundations on which such analysis is based? What are the main issues that have to be faced? What are the key variables that have to be calculated or estimated? And, finally, what determines the degree of confidence that one can have in the results?

Without a doubt, the best starting point is the ex post analysis of a pure business project. Here we are interested in its financial profitability over its effective lifetime. To estimate this, we build up a profile of the "cash flows" associated with the project during each year (or quarter, month, or other period) of its life. This profile would start with recording the outlays involved in designing and planning the project, and would then turn to the disbursements for construction and equipment purchases. Then would come the whole history of the years of operation of the

**This paper was prepared under the auspices of USAID.

project. For each year, we would record the net cash flow linked to that year. This would cover the sale and other possible sources of cash inflows. From these inflows we would deduct all cash outflows -- labor and materials costs, taxes, costs of maintenance, repair and replacement of capital equipment, insurance costs, etc. In the end we would record for each year, starting with the design and planning stage and ending with the closing down of the project -- the net cash flow (i.e., inflows minus outflows) corresponding to that year. The end result of all this is a project profile. Such a profile would typically start with one or more years of net outflows, followed by a operating period in which the net flows of each year were (typically but not necessarily) positive. This project profile summarizes the key facts needed as inputs into the analysis.

But wait!! A little reflection should alert one to an additional issue that must be addressed -- the phenomenon of price inflation (or, conceivably, deflation). We all have experienced important price level movements and are fully aware that the dollar, the peso, or the rupee of one year hardly ever has the same real purchasing power as that of other years. So if we want to avoid fooling ourselves about the economic worth of the project, we have to convert our project profile into real terms. Standard practice is to choose a numeraire -- usually either the country's consumer price index (CPI) or its GDP deflator, and to divide each year's net cash flow by this numeraire, thus expressing it in real terms. If we are using the CPI, the resulting profile is expressed in consumer baskets -- how many net CPI baskets were expended during each of the investment years, and how many net CPI baskets were taken in during each of the productive years. If we use the GDP deflator as the numeraire, we equivalently express each year's outflow or inflow in terms of "producer baskets". That is to say, our resulting, inflation-adjusted profile would end up measured in dollars (or pesos, or rupees) of constant purchasing power, with that purchasing power being sufficient to buy a given bundle of either consumer or producer goods in each year of the project's life.

To analyze the profile further we need a discount rate. In principle this discount rate should reflect the genuine real opportunity cost of funds for the entity that is undertaking the projects. For a typical U.S. firm that real opportunity cost might be something like 8-10 percent per year. For businesses in most developing countries it would typically be significantly higher. It is important to recognize that in order to evaluate a profile expressed in real terms, one needs to utilize a discount rate that is also expressed in real terms. For a business firm one might consider the real opportunity cost of funds to be equal to a weighted average of its real cost of equity capital and its real cost of debt capital, the weights being the shares in which new investments by that firm typically financed by equity and debt.

The key thought that lies behind the idea of economic opportunity cost is that it reflects the true cost of raising the money used in this project and/or the true yield that the project operator would normally have obtained in the likely alternative use of that money. If these two numbers differ, then the relevant opportunity cost is the higher of the two.

We will here work through an example of a simple project of, say, a cattle-feeding operation. In period zero young animals are purchased for, say 1000. During periods one and two, the maintenance and feeding costs are 300. In the period 3, the animals are sold for 2520. All these numbers should be thought of as being already expressed in real terms and as representing the net cash flows corresponding to each year. The profile of this project is thus -1000, -300, -300, +2520.

In the top panel of Table 1, a calculation is made of the net present value of this project, using an opportunity cost rate of 10% per year. The project starts out in period zero using

TABLE 1

Net Present Value and Internal Rate of Return

Net Present Value

	0	1	2	3
Period:	0	1	2	3
Profile:	-1000	-300	-300	-2520
Net Present Value @ 10%				
	-1000			
	x 1.10 =	-1100		
		-1400		
		x 1.10 =	-1540	
			-1840	
			x 1.10 =	-2024
Net Present Value at 10% Calculated to Period 3 =				+496

Internal Rate of Return

	0	1	2	3
Period	0	1	2	3
Profile	-1000	-300	-300	+2520
Net Present Value @ 20%				
	-1000			
	x 1.20 =	-1200		
		-1500		
		x 1.20 =	-1800	
			-2100	
			x 1.20 =	-2520
Net Present Value of Zero Confirms IRR = 20%				0

1000 of resources, thus “owing” the owner the sum of 1000. But since the appropriate opportunity cost rate of return is 10%, by period 1 this 1000 grows to 1100. Add to this the 300 of feeding costs in period 1 and the project now “owes” its owner 1400. This, in turn, grows (at 10%) to 1540 by period 2, to which must be added the 300 of feeding costs in that period, yielding a total of 1840 of what we call “capital-at charge” at that time. This in turn again grows by 10%, to 2024 in period 3, when the sale takes place. This sale, for the sum of 2520, yields the owner a profit of 496 (= 2520 minus 2024), which is the present value (as of period 3) of the gain from the project, over and above (a) the 10% return that the owner could have gotten via a normal alternative investment or alternatively (b) the 10% real interest rate that the owner actually had to pay for the funds used in the project. If (a) and (b) are not the same, then our assumed 10% opportunity cost should be taken to represent the higher of the two.

Often one finds it difficult to pin down precise numbers for the opportunity cost concepts (a) and (b). In that case one can work with upper and lower bounds for the discount rate. This is an easy route to take when a project passes the profitability test at both the upper- and lower-bound discount rates, or when it fails to pass at either of these rates. But one must recognize that some projects will pass at the lower-bound rate and fail at the upper-bound rate, leaving unsettled the decision of whether or not to accept the project.

One statistic that can be useful in such situations, and in many others as well, is the internal rate of return of a project. The internal rate of return is like the yield-to-maturity of a bond. It tells you the precise rate of yield of the project, compounded over its whole life. Thus one can talk about a project A having a yield of 12%, project B having one of 9%, etc. The internal rate of return (IRR) is thus an extremely useful piece of information, but readers should be warned that the IRR should not be thought of as a determining criterion of project choice.

That is, one should not, when given a choice between projects A and B, always choose the one with the highest IRR. (This should not detract from the general utility of the IRR as a summary statistic. After all, intelligence is a relevant factor in one's decision of whom to marry, but that does not say that from among her suitors, a girl should always choose the one with the highest IQ.)

A particular advantage of the internal rate of return is that it is an attribute of the project profile itself. It can be calculated directly from the profile data. And for this purpose one doesn't have to know what the relevant opportunity cost of capital is. That is why many boards of directors (including, for example, the executive boards of the World Bank and of most regional development banks, ask that the IRR be calculated for each project that is submitted for their approval. Once again, it is not a determining criterion,[1] but it is a very useful and informative statistic.

The lower panel of Table 1 illustrates the calculation of the internal rate of return of the cattle-feeding project. In point of fact one finds the IRR by trial and error, or more realistically finds a computer that is programmed for this purpose. In this case the example was created so as to have an "easy" IRR of 20%. Thus, the initial capital-at-charge of -1000 in period zero grows

---

[1]Suppose you are asked to choose between project A, with a profile of -100 + 130, and project B, with a profile of -1000 + 1200. The internal rate of return of A is 30% while that of B is only 20%. Yet if the relevant opportunity cost of capital is 10%, the net present value of B calculated as before to the closing period is 100 (= 1200-1100), while that of A is only 20(= 130-110). This is a very simple example of why the IRR is not a reliable sole criterion for project choice, but the principle it illustrates extends to a wide range of real-world projects. The principle states that the larger of two projects may be preferable to the smaller one, even if it (the larger) has a lower IRR -- because, in spite of its lower IRR, its net present value can end up being significantly higher than that of the smaller project. This principle assumes that the relevant choice is between only one of each of the alternatives, but that is the real situation in a great many cases -- a high dam (more expensive) versus a low dam (cheaper), a concrete

to -1200 by period 1, and is augmented by costs of -300, leaving a capital-at-charge of -1500 in period 1. This grows to -1800 in period 2 and is again augmented by -300 of feeding costs. The resulting capital-at-charge of -2100 in period 2 grows (at 20%) to -2520. This accumulated cost is precisely canceled by the sale price of +2520 in period 3, leaving a net present value of zero. This is how the IRR is defined -- it is that rate of return which, when applied to a given profile, yields a net present value of zero.

"Economic" Cost-Benefit Analysis

What makes the analysis of a pure business project so easy is the fact that the profile of benefits and costs is so easily defined. "Money coming in" is good, a plus; "money going out" is bad, or minus. Some public projects may be, effectively, business ventures that can be analyzed in this way. But this is certainly not the case for the vast majority of public projects. Here the benefits may take the form of a saving of travel time for road users, the improvement of farmers' crop yields, a better-educated labor force, a healthier and longer-lived population, etc. Benefits of these kinds do not come as increased cash flows to the project, yet they are clearly benefits from the point of view of society, and should undoubtedly be taken into account in the evaluation of the projects that generate them. This reflects the broader vision that distinguishes an "economic" from a purely business or financial analysis.

Readers will easily appreciate how difficult it can be to place a monetary value on some of these non-cash benefits. Such valuation has been one of the major challenges that economic project (and program) analysis has had to face. It is a struggle that is ongoing, in which victories are hard to achieve, and typically only clear up small patches of a large and cloudy panorama. Highway benefits are easier to evaluate than those of irrigation projects; irrigation benefits are

---

highway (more expensive) versus a gravel road between the same points, a bridge to a nearby easier to quantify than education benefits; the latter, in turn, are easier to handle than health benefits.

In light of the above-mentioned difficulties, this introduction to the "economic" facet of cost-benefit analysis will focus on the most straightforward type of externalities -- namely, taxes and subsidies. These come in the form of cash, but they represent benefits of costs that accrue to the government rather than the project entity. Thus, for example, we can think of undertaking an economic cost-benefit analysis of a textile mill or a cattle ranch or a private electric power plant. The economic analysis would differ from the business or financial analysis of these firms by recognizing that the taxes they pay, while a cost to the firm, are a benefit to the government, and that any subsidies they receive, though properly counted as a benefit from the business point of view, have also to be counted as a cost to the government when the analysis is undertaken following the broader "economic" point of view.[2]

Typically, the economic profile of a project will show higher benefits than the corresponding financial profile. Most projects involve the use of materials and other inputs that are internationally traded, and to which import tariffs and possibly other taxes apply. The financial analysis of such a project would count those taxes as a cost, thereby reducing the net

---

island (more expensive) versus a ferry project.

[2]Cost-benefit terminology has changed somewhat over recent decades. What we now call economic benefits and costs were formerly labeled "social" benefits and costs. As late as 1985, I entitled a major paper "Reflections in Social Cost-Benefit Analysis", yet if I were to do it today, the title would refer to "economic" cost-benefit analysis. The reason for the change is that many people interpreted the term "social" to refer to the items usually dealt with in what are labeled social programs -- inoculations for children, unemployment benefit payments, government health care programs and subsidies, poverty relief, etc. These programs can be subjected to an economic cost-benefit analysis, alright, but there is nothing about economic C-B analysis that is in any way limited to such programs. It was in order to avoid inaccurate interpretations that our professional terminology has now for some time been using the term "economic" in places where "social" would earlier have been employed.

flow of benefits that is calculated for the project. In the project's economic profile, such taxes are not counted as costs, leading to a higher flow of net benefits.[3]

The same story applies with respect to all taxes that are either directly paid by the project or embodied in the costs of items that the people purchase. This includes corporation income taxes, sales and excise taxes, franchise taxes, etc.

Economic Opportunity Costs

The concept of economic opportunity costs (sometimes called "shadow prices") is quite central to economic cost-benefit analysis. It is also a concept that is unfamiliar (and hence often quite puzzling) to many non-specialists. In this section we will try to convey the basic idea at an intuitive level, and with the least possible complications. We will consider here the economic opportunity costs of foreign exchange and of capital.

Consider a case in which a country has an average import tariff of, say, 20%. When our project goes into the market to buy foreign exchange, that foreign exchange will ultimately come from a combination of "displaced other imports" and "newly stimulated exports". Suppose that the project buys $100 of foreign exchange, of which $60 comes from displaced imports and $40 comes from newly stimulated exports. Suppose, too, that the project is in a "peso" country, with

---

[3]The easiest way to think of the adjustment is to note that the financial profile does not count the benefit to the government that these taxes represent. This benefit is an "externality" so far as the project itself is concerned, and therefore should be added as we pass from its financial to its economic profile. At a perhaps more subtle level, thinking all the time about the benefits and costs of a project from the point of view of society as a whole, the taxes are a transfer payment, being a cost to the project itself, and a benefit to the government. Thus if we start by counting benefits and costs other than taxes, we can proceed by simply not making any adjustment -- i.e., treating taxes as a transfer payment, a non-cost. When so treated they are a part of economic "profits", as they have not been deducted from the sales or other gross benefits of the project. But if we take the step of counting the taxes as a cost (in creating the project's financial profile), we then have to count them again, but as a benefit, when moving from the financial to the economic profile.

an exchange rate equal to 10 pesos per dollar. The project pays 1000 pesos for the $100 it buys, but this cost does not include the 120 pesos of tariff revenues that are lost as $60 worth of imports are displaced. The total economic cost of the $100 of foreign exchange is thus 1120 pesos, not the 1000 pesos paid by the project. Thus, the economic opportunity cost of a dollar of foreign exchange would be 11.2 pesos, not the 10 pesos that the project had to pay when it bought the dollars in the foreign exchange market. The extra cost of 120 pesos (= 1.2 pesos per dollar) is an externality, represented by the import tariff revenue that the government forgoes, as a consequence of the displacement of other imports, as our project enters the market to buy foreign exchange.

The concept of economic opportunity cost would be useless, for all practical purposes, if it did not refer to repetitive operations. But in point of fact the foreign exchange market is totally impersonal -- its participants do not know who is behind each $100 increment in demand, nor do they know for what purpose that $100 will be spent. All the market senses is the added pressure coming from $100 of additional demand. Hence, so far as the market's reactions are concerned, of import displacement and export stimulation, every extra $100 is like every other extra $100. We must therefore consider that the division (60-40 in our case) of the effects of an incremental demand does not depend on who is the demander or the purpose for which the foreign exchange will be used -- it depends instead on the conditions of supply and demand in the foreign exchange market.[4]

---

[4]The simple, traditional example here is that where the elasticities of imports demand and export supply are the same, the division is 50-50. If the elasticity of import demand is twice that of export supply the division will be 2/3 vs. 1/3. If the export elasticity is twice that of imports, the division will be 1/3 vs. 2/3.

Thus, the economic opportunity cost of foreign exchange can be applied to all foreign exchange transactions. In our case, any purchase of foreign exchange will represent an economic cost of 11.2 rather than10 pesos per dollar and every act of generating new foreign exchange revenues (e.g., via incremental exports) will bring about an economic benefit of 11.2 pesos per dollar. This means that we can amalgamate all of a project's purchases and sales of foreign exchange into a single net dollar value, and make the necessary adjustment (moving from the project's financial to its economic profile) by augmenting the net dollar cost or the net dollar benefit by the "foreign exchange premium" of 12% (in our example).

If the operations involved were not repetitive -- with all of the increments to demand for foreign exchange having a similar effect, this whole simplifying exercise would not be possible. We would then have to treat each act of buying foreign exchange as a separate exercise, finding for each of them its own fraction of import displacement and export stimulation, and maybe even end up with different purchases of foreign exchange displacing different imports with different tariff levels, etc. This would be an utter nightmare for the analyst, but we are lucky -- instead of calculating the economic opportunity cost of foreign exchange (EOCFX) 1000 different times for 1000 different projects, we only have to calculate it once, for all of them. This is ideally, therefore, a job for each country's project authority itself, or perhaps for an international agency like the IMF or the World Bank to do for small countries with limited technical expertise. Since EOCFX is an important parameter, and since it only has to be calculated once for each time period, it pays for those responsible to do a careful, professional job when carrying out such a calculation.

