

ILR Collection
Articles & Chapters

Cornell University ILR School

Year 2009

Segmented Labor Market Models in
Developing Countries

Gary S. Fields
Cornell University, gsf2@cornell.edu

Segmented Labor Market Models in Developing Countries

Gary S. Fields *
Cornell University
gsf2@cornell.edu

This Version: March, 2008

Forthcoming in English in Harold Kincaid and Don Ross, The Oxford Handbook of the Philosophy of Economic Science, Oxford University Press.
Forthcoming in Spanish in El Trimestre Económico.

Abstract

Labor markets are important, because most people, especially the poor, derive all or the great bulk of their income from the work they do. This paper approaches labor markets through models of segmented labor markets.

The first main substantive section presents the essence of segmented labor market modeling, in particular, the role of labor market dualism. Given that labor markets often consist of quite distinct segments, a useful and insightful analytical approach is to start with just two interrelated segments, which here are termed formal and informal. Accordingly, the next sections present models of wages and employment in the formal sector, the informal sector, and the linkages between the two respectively. The final substantive section shows the contributions that these models make to understanding and policy analysis in labor markets.

It would not be expected that the same model would fit East Africa and East Asia or South Africa and South Korea. Surely, the “correct” model is context-specific. Blending empirical observation and analytical modeling has yielded great advances. Sound labor market policies require sound labor market models.

* I am grateful to Mabel Andalón, Gordon Betcherman, Shanta Devarajan, Lisa Dragoset, Louise Fox, Robert Duval Hernández, Maria Laura Sánchez Puerta, and three anonymous referees for helpful comments on an earlier draft.

This is a paper on labor market models. The aim is to construct models that are as simple as they can be but as complicated as they need to be. Such models, if carefully done, can contribute to an understanding of observed labor market phenomena and to the formulation of sound labor market policies.

Some branches of economics work with models that assume that everybody who works participates in a single, undifferentiated labor market. I regard such models as grossly unrealistic. A better description, I would maintain, is that jobs differ in quality, these different groups being called “segments” or “sectors”. Thus, labor market segmentation is said to exist if 1) Jobs for individuals of a given skill level differ in terms of their pay or other characteristics, and 2) Access to the more attractive jobs is limited in that not all who want the better jobs can get them.

The notion of labor market segmentation can be stylized most simply by maintaining that there are two labor market segments. A realistic assumption is that all who participate in the labor market want the better jobs, but good jobs are available only for a fraction of the labor force. Those who do not get the good jobs must either take up a bad job or remain unemployed. Models with two labor market segments prove to be both tractable and insightful, and so are used here.

In this paper, labor markets should be thought of as consisting not only of wage and salaried employment but also self-employment. All who work or seek to work in labor markets are termed “workers.”

Why are labor markets important to economic development? Many individuals and institutions, including the World Bank and the regional development banks, seek “a world free of poverty.” Broadly speaking, those who are poor are poor because 1) they earn little from the work they do, if indeed they have work at all, 2) the societies in which they live are too poor to provide them with substantial goods and services by virtue of their citizenship or residency, and 3) the poor are not permitted to move to richer countries. Thus, anti-poverty efforts can be focused on 1) helping people as workers by

creating more and better-paying employment,, 2) helping people as citizens/residents through publicly-provided goods and services, and 3) striving for freer movement of labor from poor to rich countries. This paper is concerned with the first channel: helping improve labor market opportunities for workers.

The importance of labor markets for anti-poverty efforts is underscored by research studies using decomposition methodologies. These studies have shown that labor income inequality is as important or more important than *all other income sources combined* in explaining total income inequality; see Ayub (1977) for Pakistan, Fields (1979a) for Colombia, and Fei, Ranis, and Kuo (1978, 1979) and Fields and Mitchell (1999) for Taiwan. The reason that labor income is so important is, in the words of the 1990 World Development Report, that “the poor’s most abundant asset [is their] labor.” (World Bank, 1990, p. 3). It follows that a very important factor in explaining family income inequality is that some people earn very large amounts for their labor while a great many earn very little. Thus, it is the inequality of labor incomes that accounts primarily for the inequality of total incomes.

Labor income also plays a predominant role in income mobility research. In much of this literature, economic welfare is gauged by household per capita income (PCI) or household per capita consumption (PCC). Research on changing PCI in Indonesia, South Africa, Spain, and Venezuela has shown that household per-capita income changes are determined much more by changes in household income (the numerator) than by changes in number of household members (the denominator) and that changes in labor income far outweigh changes in other sorts of income (Fields et al., 2003).

This paper approaches labor markets through segmented labor market modeling. Such models start with the recognition that in many countries, the labor market consists of quite distinct segments that are linked with one another. When there are just two segments, these models are called “dualistic” labor market models.

As explained below, segmented labor market models are valuable because they can explain a number of phenomena that simply do not make sense in a single market setting. To develop a framework/typology and lay out the main issues in segmented labor market modeling, the paper proceeds in five stages.

The first main substantive section (Section I) presents the essence of labor market dualism. I maintain that labor markets often consist of quite distinct segments and that a useful and insightful analytical approach is to start with just two.

The second main substantive section is on models of wages and employment in the formal economy. To be reviewed here are 1) the market-clearing labor market model and the presumed equilibrating forces, 2) above-market-clearing wages set institutionally, 3) above-market-clearing wages set by efficiency wage considerations, and 4) above-market-clearing wages set by worker behavior.

The third main substantive section is on wages and employment in the informal economy. This section presents three characterizations of informal sector labor markets: 1) the informal economy as a free-entry sector that prospective workers enter only as a last resort, 2) the informal economy as a desirable sector that workers choose in preference to formal sector work, and 3) the informal economy with its own internal dualism, combining 1) and 2).

The fourth main substantive section is on formal-informal linkages. The models here are: 1) the integrated labor market model with full market clearing, 2) crowding models, and 3) the Harris-Todaro model.

Finally, the fifth main substantive section discusses five contributions that these models make to understanding and policy analysis. These issues are 1) why an increase in productivity might cause wages to fall, 2) why Taiwan's economic growth led first to falling unemployment at constant wages and then to economy-wide wage increases at full employment, 3) why the solution to urban unemployment in Kenya was not urban employment creation but rural development, 4) why poverty is so severe in the urban

informal sector of many developing countries, and 5) why expanding formal sector employment sometimes improves labor market conditions and sometimes does not.

While the models presented here differ from one another in important respects, they all share certain common features of which the reader should be aware from the outset. First, firms in these models are assumed to be maximizing profits. This means that they hire workers, raise wages, and improve worker quality if and only if it is in their profit-maximizing interest to do so. Second, workers in these models are assumed to be maximizing utility. Especially in poor countries, in which large numbers of people value additional goods greatly compared to leisure, the utility-maximization assumption may often be fruitfully replaced by an income-maximization assumption. Third, the notion of “market equilibrium” used in this paper needs to be clarified. “Market equilibrium” is a state toward which a market tends and, once there, it tends to stay. “Market-clearing” is a state in which the quantity of a good or service supplied equals the quantity demanded. Some of the equilibria we shall deal with in this paper are market-clearing and others are not. And finally, the objective of this paper is to make labor outcomes more understandable; much of what is understandable is not pretty.

I. The Essence of Labor Market Dualism

At the core of segmented labor market modeling is the distinction between one part of the labor market and another. In the literature, one sector is alternatively called “formal,” “modern,” “industrial,” “good jobs,” or “urban,” while the other is alternatively called “informal,” “traditional,” “agricultural,” “bad jobs,” or “rural.” (At one point, I even called this latter one the “murky” sector.) Throughout this paper, the formal/informal terminology shall be used.

Labor market dualism is a useful stylization of what has been called “labor market segmentation” or “labor market fragmentation.” Dixit (1973) explained why:

The dual economy has, over the last decade, proved itself to be a useful conceptual framework for analyzing several problems of economic development. . . Dual economy models provide a significantly better description and understanding of the problems of development than any aggregate model . . . because the sectoral division chosen reflects several vital social and economic distinctions in the type of economy being analyzed.

Why have just two sectors? Basu (1997, p.152) put it thus:

The assumption of duality is merely for analytical convenience. If fragmentation – irrespective of the number of parts – in itself causes some problems and we wish to examine these, then the simplest assumption to make is that of dualism.

Unfortunately, international practice has been quite ambiguous about the feature distinguishing the two sectors. The International Labour Organisation and the Economic Commission for Latin America and the Caribbean have defined the informal sector as the sum of non-professional self-employed, domestic workers, unpaid workers, and workers in enterprises employing five or fewer workers. In Brazil, the formal sector consists of workers who hold labor cards entitling them to various benefits and protections and the informal sector of those who do not. In other contexts, the formal sector is distinguished according to whether the firm is registered with the government and pays taxes. Yet others equate the informal economy with drugs, prostitution, and other illegal activities. For alternative definitions and operationalizations, see ILO (2002) and Jhabvala, Sudarshan, and Unni (2003).

The distinguishing feature used by Nobel laureates Arthur Lewis (1954) and Simon Kuznets (1955) as well as other dual economy modelers 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.” Lewis explained that although part of the gap is “illusory” because of the higher cost of living in the capitalist sector, 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.”

Kuznets (1955) further developed the model of wage dualism and intersectoral shifts by exploring how various measures of income inequality (including the income share of the lowest income quintile, the income share of the richest income quintile, and the range) would change as the high-income sector comes to employ an increasing share of the population. All of the inequality measures used by Kuznets (the income share of the poorest quintile, the income share of the richest quintile, and the interquintile range) exhibited an inverted-U pattern, which later came to be known as the “Kuznets Curve.” Subsequent research examined inequality further (Knight, 1976; Robinson, 1976; Fields, 1979a; Anand and Kanbur, 1993) and also examined poverty (Fields, 1979a ; Anand and Kanbur, 1985) in the Lewis-Kuznets process of intersectoral shifts.

