Return and Lifetime Migration in Low Income Countries

The Impact of Return Migration and the

Determinants of Lifetime Migration

# in Low Income Countries

By

Carlos A. Leite. M.A.

A Thesis

Submitted to the School of Graduate Studies

In Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

© Copyright by Carlos A. Leite, June 1995

DOCTOR OF PHILOSOPHY (1995) (Economics) McMASTER UNIVERSITY Hamilton, Ontario

TITLE: The Impact of Return Migration and the Determinants of Lifetime Migration in Low Income Countries

AUTHOR: Carlos Alberto da Cunha Leite, M.A. (University of Waterloo) B.A. (University of Waterloo)

ii

SUPERVISING COMMITTEE:

Stephen R. G. Jones (Supervisor) Martin Browning David H. Feeny

NUMBER OF PAGES: xi, 290

### ABSTRACT

To study the impact of return migration on the urban labour market, we extend the basic monitoring model of efficiency wages to include two urban sectors (formal and informal) and a rural sector, and a labour force composed of permanent (or resident) and temporary (or migrant) workers. The proposed model provides a micro-economic foundation for a rural-urban wage gap as well as an intra-urban formal-informal sectoral wage structure. Unlike previous models, our explanation is based on imperfect information on the part of the employer instead of the employee. The proposed model appears to conform closely with stylized facts as discussed in the first chapter.

The difference in expected tenure between resident and migrant workers provides a rationale for discrimination by formal sector employers in favour of resident workers, who then have an incentive to engage in signalling, usually by choosing unemployment over informal employment. In this way, the model suggests a non-human-capital based explanation for the incidence of luxury unemployment and predicts that migrants tend to be less likely to endure periods of unemployment. Under certain specified conditions, informal sector wages may be higher than formal sector wages, and informal sector wages may be negative.

Using the 1988 Malaysian Family Life Survey (MFLS-2), we provide a description of the migration process in Peninsular Malaysia. The data reveal that

three quarters of the moves are other than the standard rural-urban move, and that return migration is widespread, particularly for rural workers. Using cross tabulations, we find that geography and ethnicity variables and the level of education are important determinants of migration patterns. We also find the presence of life cycle effects with a common starting age of between eighteen and twenty.

To study the determinants of lifetime migration in a multivariate context, we estimate double hurdle event count models using MFLS-2. Double hurdle models distinguish between factors which affect the probability of migration from factors which affect the number of migrations chosen over a lifetime. Specification testing and in-sample forecasting confirm that double hurdle models perform significantly better than the standard Poisson and Negative Binomial models.

The availability of information about different facets of the individual's background allows us to control for factors overlooked in most studies. Specifically, we are able to disentangle the effects of parent's and own education and we find that both effects induce both higher participation rates and a greater number of moves. Studies which do not control for parent's education levels would then be likely to overestimate the effect of own education. We also find support for the effects of location specific capital, particularly in rural areas.

### ACKNOWLEDGEMENTS

I would like to thank my thesis committee, Drs. Stephen R. G. Jones, Martin Browning and David H. Feeny for their supervision. I appreciate their time, patience and understanding in helping to shape this dissertation.

I would like to thank Dr. Feeny for a memorable course on economic development, out of which sprang the second chapter, and for his continuing encouragement and support.

I would also like to extend my gratitude to Dr. Michael Veall who provided advice and support throughout my PhD program, and to Mrs. Aloma Munian in the Economics Department who was immensely helpful with everything administrative and who always had a smile, and sometimes a doughnut, to share.

Para os meus pais, que sempre encorajaram e suportaram as minhas actividades escolares, ofereço um enorme abraço e a minha eterna gratidão. Para a minha mãe e primeira professora, com quem eu aprendi as lições mais importantes, a minha dívida é incalculável. Obrigado também para os meus tios, António e Alberto, que nos deram a mão quando o mar estava capelado e que me ofereceram a oportunidade de chegar a esta meta.

Special thanks to my wife who withstood four (more) years of schooling and whose courage was instrumental in our triumph over the inevitable struggles and difficulties. Without your influence, it is unlikely I would have even embarked on this challenge. Without your rock-solid support, it is quite likely that the end result would have been considerably weaker. This dissertation is dedicated to you and your strength.

Special thanks also to Nicole and Jeremy, who learned what a thesis was before they themselves went to school and who endured many evenings of dad distracted by his work. I hope you too have the opportunity to reach your academic goals.

### TABLE OF CONTENTS

### Chapter 1. Migration and Economic Development: Analytical Insights and Empirical Evidence From the Literature

1. Introduction
2. Rural-Urban Migration in Sub-Saharan Africa
2.1 Migratory Patterns
2.2 Volume and Direction
2.3 Demographic Characteristics of Migrant Flows
2.4 Specification of the Utility Maximizing Unit
2.5 Reasons for Return Migration16
3. Income Differential Model
3.1 The Human Capital Approach
3.2 The Harris-Todaro Specification
4. The Risk Pooling Approach to Migration
5. Labour Market
6. The Harris-Todaro Model
7. Proposed Model
7.1 Stylized Migration Facts
7.2 Support from the literature
APPENDIX: THE 'TODARO PARADOX'
BIBLIOGRAPHY

### Chapter 2. Rural-Urban Migration With Efficiency Wages

1.	Introduction	7
2.	The Monitoring Model of Efficiency Wages	3
3.	Adding a Rural Sector	9
	3.1 Comparing the Two Models	3
	3.2 Comparative Statics	5
4.	Adding An Informal Sector	9
	4.1 Solving the Model	0
5.	Adding Return Migration	7
	5.1 Solving the Model	9
	5.2 Complete Screening Equilibrium	2
	5.3 Incomplete Screening Equilibrium	2
6.	Summary	9

APPENDIX A	134
APPENDIX B	136
BIBLIOGRAPHY	140

### Chapter 3. Migration Patterns in Malaysia

1. Introduction
2. Malaysia
3. Malaysian Family Life Survey-2 149
4. Demographics
4.1 Ethnic Composition
4.2 General Area of Origin
5. Migration Patterns
5.1 Rural-Urban Distribution
5.2 Ethnic Composition
6. Migration Spells
6.1 Rural-Urban Distribution171
6.2 Ethnic Composition
7. Age Distribution
8. Education
8.1 Educational Selectivity
9. Life Cycle Effects
10. Conclusion
APPENDIX A
APPENDIX B
APPENDIX C
APPENDIX D
APPENDIX E
APPENDIX F
APPENDIX G
APPENDIX H
APPENDIX I
APPENDIX J
APPENDIX K
APPENDIX L
BIBLIOGRAPHY211

Chapter 4.	Lifetime	Migration in	1 Malavsia:	An Event	<b>Count Approach</b>

1. Introduction
2. Literature Review
2.1 Limitations of Previous Studies
3. Methodology
3.1 Event Count Models
3.1.1 Poisson Regression Model
3.1.1.1 Unequal Periods of Observation
3.1.1.2 Poisson Restriction
3.2 Relaxing the Poisson Assumptions
3.2.1 The Zero Altered Poisson Model
3.2.2 The Negative Binomial Model
3.2.3 The Zero Altered Negative Binomial Model
3.3 Specification Testing
4. Econometric Model
4.1 Theoretical Considerations
4.2 Parameterizing the Participation Decision
4.3 Estimation Results
4.3.1 Parameterizing the Participation Decision
4.3.2 Zero Altered Models
4.3.3 Zero Altered Poisson Model: The Participation Equation 256
4.3.4 Zero Altered Poisson Model: The Frequency Equation 260
5. Further Research
6. Summary
APPENDIX A
APPENDIX B
APPENDIX C
APPENDIX D
APPENDIX E
APPENDIX F
APPENDIX G
APPENDIX H
APPENDIX I
APPENDIX J
BIBLIOGRAPHY

viii

### LIST OF FIGURES

### Chapter 1. Migration and Economic Development: Analytical Insights and Empirical Evidence From the Literature

Figure 1.	The Harris-Todaro	Model Accordin	ng to Corden (1	.974, p. 152)	36
Figure 2.	Effect of an Urban	Wage Subsidy			39

### Chapter 2. Rural-Urban Migration With Efficiency Wages

Figure B-1.	Complete Screening	137
Figure B-2.	Incomplete Screening	139

### Chapter 3. Migration Patterns in Malaysia

Figure 1. Age Distribution of Panel Sample	171
Figure 2. Cumulative Age at First Move	183
Figure A-1. Malaysian Regional and State Divisions	191
Figure L-1. Lifetime Migration Patterns for Rural-at-Age-15 Workers	209
Figure L-2. Lifetime Migration Patterns for Urban-At-Age-15 Workers	210

### Chapter 4. Lifetime Migration in Malaysia: An Event Count Approach

Figure 10.	Decision Process for Poisson Model	238
Figure 11.	Decision Process for ZAP Model	238
Figure C-1	· · · · · · · · · · · · · · · · · · ·	274
Figure G-1		282
Figure H-1		283

### LIST OF TABLES

### Chapter 3. Migration Patterns in Malaysia

Table 1. Gross Domestic Product Annual Growth Rates in Percentage (constant
prices)
Table 2. Follow-up Rates for Potential Panel Sample Members (MFLS-1
Women) by Ethnic Group and Urban/Rural Residence in 1976
Table 3. Rural-Urban Composition in Panel Sample of MFLS-2    155
Table 4. Ethnic Distribution by Rural/Urban 156
Table 5. Ethnic Distribution by Region of Residence at Age 15    157
Table 6. Destination and Origin of First Migration Moves    159
Table 7. Destination and Origin of Second Migration Moves      160
Table 8. Incidence of Return Migration (Second or Higher Migration Moves)
Table 9. Incidence of Return Migration (Second Migration Moves Only)
Table 10. Number of Migrations by Rural/Urban 163
Table 11. Spatial Distribution of Migrations by Area of Origin      164
Table 12. Rural/Urban Distribution by Ethnic Group 165
Table 13. Migration Participation Rates by Ethnic Group      166
Table 14. Participation Rates Among Malays 167
Table 15. Number of Migrations for Malay Workers 168
Table 16. Participation Rates by Ethnic Group for Urban Workers      169
Table 17. Return Migration Participation Rates by Ethnic Group      170
Table 18. Migration Spells by Area of Origin 171
Table 19. Censored Migration Spells by Rural/Urban Origin    172
Table 20. Age at Last Move by Rural/Urban Origin 173
Table 21. Average Age by Ethnic Group 175
Table 22. Average Age by Migration Status 176
Table 23. Average Age for Non-Malays by Migration Status 177
Table 24. Average Age for Indians by Migration Status 178
Table 25. Propensity to Migrate by Educational Group 180
Table 26. Direction of First Migration by Educational Group for Rural
Respondents
Table 27. Percentage of Migrants Making Their Last Move Within an Age Range
Table A-1. Comparative Geographic, Social and Economic Indicators I 192
Table A-2. Comparative Geographic, Social and Economic Indicators II      193
Table A-3. Comparative Geographic, Social and Economic Indicators III     194

Table B-1. Migration Participation Rates by Rural/Urban	195
Table B-2. Number of Migrations by Rural/Urban	195
Table B-3. Number of Migrations by Malay/Non-Malay	196
Table B-4. Distribution of the Number of Migrations by Ethnic Group	196
Table E-1. Average Age for Urban Workers by Ethnic Group	200
Table E-2.      Average Age for Malays by Rural/Urban Origin	200
Table F-1. Education Effects by Rural/Urban	201
Table G-1. Comparison of Years of Education by Residence at Birth	202
Table G-2. Comparison of Years of Education by Residence at Age 15	202
Table H-1. Comparison of Years of Education by Migration Status	203
Table H-2. Comparison of Years of Education by Migration Status for Ma	lays
	203
Table H-3. Comparison of Years of Education by Migration Status for No	n-
Malays	204

## Chapter 4. Lifetime Migration in Malaysia: An Event Count Approach

Table J-1.	Number of Migrations by 'WRK YONG	'
Table J-2.	Number of Migrations by 'U MV15'	

### CHAPTER 1.

# MIGRATION AND ECONOMIC DEVELOPMENT: ANALYTICAL INSIGHTS AND EMPIRICAL EVIDENCE FROM THE LITERATURE

### 1. Introduction

The labour market provides the setting for what certainly is one of the fundamental challenges facing developing nations: employment and wage policy. The backdrop is quite common across Africa, and much of Asia and Latin America: a substantial rural-urban wage gap, high rates of migration to urban areas in spite of large numbers of apparently unemployed urban dwellers, and an urban sector itself characterized by a sharp division in working conditions and wages between the formal and the informal sector.<sup>1</sup>

<sup>&</sup>lt;sup>1</sup> Kahnert (1987) provides data on the rural-urban income differential on page 62 and although a relatively large gap in wage rates has long been assumed by most theorists, there is actually a dearth of careful econometric evaluations which control for worker attributes such as skill differentials. See also Knight (1972) and Squire (1981).

Yap (1977) and Kahnert (1987) report that the percentage of urban growth attributable to migration is in the order of 35-60%. At the same time, the rate of growth of urbanization, persistently high in previous decades, almost doubled in the 1980s, as reported by World Bank (1990).

Kahnert (1987) reports that the rate of open unemployment has consistently been in the order of 7-10%. It bears emphasizing that such calculations are based on data which is often of less than sterling quality and compiled using less than fully comparable definitions. In particular, Knight (1972) notes that coverage is (continued...)