Readers should recognize that the EOCFX applies both to the acquiring of foreign exchange that the project will then spend, and to the disposition of foreign exchange that the project may have generated. It deals thus with half-a-picture, not the whole picture of a foreign exchange operation. The other half of the picture concerns how the project spends the foreign exchange that it buys, and how the project generates the foreign exchange that it sells. This part of the story is very clearly not repetitive. One project may use its foreign exchange to buy airline tickets, another to import wheat (with a zero tariff) and yet another to import a BMW car (with a 50% tariff). The tariffs and taxes paid in these project-specific operations have to be counted (as indicated in the previous section) as external benefits and costs of the project. The foreign exchange premium (or EOCFX) does not come into play on this side, but the project-specific tax and tariff externalities do operate as an offset, or balance wheel, to the extra costs involved when the foreign exchange premium is applied. Thus, in our example, we assigned a foreign exchange premium of 120 pesos to the $100 our project bought, but when that foreign exchange was used it might have brought an extra benefit of 200 pesos (if the imported items carried a tariff of 20%) or of 500 pesos (if the tariff on this item was 50%) or of zero (if the imported item entered duty free).

It is easy to see that there is nothing repetitive about the spending of foreign exchange that is bought by the project. But fortunately, project analysts should have readily at hand the project's planned purchases, so the "other half of the picture" is typically relatively simple to compute, where project imports are concerned.

The export side should be even easier, for most projects generate no export revenue at all. Those projects that do generate export receipts will typically have only one or two or at most a very few export products, most of which will carry no export tax or export subsidy. In such cases no second-side adjustment has to be made. One simply applies the foreign exchange premium as an extra benefit on the foreign exchange that the project generates.[5]

Before leaving the subject of the foreign exchange premium, we should stop to explain why we say that each purchase of foreign exchange is in some fundamental sense "sourced" either from displaced imports or newly-stimulated exports. This principle is derived from the idea of exports being the main source of a country's foreign exchange earnings, and imports being the main use to which these earnings are put. Obviously, if these are the only source and the only use, the principle becomes a virtual tautology. In point of fact, however, borrowing can be thought of as an additional source of foreign exchange for a country. But -- and here's the rub -- one should operate on the presumption that such debts will be repaid, with interest. Thus in a present value sense one is still driven back to the notion that sooner or later the foreign exchange that we extract from the market will be reflected in lower imports and more exports than would otherwise appear. And the same thought works in reverse -- if today's export earnings are not directly reflected in more imports and less other exports that means they are being lent or invested abroad. But when these loans are repaid with interest or those investments are repatriated along with the dividends they have accumulated, they will at that time be covering an excess of the country's imports over its exports that would not otherwise exist.

Of course there remains the possibility that gaps between imports and exports will be covered by foreign aid, by charitable donations, or by remittances from emigrants. These are

[5]The way this would work, following our example, would be that as $100 of export receipts is sold on the foreign exchange market, $60 of extra imports are stimulated, and $40 of other exports are displaced. On the $60 of newly-stimulated imports the average 20% tariff would apply, so there would be an external benefit of 120 pesos, to be added to the 1000 pesos received from selling the $100 in the foreign exchange market.

recognized parts of the total picture, but they are rightly considered to be determined quite independently from any given project's purchases or sales. Thus they are not considered to play any role in covering a project's demand for foreign exchange or in absorbing its supply. This, I hope, helps explain why we are so adamant about insisting that a project's effects with respect to foreign exchange take place in the world of imports and exports, and not somewhere else.

Having come this far, we still have to take one further important step before this exposition of the foreign exchange premium is complete. This next step entails shifting our focus from a project's direct purchases and sales of foreign exchange to that project's purchases and sales of internationally traded goods and services. We start with the notion that the prices of internationally traded goods and services are fundamentally determined in the world marketplace. With few exceptions, individual importing and exporting countries have little or no influence on the world prices of the tradables that they buy and sell. This has an important implication for our analysis. Argentina is an exporter of beef, so if a project located there buys beef in the local market that means that Argentine exports of beef will fall by an equivalent amount. Thus, Argentina's export earnings will fall, leading to reduced imports and newly stimulated other exports to fill the gap.

Similarly, the U.S. is an importer of copper. So even if a U.S. project buys additional copper from the U.S. producer located in Montana, that additional purchase of local copper will mean that a like amount will end up being imported (at the given world price) by other domestic users of copper. That is, it will have the same effects on the foreign exchange market as would have occurred if the project had gotten its copper via direct imports.

The bottom line of all of this is that we should apply the foreign exchange premium to the project's purchases or sales of all internationally traded goods and services -- not just to its purchases of imports and sales of exports. This extension conforms to the lessons of modern open-economy macroeconomics, which focuses sharply on the dichotomy between “tradables” and “nontradables”. Tradables in turn are broken down into “importables” and “exportables”. And, finally, imports represent the excess of demand for importables over their supply, and exports represent the excess of the supply of exportables over the demand. From these simple relationships we can easily derive that if a country’s exports exceed its imports by a given amount; its supply of tradables will exceed its demand for tradables by that same amount. And similarly, when a country’s imports exceed its exports, precisely the same gap will exist between its total demand for tradables and its total supply of them. This is the basic foundation for the rule, in cost-benefit analysis, that the foreign exchange premium should apply to a project’s total demand for tradable goods and services and to its total supply of them

Implicit in the above but not yet mentioned, the foreign exchange premium should also apply to any direct purchase of foreign currency for investment abroad and to any dividend or interest or capital repatriation flows that the project may receive as net inflows of foreign currency, to be sold on the national foreign exchange market.

The Economic Opportunity Cost of Capital

For simplicity, we will start this exposition by assuming what we call a “closed economy”, without international trade or capital movements. Later, of course, we will extend the analysis to cover the real-world case of the open economy. Most readers will have learned, in some early economics course how, when the accounts of a year are drawn up for a national economy, it has to be true that its investments have to equal its savings. The total of what is produced ends up as either consumption or investment, since national savings are defined as total production minus total consumption.

We build on the equality of saving and investment to draw the conclusion that just as a demand for foreign exchange has to come for either displacing other imports or generating newly-stimulated exports, so in the capital market a project's demand for funds must come either from displaced other investments or from newly-stimulated savings. The general picture is just the same, whether we are talking about the foreign exchange market or the capital market. It's just that in the one case the sources of funds are displaced imports and new exports, while in the other case the funds obtained from the project come from displaced investments and from new savings.

In the early years, when formalized work on cost-benefit analysis was just getting underway, the most common procedure was to consider that the alternative to "this" project's (the one being studied) use of capital funds was a "standard" investment in the rest of the economy. The investment "here" was thought to be a good one if the capital used "here" would be more productive than the same capital would be if used "there". This way of thinking naturally led to a number of attempts to measure the economic productivity of capital in the overall economy, or in its private sector, or in its "business" sector.

As thinking in this area evolved, it was soon realized that real-world mechanisms for raising money for a project typically entailed drawing those funds at the expense of both consumption and investment. In a capital-market model, one key element was a demand curve representing the demand for funds for investment in the rest of the economy (not counting "this" project). That demand curve was then juxtaposed to a supply curve of funds; both demand and supply being expressed as functions of the rate of interest. When analyzed in this way, the extraction of project funds from the capital market had two sources -- some of the funds came, in effect from displacing other investments, while the remainder came from moving up along the savings curve -- i.e., stimulating additional savings in the economy. Obviously, additional savings means reduced consumption so that project funds were seen coming part from displaced investment and part from displaced consumption.

That same dichotomy applied in the case where project funds were viewed as having come from incremental tax revenues. Tax money also comes from somewhere!! Investment and consumption were again the relevant sources, but of course the proportions in which they would come at the expense of consumption and investment would be different when the money came from taxes vis-a-vis when it came from the capital market.

The thought of considering taxes to be the relevant source of funds for a public project has a lot of immediate appeal, since most governments raise most of their money through taxes. Yet this pathway led quickly to a swamp -- different taxes had quite different effects on consumption and investment (consider personal income versus corporate income taxes or capital gains taxes versus excise taxes, or estate or inheritance taxes versus consumption taxes). Moreover, the changes in tax laws passed in one year bear little relation (in terms of their effects on consumption and investment) to those that passed in other years. All of this led to the conclusion that there is no “standard” way of extracting additional resources via taxation.

Things look much better when one considers the capital market as the source of funds. Here we see that any new demand for funds just adds an additional demand to the original picture juxtaposing investment demand against the supply of savings. That market, like the foreign exchange market mentioned earlier, is fundamentally impersonal, it “feels” additional pressure as new demands are added, and it reacts to that pressure -- not to the specific purpose for which these funds will be used. This gives us a good reason for treating the capital market as the standard source of project funds, but that would be just wishful thinking if it bore little or no relation to reality.

Luckily, we have a sound basis for treating the capital market as the "standard" source of project money. For the capital market is in fact the "sponge" that absorbs unexpected extra tax revenues as they appear from month to month, and that is indeed the source to which governments turn when they face a shortfall of monthly revenues vis-a-vis their budgeted outlays and/or budgetary overruns. This reality simply confirms that the capital market is indeed the marginal source of government funds.

Our picture, then, is one of taking funds out of the capital market, displacing investment and stimulating savings in the proportions that are determined by the supply-and-demand picture in that capital market. Fortunately we know quite a bit about how the demand for investment funds and the supply of savings respond to pressures in the capital market. Briefly, we know that the demand for investment funds is much more elastic than the supply of savings. It is quite within the range of plausibility that incremental funds drawn from the capital market would be divided 90-10 between displaced investment and newly-stimulated savings. On the other hand, it would stretch plausibility a bit to assume that this division was 75-25 rather than 90-10. Thus, the old literature that assumed that 100% of the funds came at the expense of investment was not absolutely right, but was pretty close in quantitative terms.

A simple example may help readers see the essential structure we are working with. Assume that 1000 of funds are raised in the capital market, 750 of which come at the expense of displaced investment, and 250 of which derive from newly-stimulated savings. Assume, too, that the displaced investments would have yielded a 12% real rate of return, and that the people doing the extra saving require a 4% rate of return (i.e., 4% is their supply price of saving, at the margin). Under these circumstances we would estimate the economic opportunity cost of capital (EOCK) at (.75) (12%) + (.25) (4%) = 10%.

This gives the basic idea behind the EOCK but it does not go deep enough. Readers should realize that, just because we pulled that 1000 out of the capital market the economy has really lost the product that the 750 of displaced investment would have produced. At 12%, that amounts to a stream of lost benefits equal to 90 per year, for an extended future period. This 90 per year can be seen as a "debt" that our project "owes" to the economy -- if that debt is not "paid", our project is not justified. Likewise, the 4% return (required by savers to leave them just barely indifferent versus not having saved an extra 250) is another "debt" our project has to "pay" in order to be worthwhile. This amounts to an added 10 per year.

Taken together, the 90 plus 10 per year is a challenge to the project. If it produces net benefits of exactly 100 per year, it will have just barely covered the true costs involved when the original 1000 was raised in the capital market. Only after these costs have been covered can it be said that our project was economically better than the standard alternative that would have emerged, had the 1000 of funds simply been left in the capital market (i.e., had we simply not embarked on the project).

In these circumstances, the whole story just told is automatically accomplished simply by using a 10% discount rate in analyzing the economic profile of the project. If this process yields a yet present value that is positive, the project is worthwhile, if negative, the project is not worth doing based on strictly economic criteria. If there are two projects (A and B) that are alternatives, so that only one of them can appropriately be carried out, then the choice should favor the one that yields the higher net present value, using a discount rate of 10%.

On Drawing Funds From Abroad

Our discussion so far has implicitly assumed a closed economy, in that we have only considered displaced domestic investment and newly-stimulated domestic savings as the ultimate sources of the money we extract from the capital market. In short, what we have done so far is represented below:

ρ = marginal productivity of domestic investment = 12%

r = supply price of domestic savings = marginal rate of time preference = 4%

f_1 = fraction of project funds coming at the expense of displaced investment = (750/1000) = .75

f_2 = fraction of project funds coming from newly-stimulated domestic savings = (250/1000) = .75

EOCK = economic opportunity cost of capital

$= f_1\rho + f_2 r$

= .75(12%) + .25(4%) = 10%

Now we will make the exercise more realistic by incorporating the linkage to the world capital market. The essence of this operation can be easily seen. We now might have:

f_1 = fraction of project funds coming at the expense of domestic investment = (700/1000) = .70

f_2 = fraction of project funds coming from newly-stimulated domestic savings = (100/1000) = .10

f_3 = fraction of project funds represented by a net increase in "foreign savings" coming to the country in question = (200/1000) = .20

We keep: ρ = 12%, as before.

$r = 4\%$, as before.

and we add:

MCFF = marginal cost of foreign funds = 8%.

Under these assumptions we have:

$$\text{EOCK} = f_1\rho + f_2 r + f_3\text{MCFF}$$

$$= .7(12\%) + .1(4\%) + .2(8\%) = 10.4\%$$

The marginal cost of foreign funds is very difficult to estimate, and so too is f_3, the fraction of incremental funds coming from abroad. What we know is this regard is:

a) The capital market linkage of a country to the rest of the world is not "perfect". The country can normally draw additional funds from abroad, but not at the same price. The supply curve of foreign funds available to a given country is not infinitely elastic -- instead, it has an upward slope.

b) The upward slope of any supply curve means that the marginal cost (of getting an extra unit as one moves along that supply curve) is higher than the supply price.[6]

When considering foreign sourcing of funds we have to recognize that the behavior involved is not as straightforward as the domestic demand for investment and the supply of domestic savings. That is to say, the supply of "foreign savings" shifts up and down more than

---

[6]From standard economic textbooks:

$$d(pq) = pdq + qdp$$

$$d(pq)/dq = p + q(dp/dq).$$

$$d(pq)/dq = p(1+1/\varepsilon)$$

$$\text{marginal cost} = \text{average cost times } (1+\frac{1}{\varepsilon}),$$

where ε = elasticity of supply.

Since the supply of foreign funds is upward sloping, ε is positive, and the marginal cost of funds must be greater than the average cost.

that of domestic savings, and its elasticity is also less easy to pin down to a narrow range. But, at least for many countries, it would be totally inappropriate to neglect the supply of foreign savings when trying to estimate the economic opportunity cost of capital.

In dealing with these uncertainties one looks for ways, at least of pinning down some limits for the marginal cost of foreign funds. The best approach that I know of is to try to estimate the average cost (in real terms) of foreign funds to the country in question. This number should incorporate both equity and debt financing, and should be expressed in real terms. Suppose this rate is 6%. One then estimates the corresponding "riskless" rate in the world capital market. Suppose this rate is 4%. Under these circumstances a plausible band for the MCFF would be from 8 to 12 percent. The 8% figure is found by taking the excess of the country's average rate (6%) over the world riskless rate (4%), and adding this excess (2%) to the average rate (6%). The 12% figure is obtained simply by doubling the country's average rate (6%).[7]

In advanced countries like the U.S., Canada and Western Europe, estimates of the economic opportunity cost of capital have ranged around 8 to 10 percent. This rate can be quite a lot higher in developing countries, particularly those with vibrant economies and ample investment opportunities. One guidepost can be derived from the policies and practice of the World Bank over its 60-odd years of history. That organization has faced a tricky problem in the sense that it does not seem "correct" for such an international organization to insist on a 14%

---

[7]Both of these figures follow from assuming a tangent line to the average cost curve (at its current equilibrium position) equal to $p = a + bq$. Marginal cost is then equal to $p + bq$. If a equals the riskless rate, then $bq = p - a$ (2% in our example, and this has to be added to the average cost (6%) in order to get marginal cost (8%)). If the tangent line goes through the origin, then a is zero, and we have to add another "p" to the average cost, yielding MCFF = 12%.

criterion in one country while accepting, say, an 8% criterion in another. Probably motivated more by this dilemma than by purely technical considerations, the World Bank has for decades used a real rate of 10% per annum as the "standard criterion rate" for judging whether a project merits a World Bank loan. This by itself does not reveal much about the true opportunity cost of funds in different countries around the world, but the actual experience of World Bank projects can be generally instructive. We have the fact that the World Bank has faced no shortage of projects, over the years,that were able to pass the 10% test ex ante. Moreover, the World Bank has done extensive examinations ex post, in order to assess whether its previously-financed projects had performed up to expectations. These reassessments have regularly been quite favorable, with many projects having estimated real rates of return (ex post) of 15% or more.

All of this suggests that it is not a mistake for people to think of 10% as a sort of plausible benchmark real rate of return that one should normally expect public sector projects to achieve.

A particular problem sometimes emerges when a proposed project has the prospect of outside financing at a very favorable rate. The accepted wisdom on this topic is that financing at a cheap rate should not be taken as justification for accepting projects of correspondingly low expected benefits. The country really gains when it generates a 10% yield, when the project is fortunate enough to be financed by a 2% loan. Obviously, the country would not gain if the total benefits of such a project only amounted to a 2% real return.

