Labor Market Policy in Developing Countries: A Selective Review of the Literature and Needs for the Future

Gary S. Fields
Cornell University, gsf2@cornell.edu
Abstract
[Excerpt] This is a paper about modeling labor market policies in developing countries. The paper reviews labor market models from both a positive and a normative point of view. The survey is selective in that it exposits only some of the more significant research contributions and highlights the most important gaps and analytical needs.

Keywords
labor market policy, developing countries, labor markets

Comments
Suggested Citation

Required Publisher Statement
Copyright by the author.
Labor Market Policy in Developing Countries: A Selective Review of the Literature and Needs for the Future

Gary S. Fields
Cornell University and IZA *

September, 2007

* I am grateful to Jean-Jacques Dethier and the World Bank’s Research Department for their financial support of this research, to many World Bank staffers for their helpful comments on an earlier draft, to Rolph van der Hoeven and Raymond Torres for useful suggestions for improvements, and to María Laura Sánchez Puerta for preparing a companion paper (Sánchez Puerta, 2007) that contributed a great deal to my review of the empirical literature and to which the interested reader is referred for additional empirical results.
Outline

Introduction

The Welfare Economics of Policy Analysis in the Labor Market
- Social cost-benefit analysis: an introduction
- Specifying the ultimate objective or objectives
- The Pareto criterion
- The quasi-Pareto criterion, first order dominance, and rank dominance
- Abbreviated labor market well-being functions
- Minimizing a low-earnings index
- Traditional social cost-benefit analysis
- Social cost-benefit analysis in practice

Theoretical Models
- The overall framework
- The need for multisector models
- Modeling the formal sector labor market
- Modeling the informal sector labor market
- Modeling intersectoral linkages in the labor market
- Models with skill differentials
- Modeling social protection

Empirical Evidence
- Empirical research on labor markets and labor market policies: six kinds of data
- Lesson learned: finding the right model
- An example of good empirical policy analysis in practice: assessing the effects of India’s employment protection laws
- Limitations of existing empirical research: the case of social returns to education
- A cautionary note

Conclusions and Needs for Future Research
- Main messages on how to conduct labor market policy research
- Five pitfalls to be avoided
- The main messages and pitfalls in action: minimum wage policy
- Topics for future research
- Some final word
Introduction

This is a paper about modeling labor market policies in developing countries. The paper reviews labor market models from both a positive and a normative point of view. The survey is selective in that it exposits only some of the more significant research contributions and highlights the most important gaps and analytical needs.¹

Why study labor markets and labor market policies? A person’s economic well-being is determined by how much s/he and other family members earn for their labor and by what goods, service, and cash they receive from the government, the community, and others. Even with multilateral and bilateral assistance, developing country governments are too poor to be able to make a significant dent in poverty through government spending, and families are too poor to make large enough transfers to the less-advantaged among them. This means that creating more and better earning opportunities for the poor is the only other option available.

In labor economics, the “labor market” is defined as the place where labor services are bought and sold. Two types of employment are distinguished: (1) wage and salaried employment, in which workers sell their labor services to others, and (2) self-employment, in which workers sell their labor services to themselves. The poorer the country, the higher the proportion of workers in self-employment and the smaller the proportion in wage and salaried employment. Provided that the idea of “labor market” is understood as comprising both wage and salaried workers and the self-employed alike, it is as relevant to analyze labor markets in the low-income countries of Africa and Asia as it is to analyze them in the middle-income countries of Latin America, the Middle East

¹ The definition of research used in this paper is the same as that used in a recent report by Deaton et al. (2006). Research includes papers, books, and reports that are published either a) in leading field journals in development or their equivalent, or b) as flagship reports by development organizations and disseminated around the world. In the case of the World Bank, work that meets these criteria can be classified as research regardless of whether it is performed by analysts in the research group in the Development Economics Department, in the regions, or in the networks.
and North Africa, and Eastern and Central Europe or in the high-income countries of Europe and North America.

Just as “labor market” is defined broadly in labor economics, so too are “labor market policies.” The core model of a labor market involves labor demand, labor supply, and determination of wages and other conditions of employment. It is useful to think of policies that have a bearing on the operation of a single labor market and those that involve the connection between labor markets. Accordingly, in any given labor market, there are those policies that have a bearing on the buying and selling of labor and that are usually called labor market policies (wage-setting mechanisms, industrial relations systems, and the like) as well as those policies that impinge upon the labor market (macroeconomic policies and investment policies which affect the position of the labor demand curve, education and training policies which affect the position of the labor supply curve for different skill levels, fiscal policies which affect the size of the tax wedge, etc.). And then there are those policies that affect the movement of workers and firms between labor markets (hiring and firing regulations, labor market information systems, and so on).

Some units within the World Bank use the MILES framework:

<table>
<thead>
<tr>
<th>Policy issues</th>
<th>Macroeconomic conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Policies for growth</td>
</tr>
<tr>
<td></td>
<td>Macroeconomic stability</td>
</tr>
<tr>
<td>Investment climate</td>
<td>Regulatory environment</td>
</tr>
<tr>
<td></td>
<td>Government transparency</td>
</tr>
<tr>
<td></td>
<td>Taxes</td>
</tr>
<tr>
<td></td>
<td>Financing</td>
</tr>
<tr>
<td></td>
<td>Infrastructure</td>
</tr>
<tr>
<td></td>
<td>Legal environment</td>
</tr>
<tr>
<td>Labor market policies and institutions</td>
<td>Labor market regulation</td>
</tr>
<tr>
<td></td>
<td>Wage-setting</td>
</tr>
<tr>
<td></td>
<td>Non-wage costs</td>
</tr>
<tr>
<td>Education and skills</td>
<td>Basic education</td>
</tr>
<tr>
<td></td>
<td>Higher education</td>
</tr>
<tr>
<td></td>
<td>Training and lifelong learning</td>
</tr>
<tr>
<td>Safety net for workers</td>
<td>Income support</td>
</tr>
<tr>
<td></td>
<td>Active labor market policies</td>
</tr>
</tbody>
</table>
All of these aspects fit into the purview of labor market policies as defined in this review.

The status of labor market analysis and labor market policy in the development economics community now is similar to the status of poverty analysis and anti-poverty policy two or three decades ago. At the time, the profession knew that it wanted to take on poverty more fully but most in it didn't know how. Knowledgeable insiders in development institutions joined with outsiders in universities and think tanks to set the direction for poverty analysis and anti-poverty policy and show those who were not poverty specialists how to do good poverty work. Today, anyone who is not already an expert in poverty analysis and anti-poverty policy can acquire such expertise via a rich array of books, documents, articles, and training courses.

A parallel task is still to be performed for labor market work in developing countries today. This paper endeavors to take a small step in that direction by discussing the major building blocks of policy-relevant research on labor markets and to suggest some priority areas for future labor market research. The three building blocks are the welfare economic criterion or criteria, the theoretical model of the labor market under consideration, and empirical evidence that informs the theoretical models.

The following policy evaluation framework has proven useful in a variety of contexts and forms the basis for the analysis that follows:

Policy instruments → Model → Outcomes →

Policy evaluation criteria → Policy evaluation judgment

The first step in the framework is to clarify the policy instrument or instruments under consideration: a law that might be passed, a regulation that might be imposed or removed, a tax, a public expenditure, and the like. This is the action that has been undertaken or that is under consideration.
The second part of the framework is to specify the model. Good models endeavor to capture the essential aspects of reality while leaving aside the less essential ones. Models can be theoretical or empirical; the best ones involve an interplay between the two.

Once the policy instrument(s) has/have been specified, the third step is to use the model to predict the likely outcomes. Sometimes these outcomes are theoretical ones, e.g., the policy action in question would increase outcome variable Y if parameter $\alpha$ is greater than a critical value $\alpha^*$. Other times, the outcomes are empirical ones, e.g., those countries that have better social protection programs have lower wage inequality (Rama, 2003a).

The policy evaluation is not yet complete. The fourth step is for the evaluator to specify the policy evaluation criteria. (I use the plural, because a full-fledged policy evaluation usually relies on multiple policy evaluation criteria than a single criterion.) Some evaluators are comfortable using efficiency and equity as their criteria. Examples of other criteria are discussed further below.

Fifth and finally, the outcomes are then evaluated in terms of the evaluation criteria and an overall judgment made.

Imagine that you are a researcher or a country economist formulating a labor market policy proposal which you propose to take to a client country's government. In a completely-formulated proposal, you would answer the following three questions to the best of your ability:

- What specific labor market objective or objectives are you trying to achieve and by what welfare economic criterion or criteria will you decide if your objective(s) is/are being achieved?
- What is the labor market model you are using to analyze the effects of the proposed policy?
• What is the empirical evidence favoring one view of labor market functioning over another?

These three questions are the ideal. They define what we want to strive for. They are not necessarily what we must have.

Much useful policy-relevant research has a more limited objective. One such limited objective is to find out more about what the right model is. Are wages set according to supply and demand, by efficiency wage considerations, or by institutions? Do better-educated workers in a country earn more because of human capital created in schools, because of signaling, or because of screening? Why are people working informally selling their labor to themselves? Another limited objective is to learn more about a single outcome or parameter. How elastic is the demand for labor in a country? If more jobs are created in one area of a country, how much migration into that area is induced? How important do firms judge various aspects of the investment climate to be? Policy research aimed at these limited objectives is valuable too, and in this paper I present examples of empirical research that has led to interesting policy conclusions.

In this introduction, I have emphasized that comprehensively-formulated labor market policies require explicit welfare economic analysis, well-specified theoretical models, and appropriate empirical evidence. In the next three sections, I elaborate in turn on each of these three components. The paper concludes by summarizing the lessons learned and priorities for future research.

The Welfare Economics of Policy Analysis in the Labor Market

Social cost-benefit analysis: an introduction

Economics is often defined as the science of the allocation of scarce resources among alternative uses. In the current context, the scarce resources are the funds available for development purposes, for example, in a World Bank budget for a country.
The alternative uses are many: labor market interventions, social sector programs, infrastructure activities, and the like.

Consider the following statements on cost-benefit analysis by the last three chief economists of the World Bank:

“The progressive decline of cost-benefit analysis . . . may be explained by the formidable demands it places on data availability, on the exhaustive understanding required of economic mechanisms in a non-competitive model of the economy, and on the knowledge demanded regarding institutions and governance parameters. But as our discussions have revealed, these issues still loom large . . . – we just are not dealing with them in a comprehensive framework.”
François Bourguignon, Chief Economist, The World Bank (Bourguignon, 2006)

“The purpose of cost-benefit analysis is to provide a consistent procedure for evaluating decisions in terms of their consequences. This might appear as an obvious and sensible way to proceed, but it is by no means the only one (examples of alternative procedures are majority voting, collective bargaining, the exercise of power, or the assertion of rights).”

“Around the world, many governments find it necessary, and worthwhile to subsidize artistic enterprise such as opera and theater, and some, including France and Morocco, provide subsidies for cinema production as well. . . What I find so striking about this example is the social cost-benefit analysis. . . There is little chance that French-language films, subsidized or not, will make a major dent in Hollywood profits. Whether subsidizing them is a good way for the French government to spend its money should be a matter for the French people to decide. If they spend it well, not only those in France but filmgoers around the world will benefit.”

As these quotations indicate, the World Bank’s last three Chief Economists all maintain that a useful approach for deciding how to allocate scarce resources among alternative uses is social cost-benefit analysis. In a comprehensive survey article, Prest and Turvey (1965) described social cost-benefit analysis thus: “Cost-benefit analysis is a practical way of assessing the desirability of projects, where it is important to take a long view (in the sense of looking at repercussions in the further, as well as the nearer, future)
and a wide view (in the sense of allowing for side-effects of many kinds on many persons, industries, regions, etc.), i.e., it implies the enumeration and evaluation of all the relevant costs and benefits.” See also Atkinson and Stiglitz (1980), Gramlich (1981), and Rosen (2005), among others, for further elaborations of the social cost-benefit analysis approach.

The cost-benefit criterion may be stated as follows: A project is socially profitable if the marginal social benefits (broadly defined and appropriately discounted) exceed the marginal social costs. The larger the difference between discounted benefits and costs, the more worthwhile the project.

This section begins with some remarks on the setting of objectives for policy decisions. Several types of social cost-benefit analysis are then discussed: the Pareto criterion; the quasi-Pareto criterion, first order dominance, and rank dominance; abbreviated labor market well-being functions; minimizing a low-earnings index; and traditional social cost-benefit analysis. The section concludes with some comments on social cost-benefit analysis in practice.

**Specifying the ultimate objective or objectives**

When social cost-benefit analysis is used as a basis for allocating scarce resources among alternative uses, it is essential that an objective or a number of objectives be specified. These need to be the ultimate objective or objectives of policy, not intermediate objectives.

Think of familiar maximization problems in economics. An individual maximizes utility by choosing how much money to spend on each of a number of alternative products or by allocating one’s time between work and leisure. A firm maximizes profits by choosing how much output to produce, what production technology to use, and how much of each input to employ. A government or international agency seeks to minimize poverty by allocating an anti-poverty budget among the populace. In these and other
decisions like them, what is being maximized or minimized - technically speaking, the maximand or minimand - is made explicit.

A great many studies purporting to offer “policy implications” from labor market research do not have a well-formulated maximand or minimand. Typically, inputs to or components of the maximand or minimand are analyzed but not the maximand or minimand itself. Minimum wage policy is sometimes analyzed in terms of the disemployment caused, active labor market policy by the penetration rate of the national employment service, pension policy by the social security replacement rate, and unemployment insurance by the percentage of workers who are eligible for benefits in the event of job loss.