Later writings on labor market dualism were grounded in human capital theory as developed by Schultz (1961, 1962), Becker (1962, 1964), and Mincer (1962, 1974). This later literature on labor market dualism stressed that for dualism to exist, different wages must be paid in different sectors to *comparable* workers. Subsequently, many researchers reported empirical evidence showing such dualism or segmentation for observationally equivalent workers.

The idea that different wages are paid to comparable workers has been incorporated, largely without question, into job search theory. Since the late 1960’s and early 1970’s, a whole class of models has arisen in which a wide variety of wages exist in the labor market, and workers are presumed to search among employers for the best possible opportunities. See, for instance, the textbook treatments of job search in Ehrenberg and Smith (2003) and Cahuc and Zylberberg (2004) and also the work on equilibrium wage distributions by Stiglitz (1985) and Burdett and Mortensen (1998).

Dualistic labor market models have been criticized on a number of grounds. One critique is that offered by Rosenzweig (1988). Noting that empirical studies often show that workers with given measured human capital characteristics have systematically different wages or earnings depending on the type of employment in which they are working, he asked (p. 756): “Do [these differentials] suggest barriers to mobility – non-competing groups – or do they merely reflect compensatory differentials, rewards for unmeasured skills or compensation for unmeasured differences in the disutility of the workplace?” Favoring the latter set of explanations, Rosenzweig found the dualistic labor market literature unconvincing.

Dualistic labor market models, and segmented labor market models more generally, have been criticized on other grounds as well. An Inter-American Development Bank report (IADB, 2003) put it thus: “According to [the dualistic view of the labor market], the formal and informal economies operated in segmented labor markets and there is limited mobility between the two. Nothing could be further from the truth . . . In a given six-month period, about 16 percent of workers in Mexico and 11 percent of workers in Argentina move either in or out of an informal job.” Nonetheless, the fact is that most workers *remain* in the sector in which they began.

These critiques notwithstanding, it appears to me that luck plays an important role in the sense that the very same individual with his or her own measured and unmeasured skills will do better or worse in the labor market over the long run if a job vacancy exists in a more attractive labor market sector than if such a vacancy does not exist. Thus, labor markets are better characterized as being segmented in the sense of cumulative advantage and low-level traps (Nelson, 1966; Merton, 1968; Doeringer and Piore, 1971; Boudon, 1973; Meade, 1976) than as being unified in the sense that the next-best employer is essentially indistinguishable from the current one.

To conclude this section, in most settings, I find that it is more useful to think of developing countries’ labor markets as being fragmented or segmented than to think of

all workers and firms in a country participating in one single labor market. When possible, Occam's Razor suggests limiting the analysis to two sectors. Labor market dualism was a most useful starting point for analyzing some economies when it was first presented decades ago, and it remains a useful characterization of some economies today. But when two sectors are simply not enough, three-sector or n-sector models can prove insightful, a point to be developed further in the sequel.

We turn now to the formal sector labor market, the informal sector labor market, and the interactions between them.

II. The Formal Sector Labor Market

This section presents four alternative models of wages and employment in the formal sector: the market-clearing labor market model, models with wages set above market-clearing levels for institutional reasons, models with wages set above market-clearing levels for efficiency wage reasons, and models with wages set above market-clearing levels because of supply-side considerations.

A. The Market-Clearing Labor Market Model

The market-clearing labor market model is so well-known that it can be presented quite concisely. Figure 1 displays the three essential features. First, the amount of labor demanded 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. Second, the amount of labor supplied 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.)

Let us now consider three groups of models with different causal structures and different market outcomes.

B. Above-Market-Clearing Wages Set Institutionally

An important class of models in the labor market literature holds that wages in the formal sector are set by a set of forces *different* from supply and demand. In the models reviewed in this section, the defining feature is heavy reliance on “institutional” forces. (“Institutional” forces are those other than the profit-maximizing behavior of firms and the utility-maximizing behavior of workers.)

It is useful to distinguish five institutional features that may be important in different settings (e.g., Fields and Wan, 1989; Fields, 1999). They are: minimum wages; trade unions; public sector pay policies; multinational corporations; and labor codes.

Minimum wages aim to assure workers an "adequate" standard of living. Minimum wages have long been on the books throughout most of the developing world (Starr, 1981; World Bank, 1995; Inter-American Development Bank, 2003), but their effects differ. In some countries, such as Bangladesh, Côte d'Ivoire, and India, minimum wages are said to be binding and enforced (World Bank, 1995, p. 75). But in others, minimum wage laws make little difference. South Korea introduced a minimum wage system only in 1988, and the minimum wage levels have been set so that they prove to be a binding constraint for only about 2% of Korean workers (Park, 1991; Lee, 2002). In the case of Taiwan, although a minimum wage law has been in force for decades, no company has ever been fined for violating the law (Chang, 1989).

Trade unions are often encouraged by government policy as a means of entitling workers to a "just" share of the fruits of their labor. One early theory of trade union behavior is that unions have a variety of objectives, including both higher wages and greater union membership (Dunlop, 1944). More recent approaches have stressed that those already employed in unions (the “insiders”) may be more concerned about raising their wages than about increasing membership (Blanchard and Summers, 1986; Lindbeck and Snower, 1988). Indeed, unions have raised the wages of their members by as much as 150% in Jamaica, 31% in Ghana, 24% in South Africa, and 20% in Malaysia (Tidrick,

1975; World Bank, 1995, Table 12.2; Aidt and Tzannatos, 2002). By contrast, in South Korea and Taiwan, unions have been repressed and the union wage premium is at most two to three percent. (Lin, 1989; Park, 1991; Yoo, 1995).

Public sector pay policies often result in substantially higher wages being paid to government workers than to their private sector counterparts. Costa Rica is an example of this, and as a result, "everybody" there tries to work for the government (Gindling, 1991). In East Asia, the public sector pays what it has to in order to compete with the private sector -- neither more nor less.

Multinational corporations frequently pay above-market wages in sub-Saharan Africa and elsewhere (Squire, 1981). Wage levels and working conditions tend to be higher in export-oriented firms than in firms producing for the domestic market (ILO, 1998; Moran, 2002; Ghose, 2003). Although export-oriented multinationals offer higher wages and better working conditions partly for efficiency wage reasons (see subsection C below), they also do this because some governments have "encouraged" them to do so by not so subtly threatening expulsion or expropriation if they do not (Eaton and Gersovitz, 1984).

Finally, *labor codes* in some countries regulate hiring and firing, impose payroll taxes on firms, and mandate that employers provide certain benefits to their workers. Panama had such a labor code, and it was estimated to have raised labor costs by 90% (Spinanger, 1985) before it was finally abandoned as unsustainable. Likewise, Bolivian employment legislation raises labor costs by an estimated 90-110% (Bravo, 1995). Larger firms in India and Zimbabwe are not permitted to dismiss workers; employment levels and economic efficiency have been found to be artificially low as a result (Fallon and Lucas, 1991, 1993; Besley and Burgess, 2004).

Higher-than-market-clearing wages for institutional reasons in the formal sector are at the core of many economic models. They include the Keynesian macroeconomic model, Lewis's classical development model, Harris and Todaro's dualistic labor market

model, and many others. Virtually without exception, economic models regard formal sector employment as being determined in a very neoclassical way: given the wage and the capital stock, employment is set according to the marginal revenue product of labor.

Ample research has shown that labor demand elasticities are significantly negative; Hamermesh (1993) provides a comprehensive review of the empirical literature. In South Africa, for example, various researchers have produced estimates of the wage elasticity of employment in that country's formal sector (Bowles and Heintz, 1996; Fields, Leibbrandt, and Wakeford, 2000). Most estimates range from -0.5 to -0.7. While these studies differ in terms of their precise estimates, what they agree on is that 1) the wage elasticity of employment is significantly negative and 2) the wage elasticity of employment is significantly less than one in absolute value.

Given significantly negative labor demand elasticities, higher-than-market-clearing wages would be expected to reduce formal sector employment below what it would have been otherwise. Unemployment will result in the economy unless all of the workers not employed in the formal sector take up employment in the informal sector. Whether they do or not is the subject of Section IV.

Given these research findings, the five labor market interventions reviewed above need to be considered carefully. Their aims are laudatory - to raise earnings and reduce poverty - and they do indeed benefit the workers who are fortunate enough to work in covered sectors of the economy. However, they appear to have had adverse employment and efficiency effects and to have contributed to the informalization of the economy, as employers evade the regulations by not engaging workers as regular employees or by not even appearing as official companies (DeSoto, 1989; Turnham, 1993; Maloney, 2003; Levy, 2008). Helping formal sector workers may or may not be the best tool for fighting poverty in any given context.

C. Above-Market-Clearing Wages Set by Efficiency Wage Considerations

An old and well-established idea that commands nearly universal agreement not only in economics but in human resource management is that a firm can raise its labor productivity by paying a higher wage. Credit is usually given to Leibenstein (1957) for originating this idea in the economics literature. See also Stiglitz (1974, 1976), Mirrlees (1975), Bliss and Stern (1978), Shapiro and Stiglitz (1984), Akerlof and Yellen (1986), Dasgupta and Ray (1986), and Weiss (1990) for further developments. But it goes back much further than that to Henry Ford, who pioneered the radical practice a century ago of offering his workers \$5 a day, which was twice the going wage at that time (Raff and Summers, 1987).

Efficiency wage theory incorporates the proposition that higher wages can result in higher productivity but goes beyond it in a fundamentally important way. According to the core microeconomic model of firms, firms are trying to achieve higher *profits*, which may or may not be enhanced by higher *productivity*. Thus, the basic postulate of efficiency wage theory is that profit-maximizing 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*. In other words, it is not enough simply to maintain that paying a higher wage generates benefits. It must be that *the benefits exceed the costs*. Much that is written about “high road” labor relations practices ignores this fundamental truth; see, for example, Ulrich (1997) and Noe et al. (2000).