More important in this area is the general "fungibility of funds". Certainly, it is better to use cheap money for good projects than for bad ones. But it is highly likely that doing a bad project using cheap money will in fact end up squeezing out good projects that could otherwise be financed. Cheap money is almost certain to be "inframarginal" -- to be located in the "early" reaches of the supply curve of funds to the country,. Its most natural economic use would thus be for the most productive projects, not for ones of patently low rates of net return. In short, if a lender expresses willingness to provide money at 2%,that should always be welcome. But if that lender tries to convince you to invest that money in projects that yield only 2 or 3 percent, every effort should be made to display the availability and to argue the desirability of higher-yielding projects reflecting at least the true economic opportunity cost of capital for the country.[8]

[8]The whole exercise of calculating the EOCK can be regarded as a 2-way street. If going into the capital market generates costs which can only be covered by a project yield of, say, 10.4%, then the simple act of dumping new funds into the country's capital market will have an economic yield of 10.4%. If nothing else, then, money borrowed at a rate of 2% could buy a big benefit to a country by the simple act of using those funds to feed the country's capital market.

INTRODUCTION TO COST-BENEFIT ANALYSIS

P A R T I I

LABOR MARKET ISSUES

Arnold C. Harberger

University of California, Los Angeles

December 2008

Background

There is probably no part of cost-benefit analysis that is the source of more grievous mistakes than the labor market. And understandably -- people respond to claims like "job creation" and "employment impact" in a way that has no counterpart concerning benefits linked to the generation of foreign exchange, or to high rates of real return on capital invested in a project. More's the pity, because most of labor-oriented claims are specious and ill-founded -- they represent sloppy thinking at best.

If there is one simple lesson that is more important than all others in the labor area it is that, like the use of capital, materials, and any other factor of production, the employment of labor represents a cost to a project. It may be that this cost is lower than the wage or salary paid by the project, thus generating an external benefit, but taking both direct costs and external benefits into account the labor used in a project is virtually certain to entail a net cost.

How can one best dispel the false notions evoked by the idea of "job creation" by a project. I think the best single response is that we are not comparing our project with "nothing". Instead we are comparing it with the alternative use of the resources it occupies. Thus, if one thinks of our project's use of capital as generating a number of new jobs, we have to ask what the

**This paper was prepared under the auspices of USAID.

plausible alternative use of such capital funds would have generated. Typically, a given public-sector project will be one of a number of alternative uses of public funds that are being considered; in this case each project has direct competitors under scrutiny at the same time. But even when there are no direct competitors the project will be drawing, from the capital market, funds which would otherwise be used for investment or consumption.

The preceding paragraphs look upon jobs being generated by capital investment, which is the way those who use the term “job creation” typically think about the issue. However, we must recognize that many new jobs arise that are not connected to new capital investment, and that firms often lay off workers without there being any corresponding disinvestment in buildings, machines or other capital assets.

This is why economists greatly prefer to view the project’s use of labor as a labor market phenomenon. If a project hires new workers, the costs involved in that particular action will be essentially the same, regardless of whether the project is at the same time acquiring capital goods, or not.

When we take this view, and see the project entering the market with a new demand for labor, we see it as similar to other markets. Just as an added demand for foreign exchange is in the final analysis met by some combination of displaced imports and newly-stimulated exports; just as an increased demand for capital is ultimately filled by some combination of displaced other investments and newly-stimulated savings; just as a new demand for wheat will in the end come from a combination of increased world production and the displacement of the demand of other users, so it is with labor. An added demand for labor should be seen as being filled in part by displacing the demands of others and in part by stimulating an increase in labor supply.

Measuring the Economic Opportunity Cost of Labor -- The Standard Case

For the labor that the project gets, the normal assumption is that the project pays these workers a wage equal to their supply price. This in turn is based on the idea that these workers are getting a "market wage" for the particular type of work and in the particular area of their employment.

It is not true that the market wage is the same for all truck drivers or for all barbers or for all secretaries, throughout the whole labor market of a country. For years I have told my classes of seeing ads in the Chicago papers, offering $200 per day for truck drivers to come to northern Alaska to work on the Alaska oil pipeline, which was then being constructed. This wage seemed surprisingly high, for at that time the typical wage for truck drivers in Chicago was around $50 per day. Yes, the wage in Alaska was a lot higher, but why did these ads continue month after month? The answer is that the ads did not precipitate a mad rush of applicants, leading to a long waiting list of willing workers. This is what would have resulted if the Alaska offer were really at a significantly "above-market" (for Alaska) wage. The fact is that it required a huge premium to attract drivers from a pleasant, "normal" Midwestern routine to a place where below-zero temperatures and near-total darkness reigned for half the year, where mosquitoes swarmed in the outside air for the other half, where separation from family and friends was understood to be part of the job, where most customary forms of entertainment had no place. Economists conclude in such a case that the huge premium embodied in the $200-a-day wage was an "equalizing differential" -- a compensation that was required in order to stimulate a supply of truck drivers that was sufficient to meet the pipeline project's demand. Thus, if we had a trucking project in Chicago, we would consider the market wage to be $50 per day, but if we had one in northern Alaska, we would take the market wage to be $200 per day. And if our project were located along the beaches of Hawaii, maybe we would find the market wage there to be only $30 per day.

What, then, is the economic cost of labor in such cases? Following our earlier line, we look to taxes as the main source of divergence. If workers move from Chicago, they generate a reduction in the taxes the government receives on the basis of their earnings. These include the payroll taxes linked to their earnings, the personal income taxes they pay, plus (where applicable) the sales and consumption taxes paid as the workers spend their income. The amount of these lost taxes would be the same, regardless of whether the workers moved from Chicago to Alaska (where they would earn $200 a day) or to Hawaii (where they would earn $30 a day). Suppose these lost Chicago taxes were $5 for payroll taxes, $6 for income taxes and $4 for sales taxes, for a total of $15. On this basis, it would seem appropriate for us to take the economic opportunity cost of truck driver labor to be $215 (= $200 + $15) in Alaska and $45 (= $30 + $15) in Hawaii.

However, this is not the recommended procedure. The question is, what to do about the taxes that these workers or their employers will pay, on the basis of their wages, as they take up their project employments? In the case of the economic opportunity cost of foreign exchange, we built into the opportunity cost the tariff revenue that was lost as imports were displaced, but counted the tariff or excise taxes paid by the project on its tradable inputs as external benefits of the project itself. We did this because of the great variability of possible tax treatments among the many different tradable items the project might buy. The lost tariff revenue from the "sourcing" of the foreign exchange was standard, regardless of what the money was then spent on. But the spending of this money carried very different tax effects depending on the specific items being bought. In the case of workers of a given type being hired at a given place (e.g., truck drivers in Alaska), the amount of taxes that will likely be paid in that place, on the basis of their wages, is just as predictable as the taxes lost at the source. Hence it is quite appropriate to include these "destination taxes" as benefits, operating to offset the cost of the taxes lost at the source.

Thus, suppose the estimated taxes in Alaska were $20 for payroll taxes, $40 for income taxes and $10 for sales and excise taxes, for a total of $70. Then our figure for the economic opportunity cost (EOCL) for truck drivers in Alaska would be $200 + 15 - $70, or $145, significantly less than the $200 prevailing market wage.

In contrast, the EOCL of truck drivers in Hawaii would tend to be greater than the $30 market wage being paid. This is because the taxes newly paid in Hawaii in the basis of a $30 wage are likely to be significantly less than the $15 of taxes lost at the point from which the labor was sourced. Taking the Hawaii taxes to be $2 for payroll taxes, $3 for income taxes, and $2 for excise taxes, for a total of $7 a day, we have EOCL = $30 + $15 - $7 = $38 per day.

<u>On The "Sourcing" of Project Labor</u>

In the above examples, we have taken Chicago to be the source of the truck drivers drawn to either Alaska or Hawaii. In actuality, workers are drawn from many different sources, the identification of which is much more subtle than it initially appears.

One's initial tendency concerning the sourcing of project labor is to think like a pollster -- if the project already exists, just go and ask the workers what their last job was, where it was located, how much they earned there, what taxes they paid, etc. This, however, is not what is needed. When project workers leave their last job to come to this one, they are very likely going to be replaced by other workers, who in turn will leave other jobs where they again might be replaced, etc., etc. When we seek the sources of project labor, we are looking for the ultimate sources, not the immediate ones.

The idea of the "ultimate sources" is in reality quite simple, and it is based on an absolutely fundamental element in project evaluation, in policy analysis, and in the whole field of applied welfare economics. What we are doing in such analyses is comparing an economic equilibrium that exists in the presence of a given project (or policy) and comparing that equilibrium with the one that would exist in the absence of that project (or policy). When we analyze the effects of a given tax, we usually employ a simple supply-and-demand framework, and insert into that framework first a zero tax, and alternatively a tax of size T. We then compare the equilibria that result from these two exercises to estimate the efficiency cost of the tax, or its incidence on different groups, or its revenue yield.

When we deal with project analysis the story is similar, but more complicated. For here we are looking at the profile of the project's costs and benefits, over its entire life. We thus have to think of comparing two moving pictures -- one depicting how the world would evolve (in its relevant aspects) in the absence of a project, the other showing a similar evolution in its presence. Our project profile is then obtained by each year's difference between these two moving pictures, with respect to the project's direct inputs and outputs. The external effects of the project are similarly obtained by taking, for each period, the difference with respect to government tax receipts and possibly other relevant distortions.[1]

---

[1]Distortions play a very important role in applied welfare economics. Without entering into a full exposition, I hope I can convey the key point, that we are not concerned with changing prices and quantities in other markets when those markets are not distorted. In an undistorted market, the market price measures the economic opportunity cost from the side of the suppliers, and it also measures the benefit of added units from the standpoint of the demanders. Movements up and down of the equilibrium price and quantity thus have incremental costs equal

We have to think of these same two moving pictures as we contemplate the sourcing of labor for a given project. Rather than try to expound this subject in abstract terms, I will try to convey its essence by describing a real-world study in Mexico, on which I served as an adviser. The basic data for this study were labor market surveys in around 40 different labor market areas of the country. In each of these areas the surveys provided data on wages, average hours, average earnings, etc. for some 100 different occupations, with separate data for men and for women.

The first task was to simplify the problem by amalgamating the 100 different occupations into meaningful and manageable groups. Thus we had common, semi-skilled, and skilled labor, ordinary clerical and skilled clerical workers, technicians, professionals, and perhaps one or two more categories. The rule for putting several occupations into one category was that these occupations had to pay similar wages.

So now one had the average wage rates of, say, 10 labor categories, in 40 labor markets covering basically the whole territory of Mexico. For each category and each location, the tax payments linked to that category in that place were estimated (analogous to $70 in Alaska, $15 in Chicago, and $7 in Hawaii).

Next, we had to make assumptions about the geographical sourcing of incremental labor for each of the 40 centers. In all cases it was assumed that half the labor came at the expense of other jobs in that same labor market area. Different assumptions were made about the sourcing of the other half of project labor. Under the first alternative -- called the donut -- the second half of project labor was assumed to come from the contiguous labor market areas, in proportion to

---

to incremental benefits. When distortions are present, there is an economic gain when quantity goes up and demand price (benefit) exceeds supply price (cost), and an economic loss under the same concept if the equilibrium quantity goes down (as between the two moving pictures).

their existing numbers of workers in the category (e.g., semi-skilled). Under the second alternative -- called the nationwide -- the second half of project labor was assumed to be drawn from all 39 other labor market areas, again in proportion to their existing numbers of workers in the category.

So, if w_{ir} is the monthly wage of workers of category i in region r, one estimated the amount of T_{ir} -- equal to the monthly taxes paid, linked to the wage w_{ir}. Then the economic opportunity cost of labor of category i in region s would be

$$w_{is} + [T_{is} - \sum_{r} a_{ir} T_{ir}].$$

Here a_{ir} for region r would be the fraction of workers in category i (in this case semi-skilled workers) coming from each sourcing region. In the Mexican study, a_{ir} was always 1/2 for the region in which the project was located. The remaining half would be distributed across the other sourcing regions in proportions of their numbers (N_{ir}) of workers in the category in question. That is, a_{ir} for regions other than s would be equal to $(N_{ir} / \sum_{r \neq s} N_{ir})(1/2)$.

This may look more complicated than it really is. What we are doing is taking as a benefit the taxes to be paid by (or on account of) workers in the project, and as a cost that taxes that would have been paid by (or on account of) them in the employment they likely would have had in the absence of the project.

Where one has sourcing among major cities in a country, one typically does not expect to find significant difference in the taxes paid (per worker) as between source and destination. However when a project is in the main metropolitan area of a country, and is likely to be sourced largely by migrants from the rural hinterland, then there can be large differences in taxes paid, which will tend to lead to an economic opportunity cost of labor somewhat lower than the market wage. This is likely to be the case for low-skilled labor in many developing countries (notably China and India). There the wage paid in the urban destinations of migration tends to be very substantially higher than the wage in the rural sources of migration. This is taken as a genuine economic cost, as it reflects the fact that labor probably would not move to the cities in the absence of a significant premium (as in the case of truck drivers in Alaska, but with a more modest premium (e.g., city wage perhaps double the rural wage).

The bottom line of this section is simply that most of the time one takes the wage actually paid (or to be paid) by the project as the first approximation measure of the economic opportunity cost of labor. This is then adjusted by a difference in taxes linked to the wage (payroll taxes, personal income taxes, and sales and excise taxes) as between employment on the project in question and the alternative employment "at the source" (conceptualized as the difference in employment at each source as between the two moving pictures -- "with" and "without" the project in question).

Dual Labor Markets -- Protected vs Unprotected Sectors

In many developing countries we observe the phenomenon of "dual" labor markets. Sometimes these are referred to as "formal" and "informal" sectors; sometimes as "modern" and "traditional" sectors. I like to use the terms "protected" and "free" sectors, because these terms most readily reflect the theoretical underpinnings of a dual labor market. The easiest way to visualize this syndrome is to think of everybody in the labor force wanting to have a job in the protected sector, but only a fraction of them are lucky enough to get one. The rest fall into two categories -- those who take a job (at a significantly lower wage) in the free sector, and those that I call the quasi-voluntary unemployed -- who are eager to get a job in the protected sector but who are not willing to accept the wages and working conditions of the free sector.

To put a little meat on these conceptual bones, let me recount some relevant experiences from India. I spent the academic year 1961-62 working on a project of the MIT Center for International Studies, which served in a collaborative and advisory capacity with the Planning Commissions of India. The Center's headquarters were in the diplomatic section of New Delhi, where most embassies were located. And the project's financing came mainly from the Ford and Rockefeller Foundations. It should come as no surprise, then, that the wages and salaries paid by the project were significantly above the free market level -- being in this respect similar to those paid by the U.S. Embassy, the U.K. High Commission, and nearly all the other missions representing modern industrial countries.

It should come as no surprise, then, that just about everybody in the Delhi labor force would like to have a job in this sector. Many embassies had waiting lists of hundreds of applicants. In our project's case, we didn't keep a waiting list, but a steady stream of applicants came to our office door. Sensing the relevance of our situation, I asked Mr. Lakshman, our office manager, to peek into my office when he finished interviewing each applicant. If I was not too busy, he would then pass that applicant on to me, for a second session. My practice in such cases was to pull out the "help wanted" page of the daily paper, and go through the ads with the applicant looking over my shoulder. Quite regularly, we would find ads for salesmen, particularly for travel to rural areas. People with middle-school education were typically wanted, for example, as salesmen to sell pump-irrigation equipment to farmers. The pay, as I remember was quite good, but not a single one of my interviewees expressed any interest in a pump salesman job. A number of them expressed actual disdain for such a job, thinking it to be far below their own social status. When pressed as to what kind of jobs they would like, they almost always mentioned the U.S. Embassy and the British High Commission, plus other diplomatic representations. Usually, too, they mentioned the Indian government as an employer they would like. In such cases, I would turn again to the want ads, and usually found some that had been placed by the Indian government. Typically, these might be for educated Indians to serve as public sector representatives and technical advisers in rural villages. Once again these jobs were spurned, usually with comments like "What would I ever be able to do in a village?" The fact that these types of ads came up repeatedly convinced me that they did not represent protected sector jobs. What I was seeing was a whole series of people who were eager candidates for genuine protected sector jobs, but who had no interest in the available, specific, free-sector jobs.