Let us consider some more comprehensive alternatives. The interested reader is referred to Fields (2001, Chapter 8) for further details.

**The Pareto criterion**

The Pareto criterion maintains that a policy is beneficial if it makes some people better off and nobody worse off. “Better off” and “worse off” are defined in utility terms. A policy that makes some people better off and none worse off is said to be “Pareto-improving.”

Seeking Pareto-improvements and implementing them when they are found is indeed a sound criterion for social decision-making. However, implementing the Pareto criterion raises three practical problems.

First, the analyst must know the utility functions of the persons involved. Hardly ever is this the case.

Second, the analyst must accept the utility functions of the persons involved, even if they are nasty, malevolent, or weird. In practice, we may well find such judgments objectionable and might therefore want to “launder” them (Harsanyi, 1977, Goodin, 1986), for example, by paying no heed in our policy decisions to any malevolence a person may feel.
And third, we must have data on the actual or prospective changes in utility for each specific individual; rarely are prospective models or panel data on actual outcomes available.

Suppose though that these practical problems could be overcome. There remains the fact that it would be an unusual policy indeed in which nobody is made worse off. Typically, there are losers and rarely are the losers compensated fully for their losses.

Although the Pareto criterion is appealing conceptually, the practical issue is the general absence of Pareto-improving policy interventions – that is, the Pareto criterion typically is too demanding to be useful. In the great majority of circumstances, other more applicable policy criteria must be sought.

**The quasi-Pareto criterion, first order dominance, and rank dominance**

Two of the problems with the Pareto criterion – its reliance on utility information and its need to compare the same individuals before and after a proposed policy intervention – can be overcome by switching to a less demanding policy criterion. The problem of not knowing utilities can be overcome by working instead with incomes. If we know what a particular policy decision would do to the incomes (not the utilities) of each individual in the economy, then we can adopt the quasi-Pareto criterion: a policy is quasi-Pareto-improving if some incomes go up and no incomes go down.

However, even the quasi-Pareto condition may be too demanding. Sometimes we know that some people will be affected but we do not know which ones – for example, some but not all people will get additional educational spaces or better jobs, but it is up to admissions officers or hiring officers to decide who these individuals will be. In such cases, the problem of not being able to compare the same individuals before and after a policy intervention is overcome by looking instead at anonymous individuals – for example, those in the lowest income quintile, those in the second income quintile, etc. First order dominance and rank dominance are tools for making such comparisons.
Specifically, first order dominance is the following idea. One income distribution is said to first-order-dominate another if, at each income level, there are fewer people below the specified income amount in the first distribution compared with the second. First order dominance therefore overcomes the two aforementioned problems by looking at the incomes (not utilities) of anonymous (not identified) individuals.

Practically speaking, what needs to be done to implement first order dominance is both quite simple and quite difficult. The simple part is when you have the distributions to be compared. Draw two cumulative distribution functions, which virtually any statistical software package can do. If one cdf is below the other everywhere and strictly below it somewhere, then the first distribution first-order-dominates the second.

First order dominance is equivalent to rank dominance. Rank dominance looks at the income received at each rank in the income distribution. If the income is everywhere at least as high in one distribution compared with another and strictly higher somewhere, then the first distribution rank-dominates the second. Practically speaking, to implement rank dominance, draw exactly the same graphs as with first order dominance but look at them differently. For rank dominance, draw the two cdfs, and if one cdf is found to lie to the right of the other one everywhere and strictly to the right somewhere, then the first distribution first-order-dominates the second. Of course, if one curve lies to the right of another, then it necessarily lies below the other one also, which is why rank dominance and first order dominance are equivalent.

The hard part is deriving the distributions to be compared. Implementing first order dominance requires that the analyst compare the entire income distribution vectors before and after a policy is put into effect. Careful modeling is required to derive the counterfactual vector of incomes under the proposed policy. This is demanding informationally, but no more demanding than the other policy evaluation methods that follow.
Finally, it bears mention that when the pre- and post-policy cdfs are drawn, they may cross. In such a case, the post-policy distribution neither dominates nor is dominated by the pre-policy distribution. The analyst must then rely on some other criterion for evaluating the policy.

**Abbreviated labor market well-being functions**

Lambert (1989) coined the term “abbreviated social welfare function” to indicate that social welfare (or, equivalently, “economic well-being”) is a function of statistics which are themselves calculated from the before and after income distribution vectors. Examples of such statistics are GDP per capita, inequality, poverty, total wage bill, and the unemployment rate. Each of these components must itself be carefully defined and measured.

In addition to specifying the arguments of the abbreviated social welfare function, the analyst must also specify the direction of the effect: for example, that an increase in GDP per capita is judged favorably and an increase in inequality negatively. Social welfare is then said to have increased for all functions belonging to a broad class of abbreviated social welfare functions if and only if all of the components that have been deemed to be good are found to have increased or remained unchanged and all of those deemed bad are found to have decreased or remained unchanged. If even one component goes in the “wrong” direction from the others, that broad class of abbreviated social welfare functions produces an ambiguous evaluation. Note that the evaluations we reach are contingent: all evaluators who adopt the particular class of social welfare functions in question would reach the same judgment, but evaluators who adopt different social welfare functions might reach different judgments (Atkinson and Stiglitz, 1980; Atkinson, 1983).

An example of such an abbreviated social welfare function is one that I used in Fields (2005b) to evaluate alternative labor market policies in one of the theoretical
models reviewed below, the Harris-Todaro labor market model. The welfare function chosen was

\[ SW = f(\text{Total labor earnings, unem, ineq, pov}), f_1 > 0, f_2 < 0, f_3 < 0, f_4 < 0 - \]

that is, social welfare is a function of total labor earnings, unemployment, inequality, and poverty, with the judgments that other things equal an increase in total labor earnings increases social welfare and increases in unemployment, inequality, or poverty lower social welfare. Three policies were evaluated: modern sector enlargement, which increases the number of modern sector jobs holding wages in the two sectors constant; modern sector enrichment, which increases wages in the modern sector holding the number employed in the modern sector and wages in the traditional sector constant; and traditional sector enrichment, which increases wages in the traditional sector holding wages and employment in the modern sector constant. For the class of social welfare functions given by (1), a policy of traditional sector enrichment was shown to be unambiguously welfare-improving, whereas policies of modern sector enrichment and modern sector enlargement were shown to be welfare-ambiguous.\(^2\)

In similar fashion, in analyzing labor market policies, we can move to what might be called an “abbreviated labor market well-being function.” What should be included in such a function? Today, four components would be regarded as essential.

First, following traditional practice, an abbreviated labor market well-being function would presumably include the unemployment rate. Even though the

\(^2\) Abbreviated social welfare functions have been used both explicitly and implicitly in World Bank research. As one example of explicit criteria, de Ferranti et al. (2000) assess social protection options in Latin America using coverage, cost, incentives, and insurance. As another example of explicit criteria, Vodopivec (2004) identifies two main classes of performance criteria: distributive effects and efficiency effects. Distributive effects include coverage, adequacy of support, and income distribution. The efficiency effects include job-search efforts, post-unemployment wages, equilibrium labor market outcomes (employment, unemployment, labor force participation), restructuring and overall economic adjustment, labor supply of other family members, taking jobs in regular versus informal sector, and aggregate output and growth.
unemployment rate should probably receive some weight, how much weight it should receive is open to discussion (Peek, 2006).

The abbreviated labor market well-being function would also recognize that in addition to how many workers have jobs, what also matters is how much workers earn in those jobs (Fields, 1980b; Squire, 1981; Turnham, 1971, 1993; Somavia, 1999). Thus, a second argument of an abbreviated labor market well-being function would be a measure of labor market earnings such as the average earnings of those employed, the incidence of low earnings, or the extent of earnings deprivation.

A third component would be social protection. For example, Bourguignon and Dethier (2006) have written, “The term ‘good job’ has two connotations: higher wages and more protection against dismissals, earnings variation and several risks linked to health, old age and unemployment.”

A fourth component would be core labor standards. The ILO regards the Declaration on Fundamental Principles and Rights at Work as a central feature of its drive for Decent Work. A job in which core labor standards are violated cannot be regarded as decent.

Taking these four components into account, we arrive at an abbreviated labor well-being function of the form:

(2) Labor market well-being = g(level of employment, earnings of those employed, extent of social protection, adherence to core labor standards),

\( g_1, g_2, g_3, g_4 > 0 \).

Four points may be noted about the use of such an abbreviated labor market well-being function.

First, in the same way that the abbreviated social welfare function given by (1) requires careful specification of the total earnings, unemployment, inequality, and poverty arguments, so too does the abbreviated labor market well-being function given
by (2) require careful specification of the employment, earnings, social protection, and core labor standards arguments.

Second, the \( g(.) \) function in (2) requires that at least one component increase and no component decrease in order for an improvement to be judged to have taken place. Any policy that increases one argument while decreasing another cannot be judged to be unambiguously desirable or undesirable.\(^3\)

Third, in the event of both increases and decreases, \( g(.) \) must be further particularized in order to weigh the tradeoffs and reach a judgment. The low-earnings index discussed in the next sub-section is designed to do exactly that.

And fourth, well-being in the labor market is, of course, only part of economic well-being in the economy as a whole. Analyses of the type given by (2) can be supplemented by additional considerations outside of the labor market, but in my view labor market considerations are the best place to start.

Much labor market policy research analyzes only one or a small number of effects of a given policy, and therefore fails to fulfill the abbreviated social welfare function criteria outlined here. For example, Coudouel and Paci (2006) reviewed three areas of labor market policy. Studies by various researchers in various countries have analyzed the effects of minimum wages, employment protection legislation, or unemployment benefits on: the level of employment; the wage level, the average wage, and the wage distribution; inequality and poverty; employment flows; types of employment; the demographic structure of employment; formal employment; unemployment; productivity; and poverty. However, none of the research reviewed has examined all of these effects or even a substantial subset of them. It is useful for policy-makers to know about some of

\(^3\) In an otherwise highly informative report, the World Bank (2005) expressed its policy challenge in Eastern Europe and the former Soviet Union as “fostering the creation of more and better jobs.” However, the report calls for some policies, including liberalizing employment protection legislation and reducing taxes on labor (and hence the benefits those taxes pay for), that may make for more jobs but not better ones. Such conflicting policy recommendations are also found in other World Bank reports that are labeled “Not for Circulation” and which therefore cannot be discussed here.
the effects that their policies have, but they are in no position to make wise policy judgments unless they know about most or all of the policy effects.

**Minimizing a low-earnings index**

An attractive way of combining the level of employment with the earnings of those employed and to allow for tradeoffs between the two is to draw a parallel with poverty measurement. A poverty index widely used today is the measure developed by Foster, Greer, and Thorbecke (1984):

\[
P_\alpha = \sum_{i=1}^{q} \left( \frac{y_i}{z} \right)^\alpha,
\]

where \( q \) is the number poor, \( z \) is the poverty line, and \( y_i \) is the income of recipient \( i \). To allow for diminishing marginal utility of income, the parameter \( \alpha \) must be chosen to be greater than one. The value \( \alpha = 2 \) is conceptually appealing and commonly utilized. The \( P_2 \) index is sometimes called the “severity of poverty” (Ravallion, 1994). In similar fashion, an index of the severity of low-earnings might be defined as

\[
P_2 = \sum_{i=1}^{q} \left( \frac{z - y_i}{z} \right)^2,
\]

where now \( q \) is the number of low earners, \( z \) is the low-earnings line, and \( y_i \) is the labor market earnings of worker \( i \). Observe that the index of severity of low-earnings given by (4) treats the unemployed as low-earners: the earnings deficit of an employed worker with earnings 10% below the low-earnings line is counted with a weight of 10%, the deficit of one with earnings 50% below the low-earnings line is counted with a weight of 50%, and the deficit of a potential earner who is unemployed and who therefore has zero earnings is counted with a weight of 100%.

The approach used in (4) is almost utilitarian but not quite. As described in Foster and Sen (1997), simple utilitarianism can be factored into three components: (i) consequentialism (the well-being of an action or institution is judged only by the goodness of the consequent state of affairs); (ii) welfarism (judging the goodness of states
of affairs only by utility information); and (iii) sum-ranking (judging utility information for a given population simply by summing utilities). Let \( u_i(.) \) denote the actual utility function of individual \( i \) as a function of his/her labor market earnings and let \( v_i(.) \) denote how the evaluator evaluates \( i \)'s labor market earnings. Denote by (ii') what we get if we use \( v_i(.) \) in place of \( u_i(.) \) in (ii). The combination of (i), (ii'), and (iii) might be called “quasi-utilitarianism.” Quasi-utilitarianism is similar in spirit to the approach taken by Atkinson (1970) in his famous paper on the meaning and measurement of inequality. As he later wrote (Atkinson, 1983, p. 5), “It should be stressed that the [welfare] function
\[
W = \int_0^7 U(y) f(y) dy
\]
is defined in terms of individual incomes, and not individual utilities . . . In retrospect, it would have been wiser to use a letter other than \( U \).” [Emphasis in the original.] As Foster and Sen (1997, p. 116) later put it, “Taking \( U_i \) to be individual utility . . . would be one possible interpretation, but in general \( U_i \) would stand for the individual component of social welfare that is associated with person \( i \).” [Emphasis in the original.]