Efficiency wage theory has also contributed usefully to analyzing the mechanisms by which productivity gains are realized. These fall into two major categories.

One set of explanations is that higher wages enable firms to hire better-quality workers from a heterogeneous labor pool. They may, for example, hire workers who have more education and who for this reason are expected to be more productive.

Alternatively, they may administer tests of potential job performance and hire those workers who perform the best on these tests.

The other set of explanations is that higher wages induce workers of a given skill level to perform in a more productive manner. The mechanisms analyzed here include better nutrition, improved morale, reduced shirking, lower labor turnover, reduced absenteeism, and greater discretionary effort.

Where the efficiency wage models come out, then, is that wages remain above the market-clearing level because firms in the labor market *find it in their profit-maximizing interest to keep wages above the market-clearing level*. Put differently, a firm that is paying efficiency wages would hurt its profits if it *lowered* wages.

As in the models reviewed in the last subsection, when wages are higher-than-market-clearing for efficiency wage reasons, we also have unemployment as an equilibrium outcome. However, the unemployment that arises here occurs for a very different reason from the institutional wage case. In the efficiency wage models, it is *firms* that do not want to reduce wages, even though at least some of the unemployed would be willing to work for lower wages rather than remain jobless. This contrasts with the institutional wage case, in which it is *employed workers* who want the wage to remain where it is.

D. Above-Market-Clearing Wages Set on the Supply Side

Another explanation for wages remaining above the market-clearing level has been suggested and modeled by Bardhan and Rudra (1981), Drèze and Mukherjee (1989), Solow (1990), and Osmani (1991). Suppose that daily wages in a labor market start out initially above the market-clearing level for some reason – for example, because the wage was set in the peak season and the economy is now in the slack season. According to the standard account of equilibrating forces in labor markets, when the wage is higher than the market-clearing level, unemployed workers would offer to work for lower wages rather than remain unemployed.

However, in the models of Bardhan and Rudra, Drèze and Mukherjee, Solow, and Osmani, workers' supply side behavior may differ from the standard account, as follows. Suppose that the labor market is a casual one in which hiring takes place afresh each day. If the demand for labor is inelastic, the total wage bill paid to labor over a longer period such as a month or a year will be higher the higher is the daily wage. Each of the unemployed knows that he or she will earn more on average over the course of many days if s/he does *not* undercut the established wage and therefore will not do so. Wages remain above the market-clearing level as a result.

In this class of models, unlike the models in the earlier subsections, wages are kept above the market-clearing level by the behavior of the *unemployed*. In this way, the wage remains above the market-clearing level, and unemployment persists as a result.

III. The Informal Sector Labor Market

The crucial feature of labor market dualism described above is that the formal sector offers relatively attractive wages and other terms and conditions of employment while the informal sector offers relatively unattractive ones. This leads to the first characterization of the informal economy: workers prefer formal sector jobs and enter the informal sector only as a last resort. More recently, though, a different view has been put forth: that the informal economy is a desirable sector that workers choose in preference to formal sector work. A third view is that the informal economy has its own internal dualism, combining these two characterizations. A current resource on the informal economy is the ILO's Informal Economy Resource Database, available at http://www.ilo.org/dyn/dwresources/iebrowse.home?p_lang=en .

A. The Informal Economy as a Free-Entry Sector of Last Resort

Most of the poor in the world are working poor. New ILO data show that while open unemployment throughout the world is 6.2% of the labor force, another 42.5% of those in the labor force are working but earning less than US\$ 2 per day (ILO, 2005). It has long been recognized that open unemployment is the tip of the proverbial iceberg: the greater part of the employment problem in developing countries consists of workers who earn so little when they work that they and their families are poor (e.g., Turnham, 1971; Squire, 1981). The working poor are found disproportionately in the informal sector.

Ample empirical research has shown that labor earnings in the informal sector are low, lower even than in the formal sector in a large number of countries. For example, Sudarshan and Unni (2003) see informal work as “a survival activity of the very poor,” noting that the dimensions of informal activity are large: 35-85 percent of non-agricultural employment in Asia, 40-97 percent in Africa, and 30-75 percent in the Latin America-Caribbean region..

In the cities of developing countries, we see large numbers of people engaged in work that earns them some cash each day or week. These include hoards of shoe shiners clustered in the town square, lottery ticket vendors seemingly every few feet, would-be construction workers clustered at a particular street corner awaiting the daily round-up, newspaper vendors approaching stopped cars at virtually every traffic light, and (sadly) groups of women, and sometimes men and children, gathered in the red light district. Lewis (1954, p. 141) referred to “the whole range of casual jobs – the workers on the docks, the young men who rush forward asking to carry your bag as you appear, the jobbing gardener, and the like. *These occupations usually have a multiple of the number they need, each of them earning very small sums from occasional employment;*

frequently their number could be halved without reducing output in this sector.”

(Emphasis added)

Subsequent investigations into these people’s lives as well as casual empiricism led analysts to view these types of jobs as having free entry. In a pathbreaking ILO report on Kenya (1972, p. 6), the criteria defining the informal sector were:

- i) ease of entry;
- ii) reliance on indigenous resources;
- iii) family ownership of enterprises;
- iv) small scale of operation;
- v) labour-intensive and adapted technology;
- vi) skills acquired outside the formal school system; and
- vii) unregulated and competitive markets.

The essence of free entry is that all who want a job can get one. (“Job” here is defined to include both self-employment and wage employment.) Barriers to entry into such occupations are small or non-existent. In some contexts, primarily urban, all that would-be workers need to do is make a minimal investment in the product or service to be sold. In rural contexts, it is obligatory for the family or community to take back into the home those who find such work the best of a bad set of alternatives. One is reminded of Robert Frost’s immortal words in his poem “Death of the Hired Man”: “Home is the place where, when you have to go there, they have to take you in.”

The existence of free-entry employment opportunities in the informal sector helps explain why open unemployment rates in developing countries are comparable to those in developed countries, and often considerably lower (Turnham, 1971, 1993; World Bank, 1995; ILO, 2003). The standard ILO definition of unemployment is a person who did no work for pay in the preceding week, not even for one hour. In poor countries lacking systems of unemployment insurance and cash assistance allowances, the great majority of poor people cannot afford to be without income for as long as a week. So to the extent

that the poor can quickly find an opportunity to earn some cash in an informal job, they take it. Open unemployment in their economies is low as a result.

Because of easy entry into economic activities of such kinds, a different wage determination process from the standard marginal productivity rule must be found. Lewis posited income-sharing, a feature taken up by others (e.g., Fei and Ranis, 1964; Harberger, 1971; Fields, 1975). As viewed today, what matters, writes Ranis (2006), is “that the marginal product is low, and sufficiently low to fall below the bargaining wage or income share.”

How, then, would we now want to model informal sector wage determination? Essentially, there are four tacks that might be taken, the first two for analytical simplicity and the second two for greater comprehensiveness.

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. For example, there may be a fixed number of newspapers to be sold regardless of the number of newspaper vendors. How is the fixed income from newspaper vending to be divided? The easiest simplifying assumption here is full income-sharing among the informally employed, so that each earns the average product. The average product is not constant, though – it varies inversely with the number of people in the informal sector. This was the way the urban informal wage was modeled in Fields (1975, 1989).

A second approach is to regard a part of the informal sector as facing, instead of zero marginal product, *constant* marginal product. The dual economy model developed by Harris and Todaro (1970) was formulated to fit the East African case, which they and others regarded as a land surplus economy at the time. Harris and Todaro assumed that anyone who wanted to work in agriculture could find a plot of land, cultivate it, and earn the marginal product from his or her efforts. Agricultural wages were equated to marginal product, not average product as in Lewis. If the marginal worker and the marginal land

are assumed to be as productive as preceding inputs were, a convenient simplifying assumption would be to regard the *marginal* product of labor in agriculture as constant. 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) 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.

Models with positive but variable marginal product are harder to work with than either of the two preceding ones.

A fourth approach is to model a full demand system for agricultural and non-agricultural products and workers. This was done by Bourguignon (1990). The equations of such systems are so complicated that they are best left to microsimulation and computable general equilibrium exercises.

To conclude, the most common characterization of the informal sector is that it is an easy-entry sector that workers can enter to earn some cash in preference to earning nothing. An alternative view has been gaining popularity in recent years. Let us now turn to it.

B. The Informal Economy as a Desirable Sector

A very different view of the informal sector also appears in the literature. It is the idea that a large number of those working in the informal sector are there *voluntarily*. This view has a long history in the literature (e.g., Hart, 1973; Balán, Browning, and Jelin, 1973). Fields (1990, p. 66) put it thus:

Many people are in informal activities by choice. When asked their reasons for doing what they were doing, many informal workers in Costa Rica gave the following answers most frequently: i) They feel they could make more money at the informal sector job they were doing than they could earn in the formal sector, or ii) Even though they made a little less money, they enjoyed their work more, because it allowed them to choose their own hours, to work in the open air, to talk to friends, etc.

The choice approach to the informal sector has been developed more recently in a series of papers by William Maloney. A comprehensive summary of these arguments appears in Maloney (2003).

According to economics textbooks, workers choose among jobs and sectors on the basis of a *package* of characteristics. These include wages, benefits, the work environment, and so on. The variable denoted W on the vertical axis of a standard labor market diagram is ordinarily thought of as a shorthand for this package of benefits, and it is this package of characteristics which Maloney maintains are “roughly comparable” between informal self-employment and formal employment, at least in Mexico. Specifically, Maloney offers a number of reasons why workers might want to be in the informal sector: some can earn more (or at least hope to earn more) in informal self-employment than they could earn in formal sector employment; they value the independence of self-employment; they would rather use the money that formal sector protections cost them for investing in their own small informal enterprises; they do not value protections such as health insurance which formal employment offers to them, in some cases because they already have these protections; and they don’t trust the

government to deliver on promises such as future pension benefits. For any or all of these reasons, there may be a sizeable numbers of workers who prefer informal self-employment to formal wage employment.