Quasi-voluntary unemployment is quite common in developing countries, particularly among the more skilled and better educated portions of the population. In India, at least when we were there, the highest rate of recorded unemployment was among recent university graduates, who often chose to wait for a job with good working conditions, ample pay, and favorable career promise. Sometimes they would be on the labor market as long as two years before they found a job they were willing to accept. Of course, in order to be classified as unemployed they would have to respond positively to questions like "What efforts have you made during the past month in order to seek employment?" And of course their answer would be "I've been to the American Embassy, the U.K. High Commission, the Indian Foreign Office, the Delhi branches of Monsanto and General Electric, etc., etc." They would thus end up classified as unemployed. But quite clearly their unemployment was quasi-voluntary. They were involuntarily unemployed, looking up at the protected sector wage, and voluntarily unemployed looking down at the free-sector wage.

What can we say about the economic opportunity cost of labor in such cases? The answer is that the opportunity cost of labor for a protected sector job (at wage w_p) is higher than the free-market wage, and lower than w_p itself. The most fruitful assumption to make here is that the people who end up filling a set of protected sector jobs are selected by criteria other than their individual supply prices of labor for these jobs.[2] Those who are working in the free market at a wage w_f are deemed to have that as their relevant supply price (i.e., they would gain a personal surplus of the extra earnings (w_p-w_f) if they were selected for a protected sector job). Those in the quasi-voluntary category are assumed to have supply prices spread out evenly over the range between w_f and w_p, so their average supply price is taken to be $(w_f+w_p)/2$. Hence the EOCL for a protected sector job at wage w_p is estimated to be a weighted average of w_f (with a weight proportional to the number of workers of the given category who are employed in the free sector) and $(w_f+w_p)/2$, with a weight proportional to the number of quasi-voluntary unemployed in that category.

Of course, each occupation, each labor category, will have a different free market wage, w_f. And so far as w_p is concerned, there can be many different entities paying more than the free-market wage to a given category of labor. We have mentioned the diplomatic sector and certain types of government jobs. In addition we find protected sector jobs where the "protection" comes from minimum wages, or from strong labor unions. One also finds, as alluded to above, protected sector jobs in the offices and plants of multi-national companies. These companies are particularly vulnerable to labor unrest, because they pay so much more (for what appears to be the same type of work) in their home base than they do in developing countries. Their best response to ward off such unrest is to offer wages and working conditions

---

[2]Formally, the conditions are that from among the available applicants, the successful ones are chosen by criteria uncorrelated with their individual supply prices.

such that they always have a large backlog of willing applicants waiting for each kind of job.

The EOCL for a protected sector job then will be a function, both of the free-market wage, and the w_p for that particular job. In general, the EOCL will be higher for employment at a wage of $10 a day than at one of $7 a day, assuming the free-market wage to be, say $4 a day. Why is this? Simply because there are more quasi-voluntaries waiting around for the $10 job than there are waiting for the $7 job. Thus not only do we have ($4 + $10)/2 entering into the weighted average, rather than ($4 + $7)/2 in the case of the $7 job, but we also have a heavier weight in this part of the formula for the $10 job vis-a-vis the $7 one.

What should also be clear is that the presence of quasi-voluntary unemployment is different from the phenomenon of cyclical unemployment. In the latter case, there is the presumption that, the higher the observed unemployment rate, the lower is likely to be the economic opportunity cost of labor. In the case of quasi-voluntary unemployment, it will be higher, for a given labor category, the higher is the relevant rotected sector wage, w_p. Hence higher quasi-voluntary unemployment is likely to be linked to a <u>higher</u> economic opportunity cost of labor for protected-sector jobs at a higher wage.

The previous discussion, leading to the conclusion that the EOCL for labor hired at the free-market wage (w_f) is w_f itself, and that for labor hired at any given protected-sector wage w_{pj} (higher than w_f) would be a weighted average of w_f itself on the one hand and $(w_f+w_{pj})/2$ on the other, was intentionally kept simple, so as to focus directly on the essential features of a dual labor market. Thus we did not introduce the complications linked to taxes paid on the project job and taxes forgone on the locations from which the labor was sourced. Our having simplified in this way does not mean that we should neglect these tax externalities, in a real-world dual-market case. Rather, one should in such circumstances graft them on to the barebones analysis that we have presented. In this type of case, the full amount of taxes paid by a new protected sector job would be a benefit, as before. Similarly, the assumption is that no taxes would have been paid by the quasi-voluntary unemployed, and that those drawn (in net terms) from the free-market sector would have paid taxes in amounts linked to the free-market wage, w_f.

Migration-Fed Unemployment

In a number of developing countries one encounters the special case of migration-fed unemployment. Typically, this phenomenon occurs in a growing urban labor market, where migration from rural areas is an important source of that labor-force growth.

To begin, one should recognize that there is nothing at all anomalous about rural-urban migration -- a process that appears to be an integral part of economic development, no matter where. A properly-working labor market would absorb migrants as they came, allowing of course for a reasonable quota of frictional or search unemployment upon arrival. As in any supply-and-demand situation, one expects the market price (the wage in this case) to continually adjust so as to maintain approximate equilibrium between supply and demand.

The special feature of situations of migration-fed unemployment is the fact that in these cases the wage is not performing its natural function of equilibrating supply with demand. The easiest way to visualize this is to assume that something (typically a legal minimum wage) is artificially holding wages (quite generally in the urban destination) above their equilibrium level. People in the rural area -- getting news of such an attractive opportunity, are motivated to move in large numbers to the city. This process would go on and on and on, if each migrant settled quickly into a regular urban employment. But that would bring a Niagara of migrants -- something that the city couldn't absorb and, more important, something that we do not observe.

What we see is a process of migration that in a rough way keeps pace with the increasing demand at the urban destination. That is, we see something similar to what we would observe if the labor market were working as it should with the wage rate performing its equilibrating role, but with one big difference -- the unemployment rate stays high, year after year. That is the anomaly.

The explanation is not hard to find, nor hard to understand. What happens in these cases is that the unemployment rate becomes the equilibrating variable, taking over that role form the wage rate. With a very attractive, high urban wage rate a heavy flow of migrants is to be expected. But demand is not increasing as fast as this supply so unemployment grows. With each increase of the unemployment rate, the incentives perceived by potential migrants grow less attractive. At a zero rate of urban unemployment they all want to move to the city. At, say, a 30% urban unemployment rate, they all want to stay home on the farm. Somewhere between these two extremes there has to be a rate of urban unemployment that makes moving to the city just barely attractive to potential migrants. This is the rate that will tend to prevail in a situation of migration-fed unemployment. When the unemployment rate is lower than this, net migration will be stimulated; when the unemployment rate is higher than this, net migration will slow, or even turn negative, until the equilibrium rate is restored.

Table 1 gives a stylized picture of how unemployment serves as an equilibrating force. The table is based on the assumption that each worker in the urban labor force has the same probability of getting a job. This is an unrealistic assumption, but it greatly simplifies the exposition. Readers should be assured that the example can be adapted to incorporate more realistic assumptions (e.g., that migrants, when they first arrive face a very high unemployment rate, which then gradually declines to a significantly lower rate as with the passage of time they

TABLE 1

Exercises in Migration-Fed Unemployment

(Supply price of migration = $4/day)

Panel 1: Urban Wage = $5/day; Equilibrium Unemployment Rate = 20%

		Employed	Unemployed	Labor Force	Expected Earnings
1)	Initial Equilibrium	800	200	1000	$4/day
	Now Create 100 New Jobs				
2)	Short-Run Response	900	100	1000	$4.50/day
3)	Long-Run Response (New equilibrium)	900	225	1125	$4/day
4)	Change from 1) to 3)	+100	+25	+125	-0-

Panel 2: Urban Wage = $6/day; Equilibrium Unemployment Rate = 33 1/3%

		Employed	Unemployed	Labor Force	Expected Earnings
1)	Initial Equilibrium	800	400	1200	$4/day
	Now Create 100 New Jobs				
2)	Short-Run Response	900	300	1200	$4.50/day
3)	Long-Run Response (New equilibrium)	900	450	1350	$4/day
4)	Change from1) to 3)	+100	+50	+150	-0-

assimilate more fully into the urban setting). Such an adaptation process would change neither the basic structure of the analysis nor its lessons and conclusions.

Table 1 is built on the assumption that the equilibrium rate of unemployment is 20% and that the “supply price of migrants” is $4 a day. That is to say, migrants will be satisfied with their move, so long as their average earnings amount to $4 a day or more. An expectation of less than $4/day previous migrants will want to go back to their rural home; with a higher earnings expectation, a flood of new migrants would come to the city. In the initial equilibrium, we have a labor force (of a given category of labor) of 1000, of which 800 are employed and 200 unemployed. The employed workers earn the established wage of $5 a day but their expectation is to be employed one fifth of the time, so their expected earnings are $4 per day, just equal to the "supply price of migration". We thus have equilibrium, with no incentive for either out- or in-migration.

Next, 100 new jobs are created. In the immediate short run, the labor force stays the same, yielding 900 employed (at $5 per day) and 100 unemployed. Now expected earnings are $4.50 per day, well above the $4 supply price of migration. This attracts new migrants, who will keep coming until a new (long-run) equilibrium is established. That will occur when the labor force of this skill category has swelled to1125, of whom 900 will be employed and 225 will be unemployed on an average day. The expected earnings would again be $4 per calendar day, once again equal to the supply price of migration.

We now have to ask, what economic opportunity cost should be assigned to the 100 new jobs created in this example? The answer is $5, the actual urban wage rate applying in this case. Recall that this wage is higher than would prevail in a straight supply-and-demand equilibrium. Being artificially higher, it attracts migrants to the point where they end up just barely satisfied with having made the move. The real "supply price" of the 125 migrants is (125 x $4/day), or $500 per day. But this is exactly what is earned by the 100 of such migrants who will be employed on an average day. Looked at in this way, the $500 actually paid just barely covers the resource cost of the 125 migrants who were attracted as a result of the 100 new jobs.[3]

The lower panel of Table 1 reflects a case where the urban wage has been raised to $6 a day, and where as a consequence the equilibrium rate of unemployment has moved up to 1/3. The initial equilibrium thus has a labor force of 1200, of which 800 are employed and 400 unemployed on any given day. Actual earnings of the 800 are (800 x $6), or $4800 per day. Expected earnings of the 1200 are $4800/1200, or $4 a day, precisely equal to the supply price of migration.

When we add 100 jobs, the labor force initially stays at 1200, but employment goes up to 900. Expected earnings rise to $4.50, again triggering a new wave of migration. A new equilibrium is reached when the labor force has increased to 1350; employment is 900 and unemployment 450 in this new equilibrium. Thus the creation of 100 jobs has led to increases of 100 in employment and of 50 in unemployment. This is the bane, the curse of migration-fed unemployment. If one's aim is to reduce the number of urban unemployed, creating urban jobs is actually counterproductive!!

The EOCL in this case turns out to be, as before, the same as the urban wage. Creating 100 jobs has caused 150 migrants to move to the city. The economic cost (supply price) of those

---

[3]An alternative way of reaching the same conclusion is to note that the financial cost (to the project) of the 100 jobs is $500 per day. On any given day 100 migrants will tend to be employed at $5, each earning a positive surplus (above the $4 supply price of migration) of $1 per day, for a total surplus of $100 per day. But in addition we have 25 migrants earning zero, representing a deficit of $4 a day, in comparison with their supply price. These 25 cases generate a deficit of $100 per day. When we adjust the $500 financial cost to take these external effects into account, we get $500 + -100 - 4100 = $500. Thus the economic opportunity cost turns out to be equal to the financial cost.

moves was 150 x $4/day, or $600 per day. That is exactly equal to $6 per day (the actual wage) for each of the 100 new jobs.

Though our simple example may seem remote from the real world, the analysis of migration-fed unemployment has genuine roots in reality. Harris and Todaro were puzzled by the chronic unemployment they saw in Africa. I was puzzled by a constant high rate (15%) of urban unemployment in Panama during the 1960s. Harbison noted how, when a sudden large spate of job creation occurred in Kenya, new migrants came in numbers sufficient not only to fill all the new jobs, bur also to increase the pool of unemployed in the same proportion, leaving the unemployment rate the same.

To repeat the main lesson emerging from this section -- when the conditions of migration-fed unemployment exist, new projects created in the urban destinations should get no credit for absorbing unemployment. New urban jobs in such circumstances actually add to urban unemployment, when one moves from one labor market equilibrium to another.

Cyclical Unemployment

This topic was intentionally left for last because, in spite of its being a natural subject of policy interest and at the same time posing fascinating problems for the analyst, it is probably a matter of rather minor importance in a world of institutionalized, formal economic project evaluation. The reason for this is simple -- when such project evaluation is being done as a serious, real-world exercise, it typically precedes by some significant amount of time the actual implementation of a project. Most of the time, groundbreaking in a major project will not occur until three or four years after the project has been evaluated and approved. All the detailed engineering, the packaging of project finance, the assembly of leadership teams will take place during the interim.

So if the economy happens to be in a cyclical downturn at the time a project is being evaluated, it will very likely be out of those woods by the time the project actually gets under way. One then looks forward to all the years of the project's expected economic life -- very often 30, 40, or 50 years. What can one say about the projected cyclical stage of the economy in those years? The best-assumption by far, is that we will be dealing with normally-functioning labor and product markets as the economy lives through the expected lifespan of the project.

Moreover, in project evaluation we normally are pitting our project against an alternative, either explicitly or implicitly. Capital for the project is deemed to come mainly at the expense of other investment or of consumption. Labor in the project comes mainly from other employments, etc. There is a real question of how to fit such a framework into a setting of recession or depression. Certainly we would not want to count the project's absorption of some number of unemployed as a big project benefit, without recognizing that most other expenditures of the same money would also have similar effects.

Seeking to resolve the puzzles posed by cyclical unemployment in a reasonable way, I like to think in terms of the subset of projects that actually have unemployment absorption as one of their important objectives. The setting would be something like the present (December 2008) recession, or that of the mid-1970s or the early 1980s in the United States, or the 1980s debt crisis period in Latin America. Each project is then looked upon as drawing its labor force: a) from other employments, b) from new labor-force entrants and c) by drawing down the pool of the unemployed.

It is obviously item c) that makes the cyclical case differ from the norm. What we seek here are two fractions -- first, the fraction of our hirings that will be represented by the net reduction that they induce in the pool of the unemployed; and second, the amount by which the project's outlay on such workers exceeds their true economic opportunity costs.

There is no way of getting anything like a precise estimate of either of these two fractions, but a good starting point is to recognize that the fractions, as it were, move up and down together. Thus, in a mild recession, perhaps only 10% of those hired would come (in net terms) from the pool of the unemployed, and they in turn might have an economic opportunity cost only 10% below the wage they receive from the project. In a deep depression, perhaps as many as half the workers hired might represent net reductions in unemployment and the EOCL might be 40% below their project wage. In between these two extremes we might have a serious recession (not yet a depression) in which, say, a quarter of those hired would represent reductions in unemployment, and their EOCL might be 20% below the project wage.

Since years of recession and depression are not the norm, and since the calibration of new projects in terms of their unemployment absorption is only for a specific subset of evaluations (focusing on projects ready for essentially immediate adoption), one can probably do quite well testing a couple of alternative fractions -- say a 10% absorption fraction plus a 10% wage discount on the one hand, and a 20% absorption fraction plus a 25% wage discount on the other.

Some readers may be surprised at the fact that the suggested wage discounts are so modest. The answer is that we have to be modest in this regard if we want to respect the evidence. Ultimately, voluntary supply price is the best measure of opportunity cost as seen by the worker himself or herself. Many economists start out with the idea that unemployed workers make major adjustments in their supply price, bringing it down step by step as their unemployment spell gets longer and longer. Surprisingly, however, the data do not support this view. By and large the stated supply price of the unemployed remains much the same, independent of the length of the spell. Yet these unemployed workers celebrate when they finally get a job. Here we have a contradiction. If one's stated supply price is $2000 per month, economics tells us to infer that at that wage the worker is indifferent between working and not working. But the celebration on finally getting a job negates such indifference. Resolving this contribution is a standing challenge to economists -- we certainly cannot do it here. But the contradiction itself signals us to proceed with caution. Certainly assigning a zero or very low opportunity cost, as many are prone to do, is ill-advised in the extreme, in the face of the evidence.[4]

Special note should be taken of a phenomenon that can indeed lead to a low EOCL for labor that is absorbed from the ranks of the unemployed. I refer to cases where the government has a functioning program of unemployment compensation, so that a reduction in the ranks of the unemployed generates an important fiscal saving for the government (and, more fundamentally, the country's taxpayers). In this case the employment of an unemployed worker in a $3,000 a month job might mean an economic opportunity cost of only $1,200 a month -- the difference stemming from a saving of $1,500 a month in unemployment compensation payments to that worker, plus, say, $300 a month in taxes that will be paid on the basis of the $3,000 salary. This element, however, is rarely present in low-income developing countries. Serious programs of unemployment compensation are largely concentrated in the so-called "first world".