In my view, the case for choosing labor market policies according to the criterion of maximizing
\[
W = \int_0^7 U(y) f(y) dy
\]
by minimizing
\[
P_2 = \sum_{i=1}^{q} \left( \frac{z - y_i}{z} \right)^2
\]
is a strong one, deserving serious consideration in policy discussions.

**Traditional social cost-benefit analysis**

Another way of carrying out social cost-benefit analysis – actually, the more traditional way – is to total the quantifiable costs and benefits in dollars. When dollars are used as the basis for the comparison, the analyst is said to be using a “money-metric utility function.” Totaling the quantifiable costs and benefits in dollars deliberately ignores the distributional question of which people receive how much in benefits and receive how much in costs. Limiting the analysis to quantifiable benefits and costs deliberately ignores factors that may be very real but that cannot be quantified. Both of these omissions are for practical reasons, not conceptual ones.
Two methods for social cost-benefit comparisons are available. One is the social rate of return method. The other is the net present value method.

The social rate of return is found by equating the marginal social benefits of the project in question to the marginal social costs and solving for \( r \):

\[
PV_{\text{benefits}} = PV_{\text{costs}}
\]

or

\[
B_0 + \frac{B_1}{(1+r)} + \frac{B_2}{(1+r)^2} + \ldots = C_0 + \frac{C_1}{(1+r)} + \frac{C_2}{(1+r)^2} + \ldots,
\]

where \( B_t \) and \( C_t \) are the social benefits and social costs in year \( t \). To avoid confusion, let us call the value of \( r \) that solves (5) the “internal rate of return” and denote it by \( r^* \). The project is judged to be socially worthwhile by the social rate of return criterion if the internal rate of return \( r^* \) exceeds the reference rate used in other projects.

The net present value is found by specifying a social discount rate \( r' \), calculating the present value of social benefits and social costs, and comparing the two:

\[
NPV \equiv B_0 + \frac{B_1}{(1+r')} + \frac{B_2}{(1+r')^2} + \ldots - C_0 + \frac{C_1}{(1+r')^2} + \frac{C_2}{(1+r')^2} + \ldots
\]

By this criterion, a project is judged to be socially worthwhile if the net present value is positive, not socially worthwhile otherwise.

Social rates of return are commonly used to compare projects of different sizes. Nonetheless, as stressed in the public economics literature (e.g., Rosen, 2005), the net present value rule is theoretically the better of the two decision rules.

**Social cost-benefit analysis in practice**

Different types of social cost-benefit analysis have just been discussed: the Pareto criterion; the quasi-Pareto criterion, first order dominance, and rank dominance; abbreviated labor market well-being functions; minimizing a low-earnings index; and traditional social cost-benefit analysis. Sometimes these criteria are too demanding and a less ambitious goal is sought: to determine whether Policy A would get the job done at
lower cost than Policy B or that more benefits could be delivered for the same cost if Policy C is followed rather than Policy D.

A nice example of when the social costs of a project have been factored into social decision-making is Ravallion and Datt’s (1995) impact evaluation of a workfare program in India. The program was found to have substantially reduced poverty amongst the participants relative to the counterfactual of no program. But once the costs were taken account of, the authors found that an alternative counterfactual – an untargeted allocation of the same budget outlay – would have reduced poverty more.

More typically, social costs are not factored into policy evaluations. Consequently, the validity of any policy “conclusions” coming out of such research is called into question.

Four aspects of the cost-benefit criterion for social decision-making are worth highlighting.

The first is that social cost-benefit analysis in any of its forms requires that social benefits and costs be analyzed rather than private benefits and costs. Often, the distinction between private and social rates of return seems to be little more than a difference of a few percentage points after adjusting for mortality, taxes, and public subsidies. But other times, the differences may be quite large, e.g., in the case of costly higher education which is free to the student.

Second, when analysts have done cost-benefit analysis, they typically have calculated average rates of return while generally neglecting to point out that they are in fact averages and not marginals. All economists know to base decisions on marginals, not averages.

Taken together, the calculation of average private benefits rather than marginal social benefits may be seriously flawed: in some circumstances, the average private returns may exceed the marginal social returns by an order of magnitude or more (Fields, 1995).
Third, while early social cost-benefit analysis was formulated in terms of total benefits and total costs, we social scientists have become much more sensitive in the last twenty or thirty years to distributional aspects. We now ask: Who (in terms of position in the income distribution) receives the benefits? Who (in terms of position in the income distribution) pays the costs? This concern with distributional issues reflects the reformulation of the very notion of economic development as having an important anti-poverty component rather than being just economic growth itself.

Finally, non-quantifiable benefits and costs should be considered. There are some who maintain that if something cannot be quantified, it does not exist. I could not disagree more: if that were the case, there would be no such things as love, pain, or friendship, and humanity would be much the worse for it. Revealed preference is often helpful in setting a bound on the non-quantifiable benefits or costs. When I buy my wife a dozen roses, I can be understood to have judged the benefits to our relationship to be worth at least the price paid. The same can be said for accommodations made to provide the disabled with equal access in the workplace, to provide special education programs for students with special needs, or to provide pension benefits to those who have no legal entitlement to them - provided, that is, that the benefits and costs have been carefully considered and weighed.

Cost-benefit analysis may well lead to different decisions from what other criteria might. Here are three examples.

In the education area, decisions are sometimes made on the basis of human resource planning (what used to be called “manpower planning”). It is right to ask, as the human resource planner does, “What is the absorptive capacity of the economy? Will the newly-educated person be employed in an occupation for which he or she is trained?” However, this is not enough. It is also right to ask, “What is the nature of the work the graduate will perform and what benefits will this work confer on society?” and to ask too, “How much will it cost society to educate another graduate? Do the benefits outweigh the
costs?” The human resource planner does not ask these latter questions; his/her goal is to equate the supply of persons with different skills and educational attainments to the demand for them. This is the wrong criterion, one that is likely to lead to an overproduction of education. Why overproduction? Because the demanders consider only the private costs to themselves of hiring workers with certain qualifications but not the costs to society of producing these qualifications.

As another example of how the social cost-benefit approach differs from other approaches, consider the merit goods argument originated by Musgrave (1959): that a project (be it education, health care, infrastructure, or whatever) may be inherently meritorious and therefore deserving of additional public expenditures. The problem with the merit goods argument is that it fails to recognize opportunity costs: given limited budgets, spending more money in any one of these areas means spending less in another. The social cost-benefit analyst would be justified in concluding that we have too much public spending on X, because we are spending too little on Y, given the costs and benefits of X compared with Y.

A third example is information in the labor market. It is often said that better labor market information systems will help the unemployed find work and thereby reduce unemployment in the economy. The first part is probably correct: those unemployed individuals who access improved labor market information systems will have a better chance of gaining employment. The second part may well be incorrect. It is correct that unemployment in the economy will fall as a result of improved labor market information systems provided that the unemployment is of the frictional variety caused by lack of information. But if the unemployment is of the deficient demand type, no amount of information for the unemployed will increase an insufficient volume of employment. Rather, the additional information will reallocate the existing jobs among job-seekers. In this case, there may be a private benefit without a social benefit.
As the quotations at the beginning of this section indicate, Bourguignon, Stern, and Stiglitz all believe in the importance of social cost-benefit analysis. I quite agree: in my view, it is much better to ask the right questions and answer them partially than to answer the wrong questions precisely. Social cost-benefit analysis asks the right questions. It is a shame that is used less now than it used to be. I leave the last word to T.N. Srinivasan (1982, p 244): “Indeed, by forcing the project formulators and evaluators to specify in detail the probable impact of a project, on particular socioeconomic groups, decisions are likely to be made on a more informed and possibly consistent basis than would otherwise be the case.”

**Theoretical Models**

**The overall framework**

Economics is usually divided into two branches, macroeconomics and microeconomics. For present purposes, though, neither of these is the relevant level of analysis. Rather, the focus here will be the market level, specifically the labor market.

As noted above, the labor market is defined as the place where labor services are bought and sold. Wage-, salaried-, and self-employment are all included. All of the models reviewed here maintain that the aggregate labor market consists of a small number of labor market sectors (or segments) linked to one another by actual or potential mobility of workers and firms. Conditions in one labor market affect and/or are affected by conditions in other labor markets. Thus, all of these models maintain that labor market outcomes are best understood by analyzing conditions in markets, not conditions in individual firms.

For any given labor market sector, market conditions are determined by market labor demand, market labor supply, and a process determining wages and conditions of employment. Market labor demand and market labor supply underlie all of the models reviewed in this paper. However, the models differ depending on whether earnings levels
and conditions of employment (hereafter “wages”) within a sector are determined by supply and demand, by institutional forces, by efficiency wage considerations, or by something else.

The models developed below all assume a multiplicity of labor market sectors with the conditions in one often differing from the conditions in others. Dualistic labor market models maintain that there are just two labor market sectors, which I shall term formal and informal. Others might call them industry vs. agriculture, urban vs. rural, modern vs. traditional, wage employment vs. self-employment, covered vs. non-covered, or good jobs vs. bad jobs. Segmented labor market models, also called multisector labor market models, add a third sector or more - for example, distinguishing between traditional agriculture and the urban informal sector in addition to the form sector.

Different intersectoral linkages can be specified as alternative stylizations of different countries’ conditions. Among those considered are models featuring non-competing groups, full market-clearing, crowding, expected wage equalization, and restrictions on entry due to capital market constraints.

Many of the models reviewed in this section are elaborated at greater length in Fields (2005a), to which the reader is referred for additional details.

**The need for multisector models**

Leading the way in the development of multisector models were two Nobel Prize-winning economists, W. Arthur Lewis (Lewis, 1954) and Simon Kuznets (Kuznets, 1955). Though Lewis and Kuznets differed in many respects, they agreed on one fundamental one: that the feature distinguishing one economic sector from another is the fact that workers earn different wages depending on the sector of the economy in which they are able to find work. Lewis wrote (p. 150): “Earnings in the subsistence sector set a floor to wages in the capitalist sector, but in practice wages have to be higher than this, and there is usually a gap of 30 per cent or more between capitalist wages and subsistence earnings.” Part of the gap was “illusory,” Lewis wrote, because of the higher cost of
living in the capitalist sector. Yet, there remained a real wage gap due to a) the “psychological cost of transferring from the easy going way of life of the subsistence sector to the more regimented and urbanized environment of the capitalist sector,” b) the payoff to experience in the capitalist sector, and c) “workers in the capitalist sector acquiring tastes and a social prestige which have conventionally to be recognized by higher real wages.” He then went on to analyze the dynamics of economic growth as profits earned in the high-income sector were reinvested, leading to capital formation, an increased demand for labor, and continued intersectoral shifts.

At the same time, Kuznets (1955) explored how various measures of income inequality would change as the high-income sector comes to employ an increasing share of the population, a process known as “intersectoral shifts.” All of the inequality measures used by Kuznets in his calculations exhibited an inverted-U in the context of intersectoral shifts, which pattern later came to be known as the “Kuznets Curve.”

Research on labor market dualism has been developed considerably in the intervening years. Some of the subsequent writings on labor market dualism were grounded in the then-emergent theory of human capital, which also earned its developers the Nobel Prize (Schultz, 1961; Becker, 1964). Dual labor market theory received a boost through the work of Doeringer and Piore (1970). As summarized by Wachter (1974, p. 639), the dual labor market model advanced four hypotheses:

First, it is useful to dichotomize the economy into a primary and a secondary sector. Second, the wage and employment mechanisms in the secondary sector are distinct from those in the primary sector. Third, economic mobility between these two sectors is sharply limited, and hence workers in the secondary sector are essentially trapped there. Finally, the secondary sector is marked by pervasive underemployment because workers who could be trained for skilled jobs at no more than the usual cost are confined to unskilled jobs.

Wachter and also Cain (1976) stressed that for dualism to exist, different wages must be paid in different sectors to comparable workers. Empirical evidence has been compiled
demonstrating that observationally equivalent workers earned different amounts in different parts of the economy, but skeptics remain unconvinced that dualism or segmentation in fact exists. Nonetheless, the good jobs/bad jobs dichotomy remains widespread in the current literature (e.g., Inter-American Development Bank, 2003; Bourguignon, 2005; Avirgan, Bivens, and Gammage, 2005; World Bank, 2006), as I think it should.

The alternative to a segmented labor market model is a unitary labor market model in which all workers with given skills receive the same wage regardless of the sector in which they work – a view which I find to be at odds with developing countries’ reality. When possible, Occam’s Razor suggests limiting the analysis to two sectors. Models with two sectors – formal and informal – are elaborated below. But when two sectors are not enough, three-sector or n-sector models can prove insightful.

**Modeling the formal sector labor market**

In this subsection, I review three groups of models of labor markets for formal sector workers of a given skill level. They are the supply-demand model with market wage determination, the supply-demand model with institutional wage determination, and efficiency wage models.

The supply-demand model with full market-clearing (e.g., Ehrenberg and Smith, 2006) has three essential features, well-known to all economists. First, the amount of labor demanded in the market is taken as a decreasing function of the wage, other things equal. The market labor demand curve slopes downward because of diminishing marginal revenue product of labor and the associated substitution and scale effects of a wage change for the constituent firms. Second, the amount of labor supplied to the market is taken as an increasing function of the wage, other things equal. The market labor supply curve slopes upward because a higher wage induces workers to enter this labor market from other labor markets and induces non-workers to enter the labor force. And third, the wage is set by supply and demand in order to clear the market.
According to the market-clearing model, three equilibrating forces operate: behavior of firms, behavior of workers, and behavior of wages. In the model, firms are free to hire workers or not depending on what is in their profit-maximizing interest to do. If market conditions change, what is in their profit-maximizing interest to do will change accordingly, and firms are free to act on these changes. Similarly, workers are free to supply their labor in any given labor market or not depending on what is in their utility-maximizing interest to do. For them too, if market conditions change, what is in their utility-maximizing interest to do will change accordingly, and they (workers) are free to act on these changes. And finally, if supply and/or demand conditions change, real wages are free to rise or fall accordingly. (In this paper, wages should always be thought of in real terms.)