One reason that self-employment is often seen as undesirable is that microenterprises exhibit very high rates of failure. Maloney responds to the precariousness argument thus (2003, p. 77):

Small firms will have higher costs, are likely to be informal, and will have very high failure rates. Though this corresponds exactly to the standard picture of the stagnant, precarious, unproductive, unprotected informal worker familiar in the literature, it is, in fact, the opposite. It emerges naturally from the workers trying their luck at entrepreneurship (risk-taking), often failing, and not engaging in the formal institutions until they grow. In sum, there may be nothing pathological about informal self-employment, and to recover the general sense of the word, nothing obviously less decent either.

I agree with Maloney on this point, but I think he goes too far in one respect. He presents an integrated labor market model (pp. 68, 72) in which the total package of benefits is equalized between informal self-employment and formal wage employment. While this model might fit the choice between formal sector employment and informal self-employment *for those who already have the option of working in the formal sector*, this is a limited group of people. Rather, as argued above, throughout the developing world, formal sector jobs appear to be far fewer in number than the number of people who want them. Thus, in my view, Maloney's characterization applies to a subset of informal sector workers, but by no means all of them, nor probably even most.

C. The Informal Economy with Its Own Internal Dualism

The preceding subsections put forward two polar views. One is that informal sector employment is *worse* than formal sector employment but superior to unemployment. The other is that employment in the informal economy is *preferred* to formal sector employment.

A way of combining these two polar views would be to regard the informal sector as having *its own internal duality*. On this synthesized approach, some informal activities are preferable to formal sector jobs and some are not. Such a view is developed at length in Fields (1990), where the two parts of the informal sector are labeled “upper-tier” informal activities and “easy entry” ones. See also House (1984), Tokman (1987), Marcouiller et al. (1997), and Ranis and Stewart (1999).

In fact, dualism within the informal sector is a view that Maloney has come to share. Summarizing the findings of Cunningham and Maloney (2001) for Mexico, Maloney writes (2003, p. 80): “The single distribution was rejected, supporting a two-tier view, but the share of the population found in the ‘lower’ tier was only 13 percent of the sample.” Perhaps most informal entrepreneurs are in the upper-tier in Mexico, but it remains an open question whether this is the case in India, Bolivia, or Kenya.

Another way of modeling the duality of the informal sector is to specify two informal sectors that are *geographically* distinct. Todaro (1969) had three employment sectors – urban modern employment, urban traditional employment, and agricultural employment – but no unemployment. Harris and Todaro (1970) had urban modern employment, agricultural employment, and unemployment but no urban informal sector. Fields (1975) had three employment states – urban modern employment, an urban informal sector, and rural agricultural employment – plus unemployment.

If these various sector distinctions are put together, we should have four employment states - employment in the formal sector, employment in the upper-tier informal sector, employment in the easy-entry sector, and employment in rural agriculture – plus unemployment. Adding in rural off-farm employment – what is sometimes called the z-goods sector (Hymer and Resnick, 1969; Ranis and Stewart, 1993) – would introduce a fifth employment state. To the best of my knowledge, no analytical model has included all five employment states plus unemployment, perhaps because to do so would be too complicated and intractable.

Although six-state models have not been constructed, the literature offers a number of four-state models (consisting of three employment sectors plus unemployment). How the different states link to one another is open to alternative specifications. These are discussed in Section V.

IV. Formal-Informal Linkages in the Labor Market

Based on the models of formal sector labor markets and informal sector labor markets reviewed in Sections II and III, this section reviews models of linkages between the various sectors. The models reviewed here are: 1) the integrated labor market model with full market clearing, 2) models with wage dualism but no unemployment, and 3) the Harris-Todaro model, both in its original form and as extended, which features both wage dualism and unemployment.

A. The Integrated Labor Market Model with Wage Equalization and No Unemployment

The integrated labor market model, also called the unified labor market model, has as its distinguishing features that 1) each labor market clears, and 2) full intermarket equilibrium is achieved through actual wage equalization. The model may be explained with the aid of Figure 2.

Suppose for ease of analysis that the economy has two sectors, here termed "manufacturing" and "agriculture." All workers are identical, and so would be willing to work wherever the wage is higher, be it in manufacturing or in agriculture. The labor market is assumed to be an integrated one in the sense that the same wage prevails in both sectors of the economy for a given type of worker -- a realistic enough stylization in some settings, particular for countries in East Asia. For a model in which workers are not identical, but where the marginal worker earns the same regardless of whether s/he works in one sector or the other, see Roy (1951).

Let us start with a situation in which the demand for labor curve in the manufacturing sector, D_M , is downward-sloping relative to origin O_M , and likewise, the demand for labor curve in the agricultural sector, D_A , is downward-sloping relative to origin O_A . The total labor supply is represented by the horizontal distance $O_M O_A$. If the standard equilibrating forces in labor markets are free to operate, as is indeed the case in much of East Asia, wages would equalize across the two sectors at level W^* . At this wage, $O_M E$ workers would be demanded in the manufacturing sector, and $O_A E$ workers would be demanded in the agricultural sector. Furthermore, the total labor demanded in the two sectors combined would exactly equal the total labor supplied in the economy. In such an equilibrium, the marginal worker is indifferent between working in the manufacturing sector or in the agricultural sector, because the two sectors pay the same wages.

Now suppose that economic growth takes place in the manufacturing sector. Because manufacturing firms need more workers to produce the extra output, the labor demand curve in the manufacturing sector shifts rightward to D_M' . Assuming no change in the agricultural product market, the agricultural employers' demand for labor curve would remain stationary at D_A . The labor market is now in disequilibrium, because at the original wage W^* , more labor is demanded than is supplied. To resolve this disequilibrium, some manufacturing employers raise wages in order to retain existing workers and attract new ones, and agricultural employers raise wages to prevent their workers from leaving. The result is that the labor market equilibrates at a new common wage $W' > W^*$. Because of the sector-specific shift in labor demand, more of the country's workers are now in the manufacturing sector than before ($O_M E'$ rather than $O_M E$) and fewer in agriculture ($O_A E'$ rather than $O_A E$).

In the integrated labor market model, economic growth in *one* sector benefits workers in *all* sectors. Three groups of workers have been identified in this analysis: (i) Those who had been working in manufacturing and now earn higher wages than before;

(ii) Those who are drawn by higher wages into manufacturing from agriculture; and (iii) Those who remain in agriculture and earn more than they did previously. In this way, economic growth in a country's export sector reverberates throughout the labor market, benefiting those who produce manufactured goods *and* those who produce agricultural goods.

The extension of the integrated labor market model from two sectors to N sectors is immediate.

B. Models with Wage Differentials and No Unemployment

In contrast to the integrated labor market model just discussed, a number of segmented labor market models are characterized by wage differentials between segments. Models with wage differentials between segmented and no unemployment include the unlimited supply of labor model of Lewis (1954), the intersectoral shifts model of Kuznets (1955), the crowding model of Bergmann (1971), the minimum wage model with incomplete coverage of Welch (1974), and the modern sector enlargement model of Fields (1979b, 1980).

These models maintain labor market dualism in the sense that real wages are higher in the formal sector than in the informal sector. In this sense, they differ from the integrated labor market model described in the last subsection, in which wages are the same in the different sectors. The segmented labor market models described in this section also maintain a particular kind of supply-side behavior: all workers not employed in the higher-wage formal sector are assumed to take up employment in the lower-wage informal sector. These models therefore exhibit no unemployment.

Within this class of models, the most heralded version is the Nobel Prize-winning work of Lewis (1954). As discussed above, the distinguishing feature of the Lewis model was that the modern sector faces an unlimited supply of labor at wages only somewhat higher than subsistence levels. It is this that makes the Lewis model “classical,” in

contrast to a “neoclassical” model in which labor is scarce and has to be bid away from other uses. This feature of the classical model was later elaborated on by Ranis and Fei (1961), Fei and Ranis (1964), and Jorgenson (1967).

The unlimited supply of labor to the modern sector is sometimes called an “infinitely elastic supply curve of labor,” but this designation is a misnomer. By definition, a supply curve tells the amount of a good or service that is forthcoming as a function of the relevant price. For it to be a proper function, there can be only one quantity for any price. That is, given the price of labor, the supply function delivers the *unique* quantity of labor available. Thus, in the Lewis model, when the formal sector wage is above the informal sector wage, the potential quantity of labor supplied to the formal sector is *the entire labor force*. However, because formal sector employers do not wish to employ all the workers who would like to work there at that wage, they (the employers) face an effectively unlimited supply of labor. Specifically, this means that no individual employer need raise the wage to attract additional labor, nor must employers as a whole within a substantial range. Indeed, there is a horizontal curve, but that curve is the wage as a function of employment, *not* the amount of labor supplied as a function of the wage.

Over time, the process of savings, investment, capital formation, and economic growth highlighted in many growth models (both classical and neoclassical) shifts the demand for labor curve in the formal sector rightward. Workers respond to the increased demand for labor in the formal sector by taking up formal employment to the extent possible. Throughout a long range, the wage in the formal sector remains unchanged, because employers do not need to raise the wage to attract more labor. Ultimately, though, a turning point is reached once the supply of labor to the formal sector is no longer unlimited.

Despite the many insights of the original Lewis model and Fei and Ranis’s amplification of it, I find one feature of the model troublesome: the nature of the wage in

the informal sector. Lewis used the term “subsistence wage.” If the wage is literally a subsistence wage, below which people cannot subsist, then it has a natural floor. But as these models have evolved, the informal sector wage does not take on the character of the minimum needed for survival. It is, rather, more of a basic wage, lower than the real wage received by formal sector workers. The question, then, is whether this wage is a *constant* low wage or whether it *varies* (inversely) with the number of people in the sector.