---

[4]I tell my students every year that I am quite willing to assign a zero EOCL to any worker who is willing to work for zero compensation. And there are many of them -- volunteer workers in hospitals, sport leagues, churches, etc. But we don't find them very often as members of a country's regular labor force!

INTRODUCTION TO COST-BENEFIT ANALYSIS

P A R T   I I I

ADDRESSING SOCIAL CONCERNS

Arnold C. Harberger

University of California, Los Angeles

December 2008

Background

From its oldest roots in the early 19th century, applied welfare economics has focused on the goal of economic efficiency. To be sure, its approach was anything but crude -- it could deal, for example, with the supply price of a person's labor being different for different jobs in the same place or for the same job in different places (sees Part II). It could also handle the inefficiencies of consumption stemming from each household getting a given ration of milk (in spite of their having very different intensities of demand) with the same facility as it handled the inefficiency of awarding each farmer a given number of acres on which he could plant wheat (in spite of the productivity of the land being very different from one farm to another). But beneath it all was the treatment of a dollar of cost as counting the same, regardless of which person or group within the society ended up bearing that cost, and similarly for dollars of benefit accruing to different persons or groups.

This focus on efficiency is embodied in the three basic postulates on which applied welfare economics has been based. These can be summarized as: 1) benefits being measured at each step by demand price (= willingness to pay), 2) cost being measured at each step by supply price (= willingness to supply) and 3) the aggregation of these benefits and costs across

**This paper was prepared under the auspices of USAID.

individuals and groups, regardless of who within the society in question was enjoying the benefits or bearing the costs (= adding up).

It should come as no surprise that even from the earliest days it was the third postulate (adding up) that became the focus of controversy and discussion. Ask 100 people whether an incremental dollar will “do more good” if put in the hands of a poor person or a rich person, and all or nearly all of them will without hesitation side with the poor person. It is only a small step to move from this answer to the idea of “distributional weights” -- of weighing a dollar increment of benefit or cost differently, depending on the individual a family or group to whom it accrues.

But, as we will see later, applying distributional weights systematically, within the frame of cost-benefit analysis: 1) is extremely difficult to do, and 2) carries many implications for policy that most people are unwilling to accept. The standard way to escape from the problems posed by 1) and the dilemma posed by 2) has been to stick to the three postulates of demand price, supply price, and adding up; but at the same time to emphasize that “all we are doing is measuring economic efficiency”. This is a key objective in economic policy analysis, and it is something that we really do know how to measure. Thus we can with a clear conscience say that a given agricultural policy has an efficiency cost of $800 million, or that a given slum-clearance project has an efficiency cost of $50 million, and then leave it to the “authorities” to judge whether or not the non-efficiency benefits of that policy or that project are sufficient to outweigh the $800 million or the $50 million of efficiency costs that we measure.

This “efficiency-only” position is probably the safest one for a defender of cost-benefit analysis to take. It does not claim that we are able scientifically to measure the non-efficiency benefits or costs entailed in a shift of a benefit from one group to another, or in the shift by the country's military to a new weapons system, or in measures that, at the expense of economic efficiency, accommodate the policy demands of one special-interest group or another. We measure the efficiency costs, and let somebody else worry about the non-efficiency aspects of a policy or a project.

I believe that all cost-benefit professionals have to adopt something like an efficiency-only position at one level or another. We have no professional business in placing a dollar value on improved relations with India (which may be an important byproduct of a given project or program), or on many national defense outlays. But move in a little closer to our own terrain and you enter a sort of no-mans-land, where there are good arguments for applying efficiency standards, yet where in order to do so we have to place dollar values on a whole array of benefits or costs that are often very hard to quantify.

The value of a human life is a case in point. The average citizen's instinctive reaction is "no amount of money is sufficient to compensate for the loss of a human life", yet there is a myriad of policies, programs and projects that in effect embody a tradeoff between dollar cost and human life -- the setting of speed limits, the placing of traffic lights and stop signs, the building of median strips on highways, the straightening of dangerous curves are just a few examples related to roads. If we count the life-taking costs and the life-saving benefits of such decisions as being outside our purview, simply to be weighed by the "authorities" as non-economic benefits or costs, we find ourselves with a serious problem. For we actually <u>can</u> estimate with some accuracy how many lives per year would be saved by imposing a national speed limit of 55 miles per hour and we can also estimate the costs (mainly in travel time) that such a policy would entail. Relating those two, we have an economic cost per human life saved that is implicit in either adopting or rejecting a 55 mph speed limit. Then we can do the same for the placing of traffic lights and stop signs, for the introduction of median strips, and for the straightening of specific curves on specific roads. Doing all this we would find that very different implicit values of human life emerge from these different exercises.

Thus we might find ourselves with situations in which we are paying $10 million to save a human life by straightening a specific curve, but where we could save other lives at a cost of $1 million per life by the judicious introduction of stop signs. Obviously, we cannot juggle hundreds of such specific comparisons in our heads as we analyze a host of different policies, projects and programs. The way out is the introduction into our cost-benefit framework of a "shadow price" of a human life. If this price if $5 million per life, then the stop sign projects would generate a net benefit of $4 million per life, while the curve-straightening project would show a net cost of $5 million. Instead of all the projects being juxtaposed one to each of the others, each single project would be assigned a benefit of $5 million for every human life it was expected to save, and a cost of $5 million for every life it was likely to take. All of a sudden the saving and taking of human lives has entered the world of efficiency calculation!

The moment we feel ready to place a monetary value on a given "noneconomic" objective, that value opens the door to incorporating that objective into the efficiency-oriented calculus of cost-benefit analysis. What we can do with life years we can also do with the value of commuter time, with the valuation of free public services such as those of public parks. On the cost side we can, for example, introduce prices for the various pollutants that a project might introduce into our atmosphere or our waterways.

It is here that we rather quickly reach a crossroads. While we may be prepared to set a price on carbon emissions into the atmosphere or on nitrogen spewed into a river or lake, we may not be quite ready to do the same for a battalion to be added to the army or a submarine to be added to the navy. Our readiness to quantify some "noneconomic" benefits or costs varies with how well we think we can pin down those numbers. If we are confident that the value we seek lies within a range of 10% or 20%, or even 30%, then we can still feel like professional economists when we incorporate that range (or a central value within that range) into our analysis. But if our valuation is so uncertain that it spans a range of 300% or 500% or 1000%, then we are probably better off not trying to introduce such an item directly into our quantitative analysis, and simply passing the buck on to the "authorities". Put another way, where the range is very wide, it can end up useless.[1]

In sum, those who want to introduce greater rationality into the decision process on public expenditures have, in cost-benefit analysis, a very worthy product to sell. In every country there is a vast array of projects and programs to which known, readily available techniques can be applied. And one can be quite certain that with enough effort and ingenuity, we will be able to keep on expanding the scope for reasonable application of cost-benefit analysis. But we should beware of overextending ourselves -- there is plenty that we can do while still claiming our work to be "professional", and there are many interesting possibilities for extending the range of projects over which we can function as professionals, but there also is another range of projects for which our ability to quantify benefits and costs is too limited or too vague to be useful.

On Distributional Weights

One of the important areas in which "noneconomic" considerations are often broached concerns the distribution of income and/or wealth. The idea that a dollar in the hands of a poor

[1] Like telling a pregnant mother that you can predict what height her new offspring will reach at age 21 -- and then stating the range to be from three to seven feet!

person is worth more (from society's point of view) than a dollar in the hands of someone much richer -- that idea has deep roots in most people's thinking. And it also has roots in the field of economics. The notion of people's well-being being measured by their "utility" dates back at least to the early 19^{th} century, and has a long history from that point on. "Utility" appears at three levels in our literature -- "ordinal utility", meaning individuals can state their preference (or indifference) as between any two bundles of goods and services (or any two situations); "individually measurable utility", meaning that people can rate differences between bundles (e.g., saying that the difference between A and B is bigger or smaller (in utility terms) than the difference between B and C); and "measurable and interpersonally comparable utility", which says that one person is enjoying more utility than another or that an incremental dollar is worth more (in utility terms) to one person than to another. Much of economic theory can be derived just using the notion of ordinal utility (indifference curves, etc.), but the analysis of risk typically requires one to take the next step, to individually measurable utility. Neither of those provides any basis for a distributional weights framework. For that one needs to take the third step -- to measurable and interpersonally comparable utility.

Early utilitarian thinking was based on this latter assumption, but did not pursue its detailed implications. That part came later, particularly on the subject of optimal income taxation. Here we have an extensive literature stemming from the past several decades. This literature assumes that each of us has the same utility function, translating income (or wealth) into utils (the units in which utility is measured). Higher income translates into more utils, but an extra dollar contributes less and less as income rises.

In the optimal income tax literature, the problem is posed of raising a certain amount of money through an income tax, when the weight given to incremental dollars declines with income. Typically, the examples used in this literature assume the weight is cut by a quarter, a half, or three quarters, every time income is doubled. When such an assumption is applied to income distributions similar to those we observe in the real world, the resulting "optimal" income tax structure tends to have an unexpected shape -- marginal rates of tax tend to fall as income rises. This seems counterintuitive at first, but begins to make sense when we take into account that an income tax structure is composed of a series of income brackets. Raising the marginal rate for any one bracket introduces an efficiency cost by creating a new disincentive for work on the part of people in that bracket. But it produces a distributional benefit by shifting dollars to the government, not only from that bracket but also from all higher brackets. Thus, for bracket 1 out of 5 a rise in the tax rate has one efficiency cost plus 5 distributional benefits, for bracket 2 it has one efficiency cost plus 4 distributional benefits, etc. By the time we get to bracket 5 we have one efficiency and only one distributional benefit. It is the fact that fewer distributional benefits come from raising marginal rates in the higher brackets that produces the counterintuitive result of the "optimum" marginal rate declining as income rises.[2]

What is troublesome about the optimal tax literature is that its results come from a pretty fancy set of calculations using distributional weights, and seem to argue for more moderate (i.e., less progressive) income tax structures than those we actually observe in most countries. This is

---

[2]To add to the anomaly, the typical optimal pattern has average tax rates rising (up to the top bracket) at the same time as marginal tax rates fall, as income rises. This is due to the existence of an optimal exemption level. Thus an optimal marginal rate structure might be zero up to $20,000 and 30% from $20,000 to $40,000, taking a tax of 6000 (= 15%) from an income of $40,000. It could then go on to take 25% on incomes from $40,000 to $80,000, the total tax on $80,000 being $16,000 (= 20%), and (with a marginal rate of 22% in incomes over $80,000) collect $20,400 (= 20.4%) on an income of $100,000, and $42,400 (= 21.2%) on an income of $200,000. This attribute of the optimal structure being progressive in the average rate at the same time as it is regressive in the marginal rate, was sort of reassuring to readers of the optimal tax literature -- the result was not totally counterintuitive.

taken as reassuring by many people; it thus serves to foster the general acceptance of the distributional weights approach.

Unfortunately, this last step is not warranted. In order to assess the merits of any systematic approach, one has to test it throughout the relevant range. In particular, if we are to use distributional weights in the field of project evaluation we have to test that approach as it would apply to the approval of specific projects and to the problem of choosing among alternative projects.

Looking at one project alone, consider:

	Unweighted	Avg. Weight	Weighted
Present Value of			
Benefits	500	1.5	750
Costs	-1000	0.5	-500
Net Present Value	-500		+250

If we take distributional weights seriously, and use the indicated weights, we must recommend acceptance of this project, in spite of its net efficiency cost of 500.

Suppose one says no to this on the ground that the government could make a simple transfer of 500 to the beneficiaries, and thus get the same distributional benefit without the efficiency cost. Then, if such a transfer can be costlessly made, it clearly should be made. But costless transfers are hard to find in the real world. Suppose that extracting money from one group and transferring it to another entailed resource costs of extraction, of delivery, and of administration plus the efficiency cost of the taxes themselves, equal to one third of the amount collected. Then one would reject the project above, but would accept the following one.

	Unweighted	Avg. Weight	Weighted
Present Value of			
Benefits	600	1.2	720
Costs	-900	0.5	-720
Net Present Value	-300		0

One can quickly see that if the distributional weights of gainers and losers from a project have a ratio of 1.5 (= 1.2/0.8), then that project will be at the margin of acceptability when unweighted costs are 1.5 times unweighted benefits. More broadly, when the weights of gainers and losers have a ratio of $(1+\lambda)$, then projects will be acceptable so long as unweighted costs are less than or equal to $(1+\lambda)$ times unweighted benefits.

Always, there emerges a tradeoff in which at the margin society pays in efficiency costs for what it gains in distributional benefits. One doesn't run into much trouble, then, if the lowest distributional weight is 0.9 and the highest one is 1.1. But such weights would not lead to much in the way of redistributive policies or projects. It is when the weights get to be amply different that the implications of a distributional weights framework lead to policies that would be unacceptable to most people. Those who like the idea of distributional weights, however, are typically not at all happy with weights of 1.1 at the poverty line and of 0.9, say, for millionaires. But give them a span of weights that they like, and you run into the difficulties (of in effect paying huge efficiency costs for most distributional benefits) that were outlined above.

A second significant attribute of distributional weights is that, even though it is quite common for authors to assign weights which decline as income rises, the benefit that is measured under this concept is consumer or producer surplus, not income. This is exactly as it should be;

if one is trying to assess whether persons or groups feel that they are better off in situation B (say, with a project) then in situation A (say, without it). But it also leads to a situation where increments to employment are very often assigned only a very modest (or even zero) benefit. This comes straight out of standard economic theory. A rise in the wage rate, say, from $10 to $11 normally pulls into the labor force a group of people who were not willing to offer their services at $10. Their gain, from the act of going to work at $11, is at most the $1 difference between $10 and $11. Those in the group with a supply price of $10.50 gain only 50 cents per hour. Those who are precisely at the margin, with a supply price of $11, are precisely on the margin of indifference as between being out of the labor force or in it at an $11 wage.

All the above is standard economics, which has been built into cost-benefit analysis, and into applied welfare economics in general for as long as we can remember. One should not think it is a mistake, or should be replaced in a system built on “supply price, demand price, and adding up”. But what we have to recognize is that the story we have just told focuses on the individuals who supply the extra labor -- on how their utility if affected when some change occurs. There is nothing in applied welfare economics that says that society’s valuation of an increment to employment should be based solely on its valuation by the individuals concerned.

But that is precisely the way distributional weights have worked. They have often directly measured the change in utility of the economic agents concerned, based on an assumed utility function. Otherwise, they have focused on changes in consumer and producer surplus which are simply money measures of the corresponding changes in utility. Those who have qualms about this treatment giving so little weight to increases in employment are probably thinking in different terms -- very likely in terms that fit nicely into the concept of basic needs externalities, which are the subject of the next section.

Basic Needs Externalities

Some years ago (1984) I published an article under the title "Basic Needs versus Distributional Weights in Social Cost-Benefit Analysis". The purpose of that article was to clarify that the two concepts were really quite different, and that it was wrong, as many authors were then doing, to treat them as virtual synonymous. To clarify -- distributional weights come out of an individualistic framework, with a focus on the utility of the individual consumer or worker. The concept of basic needs externalities comes out of a paternalistic framework, focusing on the willingness of other sectors of society to pay for an improvement in the economic situation of some individuals or groups. As I like to put it, basic needs externalities apply not to changes in the utility of the affected groups, but in their welfare -- welfare judged not by the affected individuals but, in some form or other, by the rest of society.

One of the standard demonstrations in elementary economics courses is why it is better for the government to give people subsidies in cash rather than in kind. This is shown to be so by demonstrating (usually using indifference curves) which bundle of goods a consumer would choose to buy, having received a given cash subsidy. Obviously that bundle -- call it A -- is the best that that particular consumer can obtain, given the expenditure limit set by the subsidy (plus other income). So if "society" decides to give that consumer a subsidy in kind -- giving a bundle of goods B instead of money, it is only by chance that society would choose bundle A. If B represented any bundle other than A, then the individual would be worse off with a subsidy in kind than with a subsidy in the form of cash. Society can make the consumer happiest by giving a subsidy of $100 in cash. If it gives an in-kind subsidy costing $100, it could in all but one case (choosing precisely bundle A) make the individual equally happy via a cash subsidy of something less than $100.