A second group of models retains the demand and supply curves but not market-clearing. Rather, in these models, wages are set by forces different from supply and demand. A number of institutional forces are identified. Minimum wages aim to assure covered workers an adequate standard of living. Trade unions are often encouraged by government policy as a means of entitling workers to a “just” share of the fruits of their labor. Public sector pay policies often result in substantially higher wages being paid to government workers than to their private sector counterparts. Multinational corporations frequently pay above-market wages. Labor codes in some countries regulate hiring and firing, impose payroll taxes on firms, and mandate that employers provide certain benefits to their workers.

What unites all of the models in the second group is that the usual kind of labor market equilibration does not take place. One of the standard equilibrating forces is that if the wage is above the market-clearing level and consequently workers are unemployed, the unemployed would offer to work for lower wages rather than remain unemployed, an offer that employers will gladly accept. However, the existence of the institutions just mentioned – minimum wages, trade unions, public sector pay policies, multinational
corporations, and labor codes - prevents the wage from falling. The equilibrium is characterized by a higher-than-market-clearing wage with consequent unemployment.

A different type of institutional force may act to maintain wages above market-clearing levels for reasons having to do with the supply side of the labor market (e.g., Bardhan and Rudra, 1981; Solow, 1990). Suppose we have a casual labor market in which fresh hiring takes place each day and suppose further that the demand for labor is inelastic. The fact that labor demand is inelastic means that the total wage bill paid to labor is higher the higher is the daily wage. The fact that the labor market is casual means that each worker will receive his or her share of the higher total wage bill if s/he does not undercut the established wage. Knowing this, each worker has an incentive to leave the wage where it is. A wage above the market-clearing level therefore persists in equilibrium.

A third group of models has firms in the formal sector setting wages above the market-clearing level for efficiency wage reasons; see Katz (1986), Akerlof and Yellen (1986), and Weiss (1990) for reviews. The basic postulate of efficiency wage theory is that firms will pay higher-than-market-clearing wages if and only if the gains in productivity from doing so outweigh the costs, so that profits are increased. Two types of mechanisms for realizing productivity gains are specified. One set of explanations is that higher wages enable firms to hire better-quality workers from a heterogeneous labor pool – for example, workers who have more education or who perform better on tests of potential job performance. The other set of explanations is that higher wages induce workers of a given skill level to perform in a more productive manner through better nutrition, improved morale, lower absenteeism, reduced shirking, diminished labor turnover, and/or greater discretionary effort.

Finally, mention should be made of models which, although not intended specifically to characterize the formal sectors of developing countries’ labor markets, nonetheless can be used to enrich our understanding of what goes on there. These include
the matching models of Mortensen and Pissarides (1999), the models of job creation and
destruction by Davis, Haltiwanger, and Schuh (1996), the ranking models of Blanchard
and Diamond (1994), and the imperfect information models of Stiglitz (2002).

Let us turn now to models of the informal sector.

**Modeling the informal sector labor market**

As described above, the essential feature of labor market dualism is that one sector offers relatively attractive wages and other terms and conditions of employment while another sector offers relatively unattractive ones. In accordance with the most common way of viewing these sectors, the more attractive sector is called here the “formal sector” while the less attractive sector is called the “informal sector.” Thinking of the labor market in this way leads working in the informal sector to be thought of as a type of employment of last resort. More recently, though, a different view has been proposed: that workers choose to work in the informal sector in preference to formal sector work. A third view is that the informal sector has its own internal dualism, combining the first two characterizations. Finally, the notion of informal work has broadened from thinking about workers in the informal sector to also include workers engaged in informal jobs outside the informal sector. These four views are taken up in turn.

Developing country labor markets exhibit low unemployment and substantial working poverty (ILO, 2007). The usual explanation offered for these facts is that workers in poor countries cannot afford to remain unemployed for very long, so unless they have a better option, in order to earn cash quickly, the majority of them take up wage-employment or self-employment in the informal sector. I believe that this explanation, proposed long ago in ILO (1972) and Hart (1973), is essentially correct. The informal sector thus plays the role of a free-entry or fallback sector for those who cannot obtain formal sector jobs. As Reynolds (1969, p. 91) put it early on, “The urban trade-service sector . . . is a relatively open sector in the sense that, with little skill and little
initial capital, a newcomer can crowd his way into employment. It is thus a natural entry point to the urban economy for migrants from the countryside.”

The earliest dualistic labor market models (e.g., Lewis, 1954; Harris and Todaro, 1970) had a single informal sector. Later models built in two informal sectors - an urban informal sector and a rural agricultural sector – to go along with an urban formal sector and unemployment (Fields, 1975). In these models, the best jobs were assumed to be located in the urban formal sector, in view of which job-seekers had an incentive to locate in the urban areas to improve their chances of being hired for the better jobs. Those choosing to work in either of the informal sectors were enabled to eke out a meager existence. This is not to say that they were happy about what they could earn when informally employed, but working and earning in the informal sector was better for most of them than being openly unemployed. Once this was recognized, the view of the employment challenge in developing countries changed from a concern with employing the unemployed to the additional concern of raising the earnings of those employed (Turnham, 1971, 1993; Squire, 1981; World Bank, 1995; ILO, 2007).

Informal sector wage determination can be modeled in a variety of ways. One is to assume that there is a fixed amount of income to be earned in the informal sector regardless of the number of people working in that sector - that is, the marginal product of labor is literally zero. A second approach is to regard a part of the informal sector as facing, instead of zero marginal product, constant marginal product. This assumption was adopted by many in what has come to be called the simplified Harris-Todaro model (Fields, 1975; Anand and Joshi, 1979; Heady, 1981; Stiglitz, 1982; Sah and Stiglitz, 1985; Bell, 1991). A third approach is intermediate between the first two: a positive but diminishing marginal product. Harberger (1971, pp. 574-5) put it thus:

[This] variant associates disguised unemployment not just with low wages but with situations in which the marginal productivity of labour lies below the actual wages earned. . . There are a variety of activities to which this argument applies. A classic example is that of fishermen on a lake. The
addition of more fishermen increases the total catch, but not proportionately, yet the last fisherman has an equal chance of making a given catch as the first. The expected catch is the same for all, and is equal to their average productivity. But, owing to the fact that the total catch does not increase in proportion to the number of fishermen, the marginal productivity of a fisherman is less than what he earns.

A fourth approach is to model a full demand system for agricultural and non-agricultural products and workers (Bourguignon, 1990).

More recently, one more need has become apparent. This is to give due recognition to the fundamental duality to be found within the informal sector. On the one hand, the informal sector has free-entry activities of the sort just described. On the other hand, it also has restricted entry activities that people who could be working formally choose to work in instead. In Fields (1990), I highlighted this distinction and labeled these two sectors the “free entry” part of the informal sector and the “upper tier” of the informal sector. Recently, William Maloney in a series of papers (among others, Maloney, 2003, 2004; Fiess, Fugazza, and Maloney, 2006; Bosch and Maloney, 2006) has maintained that, in Mexico at least, self-employment in the informal sector provides a package of wages, non-wage benefits, and working conditions that is at least as attractive for many people as what they could receive as wage employees in the formal sector. See also Webb (2005). Still, though, there is “no consensus” (Maloney’s term) on how many enterprises and individuals are to be found in each tier.

Finally, it bears mention that a broader notion of “informal economy” is on the ascendancy (ILO, 2002; Jhabvala et al, 2003; Chen et al, 2005). The ILO now defines the informal economy as comprising: informal employment (without secure contracts, worker benefits, or social protection) of two kinds: a) self-employment in informal enterprises (small unregistered or unincorporated enterprises) including employers, own-account operators, and unpaid family workers in informal enterprises and b) paid employment in informal jobs (for informal enterprises, formal enterprises, households, or no fixed employer) including: casual or day laborers, industrial outworkers, unregistered
or undeclared workers, contract workers, and unprotected temporary and part-time workers.

**Modeling intersectoral linkages in the labor market**

In this section, I review five ways in which the various sectors are linked to one another. They are the non-competing groups model, the integrated labor market model, the crowding model, the Harris-Todaro model, and the Banerjee-Newman model.

The non-competing groups model is one I mention because it is there, not because I believe it is a particularly useful stylization. This model maintains that the various labor market segments operate entirely separately from one another. Individuals belong either to labor market segment A or to labor market segment B, and they cannot or will not switch from one to the other regardless of differences in pay, prestige, or anything else. It is conceivable that such a model might apply to labor markets separated by caste in India, tribe in Africa, or gender in more traditional societies. On the other hand, it is not very likely to apply even in these contexts, because even if workers cannot move to the more attractive labor market segments, employers can switch to the workers who are more efficient relative to their cost. The outsourcing of jobs from developed to developing economies that is taking place at an accelerated rate today can be understood in precisely this way.

A second multisector model is the integrated labor market model. This model starts with two or more sectors but assumes that all of the equilibrating forces that apply to a single labor market with market-clearing also apply to a labor market with a multiplicity of sectors. In particular, workers are free to move from one sector to another, firms are free to move from one sector to another, and the wage in each sector is free to rise or fall. Given such equilibrating forces, any wage differential between one sector and another would quickly be eroded. Moreover, in the integrated labor market model, a policy leading to economic growth in one sector results in higher wages in all sectors. In
my view, such a model provides a good stylization of Taiwan and other East Asian economies.

A third model, the crowding model, makes a different set of assumptions. The crowding model starts with the assumption of a real wage differential between sectors. Also by assumption, any worker who is not employed in the high wage part of the economy takes up employment in the low wage part of the economy. The larger the number of people who “crowd” into the low wage sectors, the lower the wage will be there. A development economist will recognize that the model of Lewis (1954) is a crowding model. A labor economist is more likely to associate crowding with the labor market discrimination model formulated by Bergmann (1971). In the crowding model, as in the integrated labor market model, a policy that brings about economic growth in one sector of the economy results in higher wages in the rest of the economy – in this case, because economic growth and the consequent increase in employment in the high-wage sector of the economy leads to less crowding and hence higher wages in the low-wage sector of the economy.

A fourth way in which the various sectors may be linked is the model put forth by Harris and Todaro (1970). This model has both intersectoral wage differentials (as the crowding model does) and unemployment (which the crowding model does not). The Harris-Todaro model also has a spatial dimension: the high-wage jobs are located in urban areas while the low-wage jobs are to be found in rural areas. The reason that unemployment arises is that workers are attracted to the urban areas by the relatively high wages there. The lucky ones become employed, while the unlucky ones are unemployed. A Harris-Todaro equilibrium arises when the urban search strategy (obtain a high wage with probability p, become unemployed and earn nothing with probability 1-p) has the same expected value as the rural search strategy (obtain a low wage with probability one).
All of the models reviewed thus far have just two employment sectors. Subsequent modeling efforts added in more sectors. Fields (1975) extended the Harris-Todaro model to have three employment sectors – formal urban unemployment, informal urban employment, and agricultural employment – plus unemployment. Workers entering the labor market were assumed to face a limited number of job search strategies. In Fields (1975), these were: 1) Be unemployed and search for a formal sector job full time. 2) Take up an urban informal sector job and search for a formal sector job part time. 3) Take up a rural agricultural sector job and search for a formal sector job part time. Moene (1988, 1992) built duality into the rural sector while maintaining a single urban sector. Others extended the Harris-Todaro model to allow for on-the-job search, mobile capital, endogenous urban wage setting, risk-aversion, a system of demand for goods, and many other factors; see Fields (2005a) and Khan (forthcoming) for citations.

A final multisector model is the occupational choice model of Banerjee and Newman (1993). Earlier individual-level models of occupational choice (e.g., Boskin, 1974) posited that an individual would choose that occupation that offered the highest present discounted value of lifetime earnings. These micro-level models, however, did not ask where the present discounted values in the different occupations came from. One answer was provided by Banerjee and Newman, who modeled a labor market consisting of four possible occupations: subsistence labor, wage labor, self-employment, and entrepreneurship. They assumed a large number of workers who were identical in their preferences and abilities but who differed in terms of their initial wealth. Because of imperfect capital markets, occupations that required high levels of investment were out of the reach of poor people. At any given time, individuals could invest their time and money to increase their wealth, which they then bequeathed to their children, thereby improving the children’s occupational potential. Over time, the path of economic development and occupational attainment would evolve accordingly.
Note that the Banerjee-Newman model is a market-level model of occupational choice, not merely a micro model. The important feature differentiating this model from micro models such as Boskin’s is that the time paths of economic development and the distribution of individuals across occupations are endogenous - that is, the returns to choosing a given occupation depend on how many others in the economy are choosing that and other occupations. In this respect, the Banerjee-Newman model is like the Harris-Todaro model, in which the return to choosing a given job-search strategy depends on the number of others choosing that and other job-search strategies.

The models just reviewed are distinguished by the existence of multiple employment sectors. Another important group of models focuses on skill differentials among workers. It is to these that we now turn.