The rural labour market is largely comprised of self-employed and family labour. Although recent articles<sup>2</sup> have found an active market, a significant number of rural job openings are still seasonal with some permanent employment opportunities available. The consequent lack of long-term contractual commitment, the vast tracts of land covered along with the relative inadequacy of regulatory agencies in these regions make it very difficult for policy makers to significantly alter the outcome in the rural labour market.

Instead, social planners must turn their attention to the workings of the urban labour market: a geographically concentrated exchange encompassing most of the wage labourers in the country. Significantly, heterogeneity characterizes both the supply and the demand side of the urban labour market: job seekers

<sup>2</sup> See Adams (1991) for a recent characterization of the rural labour market in the Masvingo province of Zimbabwe. As reported by Rosenzweig (1988) for rural India, the spot labour market is the dominant institution, and the competitive model provides a reasonable approximation of behaviour.

<sup>&</sup>lt;sup>1</sup>(...continued)

questionable and generally not kept up to date, not only because of the lack of financial incentive to register but also due to the fact that a large number of positions are filled through personal contacts.

Kahnert (1987) reports on page 28 that "wages in urban areas differ greatly between high and low wage sectors, large and small firms...". Although Kahnert reports that a considerable portion of such wage differentials are explained by common labour quality variables, such as education and experience, it is also noted that the remaining variance in wages is still significant and may ultimately be attributable to "such attributes of workers as reliability, motivation, stability, health status and so on". Our proposed efficiency wage model is compatible with such an interpretation.

include both urban residents<sup>3</sup> and temporary (or circulatory) migrants; and urban firms can be effectively classified into two sectors, the high-wage, relatively capital and skill intensive formal sector and the low-wage, routine task oriented informal sector. With the formal sector wage typically characterized as fixed, the equilibrating mechanism becomes the migration response, and equilibrium occurs with a positive rate of unemployment.

The literature provides two main motivations for migrating: income differential (Todaro (1969) being the original article) and risk pooling (based largely on work by Oded Stark compiled in Stark (1991)). Although it is unlikely that these are mutually exclusive explanations,<sup>4</sup> it is important for policy makers to sort them out as the policy implications are quite different: wage subsidies or rural income augmentation are usually proposed for the first case while risk pooling calls for rural income stabilization. In the third and fourth sections, we review the income differential model, its predecessor, the human capital model, and the risk pooling approach.

<sup>&</sup>lt;sup>3</sup> We use the term 'residents' to refer to both native urbanites and to permanent migrants.

<sup>&</sup>lt;sup>4</sup> Notice, for example, that both motivations can be traced back to incomplete markets. If the urban economy is doing well, a rural dweller wishing to "buy into the action" is restricted to migration **if** the absence of capital markets does not allow participation through the purchase of shares. Similarly, the risk associated with rural income being largely dominated by agriculture could be mitigated, as it is to some extent in developed countries, if insurance contracts were available. A combination of agriculture and urban-based income performs a similar function.

To explain labour market operations, the literature provides at least four models. First, the standard neo-classical approach assumes flexible wages and fully competitive markets along with perfectly mobile factors of production. Second, the labour surplus model, initially proposed by Lewis (1954), assumes family income sharing and an allocation rule based on average instead of marginal product. Third, the institutional dual economy model posits a high minimum wage (set above the market clearing wage) that is enforced for only one of the two urban economic sectors. Finally, the labour turnover model endogenizes the urban wage by making it a function of the unemployment level and other sectoral wages. The fifth section briefly reviews each hypothesis while the sixth section provides a detailed exposition of the Harris-Todaro model, the dominant paradigm in migration research over the past two decades. The seventh section concludes with a stylized description of our proposed model.

The characteristics of migrant flows have important implications for social policy, particularly education, urban services and social security. We start with a brief review of the anthropological and economic literature on rural-urban migration in Africa,<sup>5</sup> the consensus choice as the region with the highest rates of

<sup>&</sup>lt;sup>5</sup> The focus on African evidence stems from the initial objective to investigate internal labour migration in Africa. After access to the desired African data set was denied, empirical work then focussed on Southeast Asia, a region with similar migration patterns, particularly in terms of the prevalence of return migration. The relevance of the evidence gathered remains unaltered, especially for the theoretical model developed in the companion chapter on return migration and efficiency wages.

migration, in order to establish some basic stylized facts about migration patterns, their direction and volume, constituent demographics and the choice of utility maximizing unit.

### 2. Rural-Urban Migration in Sub-Saharan Africa

#### 2.1 Migratory Patterns

Anthropological studies of rural life in sub-Saharan Africa carried out over the last 40 years point to the relevance of labour migration in general, and the prevalence of circulatory patterns in particular. For southern Tanzania, Gulliver (1955) notes on page 6 that "only a small minority of men have never migrated" while "the large majority of labour migrants spend between 9 and 18 months away from home on any one occasion" with most men undertaking more than one journey and wives generally staying at home; the first migration typically takes place prior to marriage. Consistent with the variability in return dates, Gulliver notes on page 13 that "the ordinary migrant feels very strongly that he must retain his freedom to choose and change his actual job, his place of work and area of employment, and the ability to return home as and when he feels he wishes to do so" and has a strong "dislike of engaging on contract terms, especially for a long period".<sup>6</sup>

<sup>&</sup>lt;sup>6</sup> The resulting high levels of job turnover became an early topic of interest in the economic and anthropological literature and led to Stiglitz's (1974) labour turnover model, as we point out in a later section.

For Tanzania and Uganda, Elkan (1967) reports on page 108 that by the mid-1960's "there is nothing to indicate that there has been any basic change in the pattern of circular migration", although it is possible that "the period spent in town has been getting longer". Later, Elkan (1975) reaches similar conclusions for Kenya. For Mali, Lewis (1985) reports on page 66 about the predominance of migration, especially "temporary wage employment" in the urban centres. Zachariah and Conde (1981) also present evidence of circulatory migration in an extensive area of West Africa, although as the authors note on page 9, there is some evidence that the "…average length of residence of migrants at destination is …increasing".

Similar findings by more recent surveys have led to a consensus position: unlike the rural-urban movement in industrializing Europe and, it seems, in Latin America and some Asian countries, rural-urban migration in Africa is, to a large extent, temporary (high rates of return migration and multiple journeys) and not of the onward, or sequential, type.<sup>7</sup> Consistent with this characterization, the vast majority of migrants in sub-Saharan Africa, unlike the earlier European and the current Latin American migrants, retain their rural ties and, as noted in Elkan

<sup>&</sup>lt;sup>7</sup> See, for example, Nelson (1975) and Zachariah and Conde (1981). Onward migration refers to successive moves between urban centres which are, typically, increasing in size and becoming more distant as the number of moves accumulates. As an explanation, Pessino (1991) has proposed a sequential theory of migration, under which it is hypothesized that migrants move, not only in response to wage differentials, but also to accumulate information about more distant places.

(1967, p. 113), "enjoy the security of a land holding and the addition to income that it brings."

According to Gulliver the main motive for migration was the relative lack of income earning opportunities in the home village combined with relatively high wages elsewhere. In general, the findings indicated that "areas of profitable crop production have a low rate of labour migration", as noted by Gulliver (1955) on page 25, and that "by labour migration men are able to earn and bring home money and goods that under present conditions... they are unable, or feel they are unable, to obtain here", as noted on page 42 of Gulliver (1955). Given the general proximity of the areas investigated by Gulliver as well as a common risk profile, with the weather being the main source of uncertainty, we would expect similar migration rates if risk pooling was the dominant motive for migration. Instead, Gulliver finds that areas with less income, or a larger income differential with relation to the city, have a higher migration rate. These findings are unanimously echoed by other anthropological as well as economic studies. Stark (1991) notes on page 29 that estimates of models of labour migration "uniformly support the hypothesis that individuals respond to income incentives in making decisions to migrate". Yap (1977) provides corroborating evidence collected from an extensive review of migration studies, and labour surveys uniformly find significant urban-rural wage differences.<sup>8</sup>

<sup>&</sup>lt;sup>8</sup> See Gulliver (1955) for early evidence and, later, Yap (1977), Sabot (continued...)

As for the impact of migration on rural society, Gulliver (1955) notes on page 41 that "the recourse to labour migration as a source of income saps the efforts and will ... to work more diligently at home in developing the resources of their own fields". This is, of course, a reference to the intertemporal allocation of leisure in the presence of a high and a low wage area. Moreover, Gulliver reports that "the individual peasant-family economy is little affected by the ordinary short-term absence of the husband-father, nor do wives and children encounter grave difficulties. In the short view the family gains from the money and goods brought back by the returning migrant...". This is a reference to the possible independence of male and female (or household) economic activities, an important factor for our later discussion about the appropriate specification of the utility maximizing unit.

Significantly, labour surplus, or independence of male economic activities from the level of agricultural production, does not require a zero marginal product, as stipulated in the some labour surplus models.<sup>9</sup> Instead, given the well

<sup>8</sup>(...continued) (1979) and Collier and Lal (1986).

<sup>&</sup>lt;sup>9</sup> As detailed in Lal (1973), Sen (1966) has shown that, in fact, zero marginal product is not a necessary condition for the existence of surplus labour. Sen's model specified equal work and income sharing among family members with leisure appearing as an argument in individual utility functions. The necessary and sufficient conditions are "... given by a constant disutility of effort, which implies a constant marginal rate of substitution between income and leisure..." (Lal (1973), p. 115).

documented occurrence of seasonal slack,<sup>10</sup> withdrawal of labour generally leads to a "significant re-organization representing a productivity or production function improvement, so that on balance output stays constant" as noted on page 127 by Corden (1974). This is consistent with the household receiving help, perhaps by extended family members as referred to earlier, or else, by hiring labour as documented in, for example, Adams (1991).<sup>11</sup> The frequent visits home (apparently to carry out some of the more strenuous tasks) are, in fact, an indication of a positive marginal product on the part of migrants.

The observed differences in the marginal contributions of males and females to agricultural production actually allows for a lucid explanation of migratory patterns in Africa. As documented previously, migrants tend to be male, with seasonal movements dominant in those areas where co-operative labour is not an important phenomenon, and with slightly longer absences (circulatory migration at more irregular intervals) in those areas with labour cooperation types of arrangements.<sup>12</sup> Under the assumption of declining marginal

<sup>&</sup>lt;sup>10</sup> See, for example, Schultz (1964).

<sup>&</sup>lt;sup>11</sup> Notice that the existence of a significant rural labour market, to the extent that rural labourers are willing to relocate to the urban area, invalidates the original labour surplus model (Corden (1974, p. 128)). However, the documented increase in labour supply (either of family members or hired labour) means that the output forgone is less than the migrant's marginal product, and could in fact be zero.

<sup>&</sup>lt;sup>12</sup> An interesting question, which receives relatively little attention in the anthropological literature, centers around the possible effect of labour migration (continued...)

productivity and with the male marginal product apparently less than the average product, the absence of male members results in an immediate increase in the household consumption average, even before remittances from absent members. Of particular significance is the fact that we can, in this way, account both for the migration of specific household members and also the non-migration of the remainder.

### 2.2 Volume and Direction

Whereas prior to independence in Botswana, just as in much of sub-Saharan Africa, a substantial number of rural migrants found work at mines and plantations, in the period after independence, Nam et al (1991) note on page 19 that migration to towns has become "quantitatively more important". Still, quantifying the importance of migration flows in Africa is seriously hampered by the overwhelming lack of data. Census data, the best source of comparable information, tends to be of poor quality, with the latest data available generally relating to the 1970's, as noted on page 218 of United Nations (1990).

A useful compendium for West Africa is the study by Zachariah and Conde (1981) carried out under the auspices of the World Bank. Using mid-1970's census data from nine countries, the study found that in an area comprising less than 40 million people or about 12 million workers, approximately 4.4 million people migrated within their respective countries. Principally, these

 $<sup>^{12}(\</sup>dots \text{continued})$ 

on the state of labour co-operation arrangements.

people migrated to the urban areas, especially the capital city. Although a substantial portion of these migrants do eventually return home, the capital cities still tend to be the only areas of the country with a positive net in-migration rate, as noted on page 8 of the Zachariah and Conde study. In essence, municipal authorities are burdened not only with a relatively high rate of natural population increase but also with the temporary, and sometimes permanent, addition of thousands of job seekers from the rural areas. The impact on the urban labour market of both types of migration movements are considered in the sixth and seventh sections of this chapter and more carefully examined by our efficiency wage model presented in the second chapter.

The exhibited high rates of labour mobility, and the main direction of their movement, towards the single largest city, are likely to have a significant impact on national development plans. Concern on the part of policy makers is evidenced by government responses to United Nations questionnaires on population distribution patterns:

"As of 1988 only three Governments in the ECA [Economic Commission for Africa] region (Guinea-Bissau, Mauritius, and the United Republic of Tanzania) considered their patterns of population distribution to be entirely satisfactory; 71 per cent of Governments in the region considered that their patterns of population distribution required major change. In regard to spatial strategies, more than 60 percent of Governments in the region have adopted policies to slow primate city (or metropolitan) growth, even though African cities are comparatively small by international standards."<sup>13</sup>

### 2.3 Demographic Characteristics of Migrant Flows

Examination of migrant demographic data uniformly supports Gulliver's characterization of migrant flows: migrants tend to be young, generally less than 30-35 years of age, better educated than non-migrants, and mostly male in sub-Saharan Africa.<sup>14</sup> Additionally, there is evidence of both income and capital constraints on migration. Some studies in Africa and southern Asia indicate that wealthier rural dwellers are more likely to migrate and also to maintain their rural ties and return to the countryside, as noted in Adams (1991) and Rosenzweig (1988).<sup>15</sup> We would then expect that poorer workers, having little or no access to land, tend to migrate permanently, if they migrate at all. This is consistent with the locality-specific-capital hypothesis advanced by DaVanzo (1983) as an explanation of return migration.