That demonstration is perfectly correct. But note that it always focuses on the utility of the recipient. Basic needs externalities focus not on the utility of the recipient, but on the utility of the donors. I tell my classes a story of a student in his third year at Yale, writing to his father of the terrible time he is having with his studies, now even worse than in the two previous years, and how he is even more miserable than he was then. He goes on to note that tuition is around $30,000 a year at Yale, and that transportation and living expenses add up to $20,000 more. Then the student makes his final pitch, "Dad, I just heard of a wonderful island in the South Pacific, with beautiful climate, beautiful beaches protected from sharks by a reef, beautiful native girls, and a free port where liquor costs only $4 a bottle and cigarettes $3 a carton. Dad, you could send me there for less than $20,000 a year, and I would be so, so happy!!"

To conclude, I ask my students to imagine 1000 letters like that, from 1000 miserable students all across America, to each of their fathers. My bet is that, if these 1000 letters were sent, I would be surprised at the end of the year to find even two or three of those students on the island -- but I would make no claim as to how many fathers one would find!!

This tells a lot about paternalism, and I believe it has ample reflection in the behavior of governments all across the world. Of the benefits that governments give, particularly under the label of social programs, the great majority come "in kind" rather than "in cash". You don't see governments just handing out checks and telling families to send their kids to school if they choose, but to feel free to spend the money as they like. Free public education comes in kind, not in cash. Even Milton Friedman's school vouchers represent "in kind" payments because they are only valid to pay for children's schooling. In the same way, public programs of medical care provide medical services, delivered "in kind", public housing programs provide housing, or housing-specific subsidies. Nutrition programs take the form of school lunches, food stamps and the like.

Why is this? The pressure for in-kind delivery arises because people want to feel that their tax money is well-spent -- well-spent in the eyes of the taxpayers, not of the recipients. They want education money to go for education, health-care money to go for health care, housing money to go for housing and nutrition money to go for nutrition.[3]

<u>Cost-Benefit Analysis With Basic Needs Externalities</u>

To start this section, I want to recall some steps we have already taken. When dealing with a situation in which the generation of a saving of foreign exchange brings special merit to a project or program, we end up with the concept of the economic opportunity cost of foreign exchange. When there is special merit involved in the provision of jobs, we end up working with the economic opportunity cost of labor. When judging projects that have effects on the life span of people, we have the concept of the economic value of an added year of life. These "economic opportunity costs" or "economic values" enable us to compare the effects of many different

---

[3]Sometimes these desires are frustrated by the ingenuity of the recipients. Thus, when the Indian government provided free good-quality housing to some of its low-paid employees, this in many cases did not end up with those families living with a density of 1 person per 20 square meters as the government intended. Instead the old density of one person per ten square meters was restored by the families taking in tenants (usually "cousins", real or phony). This is what good economics predicts. The very analysis that tells us that recipients prefer subsidies in cash to subsidies in kind -- that analysis also tells us that people have an incentive to turn in-kind subsidies into cash, which they can then spend on what they (the recipients) rather than we (the donors) think is most important.

Thus we see in the U.S. the (illegal) use of food stamps to buy non-food items, and even the outright sale of food stamps on street corners. And Indian parents were found to convert at least part of their children's hearty school lunches into cash,, simply by having smaller breakfasts and dinners hat home, with the rest of the family eating a larger lunch than had previously been their custom.

projects, say, in terms of the foreign exchange proceeds they generate, and at the same time judge those projects in terms of their effects in extending human lives, and also at the same time judge them in terms of their employment effects. And after taking all these things into account, we can measure the economic productivity of the capital invested in those projects.

In introducing basic needs externalities into our cost-benefit calculation, we seek to do something similar to find a way of comparing what project A does for nutrition, what project B does for education, what project C does for medical care. Obviously, we want to incorporate "values" for these benefits, but how to do it?

The basic needs approach focuses, as its name implies, on valuing improvements (or the reverse) in the welfare of people in the lower socioeconomic strata of society. And, as indicated above, the welfare in question is judged according to the standards not of the recipients but of the donors (government, taxpayers, voters, society?) Recall that if society wants to follow the recipients' standards and tastes, it should always just give them money.

So what we are looking for is a metric -- a way of placing a monetary value on specific increments to welfare, in specific dimensions.

Let me start with a very practical example, based on a case I once worked on for the Philippines. The concern was with measuring the external benefit of projects that ended up enhancing the educational achievement of the children of poor families. We had census data on educational profiles, showing, of 1000 children in each income decile, how many would leave school after 4th grade, 5th grade, etc., all the way up to the university level. We also had data on the standard costs of each successive level of education. We could then say that if a family's

---

Such misfiring of subsidies in kind are obviously much harder to achieve in the cases of education and medical care, which, perhaps partly for this reason, probably account for the largest share of in-kind subsidies.

children started with the educational prospects typical of the first decile of the population, and moved as a result of a project to have prospects typical of the second decile, this move would be assigned an external benefit equal to 50% of the standard costs associated with such a move. For a move from the educational prospects of the second decile to those of the third, an external benefit of 40% would be assigned. Further increments would be assigned declining benefits until only a 10% external benefit would be assigned for the move from the prospects of the 5^{th} decile to those of the 6^{th}. Moving beyond this latter level, no basic needs externality would be assigned.

What we have here is a way of conceptualizing a basic needs benefit, and of assigning to increments of filling that basic need a set of values that start higher, then get progressively lower as more of a given need is already met, and finally reach zero at some point. Who assigns these weights -- the government? the education ministry? the planning office? the budget bureau? Obviously any of these, but most likely it would end up in the hands of the entity in charge of the government's system of cost-benefit analysis, operating with lots of advice (and pressure) from any and all of the above.

One way in which a group of beneficiaries could be assigned an education externality would be for that group to experience a move in their actual income, from, say, the first to the third decile. This could readily happen, say, if a shipyard employing 1000 workers were to be located in Fortaleza (in the middle of Brazil's poverty-ridden Northeast) instead of in Santos (the port city of Sao Paulo, Brazil's center of wealth. The project could pay government wages, which in Santos might be just equal to the alternative earnings of the workers, but could represent a big wage premium if paid to Fortaleza workers. These could then be actually lifted from the first to the third income decile, as a result of the shipyard project.

Another way in which an education externality could be generated is for a project to enhance school attendance. In poor areas one way of achieving this goal is to provide full school lunches for the students. This measure has been found to be quite effective in improving attendance and delaying school-leaving, even in cases where its effects on nutrition were dubious. More recently, positive results have been achieved by giving parents specific cash rewards based on school attendance by their children. This type of subsidy is comparatively easy to evaluate within the suggested basic needs framework. Thus, a subsidy of up to 50% of standard annual schooling costs would be justified in an area with incomes in the first decile, if it shifted that area's school-leaving profile to the one corresponding to second-decile families. It's a very good guess that this shift could be bought with a much lower subsidy -- say 10% of average school costs -- with the result that the attendance subsidy project would have an external basic needs benefit of (50-10) = 40% of average costs.

The above example shows that the implementation of a formal basic needs approach is quite feasible. But one certainly needs some sort of indicator of each need (analogous to the school-leaving pattern for different deciles). This is relatively easy for housing, where at the most basic level one uses square meters per person as an indicator, but where one can easily introduce an index of housing quality which also includes the type of floor and roof and the presence or absence of glass windows, running water, indoor toilets, etc. For nutrition, likewise, one can use a crude measure like calories, or a more sophisticated one measuring how closely the typical diet (in each decile of the population) approximated a theoretical norm. Health care is perhaps the most difficult basic need to measure. On the one hand, an index could be based on something as simple as life expectancy (an outcome). But it could also be based on the typical number of visits made by families in each decile to a medical technician, a nurse, a doctor, a basic clinic, a full-fledged hospital, etc. (i.e., an index of medical inputs). I would want also to include public health measures, like closed or open sewers, the availability of potable water, and the degree of pest control (mosquitoes, flies, rats, etc.) in a health care index.

The basic needs externalities concept has the virtue, as against a distributional weights approach, of giving a value to increments of income (of poor people) even when their behavior reflects indifference on their part as between the work they get and the leisure they give up. This advantage stems from the fact that out of an increase in income a reasonably predictable portion will be reflected in children staying in school longer, in the quality of the family's housing being raised a notch of two, in the family's being better fed, and in some improvement in the various measures of health care. The source of this advantage should be recognized to be the element of paternalism that forms the core of the basic needs approach.

INTRODUCTION TO COST-BENEFIT ANALYSIS

P A R T   I V

APPLICATIONS TO HIGHWAY PROJECTS

Arnold C. Harberger

University of California, Los Angeles

January 2009

Background

The key element to recognize not just with respect to highways but to many other categories of projects as well, is that one must try to focus on what is the essential benefit that the project confers. Once one takes this as one's task, it yields a pretty easy answer in most cases. Certainly this is true for highway projects and for transportation projects generally. To say that a transportation project enables us to move people and/or goods from point A to point B has to be at best a huge exaggeration. Right now, without the project, one can get people and goods from just about every place on earth to every other place. What transportation projects to do is make it cheaper to transport people and goods. They don't enable it because it is already possible.

This simple, almost semantic step is tremendously powerful in focusing our analysis of costs and benefits, for it tells us to ask by how much a given project will reduce the costs of transport. This in turn leads to an extremely simple frame in which to place the analysis. First we have the existing traffic going, say, from A to B. We are interested in the volume of this traffic right now, not on its own merits but as a base for projecting the future path of that traffic, if "our" project is not undertaken. Call that traffic volume V_{ot}, the amount of traffic that we expect would exist at time t, in the absence of our project. This estimate is important because it

**This paper was prepared under the auspices of USAID.

provides the base for a major component of project benefits -- namely V_{ot} $(C_{ot}-C_{pt})$. This tells us the saving in cost that we expect to be present at time t -- C_{ot} is the cost of a trip in the absence of our project, and C_{pt} is the corresponding cost in its presence. The difference $(C_{ot}-C_{pt})$ is the cost saving per vehicle, or per trip.[1]

The next major component of project benefits comes from newly-induced traffic -- that is, traffic that would not exist in the absence of the project, and hence is seen as being <u>induced</u> by the project. So if V_{pt} is the volume of traffic anticipated for time t in the presence of the project, the amount of induced traffic at that point in time will be $(V_{pt}-V_{ot})$.

The benefit that we assign to this induced traffic is $(1/2)(V_{pt}-V_{ot})(C_{ot}-C_{pt})$. Obviously, if we were to give the new traffic the same benefit as the old, we would have $(V_{pt}-V_{ot})$ $(C_{ot}-C_{pt})$. So the question is, why do we take only half of that amount? The answer, first of all, is that if someone is willing to pay more than C_{ot} to make a trip, it is taken for granted that that person will already be taking that trip in the absence of the project. Thus C_{ot} represents a maximum for the "willingness to pay", or in this case the "willingness to bear a cost" on the part of any and all of the induced traffic. A similar line of reasoning tells us that the induced traffic shows, simply by traveling on the road in the new situation, that they are willing to bear, at the

---

[1]This exposition simplifies what would ideally be done in practice on a highway improvement project. For example, one would typically have data, not only on the amount of traffic flow in the current and past years, but also on its composition. Thus, we probably would have data on heavy trucks, light trucks, buses, large, medium and small cars, etc. Probably we would make the assumption that the percentage of each vehicle type in the total would be the same in the future as in the recent past. But if we had reasons to expect it to change in a particular way, we would base our traffic projections on that expectation. Once the projects are broken down by vehicle type, we would, of course, assign to the traffic flow V_{ot} for each vehicle type its corresponding cost saving $(C_{ot}-C_{pt})$.

very least, the new cost C_{pt}. The standard assumption is that the actual new traffic is spread out between the minimum "willingness to pay" of C_{pt}, and the maximum (for induced traffic), C_{ot}. It is this "spreading assumption" which assigns a gross benefit of $(C_{ot}+C_{pt})/2$ to the induced traffic. Since the cost these people actually bear ("pay") in the presence of the project is C_{pt}, their net benefit is $\{[(C_{ot}-+C_{pt})/2]-C_{pt}\}$, which works out to $(C_{ot}-C_{pt})/2$, as stated above.

The obvious next question is how one determines the costs, C_{ot} and C_{pt}. One part of these consists of costs connected with the vehicles and their operation, maintenance and repair. Project evaluators can consult tables which tell how many miles per gallon of gasoline one can expect on different types of roads -- also how much oil, tire consumption and repair and maintenance cost per mile. Finally, they can access data on how many miles a car is likely to last, if driven exclusively on each type of road (dirt, gravel, black top, concrete, etc., also level, hilly, straight, snaky, etc.) Since these factors are given in physical terms (e.g., miles per gallon, miles per tire, etc.), they can be converted to real dollar amounts by multiplying by the expected real prices (of gasoline, tires, the vehicles themselves, etc.) that are expected to prevail in period t.

Only rarely, however, do the material costs treated above account for the majority of C_{ot} and C_{pt}. The main element of the difference between these costs is almost always the time cost borne by drivers and passengers.

Consider a typical U.S. major road improvement. It might speed up traffic on the affected road, increasing it, say, from an average 25 miles per hour to, say 50 mph. Standard estimates of driving costs for passenger vehicles now range around 40¢ per mile. It is not likely that these would change much in a case with paved roads both before and after the project. But, taking time cost at, say $10 per hour, the time cost per vehicle mile would be 40¢ before and 20¢ after the project -- a really important change!

The example above assumes one person per vehicle (the driver), and (implicitly at least, for present U.S. conditions) that the drivers and commuters are private travelers. In a real-world analysis of a road improvement we would want to value the time of truck and bus drivers at what they are actually paid (including fringe benefits but net of taxes). One would also want to make a separate estimate for the value of time for bus passengers and for the passengers in autos and other vehicles. But probably the largest category of time cost would be that of the drivers of private vehicles, most of them commuters.

This brings us to the important question of how to place a value on commuter time, for purposes of economic project evaluation. The initial reaction of most people is to assign each driver’s own hourly rate (for time actually worked) as being also the relevant value of his/her commuting time. This, it turns out, is a huge overestimate -- which cannot be reconciled with the actual behavior of commuters.

One of the key contributions to the literature on the valuation of commuter time was made by Thomas Lisco, in a Ph.D. dissertation he wrote under my direction. The methodology that he used was at the time and still remains the standard approach to valuing travel time. The basic idea is very simple; where there are two relevant ways of getting from one place to another, the one that is cheaper will typically involve more travel time. Thus, in a sense, each individual commuter faces his/her own tradeoff between travel time on the one hand and money on the other. In Lisco’s case he was working with data on the travel habits of hundreds of commuters whose point of origin was the northwest suburb of Skokie, and whose destination was the central business district of Chicago, known as “the Loop”. One alternative for the great bulk of these commuters was to drive to work. This might have taken, for a given person, 75 minutes, including walking time from the parking lot to the place of work. The most natural alternative for that person would be to go by train. This would entail walking (or going by bus) to the train station, waiting for the train, taking the Skokie Swift Express from Skokie to Evanston, changing there to the Chicago El (for elevated) train; obviously involving some transfer time, then taking the El into the loop, and walking (presumably) to the workplace. Let's say that for our given person all this involved 115 minutes of time, on average.

Hence we have the fact that going by car would save this person around 40 minutes, each way, per day. Suppose now that going by car entailed (at the time) an average monetary cost (including parking) of $16 per day, while going by train cost $8 per day. This would reveal a tradeoff of 80 minutes against $8, which means a time cost of $6 per hour. The inference is that the commuters facing this tradeoff, which choose to go by train value their travel time at less than $6 per hour, and that those who choose to go by car value their commuting time at more than $6 per hour.

Lisco had data on hundreds of Skokie commuters, facing different tradeoffs, mainly because of the distances of home and workplace from the respective train stations. The data confirmed one's common-sense expectations -- people with low incomes tended to choose the train, those with high incomes tended to go by car. But some low-income people nonetheless went by car; they tended to be those who lived far from the train station. Similarly, the richer people who nonetheless took the train tended to be those who lived close to the station.