**Models with skill differentials**

The best-known model of skill differentials in the labor market is the human capital model. As developed in the Nobel Prize-winning work of Schultz (1961, 1962), Becker (1962, 1964), and Mincer (1962, 1974), human capital theory posits that individuals differ from one another in terms of the amount of skill they possess. Skill was essentially unidimensional in the early human capital models and could be augmented in two ways: through schooling and through training.

One variant of the human capital model (e.g., Johnson, 1997) couples the preceding assumption that education and training create market-useful skills with another assumption: that the returns to skill are determined by supply and demand in what are essentially non-competing occupations. A policy of educational expansion increases the supply of relatively educated labor and decreases the supply of relatively uneducated labor. Given the human capital model’s assumption that wages are set by supply and demand for each skill category, two consequences follow. The first is that all workers are employed in occupations befitting their skills. The second is that educational expansion
lowers the market wage of educated labor and raises the market wage of uneducated labor.

As alternatives to the human capital model, other models of education have been formulated. Signaling models (Spence, 1973) maintain that workers get educated in order to signal to employers that they (the educated workers) are inherently more productive than other workers. Screening models and their close cousin bumping models examine what happens when the educational system certifies which workers are more productive than others and analyze linked labor markets in which educated workers use their education to move to the front of the queue and be hired preferentially for jobs for which education is advantageous but not required (Fields, 1972, 1974; Stiglitz, 1975). What all of these models have in common is that the social returns to education might differ substantially from the private returns, which in turn has important implications for countries’ decisions about the quantity of resources to invest in education (Fields, 1972; Stiglitz, 2002). This issue is so important that I analyze it in depth later in this paper.

Actually, an earlier model of skill differentials had been formulated by Roy (1951). The Roy model had two types of skills, which in his narrative were called hunting and fishing. Some workers were relatively better at one or the other. Workers were assumed to sort themselves between hunting and fishing according to where the payoff to their particular skills was highest. The Roy model had limited influence among economists for quite some time, although its influence later grew, primarily as a guide to empirical research. These contributions are discussed further in the empirical section.

**Modeling social protection**

Theoretical models of social protection abound. For overviews, the interested reader is referred to textbooks in public finance (e.g., Stiglitz, 2000; Rosen, 2005) and in labor economics (e.g., Cahuc and Zylberberg, 2004; Ehrenberg and Smith, 2006) as well as other sources (e.g., Ahmad et al, 1991; Bardhan and Udry, 1999; Holzmann and Jorgensen, 2001; Blanchard, 2004).
To model social protection in a manner consistent with the multisector labor market models discussed above, several features are essential. First, the model must be market-level; models at the level of the individual firm or the individual worker will not do. Second, the model must provide for covered and non-covered sectors; in developing countries, programs that offer universal coverage are rare. Third, for most types of social protection, such as employment security regulations or unemployment insurance benefits, only a stochastic approach will do; only for some types of protection, such as providing old-age pensions or banning child labor, might a deterministic approach be satisfactory. And finally, to be consistent with the five-part framework presented at the outset, the model must be evaluated using a well-specified welfare economic criterion.

From my admittedly less-than-encyclopedic reading of the literature, I cannot find a single instance of a model of social protection in a developing country’s labor market with all four features (is formulated at the market-level, allows for covered and non-covered sectors, incorporates a stochastic element, and is evaluated using a well-specified welfare economic criterion). Absent such a literature, it would be best to formulate policies cautiously.

Finally, it is worth noting that most models of social protection institutions consider single policies. An interesting exception, though, can be found in the work of Blanchard and Tirole (2005), who model the joint design of unemployment insurance and employment protection. More generally, it is to be expected that the optimal design of one institution would be dependent on how other interacting institutions are designed.

**Empirical Evidence**

There exists a huge statistical/econometric literature on labor markets in both developed and developing countries. Highlights are surveyed in various articles in the three-volume edition of the Handbook of Labor Economics (Ashenfelter and Card, 1999), among other places.
On the empirical side, six kinds of data analysis are potentially useful. They are aggregate cross-sectional quantitative data, micro data from comparable cross sections, panel data, cross-country time series data, experimental data, and qualitative data.

This review of the empirical evidence begins by presenting examples of empirical research using each of these types of data that led to better understanding of labor market functioning and/or interesting policy conclusions in developing countries. It then continues to address what we have learned about finding the right labor market model to fit a particular country’s circumstances. As an example of good policy-relevant empirical work, the research that has been conducted on the effects of India’s employment protection laws is featured. The section concludes by considering some of the limitations of existing policy research in the area of the economics of education.

**Empirical research on labor markets and labor market policies: six kinds of data**

In this section, I present examples of studies that have used each kind of data. Before beginning, two lessons presented by Ravallion (2007) should be mentioned. The first is that no single evaluation tool can claim to be ideal in all circumstances. The second is that the standard tools of counterfactual analysis for mean impacts have some potentially severe limitations.

1. **Aggregate cross-sectional quantitative data.** The analysis of aggregate data is a useful and often-neglected starting point for policy analysis, permitting answers to questions such as the following. Do workers participate proportionally in economic growth? More generally, how do labor market conditions change during times of growth and decline? How do the outcomes in one country compare with those in another? How do changes in labor market outcomes for one population sub-group compare with those of another? Is economic growth jobless, and if so, under what circumstances?

There is a long history of compiling and analyzing such aggregate data. A just-completed World Bank project has put forward a list of recommended indicators, which
were piloted in eight countries. The interested reader is referred to World Bank (2007a) for further information.

The general result that has come out of such aggregate data analysis for a limited number of countries is that economic growth has generally led to improved conditions in labor markets as gauged by such measures as the rate of unemployment, composition of employment, real earnings levels, inequality, and poverty. Conversely, when such labor market aggregates have not improved, typically it is because economic growth has not taken place.

The policy conclusion that comes out of this line of research is that economic growth should be sought as a way of improving labor market conditions, thereby enabling workers in general and the poor in particular to attain higher material standards of living. What we do not yet know is under what circumstances the improvements in labor market conditions are greater than in others.

2. Micro cross-sectional quantitative data. The literature analyzing micro data enables researchers to find out what makes a difference at the level of the individual worker or the individual household. Such data analysis forms the basis for most contemporary research studies on labor markets. Accordingly, the literature using such data is absolutely huge.

The question these studies typically answer is, which individuals or which firms have higher levels of the dependent variable Y? More specifically, the regression coefficients can be understood thus: when we compare those individuals or firms for which an independent variable X is one unit higher, on average how much higher is Y?

Quite often, though, the results of such regressions are used to draw “policy implications” which, in my view, are not implications at all. Below, I offer a detailed explanation as to why. For now, though, let me just say that it is common in the literature to deal with two problems – bias in the estimated coefficients and correlation but not causation – with one solution, namely, instrumental variables estimation. Deaton (1997)
has discussed the option of instrumenting for right hand side variables, presenting not only the advantages of instrumental variables methods but also their limitations. The limitations have been taken up further in Bound, Jaeger, and Baker (1995) and Stock and Staiger (1997). As Bound, Jaeger, and Baker put it so graphically in the title of the working paper version of their paper, “the cure can be worse than the disease.” That observation does not, however, stop labor economists from instrumenting at virtually every opportunity.

Notwithstanding the possible misuses of micro data regressions, it would be foolhardy in the extreme to dismiss them. Such regression coefficients are very useful information to have. My point is that we need to be suitably modest in drawing inferences from them.

3. **Panel data.** Panel data analysis is a relatively new and promising approach in developing countries. Panel studies follow the same individuals or families over time. Such studies enable development processes to be analyzed in ways that analysis of comparable cross-sections cannot. Two areas in which a great deal of policy-relevant knowledge has been gained are the effects of policy interventions on treated individuals and the determinants of income and earnings mobility.

Studies of policy interventions estimate such measures as the average treatment effect, the average treatment effect for those treated, and the distribution of treatment effects for different individuals. Heckman (2001) and Ravallion (2007) present comprehensive overviews of the relevant issues. See also Cunha, Heckman, and Navarro (2006).

Empirical research of this type has proved informative in the context of developing countries. One example is the study by Kugler (2005) of a policy change in Colombia. In 1990, Colombia replaced its earlier system of severance payments with a new system of severance payments savings accounts (SPSAs). Under the new system, employers were required to pay severance at the time of separation, to be funded by
employers depositing a percentage of wages into guaranteed individual accounts. Kugler’s question was whether part of this obligation was shifted to workers in the form of lower wages. Using a difference-in-difference-in-difference estimator, she concluded that 60-80% of the total SPSA contributions were shifted to workers in the form of lower wages, thus disproving the view held by some that such mandates would be “free” to the workers.

Another interesting example of using panel data to estimate the effects of a labor market policy change is the study by Ravallion et al. (2005) of Argentina’s Trabajar Program, a kind of workfare. The program provides short-term work at relatively low wages to retrenched workers. Workers who are recruited to approved projects may work for up to six months in Trabajar jobs. Using a difference-in-difference-in-difference estimator, the authors concluded that workers suffered large earnings losses from retrenchment and experienced sizeable income gains from participating in Trabajar, while those who left the program lost considerable amounts relative to those who stayed.

Another subject that has been explored fruitfully using panel data in developing countries is income and earnings mobility. (“Income” denotes income from all sources, while “earnings” denotes income from wage employment, salaried employment, and self-employment only.) One question that has been asked is, how have economic growth and economic decline affected income or earnings changes for different groups in the population: men vs. women, better-educated vs. less-educated, initially high earners vs. initially lower earners, and so on? Fields et al. (2007) studied this question for Argentina, Mexico, and Venezuela. They found that notwithstanding the high and generally rising earnings inequality in all three countries, both in times of economic growth and in times of economic decline, the largest average earnings gains in currency units were experienced by those who were initially at the bottom end of the earnings distribution. They also found that a key explanatory variable was the change in employment status between formally employed, informally employed, and unemployed. An ongoing World
Bank – Inter-American Development Bank project (Duryea et al., 2007) has looked further into employment transitions and earnings changes, not only in these three Latin American economies but also in six transition economies. Their results reveal considerable heterogeneity within the informality and self-employment sectors. The Latin American literature is reviewed more fully in Fields et al. (forthcoming).

In addition to the work done on earnings mobility, there also is a literature on income mobility in developing countries. For a sampling of this literature, the interested reader may consult the studies on South Africa, Indonesia, Venezuela, and Spain by Fields and associates (2003) and on China by Jalan and Ravallion (2000) as well as the work by the Chronic Poverty Research Centre (www.chronicpoverty.org) on countries around the world, especially their flagship publication (Chronic Poverty Research Centre, 2004).

4. Cross country time series data. The simplest kind of time series data is that which follows given variables within a country over time. Time series data allow for the determinants of changes over time to be modeled explicitly. For example, the wage elasticity of formal sector labor demand in South Africa was estimated using such data; Fields, Leibbrandt, and Wakeford (2000) present their own and earlier estimates. Despite differences in time periods and estimation methods, these studies concurred that the range of estimates was on the order of -0.4 to -0.6 – certainly not zero. This finding led policy-makers in that country to recognize that the push by the Congress of South African Trade Unions for higher wages would likely result in reduced formal sector employment.

Another useful kind of time series data is collected for a number of countries for a number of years, producing what is called cross country time series data. An insightful use of such data has been made by Rama (2003a) to address the question of whether greater integration with world markets is a source of prosperity for developing and transition economies or a source of increased hardship. Since both views are plausible theoretically and rigorous models have been devised to support them, it becomes an
empirical question as to what matters in what direction. The data set used by Rama was information on the wage in a particular occupation in a particular year for each of a large number of countries, from which wage levels and wage inequality were constructed as dependent variables. Explanatory variables included measures of economic openness, labor market policies, and social protection policies. The research produced three broad conclusions. First, the effects of globalization on labor market outcomes are neither always positive nor always negative. Second, social protection programs are effective in reducing wage inequality, but minimum wages, public sector employment, and core labor standards are not. And third, social protection programs do not adversely affect efficiency, but high public sector employment and trade union membership are associated with weaker performance in the context of adjustment. Overall then, Rama concludes: “Does globalization affect labor market outcomes? The answer is yes, but in uneven ways.”

The cross country time series methodology has been used for a cross-regional analysis of labor market policies in India. This work is reviewed in detail later in this section.

5. Experimental data.

A new development in development economics is the rise of randomized field experiments to find out what works in practice. Duflo (2006) provides an excellent overview of this literature for development economics in general. Here, I illustrate what has been learned by describing three labor economics applications of randomized field experiments.

The first was aimed at addressing a serious problem in India, that of teacher absenteeism (Chaudhury et al., 2005). It is common for salaried teachers to simply not show up, often for days, weeks, or even months at a time. Working in conjunction with Seva Mandir, an NGO in Udaipur district in India, Duflo and colleagues selected 120 single-teacher schools to participate in a teacher monitoring program. Half of these
schools were chosen randomly as treatment schools, the other half as control schools. In
the treatment schools, the teacher was given a camera with a tamper-proof date and time
function and told to take a picture of him/herself and the students at the beginning and
end of each school day. Teachers were offered a bonus based on the number of days they
actually came to school. The evaluation results (Duflo and Hanna, 2005) showed that the
absence rate in the treatment schools (22%) was half of that in the control schools (42%).