<sup>&</sup>lt;sup>13</sup> United Nations (1990, p. 196). For earlier corroborating evidence see United Nations (1980).

<sup>&</sup>lt;sup>14</sup> Squire (1981, p. 103), Yap (1977) and Sabot (1979) and Zachariah and Conde (1981).

<sup>&</sup>lt;sup>15</sup> Stark (1991, p. 50) interprets the act of migration as an indication of imperfection in rural capital markets, while Rosenzweig (1988, p. 746) views the association of temporary migration with rural landowners as an imperfection in the rural market for land.

### 2.4 Specification of the Utility Maximizing Unit

One of the usual conflicts between the income differential and the risk pooling approaches lies in the specification of the utility maximizing unit: usually the individual migrant for the first and the rural household (or family) for the latter approach. In Africa, however, the level of intricacy in the relationships among nuclear and extended family members, makes it unclear whether the appropriate unit should be the individual, the household or even the extended family.<sup>16</sup>

In rural Africa, females are very likely to be largely responsible for food production and marketing (including travelling to towns or villages to sell marketable product),<sup>17</sup> with the male economic activities being effectively independent. Under these arrangements, the relationship of members to the household is not necessarily symmetrical, and provided other household members are able to increase their own labour supply and maintain the level of output constant as in Sen (1966), (migrating) males may behave entirely as individual

<sup>&</sup>lt;sup>16</sup> "the specification of the appropriate decision-making unit for purposes of economic analysis, for instance, is a source of considerable difficulty..." (Helleiner (1975), p. 46).

<sup>&</sup>lt;sup>17</sup> "...women take the main share of the work for food production and cultivation". Although, as previously noted, the men undertake the heavier part of agriculture work "...a very great deal of heavy cultivation work, clearing and hoeing, is performed by community working parties..." and "...even when the husband is away, the wife continues to ... summon a local working party to work the fields." (Gulliver (1955), p. 33). Similarly, for Mali, Lewis (1985) reports on the importance of "cooperative labour". For a more complete reference, see Stichter (1985), especially chapter 3.

maximizers.<sup>18</sup> Certainly, given the greater marginal contribution by females to agricultural production, it would seem that family income maximization would be achieved by having the marginally less productive males fulfil the role of migrants.<sup>19</sup> Furthermore, provided some type of family income sharing is implemented, as in the case of a positive level of remittances, it is likely that individual income maximization on the part of the migrating individual leads to family income maximization.<sup>20</sup>

<sup>18</sup> There is also some evidence from the anthropological literature that polygamous wives are encouraged to "...become economically self-sufficient in order to support their children as part of a separate unit" (Jules-Rosette (1985, p. 110, fn. 13)).

<sup>19</sup> In the case of diminishing returns, the female's marginal product is likely to be greater, while the male's marginal product is likely to be less, than the average product. The migrant's absence would then ensure a higher level of consumption for the remaining family members. It might also be more accurate to assume the migrant regards the average product, rather than the marginal product, as his income forgone. Notice, however, that the argument about the difference in marginal products between males and females relies, not on intrinsic characteristics, but instead on person-hours at the margin. That is, current social arrangements and the resulting (apparent) correlation between gender and relative marginal products may be a reflection of the relative compatibility between raising children and farm labour in contrast to the relative incompatibility between urban work (and migrating) and child rearing. The obvious exception is work which is more physically demanding, such as field clearing, in which men are likely to have an intrinsincally higher marginal product.

<sup>20</sup> In general, the requirement for individual maximization to coincide with family maximization is the enforcement of some type of social contract, in the shape of either an explicit income sharing rule or some type of altruistic behaviour. Our argument for the use of an individual approach centers on the fact that the migrant's desire to maintain rural rights induces some positive level of remittances consistent with income sharing. In this case, the migrant places a higher value on the continued access to his rural assets more than on urban assets which he might acquire, implying that his actions will maximize both his

(continued...)

With the reported predominance of marriage-motivated saving patterns for unmarried migrants, individual maximization is also likely to be the appropriate choice of utility maximizing unit for these migrants. The same studies also indicate that unmarried migrants are a significant proportion of migrant flows: Gulliver (1955) reports that approximately 50% of migrants from southern Tanzania were unmarried while Zachariah and Conde (1981) report a significantly higher propensity to migrate internally among unmarried people.

Individual maximization is also consistent with certain plausible assumptions of household economic activity. As noted on page 156 in Collier and Sabot (1982): "For example, if the marginal product is half the average, but only half of the family is economically active, then the migrant consumes only what he produces and family and individual maximization converge."

Perhaps remittances are then a payback, either for the financing of the migration move (it is noteworthy that, according to Williamson (1988, p. 433), no study to date has controlled for the funds taken by the migrant at the time of migration), or else just compensation for the reduction in agricultural production. Gulliver (1955) notes on page 33 that "apart from these working parties, and other people's working parties too, a man at home generally helps his wife to hoe up

 $<sup>^{20}</sup>$ (...continued)

individual income (by choosing the individually optimal urban labour market arrangement) and family income (through income sharing and optimal spatial allocation of family members). Under such conditions, the migration process itself can be effectively modelled as an individualistic pursuit.

and plant additional land, and this latter is likely to be reduced considerably during his absence. Yet whilst he is away there is one less person to be fed from the wife's fields and consequently a smaller acreage than usual will suffice". In either case, individual and family income maximization again converge.

In summary, social organization in Africa appears to be largely consistent with individual income maximization on the part of the migrant. Whether he does indeed decide to migrate has been shown to depend on his age and education level, distance, expected wage and (perceived) employment possibilities.<sup>21</sup>

Given individual utility maximization and in the face of higher wages in the urban area, why would a migrant still return to the rural area? Some economists, for example Stark (1991), would advance explanations such as risk pooling, inheritance rules and imperfections in the market for land. In the next section, we combine the last two concepts with the idea of the imperfect substitutability of hired for family labour to provide a rationale for the migrant's incentive to (temporarily) interrupt his urban career to fulfil his rural duties.

### 2.5 Reasons for Return Migration

Anthropological evidence on institutional arrangements prevailing in Africa and South Asia,<sup>22</sup> indicates that conditions in these regions are particularly conducive to circular migration as prevailing land arrangements allow the migrant

<sup>&</sup>lt;sup>21</sup> See, for example, Sabot (1979).

<sup>&</sup>lt;sup>22</sup> See, for example, Gulliver (1955) on "working parties" as mentioned in the previous section, or the summary in Nelson (1975).

to retain title and social customs facilitate the continued cultivation of the land plot by other family members. Labour market requirements and operations provide further impetus to the existence of circular migration.

We previously mentioned that despite the ability of the farm household to compensate for the absence of male labour in the short-term, the more strenuous tasks needed to maintain output in the longer term are usually a male domain. In this case, farm households would clearly require the migrating male to return to the rural area from time to time, and this is likely a contributing factor in the observed pattern of circulatory migration.

We also referred to evidence of active rural labour markets, for example Adams (1991), which leads us to consider the possibility of substituting hired labour for male migrant labour. In fact, Adams (1991) concludes that there exists a group of people in the rural area who should be considered permanent labourers. These people own very little land, if any, and depend on labour earnings for the bulk of their income. Correlation analysis by Adams (1991) reveals that households which hired permanent workers tended to be wealthier (metal roof on the main dwelling, larger farm and more cattle), to be female-headed and to receive remittances from the male head working elsewhere – 52% of femaleheaded households hired a permanent labourer while over 68% of households receiving more than Z\$50 per month hired at least one permanent labourer. Casual work, typically harvesting and generally remunerated in the form of piecework, was hired largely in conjunction with permanent labour. The important conclusion is that a significant number of households are both hiring in and hiring out labour, which serves to focus our query on the degree of substitutability between family and hired labour.

In essence, Deolalikar and Vijverberg (1987) note on page 294 that "if family and hired labour are imperfectly substitutable, the labour demand and supply decisions of farm households are interlinked. The supply of labour by family members then depends on the on-farm..." demand for the specific tasks which cannot be performed (as well) by hired labour such as supervisory or management duties. Substitutability means that rural-urban migration of family members should be accompanied by an increase in the demand for hired labour, as is apparent in the findings by Adams (1991). Less than perfect substitutability, however, means that the migrant's labour supply while in the city is not independent of the labour demand on the farm.

Evidence on the degree of substitutability between family and hired labour is inconclusive as well as sparse. The most careful test of the hypothesis is contained in Deolalikar and Vijverberg (1987) and uses farm level data from India and Malaysia. Although the authors do find heterogeneity in family and hired labour, their inability to identify the cause means that their result could possibly be an artifact of the combination of the dominant use of hired labour at the peak season and the limited substitutability between slack and peak-season tasks. The rest of the literature contains indirect tests which consist of estimating (Cobb-Douglas) production functions with family and hired labour as two separate inputs. Although some studies do find differential impacts on output for the two different inputs, most authors find the difference to be insignificant.

It appears that a definite conclusion on the degree of substitutability between family and hired labour awaits the availability of better data. Tentatively, however, it seems plausible to assume less than perfect substitutability. Theoretical support is derived from the idea of differential incentives to supply (more) effort, and even though the empirical support is less than overwhelming, the proposed hypothesis is certainly not rejected by tests to date. In practice, less than perfect substitutability should lead to farm operations requiring the occasional presence of the migrating male, whether to carry out the heavier tasks or to provide some type of supervision for the hired labour.

Having provided a brief characterization of migration in Africa, we turn now to a review of the economic literature on rural-urban migration. The next two sections outline the two main approaches to modelling migration motivation: the income differential and the risk pooling approach. A comprehensive model of migration, however, must characterize both the labour market, urban and rural, as well as the actual migration process itself, and the fifth section presents the four main characterizations of the labour market: neo-classical, labour surplus, institutional dual and labour turnover.

### 3. Income Differential Model

#### 3.1 The Human Capital Approach

The human capital approach to migration analysis originated with Sjaastad (1962), who first enunciated the comparison of net private returns (accruing over time) to private costs (incurred at the time of the move) as the decision framework used by potential migrants. Under this approach, individuals are expected to choose the location that will maximize their present value of lifetime earnings net of moving costs, denoted by V, by evaluating

$$V = \int_{t=0}^{T} [Y_{D_j}(t) - Y_O(t)] e^{-rt} dt - C_j ,$$

where  $Y_{Dj}(t)$  is the real income in period t at destination j,  $Y_0(t)$  is the real income in period t at the origin,  $C_j$  is the cost of migration to destination j, T is the individual's working life assumed to be known with certainty and r has the usual rate of time preference interpretation. The criterion for migration becomes V > 0, in which case the agent moves to the location with the highest income gain.

The most salient predictions of the Sjaastad model are that agents should move from areas with lower incomes to those with higher incomes and that, over time, mobility should tend to narrow spatial income differentials. Additionally, the model implies that (1) younger workers, given their longer term ability to benefit from a move, have a greater incentive to migrate; (2) there is a negative relationship between mobility costs, for example distance, and mobility; (3) there is a positive relationship between education and mobility, either because of higher returns to skills in the urban area or because education indicates better information about destination conditions; and (4) for the most part, only one lifetime move is predicted (unless, as stated on page 58 of Pessino (1991), the rate of growth differs amongst the different regions).

Empirical evidence from both economists and demographers, for example Squire (1981) and Zachariah and Conde (1981), has been consistent with early anthropological findings and has generally validated the implications arising out of the human capital model. As documented by Yap (1977), migrants tend to be younger and more educated, they tend to move to areas with better economic opportunities (higher incomes and employment levels), and distance does appear to be a deterrent. However, high migration rates in developing countries (estimated to account for 30-60 percent of urban population growth rates by Yap (1977)), have not been effective in reducing differentials between unskilled urban wages and rural wages (with unskilled urban wages estimated to have remained at 1.5 to 2 times the rural wages by Squire (1981) and Rosenzweig (1985)). Additionally, the incidence of return migration and re-migration appears to be significant, particularly in Africa, where migrants have a higher tendency to maintain their rural ties. In the next section, we review the dominant migration model of the last two decades: the Harris-Todaro specification of the income differential model.

### 3.2 The Harris-Todaro Specification

Although the human capital approach has been shown to be fairly consistent with the empirical evidence, Yap (1977) reports that a majority of people move without a pre-arranged job, so that migration appears to involve some degree of uncertainty over the earnings realization. Todaro (1969) and then Harris and Todaro (1970) made the first attempts at incorporating uncertainty over the economic benefits into migration analysis by specifying a random job allocation rule and an unemployment queue. Todaro (1969) provided the basic framework by positing that the existence of a vast army of unemployed and underemployed has a negative impact on a prospective migrants' probability of obtaining an urban (formal sector) job. Harris and Todaro (1970) incorporated this basic behavioural paradigm into a two-sector general equilibrium model of migration, employment and wage determination.