The great majority of analysts regard production in the informal sector as subject to diminishing returns; see, for example, the Harberger quotation above. What diminishing returns in the informal sector implies is that when economic growth takes place and workers are drawn out of the informal sector into the formal sector, those who remain in the informal sector each receive a *higher* income than before; from my reading, this was first pointed out by Sen (1967). The informal sector wage should *not* remain constant. Indeed, the rising wage in the informal sector is a reason for the unlimited supply of labor to the formal sector to run out eventually: because the supply price of labor to the formal sector will have risen due to improved wage opportunities in the *informal* sector.

In the dualistic labor market model with no unemployment, economic growth reduces poverty in two ways. One is the increase in wages and utility of those who are able to move from the informal to the formal sector. The other is the increase in wages of those who remain informal.

Thus, we see that in these models with wage dualism and no unemployment, as in the other segmented labor market models, employment and wages in *each* sector of the economy are determined by labor market conditions in *all* sectors of the economy. Partial equilibrium analysis simply cannot explain what we see.

C. Models with Wage Differentials and Unemployment: The Harris-Todaro Model and Extensions of It

In 1970, a major alternative was developed in the context of East Africa. John Harris and Michael Todaro (1970) formulated a model characterized by wage dualism and unemployment. Wage dualism arises in their model because employers in the formal sector are compelled by unions, minimum wage laws, or other institutional forces to pay higher-than-market-clearing wages, while the wage clears the informal sector labor market. The Harris-Todaro model also featured a spatial distinction: to be hired for a formal sector job, it was necessary for a worker to be physically present in the urban areas where the formal sector jobs are assumed to be located.

In the Harris-Todaro model, more workers search for formal sector jobs than are hired. Employers hire some of the searchers but not all of them. Those not hired end up unemployed *ex post*. Open unemployment, though a feature of the world, was not a feature of the models reviewed in the last two subsections.

More specifically, the Harris-Todaro labor market operates as follows. Employers in the formal sector hire workers until the point where the marginal product of labor equals the institutionally-determined wage \bar{W}_F . On the other hand, in the informal sector, there is assumed to be free entry; thus, all persons who wish to work in the informal sector may do so. Each person employed in the informal sector earns a wage $W_I < \bar{W}_F$.

Workers are assumed to consider the mathematical expected wages from each of two search strategies: (1) Searching for a formal sector job, which pays a relatively high wage but runs the risk of unemployment, and (2) Taking an informal sector job, which offers a low wage with no risk of unemployment. Harris and Todaro's insight was that workers would be expected to allocate themselves between formal sector and informal sector search strategies *so that the mathematical expected wages from the two search strategies are equalized: $E(W_F) = E(W_I)$* . Let E_F denote employment in the formal sector

and L_F the labor force in the formal sector. In the basic Harris-Todaro model, expected wage equalization leads to the following equilibrium condition:

$$\bar{W}_F \frac{E_F}{L_F} = \bar{W}_I.$$

Because $\bar{W}_F > \bar{W}_I$, it follows that $\frac{E_F}{L_F} < 1$, i.e., the formal sector labor force exceeds formal sector employment, and therefore a Harris-Todaro equilibrium is characterized by open unemployment.

Harris and Todaro's fundamental contribution was to build a model with wage dualism *and* unemployment. The fact that the model remains part of our toolkit more than three decades later is a tribute to its basic insight and analytical power.

At the same time, some of the assumptions of the Harris-Todaro model were judged to be too restrictive, and so the model was generalized in the years that followed to nest their specific formulation within a broader framework. Their model was first extended by Fields (1975), which allowed for on-the-job search from rural agriculture, the existence of an urban informal sector, preferential hiring of the better-educated, and employment fixity. The model has subsequently been extended and generalized to allow for duality within the rural sector, mobile capital, endogenous urban wage setting, risk-aversion, a system of demand for goods, and many other factors (Corden and Findlay, 1975; Calvo, 1978; Moene, 1988, 1992; Khan, 1989; Fields, 1989; Chakravarty and Dutta, 1990; Bourguignon, 1990; Basu, 1997).

V. Contributions of these Models to Understanding and Policy Analysis

Positive economic analysis is about “what is” and “what will be.” Normative economic analysis is about “what should be.” The remainder of this section demonstrates how the segmented labor market models presented above contribute to understanding (the concern of positive economics) and to policy analysis (the concern of normative economics). What all of the cases presented here have in common is that it is difficult if not impossible to make sense of them using single sector models only.

A. Why an Increase in Productivity Might Cause Wages to Fall

The integrated labor market model maintains that the wage that any worker receives reflects supply and demand for labor in the labor market as a whole. Specifically, the wage is determined by what the last employer is willing to pay in order to attract and employ a worker and by what the last worker requires in order to be attracted and employed.

One common misperception is that the wage “should” vary directly with labor “productivity”, commonly measured as value added per worker, in a given firm or sector. According to the integrated labor market model, nothing could be further from the truth. The following example illustrates why.

Suppose that computers become available which enable half the formal sector workers to be replaced at lower cost while keeping total output constant. “Productivity,” measured by value added per worker, approximately doubles (“approximately,” because productivity would exactly double if the computers cost the same as the dismissed workers, more than double if the computers cost less). However, according to the integrated labor market model, formal sector wages will *not* double nor necessarily even increase. In fact, to the extent that wages change at all, they are likely to *fall*. Here is why.

The availability of computers that can replace workers induces a substitution effect, which would be expected to result in less labor being demanded to produce a given level of output. For this reason, labor demand would fall approximately in half. Unless the ability to produce at lower cost induces a corresponding increase in the scale of production, the demand for labor in the formal sector will be less after the productivity improvement than it was before.

When the equilibrating forces in labor markets equalize wages across the various sectors, as the integrated labor market model maintains, a leftward shift of labor demand in one sector of the economy results in lower wages in *all* sectors of the economy. The reason that an increase in “productivity” does not result in higher wages is that employers do not need to pay wages as high as before to attract the desired number of workers. In this way, higher productivity can result in *lower* wages.

The more general point of the integrated labor market model is that a worker’s wage is set not just by that worker’s own productivity nor by labor productivity just in that worker’s sector. Rather, wages are set by supply and demand for that category of labor in the labor market as a whole.

B. Why Taiwan’s Economic Growth Led First to Falling Unemployment at Constant Wages and then to Economy-Wide Wage Increases at Full Employment

The Lewis model’s characterization of intersectoral linkages generated two major predictions. The first is that as long as there exists a surplus of labor to the formal sector, economic growth would generate intersectoral shifts of employment but little or no increase in real wages. The second prediction is that once the unlimited supply of labor to the formal sector is exhausted and the turning point is reached, subsequent economic growth is marked by rising real wages economy-wide.

The model proved to be remarkably prescient. Take the case of Taiwan, where manufactured exports were the engine of growth. Data on unemployment and real wages

(monthly) are displayed in Figure 3. At the time Lewis was writing, the open unemployment rate was 6.3%, higher than the generally agreed-upon level of full employment. In the next six years of Taiwan's economic growth, unemployment fell to 4.3% and real wages in manufacturing rose by only 2% (total, not per year), consistent with excess labor continuing to be supplied relative to the amount demanded. But then, in the next decade (the 1960s), unemployment fell to 1.5% - a rate indicating severe labor shortages – and real wages shot up by 81%. Unemployment remained below 2% in the 1970s and 1980s, then rose to 3% in the 1990s. At the same time, real wages doubled again in the 1970s and 1980s and rose by another 36% in the 1990s, not only in manufacturing but throughout the Taiwanese labor market.

The two phases predicted by Lewis appear clearly in the data for Taiwan: falling unemployment at essentially constant wages, then rapidly rising real wages at full or over-full employment. The dualistic model with intersectoral linkages tells a compelling story, and it did it *before* it happened.

C. Why the Solution to Urban Unemployment in Kenya was not Urban Employment Creation but Rural Development

The Harris-Todaro model was formulated in response to the emergence of serious unemployment in urban Kenya. The government of Kenya tried a policy of urban employment creation, which appeared not to have worked. Harris and Todaro's model helped explain why.

The model produced two powerful policy results. The first concerned a policy of formal sector employment creation to employ the unemployed (who, in the Harris-Todaro model, were all in urban areas, because that is where the formal sector jobs were assumed to be located). A policy of increasing formal sector employment by $\Delta E_F = E_F' - E_F$ increases the formal sector labor force by $\Delta E_F \frac{\bar{W}_F}{\bar{W}_I}$ and increases open unemployment by

$\Delta E_F \left(\frac{\bar{W}_F}{\bar{W}_I} - 1 \right)$. Thus, the solution to urban unemployment is *not* urban employment creation.

The second policy option considered was a policy of rural development. Suppose that such a program could increase the (rural) informal sector wage from \bar{W}_I to \bar{W}_I' .

From the H-T equilibrium condition, unemployment would then *fall* from

$UNEM = E_F \left(\frac{\bar{W}_F}{\bar{W}_I} - 1 \right)$ to $UNEM' = E_F \left(\frac{\bar{W}_F}{\bar{W}_I'} - 1 \right)$. Thus, in the Harris-Todaro model, the

solution to urban unemployment is *rural* development.

Soon after the model was published, the government of Kenya followed the Harris-Todaro precepts by putting into place an integrated rural development program. Indeed, unemployment in Kenya *did* indeed fall. For a more comprehensive welfare economic analysis of various policy options in the Harris-Todaro model, see Fields (2005).

D. Why Poverty Is So Severe in the Urban Informal Sector of Many Developing Countries

Some of the worst poverty in developing countries is found in urban areas (United Nations, 2003). The extended Harris-Todaro model helps explain why this is so.

One way in which the Harris-Todaro model has been extended is to have three employment sectors (formal employment, urban informal employment, and agricultural employment) in addition to unemployment. What is important in this particular extension is that the three sectors are located in distinct locations: formal employment and urban informal employment in the urban areas, agricultural employment in the rural areas. In this extension, there are three search strategies: (1) Search for a formal sector job full-time while unemployed. (2) Give up on the search for a formal sector job and be employed in agriculture. (3) Search for a formal sector job part-time while informally

employed. (Note that search strategies (1) and (2) were those modeled by Harris and Todaro.)