A second example of randomized treatment in the labor market is the Colombian
workfare program Empleo en Acción (Employment in Action). Within each community,
the local authorities were directed to choose randomly which applicants would participate
in the program and which would not. In an evaluation of this program, Attanasio, Meghir
and Vera-Hernandez (2007) studied three issues: 1) whether or not the program crowds
out labor effort by members of the household other than the participant, 2) whether or not
there are gains from participating in the program six months after the program has
finished, and 3) whether or not there are gains in household consumption, so as to assess
the role of the program as an insurance mechanism. The authors found that the program
positively influenced number of hours worked and both individual and household labor
income. Household income increased more than individual’s income, indicating that the
program had some positive externality on the other members of the household.

A third application of these methods is in the evaluation of job training. One
study, by Card et al (2007), presents the results of the first impact evaluation based on an
experimental design for a job training program in Latin America. Between 2001 and
2005, the government of the Dominican Republic implemented a job training program for
low income youth population (Juventud y Empleo) consisting of classroom training
followed by an internships in a private sector firm. The eligible population was randomly
selected to training, and information was gathered 10-14 months after graduation for both
trainees and members of the control group. Training was found not to have a significant
impact on the likelihood of having a job. However, a marginally significantly impact was
detected for wages and also for having health insurance coverage conditional on being employed. The results suggest significant heterogeneity of impacts, with male teens being the group that benefits from the program; impacts were not found for women or for young adults. Regional differences also seem to be present. Although small, the impact on wages (if maintained over time) coupled with no discernible employment effect implies that the costs of the program are recovered in two years.

Another randomized evaluation of a training program is reported by Attanasio, Kugler and Meghir (2007) for Colombia. Contrary to Card et al’s findings for the Dominican Republic, this study found that job training in Colombia had large, widespread effects on women, but fewer and less pronounced effects on men.

These studies illustrate the power of randomized experimental data of labor market programs. Undoubtedly more such evaluations will be conducted in the future.

6. Qualitative data. Analysis of qualitative data has a role to play in informing us about labor market issues. Kanbur (2003) has identified five dimensions in which qualitative and quantitative approaches differ from one another:

i. Type of information on population: Non-numerical to numerical.

ii. Type of population coverage: Specific to general.

iii. Type of population involvement: Active to passive.

iv. Type of inference methodology: Inductive to deductive.

v. Type of disciplinary framework: Broad social sciences to neoclassical economics.

An excellent example of qualitative data analysis is the World Bank’s Voices of the Poor study. Sixty thousand poor women and men told their stories, which were then compiled. For purposes of the present paper, what stands out is that the number one problem expressed by the poor is the lack of income-earning opportunities. The poor themselves see clearly that the labor market plays a critical role in improving their livelihoods.
Another nice example is Pierre and Scarpetta’s (2004) study of employment regulations as viewed by employers. The authors found that not all firms are affected in the same way by onerous labor regulations and they respond differently to them. In particular, when labor regulations made hiring and firing very costly, small firms reported that they relied mainly on temporary employment, while medium and large firms relied more on on-the-job training.

And then there is the famous work conducted in the Indian village of Palanpur over the last quarter-century, which demonstrated among other things the crucial role played by non-farm employment in reducing rural poverty. See, for example, Bliss and Stern (1982), Drèze, Lanjouw, and Stern (1992), and Lanjouw and Stern (2003).

In my judgment, qualitative data analysis can supplement but in no way should supplant quantitative data analysis. The best labor market work will involve a judicious mix of qualitative and quantitative research.

**Lessons learned: finding the right model**

Where empirical labor market policy research has proven most useful is in telling us what the right model of labor markets is in a particular country context. Here are some examples where 1) important lessons have been learned, but 2) more remains to be learned.

**Are labor markets segmented?** The theoretical section of this paper maintained that developing country labor markets are segmented and that researchers therefore need to understand the workings of each segment and the linkages between segments. For labor market segmentation to exist, two conditions must hold: 1) Labor market returns (including earnings but also non-pecuniary aspects of jobs) differ across segments for identical workers, and 2) Employment in the better jobs segments is rationed.

How strong is the empirical evidence on the existence of labor market segmentation? In my view, the evidence is strong but not unassailable. Earnings differentials for observationally equivalent workers have been demonstrated repeatedly in
a great many countries. Examples are Mwabu and Schultz (2000), who find a large union wage premium in South Africa, and Assaad (1997) and Rama (2003b), who find for Egypt and Sri Lanka respectively that public sector jobs are the better ones, leading to queueing for them. The effects of minimum wages in developing countries are surveyed in Cunningham (2007), Svejnar (2007), and Freeman (2007).

Critics such as Rosenzweig (1988) are unpersuaded. He points to the possibility that such earnings differences could be due to compensating differentials, rewards for unmeasured skills, and/or compensation for differences in the disutility of the workplace. To the extent that these objections can be examined empirically, it would be useful to do so, paying careful attention to the methods used in such studies as Heckman and Hotz (1986) for Panama, Magnac (1991) for Colombia, and Pratap and Quintin (2006) for Argentina.

Why do people work informally?

The early view of why people work informally was that no other work was available to them. The informal sector was equated with the “bad jobs” part of the economy. Such a view is maintained in many quarters at present, for example, by the ILO (2002a), which maintains that those working informally are part of a “decent work deficit” which must be remedied.

Other research, initially qualitative in nature, maintained that there exists a fundamental duality within the informal sector. On the one hand, there is a segment corresponding to the traditional view: people work informally because they can get no other work. On the other hand, there is also a segment in which people work by choice in preference to formal employment. These have been labeled the “free entry informal sector” and the “upper tier informal sector” respectively (Fields, 1990). Evidence in support of the view that the informal sector has a fundamental duality within it is offered by Gunther and Launov (2007) for Côte d’Ivoire. According to their estimates, 34% of urban workers are working formally, but 61% would like to be working formally. On the
other hand, they estimate that 28% of urban workers are working in the lower tier of the informal sector, whereas only 9% would like to be.

Maloney (2003, 2004, other) has gone a step further, claiming that most people working informally in Mexico are doing so by choice. Navarro and Schrimpf (2004) also conclude that barriers to entry into formality are unimportant in Mexico. On the other hand, Duval Hernández (2006) maintains the more traditional view: that most of those working informally in Mexico are doing so because they are unable to work formally.

Other studies have been conducted investigating the voluntariness of self-employment and informal wage employment. Among them are Gasparini and Tornarolli (2007) and Perry et al. (2007) for Latin America as a whole and Arias and Khamis (2007) and Packard (2007) for Chile. What we know from this literature is that significant numbers among the self-employed and the informally-employed are in these states by choice; what we do not know is how many.

An important question to be asked about these studies is, if self-employment is voluntary, relative to what is it voluntary? Relative to informal wage employment? Relative to formal wage employment? In my view, none of these studies shows convincingly that the self-employed could have been formally employed but chose not to be. The same holds for informal wage employees.

This is another area where the state of knowledge is deeply unsettled. Two fundamental questions remain to be answered. First, how convincing is the evidence for the countries studied? And second, to what extent do the results generalize to other countries? An in-depth assessment remains to be performed.

Why is the unemployment rate so low in most developing countries but so high in a few?

By the standard ILO definition, a person is said to be unemployed if he or she 1) did not work even one hour for pay or fifteen hours not for pay in the preceding work and 2) was actively looking for work. Early studies of developing countries by Turnham
(1971) and Squire (1981) reported unemployment rates on the order of 5-10% in most of them. Even today, the ILO (2007) reports that just 15% of low earners in the world are unemployed; the other 85% are employed but fall below the low-earnings line.

A generally-accepted explanation for why the unemployment rate is so low in most developing countries is the so-called “luxury unemployment hypothesis.” On this view, only supplementary earners in well-to-do families can afford to be doing no work whatsoever for very long. On the other hand, potential workers in poor families are compelled by economic necessity to take up whatever form of wage employment or self-employment they can get.

The preceding explanation is so widely accepted that when high unemployment is found in a developing country, researchers are sent scrambling. Such is the case in South Africa. Using the ILO’s definition, South Africa’s unemployment rate is 26%. The ILO definition, however, excludes those people who were not working, were not actively looking for work, but say they would take a job if one were offered. Using this so-called “broad definition of unemployment,” South African unemployment reaches 41%, clearly a crisis level for the society and for the government.

What is going on? Either the data are wrong or the theoretical models are wrong. The data could be wrong, because South Africans are working by the ILO definition but not according to their own self-conception. Alternatively, the data could be right and the models wrong. A leading South African labor economist, Haroon Bhorat, calls this “the international visitor disbelief” position. Kingdon and Knight (2004) and Cichello et al. (2006), among others, have analyzed the barriers to self-employment in South Africa and

---

4 On a recent visit to South Africa, I talked to many people about their employment and earnings statuses and heard similar stories from many of them. One was from a middle-aged man who said both he and his wife were unemployed. Given that they had no children or elderly in their family, they received no transfer payments from the government. If they weren’t working and not receiving government transfers, how could they be living? Bit by bit, we learned that he received money for leading the local church and doing traditional healing and that his wife received money for selling handicrafts at the city market. Yet, he said, they were unemployed. What would it take for you to be employed, we asked? His answer: You are employed when you get paid every Friday by the same employer.
found significant costs to entering the informal sector. Although costly entry exists in South African reality, it does not exist in the labor market models that we take to South Africa and other developing countries.

The challenge, then, is to make the theoretical models and the empirical findings accord with one another. This task too awaits future research.

An example of good empirical policy analysis in practice: assessing the effects of India’s employment protection laws

At its best, labor market policy research has taught us about some of the labor market and welfare effects of particular policies. A particularly insightful example is the research that has been conducted on the consequences of employment protection laws in India.

India’s Industrial Disputes Act of 1947, as amended in 1976 and 1982, requires that firms employing 100 or more workers obtain government permission for retrenchments, layoffs, or plant closures. In practice, such permission is rarely given. The research question is then whether this well-intentioned legislation has in fact had positive effects on the workers it was designed to protect.

Economic theory is helpful in guiding our expectations. Many observers have noted that as a consequence of the legislation, large firms might be reluctant to employ workers, lest they be unable to dismiss them in bad times. Other effects have also been noted including a relative price effect, an expropriation effect, a rigidity effect, an employment effect, a labor substitution effect, and an industrial disputes effect (Besley and Burgess, 2004; Anant et al, 2006; Ahsan and Pagés, 2007).

Initial research by Fallon and Lucas (1993) concluded that as the employment protection laws were strengthened, long-run demand for employees in India fell by 17.5%. Their research has been criticized, however, for being a before and after study and not holding constant other determinants of employment.
An important aspect of the Industrial Disputes Act is that it grants to the states the right to amend the central legislation. Some 113 amendments have been enacted by the states. This variation in employment protection laws across the states serves as the basis for recent research on the effects of the laws, most importantly, the work of Besley and Burgess (2004) and Ahsan and Pagés (2007).

Besley and Burgess assessed the act by using sophisticated econometric methods to estimate a number of effects. They found that output, employment, investment, and productivity in formal manufacturing firms were lower in states that amended the Industrial Disputes Act in a pro-worker direction compared with other states. They also observed that output in informal manufacturing increased in those states, implying an informalization of the economy because of the legislation. Lastly, they estimated that urban poverty was higher in those states that had enacted pro-worker legislation. According to their calculations, urban poverty would have been 11% lower in West Bengal had it not enacted substantial pro-worker legislation, while Andhra Pradesh would have had 12% higher urban poverty had it not moved in a pro-employer direction. In their words (p. 124): “The battle cry of labor market regulation is often that pro-worker labor market policies redress the unfavorable balance of power between capital and labor, leading to a progressive effect on income distribution. We find no evidence of this here – indeed the distributional effects appear to have worked against the poor.”

Despite the Besley-Burgess study’s clever design, empirical sophistication, and use of explicit social welfare criteria, it has nonetheless been subjected to a variety of criticisms. A number of these came from Bhattacharjea (2006), who raised points relating to the scoring of individual measures, the aggregation of measures, the omission of certain laws, and so on. These objections caused Bhattacharjea to take an “agnostic” position on the merits and demerits of the act. Anant et al. (2006) have raised other objections: that reading off directly from legal statutes to measure rigidities could be misleading; that the linkages between labor regulations and industrial outcomes might be
weak; and that the lack of flexibility in organized employment is attributable only partially to labor laws. They noted as well that the data for the Besley-Burgess study ended just as India’s economic reforms took effect. But after stating these criticisms, they did nothing to improve upon the Besley-Burgess calculations.

A subsequent study that did in fact present new estimates is the paper by Ahsan and Pagés (2007). Building on the Besley-Burgess methods, Ahsan and Pagés distinguished between amendments concerning laws that affect firms’ flexibility in hiring and firing labor from those concerning the procedures and machinery for resolving labor disputes, disaggregated at the two-digit level, and assessed the effect of the increasingly widespread use of contract labor. Like Besley and Burgess, Ahsan and Pagés found that de jure legal amendments and de facto use of contract labor hurt formal sector workers and impose substantial costs on society. In their words: “India must find alternative way (sic) to improve labor conditions for the majority of workers. While traditionally portrayed as labor advances against the abuse of capital, current labor regulations favor no one. The answer to this conundrum does not involve a complete deregulation of the labor market. Instead, it requires better regulations (and appropriate enforcement) so workers’ fundamental rights can be protected and jobs can be created.”

This literature is a nice example of rigorous empirical study combined with sensitive welfare economics. Perhaps the one limitation of this work is that while the studies’ authors have told readers what they think India should not have done, they have not told readers what they think India should now do in this regard.