The motivation for the Todaro (1969) extension to the standard wage differential model<sup>23</sup> was to explain conditions prevailing in Africa: high ruralurban migration rates in spite of relatively high levels of urban unemployment,<sup>24</sup>

<sup>&</sup>lt;sup>23</sup> In this model, workers are paid their marginal product in both sectors although an intersectoral wage differential does persist. The optimal labour reallocation, from the low to the high productivity sector, could be attained by a subsidy to the high wage sector (see, for example, Corden (1974), p. 129). Notice that in the original wage differential model, as well as in the labour surplus model specified in Lewis (1954), the two sectors were the modern urban (industrial) sector and the subsistence rural (agricultural) sector. While the HT model does represent the first rigorous attempt at capturing the heterogeneity of the urban sector, the publication of ILO (1973) and Hart (1973) made clear that the urban informal sector was in fact a dynamic, productive sector with (at least a significant portion of) wages higher than subsistence levels, as assumed by HT.

<sup>&</sup>lt;sup>24</sup> "... to explain the apparently paradoxical (at least to economists) phenomenon of accelerated rural-urban migration in the context of rising urban unemployment." (Todaro (1979, p. 194)).

and an expanding urban informal sector characterized by low wages and underemployment.<sup>25</sup> Under the Todaro approach, the decision to migrate depends on the evaluation of<sup>26</sup>

$$V = \int_{t=0}^{T} [p(t)Y_{U}(t) - Y_{R}(t)]e^{-rt}dt - C$$

,

where  $Y_U(t)$  is the real income in the urban area in period t,  $Y_R(t)$  is the real income in the rural area in period t, p(t) is the probability of securing a formal sector job in period t, C is the cost of migration, T is the individual's working life assumed to be known with certainty and r is the rate of time preference. A positive value for V leads to a decision to migrate as the expected urban income is higher than the rural income. Migrants are assumed to be individualistic, rational, utility self-maximizers driven mainly by economic incentives. In this context, migration serves to equalize not actual, but expected, urban wages with rural wages: the equilibrium condition can then be written as  $W_R = E(W_U)$ , with  $W_R$ generally assumed to be fixed throughout.

Economic studies, Yap (1977) and Robinson and Tomes (1982) for example, show that wage differentials, destination employment conditions,

<sup>&</sup>lt;sup>25</sup> "...upon entering the urban labour market many migrants will either become totally unemployed or will seek casual and part-time employment in the urban 'traditional' or 'informal' sector. Consequently, in his decision to migrate the individual must balance the probabilities and risks of being unemployed or underemployed for a considerable period of time against the positive urban-rural real income differential." (Todaro (1979, p. 195)).

<sup>&</sup>lt;sup>26</sup> See page 201 of Todaro (1979). Notice the assumption of zero open unemployment in the rural sector embodied by the absence of a discount factor for  $Y_R(t)$ .
education and distance are motivating factors in the original decision to migrate. Consistent with the hypothesis that education "broadens the horizons", Fields (1982), in a study of Colombia, also finds that education tends to blunt the negative effect of longer distances. This finding has been interpreted in the literature as an indication of a positive correlation between the level of education and information. It also seems reasonable, however, that if input complementarity to skills is lacking at the origin but not at the destination, a larger bundle of skills would translate into the possibility of higher income at the destination, and make it more profitable to incur a higher level of migration costs.

Assessments of income changes for migrants have uniformly found significant improvement upon migration, as detailed in Yap (1977), for example. However, Robinson and Tomes (1982) point out that such comparisons, based largely on aggregate data, may be biased because of self-selection. That is, it may not be appropriate to estimate the returns to migration by comparing the wages of migrants with the wages of similar workers who stay in the origin area. Instead, Robinson and Tomes propose the use of micro data to take account of heterogeneity and control for the sizable reverse migration flows from higher to lower income provinces, and the use of the inverse Mill's ratio in order to correct for the truncated samples on which the earnings functions are estimated.<sup>27</sup>

<sup>27</sup> The samples are truncated because for workers who migrate we only observe the destination wage, while for workers who do not migrate we only observe their origin wage. The selectivity bias correction used by Robinson and (continued...)

The results obtained by Robinson and Tomes (1982) are consistent with the predictions arising out of the general human capital approach to migration: the education level enhances mobility, the expected wage gain is an important determinant in the migration decision, and marginal propensity to migrate diminishes with age. The innovation is the treatment, and related findings, of self-selection. For the least experienced of the two groups into which the data were divided, there is evidence of positive selection into the migrant's group - at the destination location, migrants earned more than non-migrants would have earned, had they moved.<sup>28</sup> The pattern for the most experienced group is not clear. Significantly, the estimate of the expected wage gain in the structural probit (with the dependent variable denoting 'yes' or 'no' to the migration decision) turned out to be very sensitive to the inclusion of the selection bias correction term in the wage equation. In the absence of correction, the expected wage gain was not important in the migration decision (a clear rejection of the human capital approach) whereas if self-selection is taken into account, the opposite is true. These results tend to support the income differential model.

<sup>27</sup>(...continued)

Tomes (1982) was originally proposed by Heckman (1976).

<sup>28</sup> Similar findings in the literature include: Falaris (1987) for Venezuela, Pessino (1987) for Peru and Vijverberg (1989) for Côte d'Ivoire.

#### 4. The Risk Pooling Approach to Migration

Even though the importance of risk for economic decisions has been recognized and duly incorporated in mainstream economics for a relatively long time, the incorporation of risk into formal models of rural-urban migration is fairly recent.<sup>29</sup> In fact, it was not until the 1980s that a series of papers on migration and remittances<sup>30</sup> explicitly modelled the role of risk diversification. Combined with a view of the family as the economic decision unit, this approach to migration stresses the role of insurance provided by spatial allocation of labour.

In contrast to the differential income approach, the risk-theoretic approach yields a positive level of migration even in the presence of equality of earnings: as long as earnings are less than perfectly positively correlated, the family, given contractual arrangements for co-insurance and some level of risk aversion, will be better off with a spatially diversified labour force allocation. Risk considerations are then capable of accounting both for migration of some family members and non-migration of the remainder, as well as temporary migration, especially if the principal source of risk is locational, as noted by Rosenzweig (1988).

Empirical evidence, though limited, appears to be supportive of the risktheoretic approach. Lucas and Stark (1985), in a study on the pattern of

<sup>&</sup>lt;sup>29</sup> It must, however, be noted that the development literature does contain some early references to the likelihood of the importance of risk, as well as expected income, for migration decisions, for example Helleiner (1975, p. 33).

<sup>&</sup>lt;sup>30</sup> See Stark (1991) for a collection of the main works in this literature.

temporary migration and remittances in Botswana, find results consistent with (1) co-insurance agreements – controlling for average income, families at greater risk of incurring capital losses (cattle and crops in a season of drought) tended to receive higher remittances; (2) bequest incentive to remit – sons of wealthier families tended to remit more; (3) positive relationship between education (interpreted as a measure of the family's investment) and the level of remittances; and (4) higher remittances from closer family members (altruism).

Significantly, Rosenzweig (1988) notes first, that Lucas and Stark derived their hypothesis in a somewhat ad-hoc manner instead of "within the context of a rigorously formulated, integrated model of the household", and second, that the finding of a positive relationship between family wealth and the level of remittances could be entirely due to the "greater ability of households receiving remittances to accumulate wealth, rather than the bequest motive". In essence, it appears that we need "richer models of the family enterprise... to understand fully the complex nature of the spatial mobility of labor" as noted on page 753 of Rosenzweig (1988). An additional weakness of this type of study a lack of is a lack of control for the funds migrants take with them at the time of migration. Some researchers claim that "when these emigrant funds are subtracted from immigrant remittances, the net remittances turn out to be small even for Africa", as noted on page 433 of Williamson (1988).

The importance of risk considerations is likely to decrease if local occupational diversification among family members is adequate as a form of

insurance. This is likely to be the case only if agricultural production is less than the predominant source of alternative economic activities in the rural area, a condition generally not satisfied in Africa.<sup>31</sup> On the other hand, the stylized fact that in Africa "there is an inverse relationship between the number of men leaving the villages in any given year (the rate of out migration) and the size of village harvests has long been recognized everywhere in Africa" as noted on page 482 of Berg (1961) points to income augmentation as the predominant motive for migration. In reality then, both risk diversification and income augmentation are likely to be important considerations for potential migrants.

Having reviewed the two major motives for migration, we widen the context and turn to a discussion of the models that have been influential in our understanding of the labour market in developing countries. This literature is vast and we focus on models of the urban labour market, which will be the focus of the proposed efficiency wage model.<sup>32</sup> In the next section, we provide a brief review of the neo-classical, labour surplus, institutional dual and labour turnover hypotheses. We follow that, in the sixth section, with a more extensive

<sup>&</sup>lt;sup>31</sup> While diversity of agricultural output and the correlation among yields and prices of various components are also important factors, the main source of risk for agriculture in low income countries is likely to be weather patterns, and these will clearly affect all agricultural products.

<sup>&</sup>lt;sup>32</sup> For a review of models which help to explain behaviour in the agricultural sector, see Rosenzweig (1988).

presentation of the Harris-Todaro model, the dominant paradigm in migration research since its introduction in 1970.

#### 5. Labour Market

The benchmark description of labour market operations is provided by the neo-classical model. The essential assumptions are perfect competition and frictionless labour mobility, with equilibrium characterized by full employment and equality of wages. In addition, wages are fully flexible and labour is paid its marginal product. In this context, an intersectoral wage gap, for any particular skill group, generates enough labour movement such that, over time, the differential is eliminated. Clearly, long-term wage differentials in the face of large movements of labour as well as significant levels of open unemployment, as widely documented for most of the developing world, constitute fairly strong evidence against this hypothesis of labour market operations. The full implementation of this model requires similar assumptions for all segments of the labour market, ie. urban formal and informal sectors in both the urban and the rural labour markets.

An alternative hypothesis of labour market operations postulates an urban labour market operating in the neo-classical manner combined with a rural labour market best described by the labour surplus model: the original dual market theory. Under this view, the urban labour market is, as usual, governed by profit maximization and the marginal product hiring rule. The rural labour market, on

the other hand, is governed by income sharing with family members receiving their average, instead of their marginal, products. Under the usual assumption of declining marginal productivity then, the marginal members receive more than their marginal products.

To attract rural workers, the urban sector would have to at least match the alternative source of income, the average product in the rural sector, and the resulting divergence in marginal products would make the labour allocation less than optimal. Under the extreme case, the marginal product in the rural sector is actually zero – labour out-migration does not lead to a decrease in agricultural production – and the labour supply curve to the urban labour market is then perfectly elastic, at least until the entire surplus is absorbed into the urban sector.

As previously mentioned, however, subsequent theoretical work has demonstrated that zero marginal product is not necessary for the existence of surplus labour, and subsequent empirical work has led to the emerging view that any labour surplus is only seasonal.<sup>33</sup> That is, peak periods of rural activity do indeed require the efforts of the entire labour force, and extended absences, given fixed organization patterns, will actually lead to a decline in production. In the urban sector, unemployment is predicted to be non-existent while expansion of unskilled employment, given a perfectly elastic supply of labour from the rural areas, should not need the incentive of rising wages. The evidence, however,

<sup>33</sup> See, for example, references in Corden (1974, p. 126).

shows that queues do form for urban jobs and that the real wage level for workers with no education has been rising at a rapid pace.<sup>34</sup>

As noted on page 128 of Corden (1974), the applicability of the labour surplus model is further limited by:

the existence of an active rural labour market, provided these workers are paid their marginal products and are available for relocation to the urban area; and
 the family acting to maximize its total income and equating the value of the marginal product of its labour across all activities, including those engaged in by migrant members.

In both cases, urban employers need only match the marginal product in the rural sector.

A third hypothesis postulates that profit maximization and the marginal product hiring rule prevailing in the "advanced" sector are combined with legal minimum wages higher than the prevailing wages in the "traditional" sector or, technically, higher than the opportunity cost of labour. The original proponent of this model, Todaro (1969), assumed that this barrier to entry was due to unspecified institutional arrangements. Given the intersectoral wage differential, this model predicts a queue for the high paying jobs as some workers prefer to maximize their chances for employment in the advanced sector rather than (temporarily) engage in low paying work. With the usual assumption of a clearing

<sup>&</sup>lt;sup>34</sup> See, for example, Sabot (1979, pp. 59 and 134).

rural market, the model also predicts a positive level of rural-urban migration as some rural workers are attracted by the chance of winning the job "lottery". Potential migrants are thus predicted to take into account not only the income differential, but also the employment prospects at the destination.

Although evidence has generally been supportive of this model, with a majority of economic studies finding income differentials and employment prospects to be statistically significant explanatory variables, the model leaves the reason for urban wage rigidity unspecified. The proposed explanation of effective minimum wage legislation and union activism certainly does not appear very convincing under conditions currently prevailing throughout the developing world.<sup>35</sup> and it is further rejected by historical studies. In particular, Mazumdar (1990) notes on page 347 "...that wages in the large textile factories have been established at a high level compared to alternative earnings of labour coming into town from the rural areas and to other non-factory labour in the city for a long time before the era of trade unions or government intervention...". By contrast, Esfahani and Isfahani (1989) provides a cogent explanation of wage duality in terms of effort observability and worker productivity, and Mazumdar (1990), building on earlier research, clearly endorses the efficiency wage explanation. We shall follow the same approach but provide more structure to our model.

<sup>&</sup>lt;sup>35</sup> See, for example, Squire (1981), Mazumdar (1984), Lluch and Mazumdar (1983) and Montgomery (1985).