Presumably, those people located nearer to where the formal sector jobs are stand a better chance of being hired for any given job vacancy. The fact that they do has implications for urban informal wages.

For strictly positive numbers of people to choose each of the three search strategies, the extended Harris-Todaro equilibrium requires that expected wages equalize across the three search strategies. If one group of informal sector workers has better on-the-job search opportunities than another, the labor market equilibrium must be one where the group with the better on-the-job search opportunities ends up with a *lower* wage in equilibrium.

Viewed in this way, it is not surprising that some of the worst poverty in the developing world would be found in the *urban* areas: the urban poor consist at least in part of those who sought urban formal sector jobs but who were unlucky enough not to be hired for them. Of course, there is another reason for very low urban informal sector wages – lack of opportunities for wage employment and self-employment in rural areas – which the extended Harris-Todaro explanation complements.

E. Why Expanding Formal Sector Employment Sometimes Improves Labor Market Conditions and Sometimes Does Not

The same policy can have different effects in the different models. Take the policy of expanding employment in the formal sector.

In all three of the segmented labor market models, when more formal sector jobs are created, labor moves from the informal sector to the formal sector to take up the available jobs. Those individuals who are able to make the move are better off in all three models.

The models differ, however, in other respects. In the integrated labor market model, wages rise by equal amounts in the two sectors. All workers are better off by the same amount.

In the model with wage dualism and no unemployment, the wage level in the formal sector stays the same, but wages rise to some degree in the informal sector. Thus, the remaining informal sector workers benefit from the expansion of employment in the formal sector.

Finally, in the Harris-Todaro model with wage dualism and unemployment, the informal sector wage may or may not rise. What will surely rise, though, is unemployment. In this model, unlike the other ones, formal sector employment creation produces both winners and losers.

Thus, whether a policy of formal sector employment creation has favorable labor market effects depends on which labor market model best fits a particular country's institutional circumstances.

VI. A Final Word

I shall now try to summarize the main points of this review and offer some brief concluding thoughts.

First, differences between the various sectors' labor markets appear pervasive. When possible, Occam's razor suggests limiting the analysis to two sectors. But when two sectors are simply not enough, three-sector or n-sector models can and have proved insightful.

Second, formal sector labor markets can be formulated in several alternative ways: in terms of market-clearing, in terms of institutional wage-setting, in terms of efficiency wages, and in terms of worker-side resistance to wage cuts.

Third, informal sector labor markets can be modeled as a free-entry sector, as a desirable sector, or as having its own internal duality.

Fourth, the linkages between the different sectors can be modeled in a number of ways: as an integrated labor market, as a model of wage dualism and no unemployment, and as a model of wage dualism with unemployment.

Fifth, a number of external events and policies can be understood only by using segmented labor market models. In these models, employment and wages in *each* sector of the economy are determined by labor market conditions in *all* sectors of the economy. Single-market analysis simply cannot explain what we see.

The number of possible models combining these various components is enormous. Each of the three components – formal sector labor market, informal sector labor market, and intermarket linkages – has three or four alternatives. Even this relatively coarse categorization results in thirty-six different labor market models.

No analyst would expect that the same model would fit East Africa and East Asia, South Africa and South Korea. Surely, the “correct” model is context-specific. Blending empirical observation and analytical modeling has yielded great advances. Coming up with the “correct” model matters for more than understanding; it matters for policy purposes as well. Sound labor market policies require sound labor market models.

References

- Aidt, Toke and Zafiris Tzannatos (2002). Unions and Collective Bargaining: Economic Effects in a Global Environment. (Washington: World Bank).
- Akerlof, George and Janet Yellen, eds. (1986). Efficiency Wage Models of the Labor Market. (Cambridge: Cambridge University Press).
- Anand, Sudhir, and Vijay Joshi (1979) "Domestic Distortions, Income Distribution and the Theory of Optimum Subsidy," The Economic Journal, June.
- Anand, Sudhir and S.M.R. Kanbur (1985). "Poverty Under the Kuznets Process," Economic Journal 95 (supp.): 42-50.
- Anand, Sudhir and S.M.R. Kanbur (1993). "The Kuznets Process and the Inequality-Development Relationship," Journal of Development Economics 40: 25-52.
- Ayub, Mahmood (1977). Income Inequality in a Growth-Theoretic Context: The Case of Pakistan. Unpublished Ph.D. dissertation, Yale University.
- Balán, J., H.L. Browning, and E. Jelin (1973). Men in a Developing Society. (Austin, TX: Institute of Latin American Studies, University of Texas at Austin).
- Bardhan, P.K. and A. Rudra (1981). "Terms and Conditions of Labour Contracts in Agriculture: Results of a Survey in West Bengal 1979," Oxford Bulletin of Economics and Statistics 89-111.
- Basu, Kaushik (1997). Analytical Development Economics. (Cambridge, MA: MIT Press).
- Becker, Gary S. (1962). "Investment in Human Capital," Journal of Political Economy. LXX Supplement: 9-49.
- Becker, Gary S. (1964). Human Capital. (New York: Columbia University Press for the National Bureau of Economic Research).
- Bell, Clive (1991). "Regional Heterogeneity, Migration, and Shadow Prices," Journal of Public Economics, October.
- Bergmann, Barbara (1971). "The Effect on White Incomes of Discrimination in Employment," Journal of Political Economy 79: 294-313.
- Besley, Tim and Robin Burgess (2004). "Can Labor Regulation Hinder Economic Performance? Evidence from India," Quarterly Journal of Economics 119: 91-134.
- Blanchard, Olivier and Lawrence Summers (1986). "Hysteresis and the European Unemployment Problem," NBER Macroeconomics Annual 15-78.

Bliss, Christopher J. and Nicholas H. Stern (1978). "Productivity, Wages and Nutrition: 1. The Theory; 2. Some Observations," Journal of Development Economics 5: 363-398.

Boudon, Raymond A. (1973). Mathematical Structures of Social Mobility. (Amsterdam: Elsevier).

Bourguignon, François (1990). "Growth and Inequality in the Dual Model of Development: The Role of Demand Factors," Review of Economic Studies 57: 215-228.

Bowles, S. and Heintz, J. (1996) "Wages and Jobs in the South African Economy: An Econometric Investigation," Department of Economics, University of Massachusetts, processed, January.

Bravo, David (1995). "Cost of Labor Standards in Bolivia," Harvard University, processed.

Burdett, Kenneth and Dale Mortensen (1998). "Wage Differentials, Employer Size, and Unemployment," International Economic Review 39: 257-273.

Cahuc, Pierre and André Zylberberg (2004). Labor Economics. (Cambridge, MA: MIT Press).

Calvo, Guillermo A. (1978). "Urban Unemployment and Wage Determination in LDC's: Trade Unions in the Harris-Todaro Model," International Economic Review, February.

Chakravarty, Satya R. and Bhaskar Dutta (1990). "Migration and Welfare," European Journal of Political Economy.

Chang, C.-H. (1989). "A Study on the Labor Market in Taiwan," in Chung-Hua Institution for Economic Research, Conference on Labor and Economic Development, Taipei.

Corden, Max and Ronald Findlay (1975). "Urban Unemployment, Intersectoral Capital Mobility and Development Policy," Economica, February.

Cunningham, Wendy V. and William F. Maloney (2001). "Heterogeneity in the Mexican Micro-Enterprise Sector," Economic Development and Cultural Change 50: 131-156.

Dasgupta, Partha and Debraj Ray (1986). "Inequality as a Determinant of Malnutrition and Unemployment," Economic Journal 96: 1011-1034.

DeSoto, Hernando (1989). The Other Path. (New York: Harper and Row).

Dixit, Avinash K. (1973). "Models of Dual Economies" in James A. Mirrlees and Nicholas H. Stern, eds., Models of Economic Growth: Proceedings of a Conference Held

by the International Economic Association at Jerusalem. (New York: International Economic Association).

Doeringer, Peter B. and Michael J. Piore (1971). Internal Labor Markets and Manpower Analysis. (Lexington, MA: Heath).

Drèze, Jean and Anandita Mukherjee (1989). "Labour Contracts in Rural India: Theories and Evidence," in S. Chakravarty, ed., The Balance Between Industry and Agriculture in Economic Development. (London: Macmillan).

Dunlop, John T. (1944). Wage Determination Under Trade Unions (New York: Macmillan).

Eaton, Jonathan and Mark Gersovitz (1984). "A Theory of Expropriation and Deviations from Perfect Capital Mobility," Economic Journal 94: 16-40.

Ehrenberg, Ronald G. and Robert S. Smith (2003). Modern Labor Economics. (Boston: Addison Wesley).

Fallon, Peter and Robert Lucas (1991). "The Impact of Changes in Job Security Regulations in India and Zimbabwe," World Bank Economic Review. September.

Fallon, Peter and Robert Lucas (1993). "Job Security Regulations and the Dynamic Demand for Industrial Labor in India and Zimbabwe," Journal of Development Economics.

Fei, John C.H. and Gustav Ranis (1964). Development of the Labor Surplus Economy. (Homewood, IL: Irwin).

Fei, John C.H., Gustav Ranis, and Shirley W.Y. Kuo (1978). "Growth and Family Distribution of Income by Factor Components," Quarterly Journal of Economics 92: 17-53.

Fei, John C.H., Gustav Ranis, and Shirley W.Y. Kuo (1979). Growth with Equity: The Taiwan Case. (Oxford: Oxford University Press).

Fields, Gary S. (1975). "Rural-Urban Migration, Urban Unemployment and Underemployment, and Job Search Activity in LDC's," Journal of Development Economics 2: 165-188.

Fields, Gary S. (1979a). "Income Inequality in Urban Colombia: A Decomposition Analysis," Review of Income and Wealth 25: 327-341.