One could probably reach a pretty good judgment on a relevant value of travel time, simply by grouping commuters by income level and forming histograms of the time/money tradeoffs they accepted or rejected. Lisco's procedure was more precise than this, and made fuller use of the data. He used an econometric technique known as probit analysis. Probit analysis deals with the probability, in this case, of a given choice of mode. Using this device, Lisco was able to derive the implicit tradeoff between time and money that would lead people of given income levels to split 50-50 on their choice of mode. That tradeoff would then be the value of commuter time assigned to that income level.[2]

The resulting values of commuter time were far below the average hourly earnings of the commuter income groups. This result has also emerged from every similar study, and has become an accepted part of transport economics.[3]

In developing countries it is hard to find the kinds of data that one needs for careful econometric studies. However, two important facts can help surmount this obstacle. First is the fact that in developing countries we see that on-the-job drivers (of trucks, buses, taxis, etc.) represent a higher share of the traffic than in more advanced countries. For these one not only can but directly should use their hourly earnings as the value of their time. Second, we can consider that car owners typically have considerably higher income than truck or bus drivers. We combine this with the result of modal choice studies (that travel time for these people should be valued at only a fraction of their hourly earnings) to simply take the hourly earnings of truck

---

[2]If P_c is the probability of choosing to go by car, X_c is the extra cost per hour saved as a result of going by car, and Y is the level of family income, one can fit a probit regression of the form $P_c = a - bX_c + cY$. Then, setting $P_t = 0.5$, one can solve for $X_c = (a-0.5+cY)/b$. This shows how X_c, the implicit value of commuter time, varies by income level.

[3]One can make a personal check in this proposition in any place where commuters all tend to work in a given compact area, like the Loop in Chicago. What we find in such cases is that you can get all-day parking at lower rates, the farther you go from that central area. Thus, if you can save $1 by walking two blocks more, each way, and the average total walking time for four blocks is 10 minutes, then walking time is worth more than $6 an hour for those who park nearer, and less than $6 an hour for those who park farther away.

and bus drivers as also representing the cost of travel time to car owners. Not an elegant solution, perhaps, but a workable one.

Congestion Externalities

Even though the fraction of families owning cars is very much smaller in developing countries than it is in the developed world, the problem of traffic congestion seems pervasive at all levels. This is important because of the specific role that congestion externalities play in the economics of transportation.

The first lesson on this subject is that congestion externalities are present, even when we may not be conscious of their existence. How to explain this? Consider the fact that on a standard well-built concrete road, traffic will flow at something like 60 to 70 miles per hour, when traffic is light. But we all have experienced traffic flows at 20 to 30 or 40 miles per hour on these very same roads. Not only that, but we know why we see these lower speeds -- more vehicles are using the road. Thus only two points of observation -- one with high average speed and low traffic volume, the other with low average speed and high traffic volume -- are enough to reveal the basic principle. More vehicles on the road mean slower average speed. Highway engineers have plotted this relationship, and found it to be a continuous curve -- traffic does not go unimpeded at 60 mph up to 500 vehicles per hour, and then drop suddenly to 30 mph as volume exceeds 500. No, the relationship between average speed and traffic volume is a smooth curve, in which speed starts to decline long before one perceives noticeable congestion.

This is important, because it isn't just on those frustrating occasions that we feel ourselves "stuck in traffic" that congestion externalities exist. In fact, they are present most of the time that vehicles are using the roads.[4] What happens is that one extra swath of traffic brings average speed down from 60 to 58 mph, another swath brings it down from 58 to 55, etc. -- until finally, after more and more vehicles have been added to the flow, average speed gets down to 40 and 30 and 20 mph.

The externality connected with traffic flow is that each added vehicle contributes a tiny bit to this general slowing of traffic. This can be appreciated by simply thinking of what a continuous curve relating average speed to traffic volume really means. Just as a bucket is filled by a huge number of successive drops, each of which brings its own equal contribution, so it is with traffic. Here, starting with quite modest volumes of traffic, each added vehicle contributes its tiny bit to the slowing down of the rest. If an extra 100 vehicles end up slowing traffic from 60 to 58, we can infer that each one of these has slowed the whole mass of users by 2/100, or .02 miles per hour. If going at 60 mph, their time cost was 10 cents per mile, one extra vehicle now brings this cost to 10.02 cents per mile. In the end everybody -- the "old" drivers and the "added" drivers, ends up going 58 mph. The added drivers perceive this cost. But they do not perceive the cost they inflict on the 1000 or 2000 "old" drivers who, but for the 100 newcomers, would have sped ahead at 60 mph.

To put this externality in terms that are easy to understand, we can make use of a very convenient approximation, which necessarily gives a conservative estimate of the time externality. A key element in this rule of thumb is what we call the "unimpeded speed" of traffic

---

[4]When I say most of the time I do not mean most of the hours of the year. Rather, I mean most of the time spent by vehicle occupants. I once did a study of congestion externalities for five American cities. Their data revealed that something like 3/4 of the total traffic going to and from the city occurred between 6 and10 a.m., and between 3 and 7 p.m., on weekdays. During these hours, for sure, added vehicles would lead to lower average speed.

on the road. This is the average speed of traffic on that road when traffic is very light. This might be 20 mph on a dirt road, 30 mph on a gravel road, 40 mph on a blacktop road, 50 mph on a 2-lane high-type concrete road, 60 mph on a 4-lane divided highway and 70 mph on a multilane freeway. Call this unimpeded speed “a”, and let “s” be the actual speed that you observe at a given point in time. Our approximation would then say that the externality, at that point in time, would be (a-s)/s times the average time cost per vehicle mile.

To give an example, suppose we are dealing with an ordinary blacktop road, on which we know (or guess) that the unimpeded speed is 40 mph. We are going on a Sunday drive, and we notice that our average speed is about 30 (as we move with the traffic). Our formula says the externality is (a-s)/s; or (40-30)/30 times our average time cost. If the vehicle hour is worth $7.50, then the vehicle mile is worth 25 cents (= $7.50/30 mph). So the externality would be 8 1/3 cents per mile that we are driving in our car. If the traffic were going only at 25 mph the externality would be (40-25)/25 times average cost. That is, 60 percent of $7.50/25, or (.6)(.30) or 18 cents per mile that we drive our car. This is a calculation that each of us can make, as we drive on different types of road and with varying densities of traffic. It can be quite instructive, and certainly helps to hone one’s intuition about highway economics.

Let us go back to the project of improving a given road -- call it road H. Volume of traffic on H will increase as a result of the project, but we have already seen how the direct benefits on road H should be counted. We want now to look into what happens on other roads; call them A and F, Road A is an alternative (substitute) for road H (our project road). So when H is improved, part of the increment of traffic that we observe on H will have come from A. Thus road A gets to be less congested as a consequence of our project on road H. The gain on road A is measured by its own (a-s)/s, times its own time cost for the traffic diverted to H.

Road F is a feeder road into H, hence when H is improved, traffic on F increases. F becomes more congested and there is an external loss on that road, equal to its (a-s)/s times its average time cost times its increase in traffic volume.

On Critical Traffic Volumes and Stage Construction

We have seen how the benefits of a highway improvement project are linked to the volume of traffic that already prevails on that road and that can be expected to prevail, year by year in the future, in the absence of the project. If we start with a dirt road and ask when it should be turned into a gravel road, we have a pretty good idea of how much such an improvement would "typically" cost. Based on this typical cost we can make a technically founded guess as to how large a volume of existing traffic it would take, in order plausibly to justify the upgrading to gravel.

So it is with each stage. It may take only 200 vehicles a day to plausibly warrant shifting from dirt to gravel, but possibly 500 vehicles a day to make a reasonable case for converting a gravel road to blacktop. No one would argue that such numbers should be used to justify a particular gravel road or a particular blacktop project. But where they can be especially useful is in the pre-selection of cases to be studied in more detail. It is extremely easy to get a traffic count of how many vehicles pass given points on a road. The old technology of doing this by pneumatic tubes laid across the road is still perfectly functional, and very cheap. Newer technology of electronic sensors might turn out to be more convenient, and maybe even cheaper.

The idea is to use traffic counts as signals of which existing roads are good candidates for upgrading -- from dirt to gravel, from gravel to blacktop, from blacktop to concrete, from two lanes to three or four, etc. For each stage there would be a critical volume of traffic that would cause certain stretches of road to be selected for more serious evaluation. I'm told that the French highway authorities have been doing this for many decades, to very good effect.

A close cousin of stage construction is varying standards for different segments or stretches of the same road. If one travels highway 101 between Los Angeles and San Francisco, one finds that in some stretches it has four lanes in each direction, in other stretches three, in still other stretches, only two. And if my memory serves me well, there are still other stretches in which it is a standard undivided highway of two or three lanes. This is the kind of variation that will quite naturally emerge from a system of serious economic project evaluation. All it takes to generate such a result is varying volumes of traffic as between one stretch and another. In the case of highway 101, traffic is extremely dense in the vicinities of Los Angeles and San Francisco. Volumes are much lower as we get 100 miles or more away from either of these centers. Hence it is perfectly sensible, from an economic point of view to have different construction standards for different stretches.[5]

[5]Highway engineers sometimes press for absolutely uniform standards along a road. Their argument is soundest when we think of stretches of just a mile or two in length. Here, changing from two lanes to four lanes, then back to three, etc., could invite serious costs connected with auto accidents produced as drivers are "surprised" by a sudden change in pavement or in number of lanes. But that same argument has little or no merit when one considers stretches of, say, ten or twenty miles or more. Here transitions can be clearly signaled well in advance. There need be no surprises!

INTRODUCTION TO COST-BENEFIT ANALYSIS

P A R T V

APPLICATIONS TO IRRIGATION PROJECTS

Arnold C. Harberger

University of California, Los Angeles

January 2009

Background

Just as the benefits of highway projects are closely linked to the volume of traffic, so those of irrigation projects are linked -- guess what? -- To the amount of water delivered, and its value. But in both these cases claims abound for a variety of external effects that do not hold up under scrutiny. We very often see the double counting of benefits on many different types of projects, but I think that one so-called project report that I once reviewed in India probably holds the record. In that report, benefits were claimed: a) equal to the value of the water, 2) plus the increase in the value of the land that took place as a consequence of the project, 3) plus the increase in the value of crops produced on that land, and 4) plus the wages bill paid for the extra employment that emerged as a result of the project. As I said, I had seen cases of double counting if benefits quite often, but this was a case of triple, and even quadruple counting!!

We start with the value of the water delivered by the project. Water is a productive factor for agriculture, side by side with land, labor, capital (fences, buildings, cattle, farm machinery, etc.), fertilizers, gasoline and so on. Each of these factors of production has what economists call a marginal product, and, broadly speaking, a market system leads to a situation in which the value of this marginal product is brought into equality with the price that has to be paid for that

**This paper was prepared under the auspices of USAID.

factor. To illustrate, suppose the agricultural wage is $5 a day, and suppose a farmer estimated that adding one worker to his labor force would generate an increase of $6 or $7 per day in the farm's output. Obviously, it would be worthwhile to hire that worker because the benefits of that action ($6 or $7) would exceed the cost. Now the farmer's hiring one more worker is not going to have a perceptible (to him) impact on the market wage, so natural economic incentives will work to keep adding to the farm's labor force until the marginal product of an extra worker gets down to around $5 a day. The same sort of process works with respect to each and every factor of production that can be freely bought (or hired) in the marketplace at the market price. This is the case for labor, fertilizer, gasoline, farm machinery, etc.) But not for the water from an irrigation project. This water is distributed in quotas to the farmers who have "water rights" linked to that project. These rights are usually set on the basis of equal amounts of water per hectare (or acre) of the area served by the project. But no one knows in advance how much water that will be. That depends on the forces of nature (rainfall, snowfall, etc.) and on how much of the river's water is taken by others (upstream) or reserved for the use of others (downstream). The irrigation quota for a given farm might thus might be 100 cubic meters one year, 20 the next, and only 5 in the one after that. The market principle would work if farmers in the project were able to trade irrigation water among themselves, a process that would lead to a high price when water was scarce and a low price when it was abundant. Economists have long argued for freedom to trade water -- within irrigation projects and even outside them -- but progress in this direction has been very slow. The norm is still that farmers each get "their share" of the available water, a share that will be big or small, month by month and year by year, as nature and the priorities of other users dictate. Usually, the farmers have to pay something for their water rights, but more often than not it is a fixed change per hectare, rather than a price per cubic meter of water actually delivered. And when it is a price per cubic meter, it is usually a very low price, which is far below the productive value of the water. So, whereas for other factors of production there is a strong tendency to use more and more of a factor, up to the point where its marginal product matches its market wage or market price, this is rarely true for the water delivered by irrigation projects.

Thus, we cannot measure the marginal productivity of irrigation water by what the farmers pay for it. They usually pay much less than its marginal product, which leaves us with the problem of ascertaining the economic value of irrigation water (if we are doing an ex post evaluation of our existing project) or of predicting that economic value for a new project being analyzed. What we must remember here is that the value we are seeking is closely linked to the market price that would prevail, varying from month to month and year to year, if the farmers served by the project were freely able, among themselves, to buy and sell water deliveries period by period.

We are going to argue that by far the best way to estimate the likely value of irrigation water is to keep our eyes always <u>on the water</u> -- not looking at indirect ways of getting estimates. But first I want to explain the principal indirect method that is actually widely used, and that has its own underpinnings in good economics. This is often called the farm budget method, but I prefer to call it the "residual value method", because this label much better conveys how the method really works.

The basic idea of the residual value method is to build up two typical farm budgets -- one in the absence of the project in question, the other in its presence. These two budgets are often quite different -- the first involving a dryland rotation, with crops that do not need much water, and the second dealing with a very different rotation built on the presence of irrigation water.

But, also quite often, an irrigation project will deal with land that is already irrigated, drawing water from a river as it passes. Such projects typically involve building a dam on the river, leading to greater availability of water during the irrigation season, plus a degree of control over precisely when, during the agricultural year, the quotas of irrigation water are delivered to the farms. In these cases -- of dams simply enhancing the capacities of pre-existing river-irrigation projects -- cropping patterns will undergo little or no change as a result of the project.

The residual value method focuses on the typical farm's estimated profit-and-loss statement. Quantities are for the different crops produced per year, and also for labor, fertilizer, machinery and other factors of production. The estimated prices are linked to each of these quantities (except for the land itself). Then the project analysts take the sum total of the value of all of the farm's outputs, and subtracts from it the estimated costs of all of the farm inputs, other than land and water. The result is the farm's profit (including return to land, independent of whether that land is owned or rented). This whole exercise is estimated for all the future life of the project: a) assuming the project is not built, and b) assuming the project is in fact undertaken. Thus we have, year by year over the project's expected life, a residual income with the project, and without it. The difference between these two flows for each year becomes the estimate of that year's expected value of irrigation water.

	Value of Crops	Cost of Inputs Other Than Land and Water	Residual Value
With Project	1500 ± 150	1000 ± 100	500 ± 250
Without Project	1000 ± 100	700 ± 70	300 ±170
Difference In Residual Value Attributed To Water			200 ± 420

One can easily appreciate the complexity of such estimation procedures, and particularly the degree of error or uncertainty that is involved each step.

The above table reveals the Achilles' heel of the residual value method. I have made the table simple by dealing with "plus or minus" ranges, rather than standard deviations and variables, because it is easier to understand for those who may not have a background in statistics.[1]

The residual value that we calculate in this way is a combined estimate of the contributions of both land and water to farm product. The table shows an extra 200 (for a given year) to be contributed by the irrigation project. Suppose now that the irrigation authority itself collects this amount from the farmers as an irrigation charge. Then, clearly, there would be no reason why land values should rise. But if no irrigation charge were collected, land values would presumably rise by the full present value of this "difference in residual value attributed to water" for all future years of the life of the project. In the simple case of a perpetuity of 200, using a 10% discount rate, land values would rise by 2000 (= 200/.10). If the irrigation authority were to collect 80 per year, then the rise in land values (in the same simplified calculation) would

---

[1]In statistics one learns that the variance of the sum of two independently distributed variables is the sum of their respective variances. Most students are then quite surprised to learn that the variance of the difference between two such variables is also the sum of their variances. Reworking the table in these terms, suppose that the numbers indicated by ± for "value of crops" and "costs of inputs" are standard deviations. Then the variance of the "with project" residual value would be 22,500 + 10,000 = 33,500, and the same variance without the project would be 10,000 + 4,900 = 14,900.