However doubtful the institutional explanation may seem, empirical support for the Harris and Todaro assumption of urban wage duality is fairly widespread: for example, see Mazumdar (1973, 1988) for India, Sahn and Alderman (1988) for Pakistan, and Vijverberg and van der Gaag (1990) for Côte d'Ivoire. Using data on wage earners only,<sup>36</sup> Vijverberg and van der Gaag present four separate tests of labour market duality (or the idea that high wage jobs are intrinsically different than low wage jobs): the customary procedures whereby workers are sorted according to their level of wages or job characteristics, a switching regression model with unknown regime changes (thereby letting the data determine the splits) and a switching regression model with corrections for worker selectivity, and note on page 38 that "the conclusions are unanimous: segmentation exists". Although overlap between the segments implies that these findings are consistent with either segments with gaps or competing submarkets, evidence on limited mobility from the informal to the formal sector presented by Grootaert (1992) for the same country, suggests the first view is more plausible.<sup>37</sup>

In an attempt to provide a fuller explanation for the urban wage rigidity, Stiglitz (1974) advances the labour turnover model. Judging by the number of

<sup>&</sup>lt;sup>36</sup> That is, abstracting from the more typical associations of self-employment with the informal sector and wage earners with the formal sector.

<sup>&</sup>lt;sup>37</sup> Additionally, van der Gaag and Vijverberg (1989) investigate the determinants of wages for employees in Côte d'Ivoire. They find high rates of return to education plus evidence of screening or credentialism, which is clearly suggestive of heterogeneity.

references and studies of private sector firms focusing on the topic, for example Elkan (1967), labour turnover and its associated costs was an early common concern of private sector employers. Stiglitz (1974) can probably be considered as the culmination of this substantial research effort. The central notion of this model is that the level of wages influences total labour costs in two ways: first, by determining the total wage bill, and second, through an inverse relationship between the wage rate and turnover costs.

The appealing features of this quasi-efficiency wage model are the endogeneity of the urban wage and urban employment levels. The urban firm's wage offers turn out to be a function of the unemployment rate and the rural wage, with the profit-maximizing wage higher than the market clearing level and the equilibrium level of unemployment strictly positive. Rosenzweig (1988) also notes that the positive level of unemployment is optimal in the sense that the associated wage is the one that maximizes output – the additional output available from employing the currently unemployed workers is less than the output that would be foregone with lower wages. The loose ends are the assumption of a constant wage over the workers' life cycle, which is not supported in the development literature as noted on page 750 of Rosenzweig (1988), and the lack of recognition that the firm may also affect the turnover rate by sharing the training costs with its workers. Fry (1979) tests a Stiglitz-type turnover model but

he does not control for who pays for the training costs, nor does he attempt to test the wage-tenure relationship.<sup>38</sup>

Having introduced the two main motivations for the migration decision as well as the four main characterizations of the labour market, we provide now a comprehensive review of the dominant model in migration research, the Harris-Todaro model. This model combines the income differential approach to migration with an institutionally demarcated urban labour market, and provides the basic framework for our proposed model.

# 6. The Harris-Todaro Model

The Harris-Todaro (HT) model combines the institutional version of dual markets in the urban area, a market clearing rural labour market and the Todaro specification of the human capital approach to migration. Typically, the employment process is assumed to follow a lottery type of random selection (Bernoulli process), with the probability of gaining (formal) employment equal to the number of urban jobs divided by the total urban labour force<sup>39</sup>, and capital is

<sup>&</sup>lt;sup>38</sup> See Rosenzweig (1988, p. 750-1) for a brief discussion and references. Stiglitz (1982) also develops two efficiency wage models: (1) the Solow-type effort-wage version, and (2) the Weiss-type quality-wage version. The effortwage model, however, does not provide an explanation for the perceived relationship between worker effort and the wage received, and its high level of abstraction does not allow for a detailed analysis of results and implications.

<sup>&</sup>lt;sup>39</sup> Harris and Todaro (1970, p. 128).

assumed fixed and immobile – although some researchers have attempted to introduce capital mobility, for example, Corden and Findlay (1975).



Figure 1. The Harris-Todaro Model according to Corden (1974, p. 152)

The standard geometric exposition of the HT model is provided by Corden (1974). In figure 1, the horizontal axis measures the total labour force:  $L = L_R + L_U + L_{UE} + L_I$ , where UE refers to unemployed and I refers to informal sector, with  $O_R$  as the origin for rural-based labour and  $O_U$  as the origin for urban-based labour (original urban residents and migrants). UU' is the value of marginal product curve for the urban (manufacturing) sector, while RR' is the value of marginal product for the rural (agricultural) sector. We can now compare the predictions of the standard competitive model with those of the HT model.

Under the standard model, equilibrium is attained at E with  $W_R^* = W_U^*$ and full employment consisting of  $O_R L_R^*$  workers in the rural sector and  $O_U L_U^*$  workers in the urban sector. This is the usual neo-classical outcome, directly dependent on the assumption of flexible wages, and consisting of an equilibrium characterized by the intersectoral equality of factor prices.

If instead, the urban wage is (institutionally) fixed above the competitive wage level, as in the HT model, say at  $\overline{W}_U$ , then employment in the urban sector shrinks to  $O_U L_U$ . To allocate the remainder of the labour force between urban unemployment and rural employment, assume, as usual, a random lottery as the job allocation mechanism, and re-write the equilibrium condition<sup>40</sup> as  $L_U$ 

 $W_R = \frac{L_U}{L_U + L_{UE+I}} W_U$  or  $W_R(L_U + L_{UE+I}) = L_U W_U$ . Given  $\overline{W}_U$  then, the only two points at which this condition is met are Y and Z, which can be joined by a rectangular hyperbola, qq', the so-called Harris-Todaro curve.<sup>41</sup> Now, there are  $L_{UE+I}$  unemployed (or at best, underemployed) workers seeking a formal sector job in the urban area.

Consistent with empirical evidence cited in Williamson (1988, p.444) for the developing world, figure 1 assumes the elasticity,  $\eta$ , of labour demand UU' to

<sup>&</sup>lt;sup>40</sup> HT assume that the expected urban wage is the simple average of the remuneration for the two possible states: employment with a wage of  $W_U$  and unemployment with a wage of zero. The implicit assumption is that migrants are risk neutral.

<sup>&</sup>lt;sup>41</sup> The algebraic equivalent of the previously specified equilibrium condition is that the area in the rectangle  $ZW_UO_UL_R$  must equal the area in the rectangle  $\overline{YW_UO_UL_U}$ , in which case Y and Z are two points on a rectangular hyperbola with unitary elasticity.

be less than  $1.^{42}$  If  $\eta = 1$ , then UU' and qq' coincide, while if  $\eta > 1$ , UU' is shallower than qq'. Notice that an urban wage subsidy (outward shift of the Harris-Todaro curve) leads to both an increase in urban employment, and also to a decline in rural output as labour moves out of the rural sector, with the direction of both effects robust to the elasticity parameter. This serves to highlight both the double role played by the urban wage in this formulation (establishment of urban employment levels and the allocation of labour among the rural and the urban sectors), and also the fact that the value of  $\eta$  is significant for purposes of comparing gains and losses in sectoral output.

It is also interesting to ascertain the importance of the migration mechanism by comparing the policy prescriptions from the standard HT model to the special case when migrants do not respond to the intersectoral wage differential, a model that Corden (1974) labels the simple version of the wage differential model. Given a wage gap, this simple wage differential model yields clear conclusions with regards to attaining Pareto efficiency: an urban wage subsidy to reduce the urban price of labour down to its social opportunity cost and induce labour reallocation such that the marginal product becomes identical for both sectors.

<sup>&</sup>lt;sup>42</sup> We should note that the usual representation of "...a CES (constant returns to scale) production function with elasticity of substitution  $\sigma < 1...$ " combined with an assumption that "...capital is scarce relative to labour..." is consistent with  $\eta < 1$  as noted by Stark (1991).

In the HT model, however, an urban wage subsidy could actually lead to an increase in unemployment (level and/or rate) as migration responds to changes in the probability of job acquisition (the 'Todaro paradox'). Diagrammatically, an urban wage subsidy of Y'W leads to a shift in qq' to  $q^{*}q^{*'}$  and new employment levels of  $L_{U}$  and  $L_{R}$ .<sup>43</sup> Whether unemployment rises or falls, or the change in rural output,  $ZZ'L_{R}'L_{R}$ , is greater, or less than, and urban output,  $YWL_{U}L_{U}'$ , is an empirical question. In the HT model then, the first-best solution (under which labour is allocated such that intersectoral marginal products are equal and there is no unemployment) is a wage subsidy complemented by a restraint on migration.

<sup>&</sup>lt;sup>43</sup> Consistent with most analysts, we assume the rural sector to be so large relative to the urban sector that labour movements in or out or the rural sector do not affect the wage level. Notice that the steeper is UU' and the shallower is RR', the more likely it is that a given urban wage subsidy actually increases urban unemployment and lowers net output.



Figure 2. Effect of an urban wage subsidy

To summarize then, the two most striking results of the Harris-Todaro model are that:

(1) equilibrium (Z) is characterized by significant levels of urban unemployment and informal sector employment ( $L_{UE+1}$ ). Combined with the Harris-Todaro assumption of the undesirability of informal sector employment, this is clearly a suboptimum outcome for the economy as a whole. Still, the decision to migrate (and likely endure some period of unemployment) is the "rational, utility maximizing choice for individual rural migrants given the level of the minimum wage", as noted on page 131 of Harris and Todaro (1970). Essentially, potential migrants fail to take account of the reduction of income their transition imposes on other unemployed job seekers. The presence of one more migrant in the urban area reduces the expected value of the urban wage to other job seekers by decreasing the probability of finding employment. This type of search externality (the expected utility to a searcher depends on the number of searchers) means that the social cost of migrating is greater than the private cost.<sup>44</sup>

(2) an increase in the number of urban jobs (resulting, for example, from a government employment programme or a wage subsidy provided to the private sector)<sup>45</sup> induces additional migration flows and may actually lead to more unemployment (the 'Todaro paradox'). Restoration of equilibrium requires the probability of employment to be lowered which can only be accomplished by an increase in the number of people (migrants) looking for work.

In a previous section we mentioned that although the duality structure of the HT model has some support in the literature, its rationale for the urban wage rigidity does not. In fact, despite its analytical appeal and its protracted dominant

<sup>&</sup>lt;sup>44</sup> As mentioned earlier, empirical studies are uniform in finding that individual migrants improve their economic position. In the Harris and Todaro model, while the individual migrant benefits from the increase in expected wages, the income reduction imposed by his presence is spread among all other unemployed workers.

<sup>&</sup>lt;sup>45</sup> At the time the Harris-Todaro model was published, the most popular prescription for the perceived problems of unemployment, underemployment and surplus labour afflicting the developing world was a "subsidy per unit of labour equal to the wage differential..." (Hagen (1958, p. 498)). This policy prescription was advocated by the majority of the prominent development economists of the time, such as Bardhan (1964), Chakravarty (1964) and Little (1964), as referenced in Harris and Todaro (1970, p. 132).

position in migration analysis, the HT model does contain other serious shortcomings, namely: (1) its treatment of the informal sector, (2) the job assignment mechanism, (3) the assumed lack of surplus labour in the rural sector, (4) the lack of any consideration of return migration, and (5) the restrictive conditions under which the Todaro paradox applies.

First, the model only tangentially recognizes the existence of the urban informal sector, and completely ignores its contribution to urban production. Although acknowledging that agricultural labour is only fully employed at peak seasons, Harris and Todaro ignore any inducement to even temporary migration from the level of wages in the informal sector. Such an inducement would allow migration to result in an increase in urban production without reducing rural output, contrary to a main proposition in the Harris and Todaro analysis of the effects of an urban wage subsidy.<sup>46</sup>

That informal sector employment is attractive enough to be the target sector, instead of just a temporary stop, for some migrants is evidenced by findings of:

 significant levels of migration with the sole purpose of finding work in the informal sector, including migration to pre-arranged informal sector jobs as in Banerjee (1986) and Mazumdar (1981);

<sup>&</sup>lt;sup>46</sup> See Harris and Todaro (1970, pp. 132-4) for the analysis. Such criticism is also significant in view of the recognized importance of return migration (seasonal and otherwise) in Africa, and ironic given that African conditions provided the impetus for the Harris-Todaro analysis.

- low rates of active job search by informal sector wage employees as in
  Mazumdar (1981), Banerjee (1986) and Grootaert (1992);
- relatively low rates of movement of wage employees from the informal to the formal sector as in Banerjee (1986) and Grootaert (1992);

Moreover, the accepted wisdom about conditions in the informal sector was drastically altered starting with the 1973 publication of studies on Ghana by Keith Hart and Kenya by the International Labour Office (ILO). Hart emphasized the dynamism of the informal sector and postulated that workers might prefer to dabble in both the formal and informal sectors for risk sharing purposes.<sup>47</sup> The ILO study went still further and contended that the best opportunity to increase urban employment lay, not in the more capital intensive formal sector, but in the flexible, dynamic and more labour intensive informal sector.

While informality is most closely associated with self-employment and street merchants, the sector also encompasses housing, transportation and smallscale, skilled manufacturing. Perhaps the most notable, and definitely the most comprehensive, study was conducted for Peru by Hernando de Soto (1989) at the Instituto Libertad y Democracia. Defining informality as a state of 'illegality', or lack of registration with the appropriate governing bodies, the study reports on page 12 the following statistics concerning the size of the Peruvian informal (with

<sup>&</sup>lt;sup>47</sup> Hart (1973) principally mentions differences in the regularity and uncertainty of income patterns.

the working definition being operation outside the legal framework or unregistered) sector in 1984:

(1) 48% of economically active people participated in the informal sector,

(2) 61.2% of the total hours worked in Peru, were in the informal sector,

(3) informal sector output was 38.9% of reported GDP,

(4) entire towns are run informally, and

(5) 95% of public transportation in Lima is provided by the informal sector.