Fields, Gary S. (1979b). "A Welfare Economic Approach to Growth and Distribution in the Dual Economy," Quarterly Journal of Economics 372: 325-354.

- Fields, Gary S. (1980). Poverty, Inequality, and Development. (New York: Cambridge University Press).
- Fields, Gary S. (1984). "Employment, Income Distribution and Economic Growth in Seven Small Open Economies," The Economic Journal, 94: 74-83.
- Fields, Gary S. (1989). "On-the-Job Search in a Labor Market Model: Ex-Ante Choices and Ex-Post Outcomes," Journal of Development Economics, 30: 159-178.
- Fields, Gary S. (1990). "Labour Market Modeling and the Urban Informal Sector: Theory and Evidence," in David Turnham, Bernard Salomé, and Antoine Schwarz, eds., The Informal Sector Revisited. (Paris: Development Centre of the Organisation for Economic Co-Operation and Development).
- Fields, Gary S. (1994). "Changing Labor Market Conditions and Economic Development in Hong Kong, the Republic of Korea, Singapore, and Taiwan, China," The World Bank Economic Review 8: 395-414.
- Fields, Gary S. (1999). "Employment Generation and Poverty Alleviation in Developing Economies," in Ulrich Hiemenz, ed., Growth and Competition in the New Global Economy. (Paris: OECD Development Centre).
- Fields, Gary S. (2005). "A Welfare Economic Analysis of Labor Market Policies in the Harris-Todaro Model," Journal of Development Economics 76: 127-146.
- Fields, Gary S. and Walter S. Bagg (2003). "Long-Term Mobility and the Private Sector in Developing Countries: New Evidence," in Gary S. Fields and Guy Pfeffermann, eds., Pathways Out of Poverty. (Boston: Kluwer Academic Publishers).
- Fields, Gary S., Paul L. Cichello, Samuel Freije, Marta Menéndez, and David Newhouse (2003). "Household Income Dynamics: A Four-Country Story," Journal of Development Studies 40: 30-54.
- Fields, Gary S., Murray Leibbrandt, and Jeremy Wakeford (2000). Key Labour Market Elasticities. Report prepared for the Government of South Africa.
- Fields, Gary S. and Jennifer O'Hara Mitchell (1999). "Changing Income Inequality in Taiwan: A Decomposition Analysis," in Gary Saxonhouse and T.N. Srinivasan, eds., Development, Duality, and the International Economic Regime: Essays in Honor of Gustav Ranis. (Ann Arbor, MI: University of Michigan Press).
- Fields, Gary S. and Henry Y. Wan, Jr. (1989). "Wage-Setting Institutions and Economic Growth," World Development.
- Ghose, Ajit K. (2003). Jobs and Incomes in a Globalizing World. (Geneva: ILO).

- Gindling, T.H., Jr. (1991). "An Investigation into Labor Market Segmentation: The Case of San José, Costa Rica," Economic Development and Cultural Change.
- Hamermesh, Daniel (1993). Labor Demand. (Princeton: Princeton University Press).
- Harberger, Arnold C. (1971). "On Measuring the Social Opportunity Cost of Labor," International Labour Review 103: 559-579.
- Harris, John and Michael Todaro (1970). "Migration, Unemployment, and Development: A Two Sector Analysis," American Economic Review, 40: 126-142.
- Hart, Keith (1973). "Informal Income Opportunities and Urban Employment in Ghana," Journal of Modern African Studies. 11: 61-89.
- Heady, Christopher J. (1981). "Shadow Wages and Induced Migration," Oxford Economic Papers.
- House, William J. (1984). "Nairobi's Informal Sector: Dynamic Entrepreneurs or Surplus Labor?" Economic Development and Cultural Change 32: 277-302.
- Hymer, Stephen and Stephen Resnick (1969). "A Model of an Agrarian Economy with Nonagricultural Activities," American Economic Review. 59: 493-506.
- Inter-American Development Bank (2003). Good Jobs Wanted. (Washington: Inter-American Development Bank).
- International Labour Office (1972). Employment, Incomes, and Equality: A Strategy for Increasing Productive Employment in Kenya. (Geneva: ILO).
- International Labour Office (1998). Labour and Social Issues Relating to Export Processing Zones. (Geneva: ILO).
- International Labour Office (2002). Women and Men in the Informal Economy: A Statistical Picture. (Geneva: ILO).
- International Labour Office (2003). Global Employment Trends 2003. (Geneva: ILO).
- International Labour Office (2005). Global Employment Trends 2005. (Geneva: ILO).
- Jhabvala, Renana, Ratna M. Sudarshan, and Jeemol Unni (2003). Informal Economy Centrestage: New Structures of Employment. (New Delhi: Sage Publications).
- Jorgenson, Dale W. (1967). "Surplus Agricultural Labour and the Development of a Dual Economy," Oxford Economic Papers 19: 288-312.

Khan, M. Ali (1989). "The Harris-Todaro Model," in J. Eatwell et al., The New Palgrave. (London: MacMillan).

Knight, John B. (1976). "Explaining Income Distribution in Less Developed Countries: A Framework and an Agenda," Bulletin of the Oxford Institute of Economics and Statistics.

Kuznets, Simon (1955). "Economic Growth and Income Inequality," American Economic Review 45: 1-28.

Lee, Chang-Hee (2002). "The Minimum Wage," Asian Labour Update 42.

Leibenstein, Harvey (1957). "Underemployment in Backward Economies," Journal of Political Economy 65.

Levy, Santiago (2008). "Can Social Programs Reduce Productivity and Growth? A Hypothesis for Mexico," Working Paper, Inter-American Development Bank.

Lewis, W. Arthur (1954). "Economic Development with Unlimited Supplies of Labour," Manchester School, 22: 139-191.

Lewis, W. Arthur (1972). "Reflections on Unlimited Labor," in Luis DiMarco, ed., International Economics and Development: Essays in Honor of Raul Prebisch. (New York: Academic Press), 75-96.

Lin, C.-C. (1989). "The Basic Labor Standards Law and Operation of Labor Market: Theory and Partial Empirical Results," Academia Sinica.

Lindbeck, Assar and Dennis Snower (1988). Insider-Outsider Theory of Employment and Unemployment. (Cambridge, MA: MIT Press).

Maloney, William F. (2003). "Informal Self-Employment: Poverty Trap or Decent Alternative," in Gary S. Fields and Guy Pfeffermann, eds., Pathways Out of Poverty. (Boston: Kluwer).

Marcouiller, Douglas, Veronica Ruiz de Castilla, and Christopher Woodruff (1997). "Formal Measures of the Informal Sector Wage Gap in Mexico, El Salvador, and Peru," Economic Development and Cultural Change 45: 367-392.

Meade, James A. (1976). The Just Economy. (London: Allen and Unwin).

Merton, Robert K. (1968). "The Matthew Effect in Science," Science 159: 56-63.

Mincer, Jacob (1962). "On-the-Job Training: Costs, Returns, and Some Implications," Journal of Political Economy. LXX Supplement: 50-79.

- Mincer, Jacob (1974). Schooling, Experience, and Earnings. (New York: Columbia University Press for the National Bureau of Economic Research).
- Mirrlees, James A. (1975). "Pure Theory of Underdeveloped Economies," in Lloyd G. Reynolds, ed., Agriculture in Development Theory. (New Haven: Yale University Press).
- Moene, Karl Ove (1988). "A Reformulation of the Harris-Todaro Mechanism with Endogenous Wages," Economics Letters.
- Moene, Karl Ove (1992). "Poverty and Landownership," American Economic Review, March.
- Moran, Theodore H. (2002). Beyond Sweatshops: Foreign Direct Investment and Globalization in Developing Countries. (Washington, DC: Brookings Institution).
- Nelson, Richard R. (1966). "A Theory of the Low-Level Equilibrium Trap in Underdeveloped Economies," American Economic Review 46: 894-908.
- Noe, Raymond A., John R. Hollenbeck, Barry Gerhart, and Patrick M. Wright (2000). Human Resource Management. (Boston: Irwin McGraw-Hill).
- Osmani, S.R. (1991). "Wage Determination in Rural Labour Markets: The Theory of Implicit Co-Operation," Journal of Development Economics 34: 3-23.
- Park, Young-Bum (1991). "Union/Non-Union Wage Differentials in the Korean Manufacturing Sector," International Economic Journal.
- Raff, Daniel and Lawrence Summers (1987). "Did Henry Ford Pay Efficiency Wages?" Journal of Labor Economics.
- Ranis, Gustav (2006). "Is Dualism Worth Revisiting?" in Alain de Janvry and Ravi Kanbur, eds., Poverty, Inequality, and Development: Essays in Honor of Erik Thorbecke.
- Ranis, Gustav and John C. H. Fei (1961). "A Theory of Economic Development," American Economic Review, 51: 533-565.
- Ranis, Gustav and Frances Stewart (1993). "Rural Nonagricultural Activities in Development: Theory and Application," Journal of Development Economics 40: 75-101.
- Ranis, Gustav and Frances Stewart (1999). "V-Goods and the Role of the Urban Informal Sector in Development," Economic Development and Cultural Change 47: 259-288.
- Robinson, Sherman (1976). "A Note on the U Hypothesis Relating Income Inequality to Economic Development," American Economic Review 437-440.

Rosenzweig, Mark (1988). "Labor Markets in Low Income Countries," in Hollis Chenery and T.N. Srinivasan, eds., Handbook of Development Economics, Volume 1. (Amsterdam: North Holland).

Roy, A.D. (1951). "Some Thoughts on the Distribution of Earnings," Oxford Economic Papers. 3: 135-146.

Sah, Raaj Kumar and Joseph E. Stiglitz (1985). "The Social Cost of Labor and Project Evaluation: A General Approach," Journal of Public Economics.

Schultz, T.W. (1961). "Investment in Human Capital," American Economic Review. LI: 1-17

Schultz, T.W. (1962). "Reflections on Investment in Man," Journal of Political Economy. LXX Supplement: 1-8.