The estimated project benefit in the table is 200 (= 500 minus 300). The variance attaching to that 200 figure would, using the numbers just calculated, be 48,400, and its standard deviation would be 220. This is not as huge as the ± 420 emerging from the table, but the point to be made is exactly the same. When the value we calculate is based on a "difference of differences", that value is subject to a very large standard error, larger than any of the components from which the differences were derived.

be 1200 (= 120/.10).

Thus we see the folly of counting both the value of the water and the rise in land values if that were to mean counting 200 per year as the value of the water plus 2000 (= the same 200, capitalized at 10%) as the rise in the value of the land. It would not be a mistake to calculate the benefit as the actual irrigation charges collected (80 per year in the example) plus an induced rise in the land value of 1200 (based on the 120 of irrigation benefit not covered by cash payments for water).

The Indian project analysis that I mentioned at the outset also counted the increase in value of crops. This has no business appearing as an extra benefit (of 500 in the table), because there are clearly extra costs involved (of 300 in the table). When we got the residual value, we actually counted the increment to crop value as a plus, but then we deducted the associated costs, as we should do. To count the residual value and then add to it the increment in crop value is clearly double counting and is completely unjustified.

When it comes to the increment to employment resulting from the project, the basic lesson is the standard one in labor economics. The wage represents "in principle" an economic cost, not a benefit. Sometimes the wage might overestimate the true economic cost of the labor in question, in which case one would consider that, after first counting the full wage payment as a cost, we would introduce an external benefit (e.g., for taxes collected on the basis of those wages, and/or for a "producer surplus" representing the excess of the actual net-of-tax wage payments over the true supply price of the workers involved).

Working with the data of the table and summarizing the worst-case interpretation of my Indian example, we have

Value of Irrigation Water = 200 per year

Increment to Value of Land = 2000 (200 per year capitalized at 10%)

Increment to Value of Crops = 500 per year

Increment to Labor Use = 180 per year (say, 60% of the 300 increment to costs).

What we should have is only the first of these, which could be correctly represented as an 80 per year actual payment of irrigation water charges, plus 120 per year of "economic rent", generated because the full value of the water was not being collected. This 120 could be directed counted as an economic rent (the preferred way), or capitalized as an increase of 1200 (= 120/.10) in the value of land. On labor we should never count the full wages bill as a non-cost. The correct procedure would either take simply the voluntary net-of-tax supply price of the added labor as a cost (to which no externalities could be appended),or one could initially take the full outlay on labor as a cost (as in a financial analysis), and then consider external benefits equal to the extra taxes paid in connection with those wages, plus the estimated producer surplus received by those workers. In a well-functioning labor market, the project would not generate significant producer surplus because the workers would be expected to have other employment in the absence of the product.

Direct Estimates of the Value of Irrigation Water

Many years ago I was asked by the Argentine authorities to lead a very small team to do an evaluation of an irrigation project (the Ullum Dam) on the San Juan River in western Argentina. The team consisted of two former students, Lucio Reca and Juan Antonio Zapata, and myself. A major engineering firm, the Harza Engineering Company of Chicago had been active in the evaluation process for several months; they had been doing farm budget studies and were asked to make their data available to us. But before we reached that stage, we wanted to familiarize ourselves with the economics of agriculture in the region, and the role of irrigation water within agriculture.

Luckily, Reca had previously spent a couple of years in the city of San Juan, as the local representative of Argentina's Department (Secretariat) of Agriculture. He thus knew many key people in the area, so we installed ourselves on the terrace of our hotel for a couple of days, receiving a steady stream of local experts.

Early in that process, we received a briefing on the existing system of river irrigation, which had been in existence for more than half a century. During that briefing, we were told that some 120,000 hectares were covered by the project, and that the available river water for each month was distributed by two technicians (called <u>tomeros</u>), one on each side of the river. They worked their way up or down the river, opening the sluice gates of one farm after another, then closing them after each farm's water quota had been delivered.

The following day, another interviewee was talking about irrigation matters, and in the course of his conversation casually mentioned "the 60,000 hectares or so that are irrigated". That tiny phrase opened the door to our entire evaluation of the project. We asked him, is it not true that the project covers 120,000 hectares, not 60,000? Yes, he said, 120,000 hectares have irrigation rights, but the water that is delivered to each owner is actually used on about half of his eligible hectares. Why is that? Because putting 4 inches of water on one hectare produces more than one would get from putting 2 inches of water on each of 2 hectares. That is to say, the economically optimal strategy was to leave half the eligible area without irrigation water. What this meant, in this particular case, was that land was super-abundant; and its marginal productivity was essentially zero. The scarce factor was water, and if one had a residual value due to "water plus land", that value arose because of the scarcity of water alone.

This insight, which every student of elementary economics has learned (or should have learned) became the key to our entire study. Right away we asked whether water rights could be bought and sold -- the answer was no. Then we asked whether an owner of two pieces of land could use the water rights of one piece to have the water delivered to the other. The answer here was yes, provided the two pieces of land were on the same side of the river. Why was that? Because it would be too complicated to have water taken away by the left-bank tomero and then delivered by the right bank tomero. If any "transfers of water" were to take place, then, both parts of the transfer had to be handled by the same tomero.

Our next key question was, could one find among the land sales of recent years, any properties where the principal purpose of the buyer of plot B was simply to have "its" water transferred to that buyer's existing farm A? The answer was yes, there had been quite a number of such transactions. We then had these candidate transactions scanned, so as to omit any on which there were important non-land assets, such as houses or barns, etc. This culling was necessary, for our plan was to use the land prices at which our plots were sold, as estimates of the current market value of their water rights.

Next came the question, taking for granted that the buyers of these properties were really buying water rights, what kind of product were they actually paying for? Certainly they were not getting a certain number of cubic meters of irrigation water, every year, for sure. What they were really buying is what could be characterized as a series of lottery tickets, one for each month of each future year. When the previous winter's snowfall in the Andes was big, they got a lot of water, when that snowfall was small, they received little water.

We went back to the irrigation records and developed a histogram showing a frequency distribution of water deliveries in each of the past 50 years. That distribution had quite a range, with the maximum water availability being something like 10 or 15 times the minimum. It was obvious that the farmers would not value very highly amounts of water that would come only once every 10 or 15 years. On the other hand, water that they could pretty much count on was extremely valuable. If this amount were to increase by half, they could plant half again as many hectares to their main cultivations -- namely, vineyards and olive trees.[2]

Given the water delivery experience of the past 50 years, we had to somehow take account of the fact that "sure water" was much more valuable to the farmers than "occasional water". We had little time for nuances, so we adopted a quite robust scheme of "weighted cubic meters" of irrigation water. We started with an index $I_1 = .4D_1 + .3Q_1 + .2M_e + .1Q_3$, where I_1 is index #1, and D_1 is the first decile of the histogram, Q_1 the first quartile, M_e the median, and Q_3 the third quartile. This index obviously gives much heavier weight to sure as against occasional water, but it is also arbitrary, and we had no time to write a Ph.D. dissertation examining what would be the appropriate weights. So we resorted to a trick that is standard fare for cost-benefit analysis -- a "sensitivity test". In this case the sensitivity test entailed employing an alternative index, $I_2 = .33D_1 + .27D_2 + .23M_e + .17Q_3$, which gave less weight to "sure" water and more weight to "occasional" water. We carried out parallel calculations using both I_1 and I_2, all the way to the end of our study. Happily, our conclusion -- that the dam was indeed a worthwhile investment -- was the same, regardless of which of the two indexes was used.

[2]Both of these plantings last for many years, but both will die out if deprived of water. The way it works is that in years of minimal water availability, the farmers spread that water very sparingly, trying simply to keep their plants alive. Little or no harvest (of grapes or olives) comes out of such years. It makes no sense to plant olives and vines and if they are going to die off in a few years, so total plantings are largely governed by the expected bottom end of the probability distribution of water availability.

We next set 120,000 $P_H = P_{I1}I_1^0$, where P_H is the recent price (in real pesos) of land in our key transactions, and I_1^0 is the first index calculation that we derived from the histogram representing the irrigation experience of the past 50 years. This gave us a price P_{I1} for the unit of quality-adjusted water measured by I_1. A parallel calculation was then carried out to give us P_{I2}, the price of quality-adjusted water as measured by P_{I2}.

This prepared us for the next step, of trying to assess the value of the Ullum Dam project. This required us to have a reasonable projection of how water availabilities would change if the dam were built. Our procedure here was to simulate how the dam would have functioned, over the past 50 years, if it had been in existence all that time. To do this we obtained month-by-month streamflow data, plus month-by-month actual irrigation deliveries over this 50-year period. For the simulation, we followed a very simple strategy, something that was a virtual necessity due to our time constraints. We divided the year into two seasons -- irrigation and non-irrigation. The strategy was to accumulate water behind the dam during the non-irrigation season, and deliver it during the irrigation season. The amount accumulated each month (in the simulation) was the full amount by which that month's streamflow exceeded the amount that had to be left in order to cover the water rights of downstream users. This accumulation was allowed to go on, during the entire non-irrigation season. However, for past years of abundant water, the accumulation had to stop, once the dam's capacity of 440 cubic hectometers was reached.

For deliveries from the dam, we again needed a simple strategy. The one we chose was to assume that water deliveries from the dam would simply be used to proportionally increase the natural streamflow of each irrigation season, month by month. Here we reached a different limit at the point where this strategy called for deliveries over and above the delivery capacity of the canals. In such cases our simulation saved the excess water (above and beyond canal capacity) for the next irrigation season.

This simulation resulted, quite obviously, in much greater deliveries of irrigation water than what had actually occurred during the past 50 years. Then, using the simulated data, month by month and year by year, we were able to develop two new histograms I_1^* and I_2^* representing what values I_1 and I_2 would have taken, had the dam been in existence for the past 50 years. These procedures yielded one component of the benefits of the dam. This is represented by $P_{I1}(I_1^* - I_1^0)$ for the first index and $P_{I2}(I_2^* - I_2^0)$ for the second index. This measure assigns a value to the "dam water" which is equal to that which the market assigned to "river water".

Our next step was to recognize that dam water is more valuable than river water. This is because the dam managers have some degree of control over when the stored water will be delivered to the farms. Obviously they will try to time their deliveries so as to come as close as they can to giving farmers water at the times they want it most. Thus, they will certainly <u>not</u> emulate our simulation, by simply giving farmers a proportional expansion of each year's natural streamflow. They will make significantly better use of the water. How much better -- we really do not know. So we again made alternative assumptions, both of them extremely conservative. Under the first of these, dam water was assumed to be 5% more valuable than river water, and under the second it was taken to be 10% more valuable. However, this increment of value would apply not just to $P_{I1}(I_1^* - I_1^0)$ or to $P_{I2}(I_2^* - I_2^0)$, but rather to $P_{I1} I_1^*$ or to $P_{I2} I_2^*$. Why? Because once the dam is there, it can manipulate the timing of all deliveries during the irrigation season, the only limitation coming on those rare occasions where the dam is already full, in which case the natural streamflow must be delivered, and else that water would be wasted, as far as the project area is concerned.

To sum up, our measures of the benefits of the dam, up to now, are

$$B_{1a} = P_{I1}(I_1^* - I_1^0) + .05\, P_{I1}\, I_1^*$$

$$B_{1b} = P_{I1}(I_1^* - I_1^0) + .10\, P_{I1}\, I_1^*$$

$$B_{2a} = P_{I2}(I_2^* - I_2^0) + .05\, P_{I2}\, I_2^*$$

$$B_{2b} = P_{I2}(I_2^* - I_2^0) + .10\, P_{I2}\, I_2^*$$

We are still not quite finished. P_{I1} and P_{I2} were derived on the basis of the observed real P_H, which should be seen as the private discounted value of the private benefit that farmers currently get, per hectare of irrigation rights, from streamflow irrigation. Two corrections are called for if we are aiming at the overall economic benefit of the dam project. The first is to adjust the benefit stream upward to take into account the estimated increase in property and income tax revenues that will result from the project. The second is to recognize that the appropriate discount rate for calculating the economic net present value of a project is not the private discount rate but rather the overall economic opportunity cost of capital. A rough adjustment of this type is to multiply the tax-adjusted benefit figure by the private discount rate, and divide it by our best estimate of the true economic opportunity cost of capital in the country.

In our actual study we made these two adjustments. Luckily, under all the alternatives we examined -- B_{1a}, through B_{2b} -- the present value of project benefits exceeded the present value of costs. The project seemed quite definitely to be worth undertaking.

We still were not done, however. There remained the possibility that we might somehow have overestimated the true benefits of the project. This would simply be an unavoidable risk if we had no way to check on our calculations. But in this case we did have a way, about which our early interviews on our hotel terrace had informed us. The key to this check was the fact that side by side with streamflow irrigation, there was a fairly wide use of pump irrigation in the area served by the dam.

The key fact here is that (assuming the chemical characteristics of the water do not differ greatly), pump irrigation water has to be better than water from the dam, for basically the same reason that makes dam water more valuable than river water. The reason is that farmers can pump water (up to the capacity of their pipes and pumps) when they really need it most, whereas with dam water they have to wait their turn as the tomero makes his monthly rounds. Thus they might get their water two weeks before they need it most, or two weeks after, and simply have to make the best of those deliveries from the dam. Thus, if we found in that area that people could get good pump irrigation water for less than the value we estimated for water from the dam, this would mean that our benefit estimates were too high.

Pursuing this line of thought, we made a rather careful survey of the situation with respect to pump irrigation in the area. It turned out that while the land surface sloped very gently downward from west to east, the aquifer containing ground water sloped downward more rapidly. This led to a situation where the earliest wells (those farthest to the west) had reached water at a depth of perhaps 30 meters. Those wells were already put in operation before 1920. As time went on and pump technology improved, the area covered by pumps moved eastward step by step, going to 50, then 75, then 100 meters, etc. At the time of our study, the new wells that were being drilled were at a depth of about 200 meters. Our inference was that investments in these wells were just yielding the prevailing private rate of real return on agricultural investments. Happily the resulting costs estimated for pump irrigation water were significantly higher than our own estimates of the value of water from the dam. Our procedures thus easily passed this final test.

* * * * *

I presented this step by step description of an actual real-world project evaluation exercise in an effort to transmit something of the flavor of such work. It is certainly not a routine exercise that anybody with a little instruction could perform. One has to be alert to the particulars of each case. One has to mesh the real-world observation of these particulars with the lessons of economic theory. Then one has to take the result and fit it into the framework of economic cost-benefit analysis, incorporating the key places where economic opportunity costs, and economic values in general will be different from the corresponding prices that we observe in the market.

How far can the lessons from this exercise be extended to a broader range of irrigation projects? I would say quite a distance. The key observation in our analysis of the Ullum Dam project was the finding that the prices paid per hectare in a certain set of agricultural land sales could be seen as in fact paying for the rights to irrigation water that were attached to that land. This enabled us to bypass the laborious farm budget method, with its Achilles heel of a very high variance surrounding its estimates. One could contemplate doing something similar in cases where one finds irrigated land side by side with non-irrigated land of similar quality, on both of which crops are being raised (probably different crops and different rotations in each type). The difference in per hectare prices between irrigated and non-irrigated hectares could in principle play the same role as the price of our key transacted hectares -- i.e., as a measure of the present value of the “series of lottery tickets” that one is in effect getting when one acquires rights to irrigation water. If one did this, one would have to be quite sure that the two classes of land where prices were being compared really did have similar qualities of soil. Also, one would have to check concerning improvements that might tend to be present on already irrigated hectares and absent on those used for dry-land farming. Particularly important here clearly is land-leveling, this can represent a big capital investment that leaves no immediately perceptible trace. But in many cases irrigated land embodies much more of such investment per hectare than do neighboring unirrigated hectares. The presence of irrigation ditches themselves represents another capital cost not present in unirrigated lands. Also, there might well be fencing and farm buildings linked to one kind of land use and not to the other. Thus unirrigated lands used for pasture are often divided into separate fenced parcels in order to permit orderly grazing (and in particular to avoid overgrazing). In order to use the difference in land prices in the way we used the prices of our key transactions, one would have to estimate the contributions of each of these elements to the respective prices of irrigated and nonirrigated land. Only after correcting for these elements would one go on to take a price difference which we would then interpret as the private-market valuation of the present value of water rights. Obviously, if these elements get to play too big a role in the comparison one gets into the same problem as one encounters in the farm budget method -- that is, the high variance that applied to the difference between two separate elements, each of which is subject to considerable error of estimation.