The study also reports that, following current trends, the informal sector is expected to generate over 60% of reported GDP by the year 2000. Other estimates of the informal sector in developed countries are consistently around the 10-14% range of official GNP,<sup>48</sup> while developing countries are generally higher, typically in the range of 30-40% of official GNP.

Harris and Todaro invoked the standard assumptions that most migrants endure, and fully expect to endure, some time in unemployment, and that participants in the informal sector are (recently arrived) migrants scraping along until they find a higher-paying formal sector job. In contrast, it has generally been found that migrants are quickly integrated into urban labour markets: their job search tends to be of short duration and their unemployment rate is low as reported by Yap (1977) and Williamson (1988). While the evidence does support the contention that the informal sector tends to be dominated by migrant workers,

<sup>&</sup>lt;sup>48</sup> See, for example, Feige (1990) and the sources cited therein, including de Soto (1989).

there is surprisingly little evidence that they earn any less than their rural counterparts as reported by Yap (1977) and Berry (1987). However, the evidence does point to a wage differential between the formal and the informal sectors, even after controlling for human capital as reported by Sahn and Alderman (1988) for Pakistan, Banerjee (1986) for India and Grootaert (1992) for Côte d'Ivoire. As mentioned earlier, Kahnert (1987) proposes the idea of attributes such as reliability or motivation as explanatory factors for the remaining variance in earnings.

Two implications of the Harris and Todaro assumptions are that the informal sector acts mainly as a type of labour reservoir for the formal sector, with intersectoral urban flows, at least those of a voluntary nature, being entirely from the informal to the formal sector. The first point has been refuted in a recent test in Grootaert (1992), who finds little evidence of the informal sector in Côte d'Ivoire acting as a reservoir for the formal sector. Similar findings date back to Mazumdar (1973) in a study of Bombay. Additionally, where links (mainly in the form of subcontracts) exist between the formal and the informal sector, there is likely to be some scope for enterprising formal sector employees to become informal sector employers as reported by de Soto (1989).

This reverse flow by formal sector employees would likely have to be rewarded with an increase in income, and would certainly contribute to recent evidence, as in Berry (1987), suggesting that some informal sector employees have higher earnings than previously thought possible in the informal sector. Kahnert (1987) goes further in reporting evidence that some informal sector participants actually earn as much as formal sector employees. It is possible that these high earners are enterprising informal sector employers who likely use their savings and acquired skills (including contacts in the formal sector organizations) to earn profits in excess of formal sector wages foregone. Such patterns of behaviour imply that while most of the older informals are managers, the majority of the younger migrants are typically wage-earners in the informal sector, consistent with evidence reported by de Soto (1989). This is consistent with both Mazumdar's earlier, as well as Banerjee's (1986) more recent, findings that a large proportion of informal sector workers are quite happy in their jobs and have no intention of returning home soon or even moving to a formal sector job.

A second shortcoming of the HT model is its assumption of a lottery system of job assignment and its total neglect of job search considerations. Studies have consistently shown that some workers migrate to a pre-arranged job, workers with previous urban experience tend to settle into a job quicker, and that social ties and contacts are important in finding employment. Clearly then, not all workers have the same probability of finding a job, and intensity of job search is likely to be an important factor. Similarly, the assumption that all jobs are available every period is not realistic and leads to an increase in the predicted unemployment and induced migration rates for any given wage gap. That is, the probability of a migrant job-seeker actually finding employment is likely to be considerably less than one minus the unemployment rate.

For a study of the urban labour market, the clear implication is that modelling job search should be important. Fields (1975) generalizes the jobsearch process and introduces a "murky" sector to the basic HT setup. In this model, urban residents cash in on their relative proximity to job opportunities with a higher probability of obtaining a formal sector job. That is, while the probability of rural-based search is not zero (as assumed by Harris and Todaro), urban-based search is considered more intensive and is, therefore, associated with a higher probability of finding employment. Similarly, (formal sector) job search from a state of unemployment is considered more intensive than search while participating in the informal sector. The result is a generalization of HT's results, with lower predicted unemployment rates.

Unfortunately, Fields associates the "murky" sector with underemployment and assumes that the ultimate goal for migrants is really to land a formal sector job. Consequently, he arrives at two results largely at odds with evidence presented above: first, informal sector wages, given their side benefit of increasing the employment prospects for potential rural migrants, are necessarily lower than rural sector wages; and second, the informal sector acts largely as a reservoir (or a temporary stop) for the formal sector.

New york

Additionally, Fields associates the "impoverished urban class" with the roster of unemployed workers (again, similar to Harris and Todaro). This is in sharp contrast to recent evidence, including Sahn and Alderman (1988), Yap (1976 and 1977) and Papola (1981), which fairly consistently finds

unemployment rates to be highest amongst higher educated and higher class urban residents. One obvious explanation for their preference of unemployment over employment in the informal sector revolves around the idea of status signalling. Although Fields does make a very primitive attempt at modelling preferential hiring on the basis of education, his model predicts "...the equilibrium employment rate for educated workers is one", as noted on page 177 of Fields (1975). The inconsistency lies either in the inability of employers to perfectly sort out prospective workers or else in the relative scarcity of "appropriate" positions.

The higher rates of unemployment amongst workers with either more education or family support have lead to a "luxury unemployment hypothesis" originally discussed by Myrdal (1968) and Turnham (1971). The main idea is that it would be entirely rational for workers who could both benefit from an extended search period<sup>49</sup> and also afford to carry out the strategy, to be more selective. Given the association of longer, more intensive searches with higher wage offers and a dual labour market setting with relatively few high paying positions, it follows that more educated job applicants who can reduce their search costs, probably by drawing on family support, are more likely to queue for

<sup>&</sup>lt;sup>49</sup> Given their access to a wider range of jobs, these workers are likely to face a greater dispersion in wage offers. In addition, their loftier job expectations, as mentioned a little later on, are likely to be translated into higher reservation wages. The combination of these two effects is likely to lead to the observed extended search period.

a "better" offer, rather than (being forced to) seize the first opportunity. Although clearly less than conclusive evidence, surveys on job aspirations tend to confirm the emphasis of young and relatively well educated workers on avoiding "manual" work, typically associated with the informal sector.<sup>50</sup> In a recent study of Pakistan, Sahn and Alderman (1990) find two distinct groups of unemployed workers: young, more educated men waiting for a formal sector-type job and passing over informal sector type opportunities and older, less educated men with relatively substantial family support.

A third shortcoming of the HT model is the assumption of no surplus labour in the rural sector: in figure 2 above, the departure of  $L_R'L_R$  workers leads to a loss of rural output,  $ZZ'L_R'L_R$ , equal to the sum of their marginal products. As previously discussed, however, considerable evidence from Africa and elsewhere, for example Lal (1973), suggests that, first, the remaining workers increase their supply of labour, and second, rural output stays constant despite the withdrawal of labour by cityward migrants. That is, migration induced by a program of urban job creation always results in a net increase in output.

A fourth shortcoming is the assumption of a once and for all decision on the part of the potential migrant. That is, the model totally ignores the high incidence of return migration and re-migration (temporary migration) known to

<sup>&</sup>lt;sup>50</sup> See Turnham (1971, pp. 51-2). Significantly, surveys also show that, even for the younger age groups, rates of unemployment are lower for migrants than natives. This is certainly consistent with intertemporal allocation of leisure and migrants having a lower supply price.

occur in the African context. Given the impact on private economic considerations, such as intertemporal substitution of labour, provision of effort and job search strategies, and on public policy, such as the quality and quantity of housing requested by permanent versus temporary inhabitants and the differential degree of involvement in civic and political activities, this is a rather serious omission.

Cole and Sanders (1985) propose that the HT model is in fact restricted to explaining the migration of workers with enough human capital to gamble on a formal sector position, neglecting entirely those agents migrating either temporarily or who fully intend to spend their entire time in the urban area employed in the informal sector. Their analysis relies on the premise that ruralurban migration itself is a dual phenomenon, with some migrants content to take, and hold, a job in the informal sector (presumably those with less skill accumulation), while other migrants have the formal sector as their original, or ultimate, goal (presumably these would be the workers with more human capital). This modification arose specifically in response to criticisms, Fields (1975) and Rosenzweig (1988) for example, that the HT model generates levels of unemployment higher than observed in the developing world.

The Cole and Sanders duality characterization appears to fit well with African evidence. With widely documented seasonal slack in the rural area, for example Helleiner (1975), the ready availability of informal sector work might prompt a significant level of seasonal migration. These workers not likely ever to

seriously search for a formal sector job because their rural duties minimize the benefits of joining the queue for a formal sector job. Alternatively, unmarried status or the ability to use other household members or hired labour to overcome the migrant's labour withdrawal may lead to longer stays, consistent with the evidence on job tenure of approximately 2-3 years.<sup>51</sup> In this case, we would witness three types of migration: (1) seasonal. These workers migrate solely for the opportunities available in the informal sector and serve as the link between this sector and the rural area. This strategy would be consistent with the Stark idea of migration as a means of diversification in the presence of imperfect capital markets. That is, even if the informal sector wage was no higher than the agricultural wage in the slack season, as long as there is less than perfect correlation between the income outcomes, risk averse households would tend to pursue this spatially diversified strategy. If, however, as the evidence suggests, informal sector earnings are higher than rural earnings, this type of migration is also consistent with the income differential approach;

(2) circulatory. These workers tend to stay for longer than the growing season but have no real intention of making a permanent move. Even if they obtain a formal sector job, their intention to maintain rural ties is likely to lead to a return home and a possible subsequent occurrence of migration. This pattern of

<sup>&</sup>lt;sup>51</sup> This evidence arises out of cross-sectional surveys. See, for example, Grootaert (1992), along with Elkan (1956), (1960) and (1967), Gulliver (1955) and Sabot (1979).

migration is also consistent with both the income differential and the risk diversification motivations to migrate; and

(3) permanent. These workers migrate in a specific attempt to earn some of the rents available for participation in the urban formal sector: these would be the typical Harris-Todaro (or human capital model) migrants. They make a once and for all decision to migrate, fully intend to make the move permanent and so have some incentive to choose immediate unemployment in the hopes of landing a higher paying job later.

Finally, subsequent work has demonstrated that the existence of the Todaro paradox requires rather restrictive assumptions to begin with, and may actually be quite unlikely to ever arise. Initially, Mazumdar (1976) showed that induced migration leads to an increase in the number of unemployed only if the regional wage differential is exogenously set, a less than realistic assumption. Todaro (1976) then restated the paradox in terms of the unemployment rate rather than the level, and provided a simple empirical test purportedly showing its plausibility. Blomqvist (1978), however, qualified the test as relating only to the immediate impact of small changes in the rate of job creation on the rate of change in the rate of unemployment, rather than the rate of unemployment as argued by Todaro. Additionally, Blomqvist objected to the neglect of job turnover and argued that a larger change in the rate of job creation could actually reverse Todaro's conclusions. Collier (1979) empirically rejected the Todaro

paradox using data for Tanzania.<sup>52</sup> More recently, Stark (1991) showed that the Todaro paradox only holds given certain assumptions on the elasticity of urban labour demand: creation of urban jobs leads to a decrease in the level of urban unemployment in the case of isoelastic or inelastic demand, and it always leads to a reduction in the unemployment rate.

In summary, the most important contribution of the Harris-Todaro model is probably the incorporation of uncertainty over earnings<sup>53</sup> into rural-urban migration analysis. In their framework, migrants are, quite plausibly, assumed to collect information not only on the going urban wage, but also on employment conditions, particularly the probability of gaining employment. Although one can certainly quibble about the nature of the information (should the migrant not have, at best, imperfect information?), the Harris-Todaro model certainly represented a real departure from previous analysis, and the basic mechanics of

<sup>&</sup>lt;sup>52</sup> This discussion benefitted considerably from information in footnote 8 on page 106 in Squire (1981).

<sup>&</sup>lt;sup>53</sup> The migrant is assumed to evaluate his alternatives based on the "prevailing average income for his level of education or skill attainment in the urban centre of his choice" (Todaro (1977), p. 201). It thus appears that while the potential migrant might not know the actual employment conditions, he is certainly assumed to have some knowledge of the wage distribution. Expanding the analysis to several possible urban destinations, the potential migrant would have to either know his expected income level at every location (for a strict application of the Harris-Todaro methodology), or else there would be some potential gains to sequential migration in an attempt to gather information on successive destinations (see Pessino (1991) for this approach). In the African context that Harris and Todaro were writing in, however, the general lack of a multiplicity of potential urban centres means that migrants could be plausibly assumed to have equal information on all possible locations.

the model must be an important component of migration decisions. In the next section, we outline our proposed model starting with a statement of some basic stylized facts and concluding with a discussion on support for our approach provided in the literature.

#### 7. Proposed Model

## 7.1 Stylized Migration Facts

From our previous discussion of migration patterns and our review of modelling approaches to the migration decision, we are able to extract a few stylized facts. First, we note the prevalence of circulatory migration patterns (and related employer concerns with short tenure). Despite some evidence of increasing length of stay in the urban area, the consensus is that African rural migrants retain their ties to the origin location, usually by leaving the immediate family behind, and fully intend (and most actually do) to return home at some point. The evidence generally points to the migration career ending with a final return home around the age of 40, as we also show in a later chapter on migration patterns in Malaysia.