Sen, Amartya K. (1967). "Review of J.C.H. Fei and G. Ranis, Development of the Labor Surplus Economy: Theory and Policy," The Economic Journal 77: 346-349.

Shapiro, Carl and Joseph E. Stiglitz (1984). "Equilibrium Unemployment as a Worker Discipline Device," American Economic Review.

Solow, Robert (1990). The Labour Market as a Social Institution. (Oxford: Blackwell).

Spinanger, Dean (1985). "The Labor Market in Panama: An Analysis of the Employment Impact of the Labor Code," Paper presented at the Seminar on Employment Policy in Latin America, Panama.

Squire, Lyn (1981). Employment Policy in Developing Countries (New York: Oxford University Press for the World Bank).

Starr, Gerald (1981). Minimum Wage Fixing. (Geneva: ILO).

Stiglitz, Joseph E. (1974). "Alternative Theories of Wage Determination and Unemployment in LDCs" The Labour Turnover Model, Quarterly Journal of Economics 88: 194-227.

Stiglitz, Joseph E. (1976). "The Efficiency Wage Hypothesis, Surplus Labor, and the Distribution of Labour in LDCs," Oxford Economic Papers 28: 185-207.

Stiglitz, Joseph E. (1982). "The Structure of Labor Markets and Shadow Prices in LDCs," in Richard H. Sabot, ed., Migration and the Labor Market in Developing Countries. (Boulder, CO: Westview).

Stiglitz, Joseph E. (1985). "Equilibrium Wage Distributions," The Economic Journal. 95: 595-618.

Tidrick, Gene (1975). "Wage Spillover and Unemployment in a Wage-Gap Economy: The Jamaican Case," Economic Development and Cultural Change.

Todaro, Michael P. (1969). "A Model of Labor Migration and Urban Unemployment in Less Developed Countries," American Economic Review 39: 138-148.

Tokman, Victor (1987). "El Sector Informal: Quince Años Despues," PREALC.

Turnham, David (1971). The Employment Problem in Less Developed Countries. (Paris: Development Centre of the Organisation for Economic Co-Operation and Development).

Turnham, David (1993). Employment and Development: A New Review of Evidence. (Paris: Development Centre of the Organisation for Economic Co-Operation and Development).

Ulrich, Dave (1997). Human Resource Champions. (Boston: Harvard Business School Press).

United Nations (2003). The Challenge of Slums: Global Report on Human Settlements 2003. (New York: United Nations Human Settlements Programme).

Weiss, Andrew (1990). Efficiency Wages. (Princeton: Princeton University Press).

Welch, Finis (1974). "Minimum Wage Legislation in the United States," Economic Inquiry 12: 285-318.

World Bank (1990). World Development Report 1990. (Washington: World Bank).

World Bank (1995). World Development Report 1995. (Washington: World Bank).

Yoo, Gyeongjoon (1995). An Analysis and Decomposition of Changing Labor Income Distribution in Korea. Unpublished Doctoral Dissertation, Cornell University.

Figure 1.
The Standard Market-Clearing Labor Market Model.

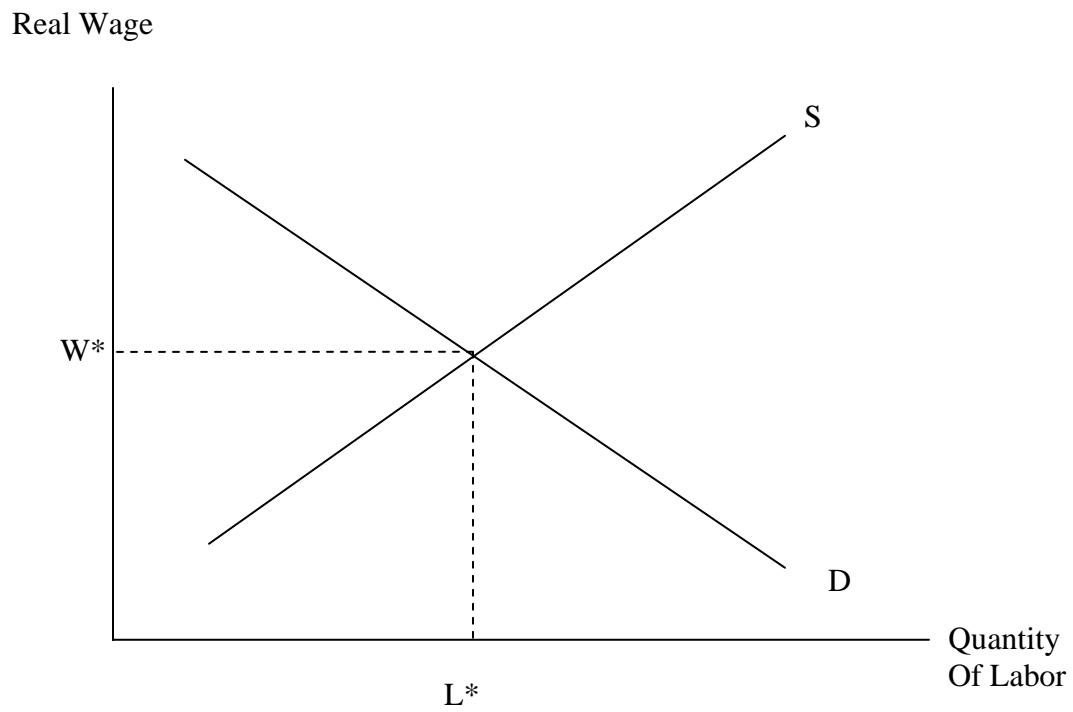


Figure 2.
The Integrated Labor Market Model:
A Higher Demand for Labor in One Sector
Raises Wages In all Sectors.

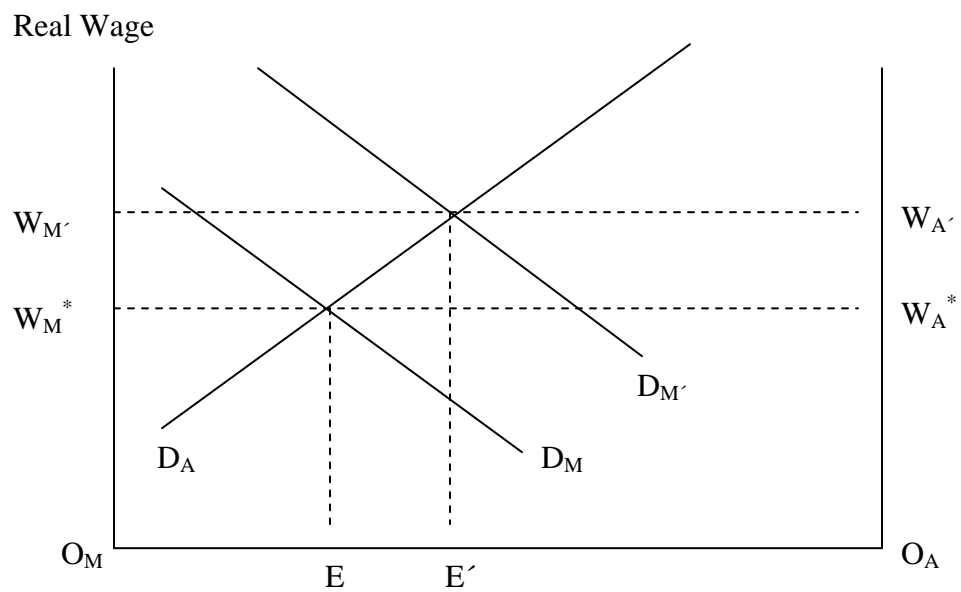
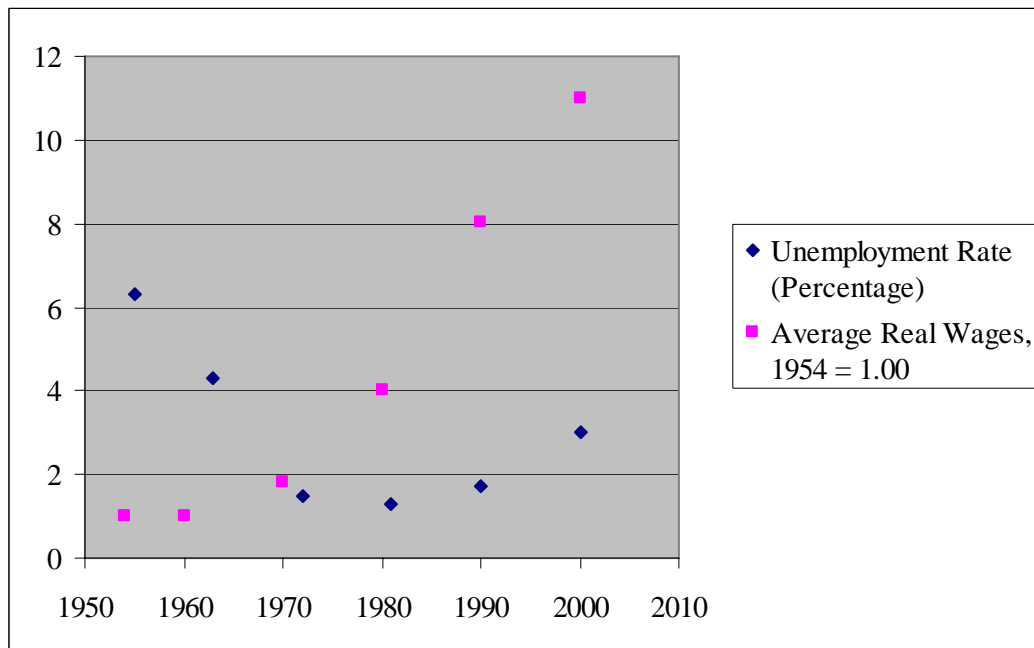


Figure 3.
Unemployment and Average Real Wages in Taiwan.



Sources: Fields (1984, 1994), Fields and Bagg (2003).