**Limitations of existing empirical research: the case of social returns to education**

As an indication of the needs in this area, it will be helpful to consider one particular policy decision: whether to invest additional social resources in education. The conclusion I shall reach is that despite the vast amount of research effort that has been
devoted to this area, we actually know quite little about what should be done in a given country context.

**Earnings functions as the analytical starting point**

The starting points for policy research pertinent to the question of whether to invest additional resources in education are the related literatures on earnings functions (also called wage equations) and rates of return to education. As reviewed by Willis (1986) and World Bank (1995), the earnings function studies show similar patterns of earnings differentials in various countries. Other things equal, earnings generally are higher for those workers with more schooling and more labor market experience, for men compared with women, for formal sector workers compared with informal sector workers, etc. Focusing on the schooling variable, George Psacharopoulos is well-known for a long series of studies summarizing studies of returns to education. His most recent compilation of evidence (Psacharopoulos and Patrinos, 2004) reaffirms three of his earlier findings about returns to education as conventionally measured: they are higher in developed than in developing countries; they are highest for primary education, next highest for secondary education, and lowest for higher education; and the private rates of return exceed social rates of return.

**Guidance from welfare economics**

These findings are **not** what we need for policy analysis. A fundamental postulate of public economics is that policy decisions should be made by comparing the marginal social benefits of a policy action with the marginal social costs. The social benefits include the gains to all members of society from the proposed action; the social costs include the costs paid by all members of society. As with all economic decisions, resources should be allocated on the basis of marginal benefits and marginal costs.

For policy purposes, two types of knowledge are needed. First, how would market conditions change if a particular action were to be undertaken? Second, what are the
marginal social benefits of the proposed action, what are the marginal social costs, and
how do the marginal social benefits and marginal social costs compare?

For purposes of answering such questions, the usual regressions are helpful but
insufficient in a number of respects. The next subsections explain why, continuing with
the illustrative case of the social returns to education.

From correlations to causations.

As noted above, the most common type of empirical study in labor economics is a
regression based on micro-data. Looking across individuals or firms, the dependent
variable Y is regressed on one or more explanatory variables X.

One limitation of regressions of this type is that while they are correlative, they
are not necessarily causal. I am as guilty of this as the next person, having asserted (but
not convincingly demonstrated) that the coefficients obtained from earnings functions can
be interpreted causally and not merely correlatively (Fields, 1980a).

In recent years, leading economists including David Card, François Bourguignon,
and James Heckman have built sophisticated econometric models of the causal effects of
education; see respectively Card (1999), Bourguignon and Ferreira (2003), Heckman
(2001), and Cunha, Heckman, and Navarro (2006). The question asked in each of these
papers is, what would be the earnings gains from acquiring additional schooling? In this
literature, the earnings gain is called the treatment effect, and various measures of
treatment effect are used including the average treatment effect, the average treatment
effect for those treated, and the distribution of treatment effects for different individuals.\(^5\)

There have been three generations of results in this literature. In the first
generation, earnings were related to education using ordinary least squares. In the second,
the estimates were adjusted for omitted factors such as ability. In the third, the estimates

\(^5\) What are here called “treatment effects” are the essence of what the World Bank calls Development
Impact Evaluations (DIMEs). The Bank’s website says, “The central impact evaluation question is what
would have happened to those receiving the intervention if they had not in fact received the program.” For
more on impact evaluations, see World Bank (2007).
were further adjusted for variations in returns to schooling, assuming that those individuals who chose to get additional education were the ones that had the most to gain by doing so.

Even the third generation results, as sophisticated as they are, do not necessarily provide the needed information. What they answer is a micro-level question: if individual i were to get additional education, by how much would i’s earnings increase? What they do not answer is a market-level question: if sizeable numbers of individuals were to get additional education, how would the labor market as a whole respond?

Let us now consider the market-level issue.

**Assessing the social benefits of education.**

In the theoretical section of this paper, I described a number of multisector models. One was the human capital model, in which it was assumed that people with more (less) human capital would all work in the more (less) skilled jobs. A second was the Roy model, in which it was assumed that there are two (or more) sectors and that all individuals are free to enter whichever sector would be the most lucrative for them, given their particular skills. And a third was the segmented labor market model, in which for a worker of a given skill level, there are good jobs and bad jobs, with the good jobs being limited in number.

Empirical researchers, even the most sophisticated ones, have typically assumed that one or another of these multisector models holds without testing if they hold. Heckman and Honoré (1990) and Cunha, Heckman, and Navarro (2006) cite a substantial number of papers that take the Roy model as a maintained structure; these authors also note that strong assumptions such as the normality of the skill distribution must be made in order for the model to be identified statistically. In the developing country context, the Measuring Income Distribution Dynamics project conducted under the leadership of Bourguignon and Ferreira (2004) is but one example of a set of studies in which the Roy model is maintained for developing countries. More generally, models in which workers
are assumed to be able to enter whichever economic sector, occupation, or employment status they wish predominate in the literature on labor markets in both developed and developing countries, and for this reason such models are called sector choice equations, occupational choice equations, and labor supply equations respectively.

From the empirical findings throughout the world that have shown conclusively that individuals with more education earn more in the labor market, it is very frequently argued that society should invest in more education. As I shall now argue, while there is no question that additional education produces benefits for the individuals who receive the education, the case for additional education producing social benefits for all members of society including those who do not get the education is much less compelling. In this subsection, I explain these doubts by considering two alternative underlying models that yield very different conclusions about the social benefits of education.

Take first the familiar human capital model (Becker, 1964; Mincer, 1974). The model makes a number of assumptions, usually implicitly: that educated workers are more productive than less educated workers; that the difference in productivity can be approximated by the difference in earnings; that better-educated workers earn more because of the additional education they have received; that there is full employment; and that if one more worker is educated, there will be one more educated worker employed in the educated workers’ labor market and one less less-educated worker employed in the less-educated workers’ labor market. The human capital model has been applied in general equilibrium contexts by, among others, Katz and Murphy (1992) and Heckman, Lochner, and Taber (1999).

An alternative to the human capital model is the screening model (Fields, 1972, 1974; Stiglitz, 1975, 2002). In screening models, the assumptions are different from those in the human capital model: the wage is set according to the job, employers seek to fill each job with the most productive worker available, and workers with more education are on average more productive than workers with less education. In such circumstances,
employers use education as a screening device, because they end up with workers who are more productive on average as a result.

A special version of the screening model is the job assignment model which Sattinger (1993) graphically termed “the dog-bone economy.” In his particular version of the screening model, jobs differ in terms of their quality according to a fixed system of rewards. Thus, the best job goes to the most highly-qualified applicant, the second-best job to the second-best applicant, and so on – by analogy, the biggest juiciest bone goes to the fiercest, most powerful dog, the second-best bone to the second-strongest dog, and so on. Just as the weakest dogs may end up with no bone, in a dog-bone economy, the least-qualified workers may end up with no job.

Screening models assume that educated workers are more productive but take no strong position as to why educated workers are more productive. This could be for any of a number of reasons or a combination of them: human capital is formed in schools; schools selectively admit the best students, who are likely to be the best workers; and workers get additional education to signal their superior ability. (This last is the foundation for the Nobel Prize-winning model of Spence, 1973.)

Consider now how the human capital model and the screening model differ in terms of the social return to additional education. The marginal social returns to education must be evaluated at the market level, not the individual level. (In both models, the individual who gets additional education will enjoy a private benefit.) Under the assumptions of the human capital model, if individual i gets more education, the effects are the following. Society will have one more employed better-educated person and one less employed less-educated person. Individual i will be more productive because of the additional education received. The amount of the productivity gain to society is the difference between the earnings of those with and without the education level in question. Given the preceding, the marginal social return to education is closely approximated by the average social return.
Contrast these effects under the human capital model with the effects of additional education under the screening model. In the screening model, the availability of one additional educated person results in one more job being filled by a better-educated person rather than a less-educated person, the reason being that employers judge that the better-educated individuals will be more productive than the less-educated ones and so hire the better-educated preferentially. However, there is no assumption now that the difference in productivity between better-educated and less-educated workers can be approximated by the difference in average earnings or that the difference in productivity is caused by human capital produced in schools rather than by pre-labor-market differences in ability. The assignment model goes one step further: in that version of the screening model, there is no productivity gain whatsoever: the now-stronger worker gets the job which somebody else does not now get. In a dog-bone economy, the distribution of bones does not change; all that changes is the distribution of dogs among bones.

Various authors have cautioned that the market-level effects of large interventions may not be even closely approximated by the effects of interventions on single individuals (Fields, 1972; Srinivasan, 1982; Heckman, Lochner, and Taber, 1999; Basu, 2005; Ravallion, 2007). In Heckman et al.’s empirical work, the general equilibrium impacts of tuition on college enrollment are an order of magnitude smaller than those reported in the literature on individual treatment effects. The easier answer may not be even approximately right.

Which model is right for China, Chile, or Chad? The answer cannot be found by running earnings functions across individuals. Instead, the answer must be found by market-level tests of the assumptions. The question then is, which set of assumptions – those of the human capital model or those of the screening model or yet some other model – best approximate conditions in the country in question?

In my view, the conditions posited by the screening model appear at least as relevant for developing economies as those posited by the human capital model, if not
more so. So called “crowding out” of less-educated workers by the better-educated has been demonstrated for Africa (Knight, Sabot, and Hovey, 1992; Bennell, 1996). Bennell’s evidence, discussed further by Pritchett (2001), shows that wage employment grew by nearly the same number of people as the change in school enrollments in just two countries (Botswana and Zimbabwe). In other countries studied, the number of newly-educated individuals entering the labor force was four times as large as the growth of wage employment in Senegal, Kenya, and Malawi, and between ten and twenty-nine times larger in Lesotho, Burkina Faso, Ghana, Uganda, and Sierra Leone. Consequently, I feel that the marginal social benefits of education may be greatly overstated by the standard methods, at least in these particular African countries.

**Taking due account of social costs**

Another limitation of standard empirical studies is that social costs are rarely given the same attention that social benefits are. The social costs of any project include the direct costs of the project plus the opportunity costs. In the case of educational projects, the direct social costs are the costs to society of providing the education, which in turn equals the direct costs paid by the student plus the subsidies provided by the government or others (e.g., the university or its alumni). The social opportunity costs consist of the value of the output that is not produced because some individuals are removed from the labor force while they are in school.

As discussed above, the social rate of return to education is found by equating the marginal social benefits of education to the marginal social costs of education and solving for r; see (5). Similarly, the net present value of the project under evaluation is found by specifying the social discount rate, estimating the present values of benefits and costs using this discount rate, and then calculating the net present value; see (6).

Following upon a long tradition originated by Mincer (1962, 1974) and continued by Psacharopoulos (1973, other), it is common to use the term “rate of return” to refer to the coefficient on schooling in a Mincer-type earnings function of the form
(7) \[ \ln Y = a + bS + c\text{EXP} + d\text{EXP}^2, \]
where \( S \) is the individual’s schooling and \( \text{EXP} \) is the post-schooling experience. Let us call the regression coefficient \( b \) in (7) the “Mincer rate of return.”

Social decisions about resource allocation should be based on the internal rate of return, or better yet, the net present value, and not on the Mincer rate of return. As noted by Mincer himself, the Mincer rate of return equals the internal rate of return if and only if the only costs of education are opportunity costs. Direct costs are simply assumed away in the Mincer rate of return calculation. But in reality, there are always direct social costs of education, even if education is entirely “free” to the student. Given that these costs are not included in the Mincer rate of return, the Mincer rate of return is not a suitable basis for social decision-making. Unfortunately, analysts often think they are estimating an internal rate of return when in fact what they have gotten is a Mincer rate of return.

The social costs of investments in education and other public projects need to be factored into social decision-making. Often, they are not. The policy suggestions emanating from such research are less than believable as a result.

**A cautionary note**

Perhaps the best statement I have seen of the limitations in going from microeconometric results to economy-level policy conclusions is that made by James Heckman at the conclusion of his Nobel address (Heckman, 2001):

Important challenges to the field include the development of a microeconomic data-based general equilibrium theory for testing theory and evaluating the impacts of large-scale policies. They also include the development of empirically credible econometric cost benefit schemes for the valuation of micro policies that link the program evaluation literature more closely to economics. I am sure that microeconometricians will rise to these and other challenges and in future years will give you updates on research in this field from this podium.

I will be interested to see what the future leaders in this field will have to teach us.
Conclusions and Needs for Future Research

Main messages on how to conduct labor market policy research

This paper has presented a selective overview of the literature on modeling labor market policies in developing countries. The paper had three major sections: welfare economics, theoretical models, and empirical evidence. The paper was organized in this way in order to highlight the three general features needed in future research on labor market policy in developing countries. One is to specify explicitly the welfare economic criteria by which labor market policy judgments are being made. The second is for fuller, more comprehensive theoretical models of how developing countries’ labor markets actually function. And the third is for statistical and econometric studies that a) are guided by and b) guide market-level models.

The following are the main points reached on welfare economics:

In labor market policy work, what is being maximized or minimized needs to be made explicit. However, it is seldom the case that policy studies articulate a well-defined maximand or minimand.

One decision rule is the Pareto criterion. The Pareto criterion is conceptually appealing but of limited practical applicability.

Some of the practical difficulties with the Pareto criterion may be overcome by using first order dominance or an abbreviated labor market well-being function. These criteria may produce ambiguous results and therefore be indecisive.