The aggregate evidence of these movements in and out of urban areas are the widely observed high rates of both in-migration and out-migration from favourable destination locations. Whereas in places like Latin America the phenomenon is more of the sequential, or onward, migration type, in Africa, as well as in some parts of South Asia, the norm is a circular pattern between home and a particular destination location. Our proposed model will consider the effects of temporary migration on the employer's wage offers and on the migrant's *effort* choice by assigning to some migrants a positive probability to return (to the low wage area). On the hand, urbanites – either urban born or permanent migrants – have a zero probability to move out of the urban area. In this way we introduce heterogeneity on the supply side of the labour market, the need for which is widely recognized in the literature.<sup>54</sup>

Second, the literature indicates that the informal sector is a productive part of the economy with numerous links to formally registered enterprises. The choice of much migrant labour, the informal sector appears to fit the perfectly competitive view of the firm, with flexible wages (generally below minimum wage standards and also below the formal sector prevailing wage) and jobs always available at the current wage as indicated by the previously mentioned propensity of recently arrived migrants to find jobs relatively quickly. Urbanites, on the other hand, tend to endure longer periods without a job, especially those with more education – what the literature refers to as luxury unemployment – as

<sup>&</sup>lt;sup>54</sup> Although we do not consider differences in skills or experience, such sources of heterogeneity would only complicate our account of the labour market a little. The higher wage in the urban formal sector would attract the more talented workers, who would then be more likely, given the larger rewards they could expect, to bypass the informal sector. Depending on the number of formal jobs then, some of the less talented amongst the permanent workers might actually end up choosing to work in the informal sector. Such considerations would likely not change our solution in any material way but would be fundamentally consistent with the luxury unemployment hypothesis as well as the findings in Sahn and Alderman (1990).

they join the queue for formal sector jobs.<sup>55</sup> We then choose to model the informal sector as a profit maximizing sector with the wage clearing the (informal) labour market.

The combination of job availability and a wage constant over the worker's life cycle makes the informal sector relatively more attractive to circular migrants: temporary absences are not penalized with the capital loss associated with re-engaging in a time-consuming job search or the capital loss associated with skipping out on the rising wage profile. In considering whether to invest in seeking a formal sector job then, a migrant balances the additional income from a higher wage, the investment return, against the longer period of unemployment, the investment cost. In deciding on whether to make the move permanent, the migrant balances the capital gain associated with the rising wage profile against the additional income from returning home and maintaining the rural ties. That is, the required investment for a formal sector job can be viewed as a destination specific investment.

Urbanites, however, are likely to find the formal sector more attractive. These workers do not have the alternative source of income associated with

<sup>&</sup>lt;sup>55</sup> Pervasive evidence mentioned by Rosenzweig (1988) also points to an increasing wage tenure profile in the formal sector. The informal sector, on the other hand, with its flexible wage and relatively high labour intensity and low training requirements is likely to exhibit a more gradual wage tenure profile.

Note that although the increasing wage tenure profile may be used by the firm as a type of bond intended to reduce the tendency to shirk, this strategy is unlikely to eliminate the need for efficiency wages, especially in developing countries where workers are much more likely to be capital constrained.

returning to a rural area, which makes them less likely to incur the capital loss associated with leaving their job but more likely to make an urban-specific investment, such as enduring a period of unemployment, in order to increase their income.<sup>56</sup> Such sectoral choices by workers would then suggest that migrants tend to have shorter unemployment spells, a conclusion supported by evidence presented in, for example, Yap (1977) and, more recently, Banerjee (1991) who finds that, on arrival in the city, 63.6% of Indian rural-urban migrants find a job within one week and 87.8% find a job within a month.

Third, empirical evidence also supports the idea of a labour market segmented along the formality/informality lines. The two most common characterizations are in terms of either capital/labour intensity or size of operation, with informal firms generally viewed as employing more labour intensive production techniques and having a smaller labour force. Significantly for our model, both characteristics associated with informal firms are typically associated with lower monitoring costs as demonstrated by Esfahani and Salehi-Isfahani (1989).

Finally, the literature indicates that family relationships are likely to be asymmetrical and that individual maximization is a plausible assumption for (male) migrant behaviour. The underlying assumption is that the institutional arrangements, either family sharing rule or some form of altruism, make

<sup>&</sup>lt;sup>56</sup> Notice that this investment is not firm-specific so that it does not prevent intra-sectoral inter-firm turnover.

individual maximization compatible with family maximization. As noted by Nelson (1975), conditions in Africa and South Asia are conducive to such outcomes as prevailing land arrangements allow the migrant to retain title and social customs facilitate the continued cultivation of the land plot by other family members.

Our proposed model takes the income differential approach in postulating that it is the possibility of earning a higher income that drives the migrant to seek work in the urban area and also assumes individual utility maximization. We assume that migrant workers have either a zero or a positive probability to return, and further, that such information is private and known only to the worker himself. There are also two types of firms, one with higher monitoring costs, the formal sector, and the other with lower monitoring costs, the informal sector – for these firms, and without loss of generality, we set the monitoring costs at zero so that their wage is fully flexible. This information is known to both the workers and the firms.

While the Harris and Todaro model solves the problem of a persistent urban-rural wage gap in the face of large migration flows by assuming an institutionally fixed urban wage, we previously mentioned that Mazumdar (1990) notes that high urban wages in the formal sector actually predate the era of trade unions or government intervention in the labour market. Additionally, widespread evidence from developing countries concerning the ability of government agencies to enforce labour regulations does not lend credibility to the possibility that high wages would be the result of legislation. We propose instead that given their relatively higher monitoring costs, formal sector firms prefer the workers with the zero probability of voluntarily returning to the home area, or quitting. However, because formal sector firms cannot distinguish between worker types, they face a moral hazard problem, and with little labour market information to use as signals, these firms must instead use their wages as a method of signal extraction. In this way, we provide the microeconomic underpinnings for the rigid<sup>57</sup> wage observed in the urban formal sector.

### 7.2 Support from the literature

The economic literature, and to a lesser extent the anthropological literature, contain persistent references to the issue of turnover costs, principally resulting from the relatively short job tenure<sup>58</sup> associated with circulatory migration.<sup>59</sup> The culmination of this research effort was, in part, Stiglitz's 1974 labour turnover model and Fry's empirical testing of that model in 1979.

<sup>&</sup>lt;sup>57</sup> Recent macroeconomic shocks have actually shown formal sector wages to be downwardly flexible, but the essential point is the continuing gap between sectoral wages.

<sup>&</sup>lt;sup>58</sup> At this point, we should note that labour turnover is an indication of one of two possible phenomena: movement into and out of the urban labour force associated with circular migration and the movement between firms associated with inter-firm mobility.

<sup>&</sup>lt;sup>59</sup> See, for example, Elkan's 1956 study of operations at the East African Tobacco Company, and Elkan (1967) and the references therein.
We provide a different explanation of the link between short tenure and the income statement. Although the obvious concern about turnover costs points to important issues of job training and securing the returns to that training, it should be remembered that employees likely to guit their job earlier (to return home to tend to their farm, for example) are also more likely to have less incentive to provide effort in their job. The reason for the different effort response is quite straightforward. The employer's inability to sort out workers perfectly (either by initially screening job applicants or by perfectly observing effort on the job) leads, in part, to a higher wage offer in order to induce permanency. Employees with long-term career ambitions (or a lower probability of leaving) are thus induced to supply effort in order to avoid the capital loss associated with a relatively long period of unemployment. For their part, workers with a higher probability of leaving the firm place relatively more weight on short-term considerations, and have, therefore, much less incentive to devote as much effort as the career-oriented workers. The essence of the story is that, given the inability of employers to perfectly screen job applicants, an assumption consistent with the relative lack of labour market information available in developing countries, it is ultimately the employer's inability to perfectly observe effort that necessitates efficiency type wages. Furthermore, workers who are more likely to leave the job will require more compensation in exchange for their effort.

The different alternatives available to urbanites versus migrants also justifies the consideration of supply side heterogeneity. That is, urbanites are likely to have little access to rural land and are then likely not to move into the countryside. Similarly, land-poor rural dwellers are more likely to make a permanent move if they migrate to the urban area. However, given the wide access to land in most of Africa and in accordance with the anthropological evidence previously presented, a considerable portion of rural-urban migrants tend to be circular migrants.<sup>60</sup> Given that migrants' intentions with respect to length of stay may be expected to affect their economic behaviour (thus the need for screening on the part of employers), the resultant supply side heterogeneity is likely to be important in explaining both urban labour market behaviour and migration.<sup>61</sup>

As noted by Nelson (1975) on page 742 however, "the consequence of variation in the permanence of cityward migration are largely unexplored". In fact, references to the need of explicit consideration of labour heterogeneity

<sup>&</sup>lt;sup>60</sup> As we mentioned previously and as echoed by Nelson (1975) on page 736: "Heavy return migration is often associated with more or less guaranteed access to land... Such conditions are common in much of Africa and South Asia ... The land represents not only a modicum of security but also a nonconvertible asset".

<sup>&</sup>lt;sup>61</sup> "...precommitments to stay or leave influence migrants' behavior in the cities, even if their intentions are not always fulfilled." (Nelson (1975), p. 741).

abound in the economic and anthropological literature,<sup>62</sup> with two proposals being advanced for the regime switch: differences in human capital and differences in the "supply price of labour". The only attempt at modelling hiring rules that discriminate on the basis of human capital was Fields (1975), discussed above. Empirically, human capital has proven to be consistently significant in explaining urban wage differentials<sup>63</sup> and is likely to be important in the migration decision as well, especially for those countries with a heavy concentration of facilities in the urban area.

The anthropological, and to some extent the economic literature, for example van der Gaag and Vijverberg (1989), also make persistent references to the phenomenon of credentialism: rising educational requirements despite the apparent stability of job descriptions. Such a practice would certainly be consistent with an attempt to use educational achievement, in a context where the great majority of educational facilities are located in the city, as a screening device given a rising level of educational attainment in the general populace. The implication for the urban labour market is that the barrier to formal sector employment is then likely to be in terms of human capital acquired; and attempts

<sup>&</sup>lt;sup>62</sup> See Nelson (1975). Additionally, Elkan (1967) states on page 106 that the temporary nature of a significant part of migration flows has "...important policy implications for all those concerned with social security and urban growth in Africa". See also chapter 8 in Squire (1981).

<sup>&</sup>lt;sup>63</sup> See, for example, Sahn and Alderman (1991), Banerjee (1986) and Mazumdar (1981).

to separate firms into a formal/informal type of dichotomy would be better served by adopting an occupational definition of the informal sector (where a busboy at a modern hotel or the dishwasher at the same hotel would all be classified as belonging to the informal sector) instead of the usual size or government registration criteria.

The second proposal, differences in the "supply price of labour", is mainly associated with Dipak Mazumdar. In fact, Mazumdar (1990) provides an "eclectic" theory of wage differentials composed of three complementary explanations. First, migrants who retain their ties to the rural sector and leave their families behind, fully intending to stay for short periods only, have a lower reservation wage than more permanent migrants. "The major reason is that the output forgone in the family farm due to the departure of one member for a short period is very low - and could, indeed, be negligible - while the migration of the entire family from the rural economy means the loss of the entire income of the farm" as noted on page 350 of Mazumdar (1990). Given that some firms, likely the larger, more modern enterprises commonly associated with the formal sector, can benefit more from a more stable work force, they will offer higher wages in the hope of attracting either urbanites or the more permanent migrant. By contrast then, the informal sector should be dominated by the less stable (or temporary) migrants.

Second, formal sector firms may be led to increase wages still further due to the usual efficiency wage argument: given incomplete information on worker types and imperfect ability to monitor effort, the employer may find it more profitable to pay a higher wage and induce higher effort than to take on more workers. This is the argument common to Shapiro and Stiglitz (1984) and Esfahani and Salehi-Isfahani (1989). These two points aim to explain differentials in entry wages.

Whereas the initial two components in Mazumdar's hypothesis aim to explain differentials in entry wages, the third element suggests that formal sector firms are likely to be characterized by steep wage-experience profiles. The reason is that large firms may be able to take advantage of internal labour markets to promote from within and effectively reduce screening and training costs. The implication is that the firm would conduct external hiring for lower level jobs. Empirical evidence appears to support the existence of steep wage profiles in the formal sector, as noted by Rosenzweig (1988).

In light of our earlier discussion, the relevance of Mazumdar's ideas seem particularly relevant in the African context. With the widely acknowledged maintenance of rural ties by African migrants, the patterns of rural work allocation, active rural labour market and the (resulting) predominance of circulatory migration patterns, African urban labour markets are likely to be populated by some workers who intend permanently to supply labour in the urban area and by other workers whose intentions are to remain in the urban area only temporarily. The latter type would be Mazumdar's so-called "less stable" labour who are more likely to work in the informal sector, while the permanent urban

residents would be the preferred choice of formal employers given the likelihood of a firmer commitment to urban employment.