For those analysts who wish to focus on the level of employment and the earnings of those employed, it may be appealing to minimize a low-earnings index. The analysis may be broadened to include social protection and labor rights, in which case a policy may be judged to be socially worthwhile if employment, earnings of those employed, social protection, and adherence to core labor standards all increase.

Finally, on welfare economics, traditional social cost-benefit analysis merits more attention than it currently gets.
Turning now to **theoretical models**, good policy work requires sound theoretical foundations. Developing countries’ labor markets are marked by distinct labor market sectors that work in different ways from one another and by complicated interrelationships among the sectors. Unfortunately, few existing labor market models begin to capture the rich empirical reality of developing countries’ labor market conditions, and no existing model captures them all.\(^6\)

In order for improved labor market policy analysis to be undertaken, policymakers and advisors need to work with models that contain enough sectors, provide realistic stylizations of the workings of the labor markets in each sector, and contain reasonable specifications of the linkages between sectors in the country in question. This is a demanding set of requirements, requiring more time than is available for many policy decisions. When time is of the essence, as it frequently is,\(^7\) I would say that it would be better to base policies on a less rigorous model with approximately the right features than on an explicit model with demonstrably wrong features.

Finally, **empirical evidence** was reviewed, producing the following highlights:

Useful insights have been obtained from analysis of six kinds of data: aggregate cross-sectional quantitative data, micro cross-sectional quantitative data, panel data, cross country time series data, experimental data, and qualitative data.

---

\(^6\) In commenting on the Deaton et al. (2006) report, former World Bank Chief Economist Nicholas Stern wrote: “The emphasis in DEC should remain in economics but too often it is of a rather narrow kind, based on the standard courses in a U.S. graduate school, with its standard models oriented largely towards rich countries, perfect markets, and embodying economic agents and institutions which show little complexity or subtlety in how they function. To overcome this, still stronger emphasis on collaboration with developing country institutions is necessary” (Stern, 2006). Stern’s predecessor, Joseph Stiglitz, made the same point thus: “Because the basic models are often so inappropriate, there is a great need for the Bank to sponsor more fundamental research into the kinds of models that are more appropriate for developing countries” (Stiglitz, 2006b).

\(^7\) And is it especially is for World Bank operations officials.
Among the questions that have been asked and partially answered are: Are labor markets segmented? Why do people work informally? Why is the unemployment rate so low in most developing countries but so high in a few?

One area in which good empirical policy analysis has been conducted is an assessment of India’s employment protection laws. Recent work is marked by careful attention to welfare economic concerns and sophisticated econometric modeling.

On the other hand, an area in which many policy conclusions have been drawn, in my view unconvincingly, has been policies with respect to investments in education. In education as in many other areas, the existing studies do not provide the information that is needed for public policy: knowledge of how market conditions would change if a particular action were to be undertaken along with knowledge of how the marginal social benefits compare with the marginal social costs.

For labor market policy modeling to be persuasive, it must have an explicit and carefully-formulated objective or set of objectives, theoretical foundations that capture the essential stylized facts in the economy under consideration, and statistical/econometric evidence drawn from careful empirical studies. To the extent that any given study falls short of these criteria, our confidence in the policy conclusions reached falls short accordingly.

Many so-called policy implications do not stand up to the most basic scrutiny. When it comes to the question of what should be done from a policy point of view, few studies have reached convincing policy implications of the type, “Do A, not B, with the available resources.” This is because no form of social cost-benefit analysis has been used.

Overall, weighing the strengths and limitations of existing labor market policy research, there is a need to be humble in drawing policy conclusions. Many empirical researchers feel obliged to add a policy conclusions section to their papers. For the reasons given, more circumspection is called for.
**Five pitfalls to be avoided**

Five pitfalls are common enough in the literature that it is worth cautioning the reader specifically against them. The first is to think of labor markets in terms of productivity rather than in terms of supply and demand. A simple example illustrates the point.

Consider a two person economy in which workers dig holes. Suppose that together, two workers working with one shovel can dig 10 holes a day, i.e., 5 holes each. Suppose one worker is replaced by a power shovel, which costs the same amount per day as the worker displaced, and that the one worker who remains employed working with a power shovel is able to dig ten holes a day. Suppose the other worker is unemployed. Productivity has doubled. “Should” the wage of the employed worker double? No – supply and demand analysis tells us that if anything the employed worker’s wage would fall, not rise.

A second pitfall is to rely on the wrong kinds of empirical studies. What we need most for policy purposes are empirical studies that give guidance on what the right theoretical labor market model should be or what the empirical magnitudes included in existing models actually are. Standard household and firm surveys are usually used to perform individual-level analysis. They can be used to perform market-level analysis instead, and indeed it would be helpful if they were used in that way more often. For example, if we want to know what the consequences of future educational expansion might be, we can derive useful policy insights by asking what were the market-level effects of past educational expansion. If, in the recent past, the supply of educated people has been increased by 100 university graduates, are 100 more people now working in “university-level” occupations? Sometimes, though, non-standard data are needed. In South Africa, the broad unemployment rate is 41%, whereas in most developing economies, unemployment rates are in the range of 5% to 10%. Why is South Africa so different? An interesting field study in South Africa (Cichello et al., 2006) asked the
unemployed why they have not entered self-employment. The principal result of this research – that the number one factor impeding entry of South Africans into informal self-employment is the fear of crime if the enterprise is successful – is more informative than any regression would have been.

Third, many policy conclusions are offered with no supporting cost-benefit analysis. As an example, consider a country in which a poverty profile shows (as most if not all poverty profiles do) that those who work in agriculture are disproportionately poor. An unwarranted policy conclusion would be to decide to invest the available development resources in the sectors where the poor are. Equally unwarranted would be the policy conclusion to invest the available resources in the sectors where the poor are not. The policy conclusion that is warranted is to invest development resources in whichever sector, agriculture or non-agriculture, produces the highest marginal social benefit compared with the marginal social cost – gauged, for example, in terms of poverty reduction.

A fourth policy pitfall is to limit one’s analysis to only a subset of the relevant goods or bads. Two examples are common in the literature on the economic effects of labor unions. One is to say that strong unions are good and should be encouraged because those who remain employed will earn higher wages. The other is to say that strong unions are bad and should be discouraged because the higher wages negotiated by unions are likely to cause higher unemployment.

A fifth pitfall is to commit a fallacy of composition. Individual i or firm j may be able to get ahead by doing more of something, but if that very action causes individual m or firm n to fall behind, the social benefits may be very much smaller than the private benefits. This is what Nobel Laureate George Akerlof (1976) called “The Economics of the Rat Race.” Tests need to be performed to determine whether it is likely that social and private benefits and costs approximate one another or not.
The main messages and pitfalls in action: minimum wage policy

I have highlighted three desirable research components (welfare economics, theoretical modeling, and empirical modeling) and five pitfalls (inappropriate use of productivity, reliance on wrong kinds of empirical studies, lack of cost-benefit analysis, attention to only a subset of the goods and bads, and fallacy of composition). As an example of how the three components might be included and the five pitfalls avoided, let us look more carefully at the literature on the labor market effects of minimum wages.

The familiar starting point for minimum wage analysis is the basic supply-demand model in which a higher minimum wage in a sector reduces the quantity of labor demanded and induces in-migration of labor, thereby increasing unemployment in that sector. Also well-known is the monopsony model in which a higher minimum wage in a sector can lead to increased employment in that sector provided that the wage increase is not too large.

Moving from one sector to two, minimum wages in labor market models in economies with covered and non-covered sectors have been analyzed by Harris and Todaro (1970), Harberger (1971), Mincer (1976, 1984), Gramlich (1976), and Fields (1975, 1997). We now know from this literature that a higher minimum wage in the covered sector may but need not result in more unemployment.

Unemployment should probably not be the only criterion for evaluating a minimum wage. When total labor earnings, inequality, and poverty are also included as components of the welfare judgment, Fields (2005b) showed that a higher minimum wage in a Harris-Todaro model could improve social welfare or not depending on parameter values.

Another example of interest would be the Bank’s efforts to develop a Social Protection, Education, Labor Market and Macroeconomy model (SELMA) for Morocco. SELMA is a dynamic general equilibrium model designed to assess the effects of education and social protection reforms and labor market policies on employment, unemployment, and other macroeconomic aggregates. That work is not reviewed here because it is still in draft form not for circulation. As with all such models, the credibility of the results hinges on the credibility of each of the constituent equations.
Finally, taking poverty as the welfare criterion, Fields and Kanbur (2007) showed in a single-sector model with income-sharing that a higher minimum wage could lower poverty or that it could raise poverty depending on parameter values which might then be estimated empirically.

Despite the progress that has been made, more work remains to be done in modeling minimum wages theoretically– in particular, enriching the model to include enough heterogeneity, most importantly, by placing some minimum wage workers in poor families and others in non-poor ones – and estimating the relevant parameters empirically. The studies reviewed by Cunningham (2007), Freeman (2007), and Svejnar (2007) offer guidance in this respect.

**Topics for future research**

Based on this review as well as other investigations, I would highlight six priority research areas. Within each, I pose a number of questions, highlighting in **bold** one question in each category which in my view merits the highest priority:

1. **Growth and labor market outcomes**
   - Using the indicators recommended in World Bank (2007a), **to what extent does macroeconomic growth improve labor market conditions?**
   - **Through what channels?**
   - How do the outcomes in one country compare with those in another? What explains the differences?
   - How do changes in labor market outcomes for one population sub-group compare with those of another?
   - There is a perception that economic growth has not necessarily translated into desirable labor outcomes. Is this perception right? Where? Under what circumstances?
   - Growth in many countries is characterized as jobless. To what extent is it the case that growth is indeed jobless for the population as a whole? Or is
it that growth has not been accompanied by growth in key sectors, e.g., the formal sector?

- Has the relationship between growth and labor market opportunities for the poor changed as the world has become more globalized?

2. Demand for labor

- **Why is the labor demand curve in a country positioned where it is?**
- Which components of the business environment are important influences on domestic and foreign investment?
- How does what goes on in other markets affect what goes on in the labor market? What role for infrastructure? For physical capital? For financial capital and credit? What determines factor pricing on these other markets?
- Which trade policies (e.g., open vs. closed economy) and commercial policies (e.g., exchange rate policy) are associated with the best labor market outcomes for the poor?

3. Supply of labor

- Why do people work informally?
- What are the educational and skill characteristics of the labor force and how have they changed over time?
- To what extent do the educational and skills supplied by workers conform to the education and skills demanded by employers?
- Is there evidence of structural unemployment and underemployment?
- **What are the effects of various policy interventions on those treated and on those not treated?**
- When educational expansion has taken place, how has employment of the educated responded?

4. Wage-setting mechanisms and institutions

- Are labor markets segmented?
• To what extent are wages and conditions of employment determined by supply and demand in various key sectors of the economy and to what extent by minimum wages, labor unions, and other institutional forces?

• **Using panel data, including new panels for additional countries:** What are the patterns of earnings changes and their determinants? Why does the qualitative picture coming out of panel data analysis look so different from that found in analysis of comparable cross sections?

• What are the determinants of self-employment earnings a) in agriculture, b) in non-farm rural employment, and c) in urban self-employment?

• Which policies have made a difference in each of these three areas, enabling workers to earn their way out of poverty?

5. Other institutions impinging upon labor market functioning

• **Are there significant barriers to labor market adjustments a) within labor markets and b) outside of labor markets but impinging upon labor markets?** Specifically:

  • To what extent are labor market adjustments impeded by aspects of a country’s labor code such as restrictions on dismissals or mandated severance pay?

  • To what extent are labor market adjustments impeded by benefits linked to employers such as employer-provided housing, health care, and pensions?

  • To what extent are labor market adjustments impeded by restrictions on worker mobility because of socioeconomic factors such as debt bondage?

  • To what extent are labor market adjustments impeded by factors outside of labor markets such as location-specific entitlements to public housing, ration cards, and other locally-provided goods and services?

• Why is the unemployment rate so low in most developing countries but so high in a few?
6. Interrelationships among labor markets

- What evidence is there on the empirical applicability of the existing multisector labor market models such as the integrated labor market model, the Harris-Todaro model, and the Roy model in a given country context?

- In some cases (e.g., South Africa), none of the existing multisector labor market models captures the essential structure. What would be a better stylization in such cases?

**Methods for future research**

In future research, some of the well-established methods may fruitfully be supplemented by some of the newer methods. These include additional work using:

- Panel data and panel data methods.
- Randomized field experiments.
- Mixed qualitative and quantitative methods.
- Models that contain enough sectors, provide realistic stylizations of the workings of the labor markets in each sector, and contain reasonable specifications of the linkages between sectors in the country in question.
- Field studies to supplement quantitative desk work.

**Some final words**

Given the development community’s focus on poverty reduction as being at the core of development policy and given also the demonstrated importance of labor earnings for poverty reduction, it is lamentable that the development banks and other development organizations have not devoted more attention to research on employment, earnings, and the functioning of labor markets. I hope that this deficiency will be remedied in the near future.

---

9 The Deaton et al. report evaluating research at the World Bank mentions the word “labor” just eight times and “employment” just seven. By contrast, “trade” and “finance” are each mentioned twenty-nine times.
I would conclude where I started: sound labor market policy requires sound labor market models. Let us be both bold enough to be explicit about our policy evaluation criteria, specific about our theoretical models, and comprehensive in our empirical evidence and at the same time humble enough to know when the best policy conclusion to draw is no policy conclusion at all. The stakes are too high for carelessness.

10 And that the World Bank will soon form a labor markets unit cutting across both the research and the operations arms of the institution.
References