Our proposed model captures this form of duality by assigning to each worker either a zero or a positive probability to leave the urban labour market and return to the rural area. Our model then goes beyond the Mazumdar idea of different supply prices of labour by also considering the (effort) response of temporary versus permanent urban labourers. Although standard economic theory correctly predicts that temporary labourers, given their positive probability to return to a (rural) low wage area, prefer to supply labour in the urban area and enjoy leisure while in the rural area, it does not differentiate between their commitment to urban employment (and possible differences in effort response) and that of less mobile urban residents. Our model predicts that, ex ante, given the permanent labourers' longer term ability to benefit from supplying more effort, they have more incentive to "work harder".<sup>64</sup>

<sup>&</sup>lt;sup>64</sup> A kinder, gentler interpretation of the difference in commitment would emphasize the incentive to acquire job-specific skills. In this case, offering a higher wage is synonymous with attracting "the people most likely to be interested in acquiring skills" or choosing "the cream of the available work force" Berg (1969) as quoted in Squire (1981, p. 114). This interpretation would appear to be consistent with a labour quality version of the efficiency wage model.

Note that a positive probability of coming back to urban area after the return to the rural area might serve to moderate the effort eliciting wage on the part of migrants. In effect, our current model abstracts from reputation effects which would obviously require a dynamic specification.

Note also that we assume risk neutrality on the part of migrants. The main impact of risk averse behaviour on the part of migrants is that the equilibrium migration condition will necessarily require the expected urban wage to be strictly (continued...)

Esfahani and Salehi-Isfahani (1989) examine the relationship between effort observability and worker productivity. By postulating a negative relationship between establishment size and observability, they argue that monitoring effort in the formal sector is costlier than in the informal sector. Their model, a rather abstract version of a shirking model of efficiency wages, is then able to account for some basic intrasectoral differences: a wage gap unexplainable by differences in skills and incompatible with high rates of unemployment; commonly unexplainable productivity differences; higher unemployment rate among formal sector job seekers and duality in terms of technology, size, factor intensity and management techniques.

Notice that our monitoring model of efficiency wages leads to disguised unemployment in the formal sector: migrants with a positive probability to return provide less than 100% effort and, as a result, their joint output could be produced by fewer, but more committed, workers. Up to now, the literature has focused on disguised unemployment either in the urban informal sector or in the rural sector, a treatment congruent with the initial Lewis assumption of income sharing in either one of those sectors.

The temporary nature (and its associated costs) of a considerable part of the labour force also runs counter to the usual assumption of a perfectly elastic supply of labour. While seasonal slack and lack of rural income earning

greater than the known rural wage instead of the customary equality.

<sup>&</sup>lt;sup>64</sup>(...continued)

opportunities may imply a perfectly elastic supply of (potential) labour hours, the work skills and habits resulting from the rural orientation of a large part of the work force imply that effectively the supply elasticity is less than perfectly elastic.

In summary, our proposed model is consistent with (1) Mazumdar's findings of efficiency-type wages long before minimum wage legislation existed, (2) repeated references in the literature to the co-existence of permanent and temporary migration (as well as to the fact that the implications are still largely unexplored), (3) Esfahani and Salehi-Isfahani's results that links larger, more capital-intensive firms with higher monitoring costs, and (4) accepted duality nature of urban firms, with the formal sector typically composed of larger and generally more capital intensive firms, co-existing with informal sector firms.

#### **APPENDIX: THE 'TODARO PARADOX'**

Todaro (1977) provides an algebraic proof of the 'Todaro paradox' on page 202: let the probability of obtaining a job in the (formal) urban sector be  $\pi = \frac{\lambda N}{S - N}$ , where  $\lambda$ =the net rate of urban new job creation, N=the number of (formal) urban jobs and S=the total urban labour force (urban residents plus migrant job seekers). By definition, the expected urban-rural real income differential is d = w·  $\pi$  - r, where w=the urban real wage rate and r=the rural real wage rate.

The inherent assumption that the level of migration responds to the wage differential means that the supply of labour to the urban (formal) sector is a function of that wage differential, i.e.  $S = f_s(d)$ . Assuming that labour demand is a function of both w and a (a policy variable such as a job promotion programme), we can write  $\lambda = f_d(w;a)$ , with  $\frac{\partial \lambda}{\partial a} > 0$ .

If the government attempts to induce an increase in the rate of growth of urban labour demand,  $\lambda$ , what will be the effect on unemployment, (S - N)? To answer this question, we need to examine the response of S to a change in a, i.e.  $\frac{\partial S}{\partial a}$ . The number of unemployed workers will increase if the increase in labour supply (new migrants) exceeds the increase in the number of urban jobs, i.e. if  $\frac{\partial S}{\partial a} > \frac{\partial(\lambda N)}{\partial a} = \frac{N\partial\lambda}{\partial a}$ . Differentiating and substituting, we arrive at:

$$\frac{\partial S/S}{\partial d/d} > \frac{w \cdot \pi - r}{w} \cdot \frac{(S - N)}{S}$$

Whether, in fact, a policy of job expansion does lead to an increase in unemployment is then an empirical question. According to Todaro, this expression can be met with quite realistic values for developing countries (e.g. w=60, r=20,  $\pi$ =0.50, (S-N)/S=0.20).

# BIBLIOGRAPHY

- Adams, J. (1991). "The Rural Labour Market in Zimbabwe", Development and Change, 22: 297-320.
- Banerjee, B. (1986). Rural to Urban Migration and the Urban Labour Market. Delhi: Himalaya Publishing House.
- Banerjee, B. (1991). "The Determinants of Migrating With a Pre-arranged Job and of the Initial Duration of Urban Unemployment", *Journal of Development Economics*, 36: 337-351.
- Bardhan, P.K. (1964). "Factor Market Disequilibrium and the Theory of Protection", Oxford Economic Papers (New Series), 16: 375-88.
- Berry, A. (1987). "The Labour Market and Human Capital in LDCs", in Gemmell, N, ed., *Surveys in Development Economics*, pp. 205-235. Oxford: Basil Blackwell.
- Berry, A. and R. Sabot (1978). "Labour Market Performance in Developing Countries: A Survey", *World Development*, 6: 1199-1242.
- Blomqvist, A. (1978). "Urban Job Creation and Unemployment in LDCs", Journal of Development Economics, 5: 3-18.
- Cole and Sanders (1985). "Internal Migration and Urban Employment in the Third World", *American Economic Review*, 75: 481-94.
- Collier, P. (1979). "Migration and Unemployment: A Dynamic General Equilibrium Analysis Applied to Tanzania", *Oxford Economic Papers*, 31: 205-36.
- Corden, W. (1974). *Trade Policy and Economic Welfare*. Oxford: Oxford University Press.
- Corden, W. and R. Findlay (1975). "Urban Unemployment, Intersectoral Capital Mobility and Development Policy", *Economica*, 42: 59-78.
- DaVanzo, J. (1981). "Repeat Migration, Information Costs, and Location Specific Capital", *Population and Environment*, 4: 45-73.

- de Soto, Hernando (1989). The Other Path: The Invisible Revolution in the Third World. New York: Harper & Row.
- Deolalikar, A. and W. Vijverberg (1987). "A Test of Heterogeneity of Family and Hired Labour in Asian Agriculture", Oxford Bulletin of Economics and Statistics, 49: 291-305.
- Elkan, W. (1956). "An African Labour Force", East African Studies, No. 7.
- Elkan, W. (1960). Migrants and Proletarians. London: Oxford University Press.
- Elkan, W. (1967). "Circular Migration and the Growth of Towns in East Africa", *International Labour Review*, 96: 581-9 as reprinted in Jolly, R. et al, eds. (1973) *Third World Employment: Problems and Strategy*. Baltimore: Penguin Books.
- Elkan, W. (1975). "Is a Proletariat Emerging in Nairobi?", *Economic* Development and Cultural Change, 24: 695-706.
- Esfahani, H.S. and D. Salehi-Isfahani (1989). "Effort Observability and Worker Productivity: Towards an Explanation of Economic Dualism", *The Economic Journal*, 99: 818-836.
- Falaris, E. (1987). "A Nested Logit Migration Model with Selectivity", International Economic Review, 28: 429-43.
- Feige, E. (1990). "Defining and Estimating Underground and Informal Economies: The New Institutional Economics Approach", World Development, 18: 989-1002.
- Fields, G.S. (1975). "Rural-Urban Migration, Urban Unemployment and Underemployment, and Job-Search Activity in LDCs", *Journal of Development Economics*, 2: 165-87.
- Fields, G.S. (1982). "Place to Place Migration in Colombia", *Economic* Development and Cultural Change, 30: 539-58.
- Fry, J. (1979). "A Labour Turnover Model of Wage Determination in Developing Economies", *The Economic Journal*, 89: 353-69.
- Gillis M., D. Perkins, M. Roemer and D. Snodgrass (1992). *Economics of Development*, Third Edition. New York: W.W. Norton.

- Grootaert, C. (1992). "The Position of Migrants in the Urban Informal Labour Markets in Côte d'Ivoire", *Journal of African Economies*, 1: 416-45.
- Gulliver, P. (1955). "Labour Migration in a Rural Economy", *East African Studies*, No. 6.
- Harris, J. and M.P. Todaro (1970). "Migration, Unemployment and Development: A Two-Sector Analysis", *American Economic Review*, 60: 126-42.
- Hart, Keith (1973). "Informal Income Opportunities and Urban Employment in Ghana", *The Journal of Modern African Studies*, 11: 61-89.
- Heckman, J. (1976). "The common structure of statistical models of truncation, self selection and limited dependent variables, and a simple estimator for such models", *The Annals of Economic and Social Measurement*, 5: 475-92.
- Helleiner, G. (1975). "Smallholder Decision Making: Tropical African Evidence", in Reynolds. L.G., ed., *Agriculture in Development Theory*. New Haven: Yale University Press.
- International Labour Office (1973). Employment, Incomes and Equality: A Strategy for Increasing Productive Employment in Kenya. Geneva: ILO.
- Jules-Rosette, B. (1985). "The Women Potters of Lusaka: Urban Migration and Socioeconomic Adjustment", p. 82-112 in Lindsay, B., ed., African Migration and National Development. University Park: Pennsylvania State University Press.
- Kahnert, Friedrich (1987). "Improving Urban Employment and Labor Productivity", World Bank Discussion Papers No. 10. Washington, D.C.: The World Bank.
- Knight, J. B. (1972). "Rural-Urban Income Comparisons and Migration in Ghana", Oxford Bulletin of Economics and Statistics, 34: 199-228.
- Lal, D. (1973). "Disutility of Effort, Migration and the Shadow Wage Rate", Oxford Economic Papers, 25: 111-27.
- Lewis, W. (1954). "Economic Development with Unlimited Supplies of Labour", Manchester School of Economic and Social Studies, 22: 139-91.

- Lewis, J. (1985). "Village-Level Restraints on the Exodus from Rural Mali", p. 64-81 in Lindsay, B., ed., African Migration and National Development. University Park: Pennsylvania State University Press.
- Lucas, R. and O. Stark (1985). "Motivations to Remit: Evidence from Botswana", *Journal of Political Economy*, 93: 901-18.
- Mazumdar, D. (1973). "Labour Supply in Early Industrialization: The Case of the Bombay Textile Industry", *Economic History Review*, 26/3: 477-96.
- Mazumdar, D. (1976). "The Rural-Urban Wage Gap, Migration and the Shadow Wage", Oxford Economic Papers, 28: 406-25.
- Mazumdar, D. (1981). The Urban Labour Market and Income Distribution: A Study of Malaysia. New York: Oxford University Press.
- Mazumdar, D. (1987). "Rural-Urban Migration in Developing Countries", in Mills, E.S., ed., *Handbook of Regional and Urban Economics*, Volume 2. Amsterdam: North-Holland Publishing.
- Mazumdar, D. (1990). "Government Intervention and Urban Labour Markets in LDCs", in Scott, M. and Lal, D., ed., *Public Policy and Economic Development: Essays in Honour of Ian Little*, pp. 345-360. Oxford: Clarendon Press.
- Meier, G. (1989). Leading Issues in Economic Development, Fifth Edition. Oxford: Oxford University Press.
- Myrdal, Gunnar (1968). Asian Drama, Volume III. New York: Pantheon.
- Nam, C. et al, eds. (1990). *International Handbook on Internal Migration*. New York: Greenwood Press.
- Nelson, J. (1975). "Sojourners versus New Urbanites: Causes and Consequences of Temporary versus Permanent Cityward Migration in Developing Countries", *Economic Development and Cultural Change*, 24: 721-57.
- Pessino, C. (1991). "Sequential Migration Theory and Evidence from Peru", Journal of Development Economics, 36: 55-87.
- Preston, S.H. (1979). "Urban growth in developing countries: A demographic reappraisal", *Population and Development Review*, 5: 195-215.

- Robinson, C. and N. Tomes (1982). "Self-selection and interprovincial migration in Canada", *Canadian Journal of Economics*, 15: 474-502.
- Rosenzweig, M.R. (1988). "Labour Markets in Low-Income Countries", in Chenery, H. and T.N. Srinivasan, eds., *Handbook of Development Economics*, Volume 1: 713-762. Amsterdam: North-Holland Publishing.
- Sabot (1979). *Economic Development and Urban Migration*. Oxford: Oxford University Press.
- Sahn D. and H. Alderman (1990). "Formal and Informal Sector Wage Determination in Urban Low-Income Neighborhoods in Pakistan", Living Standards Measurement Study Working Paper No. 65. Washington: The World Bank.
- Schultz, T. (1964). *Transforming Traditional Agriculture*. New Haven: Yale University Press.
- Sen, A. (1966). "Peasants and Dualism With and Without Surplus Labour", Journal of Political Economy, 74: 425-450.
- Shukla, V. and O. Stark (1985). "On Agglomeration Economies and Optimal Migration", *Economic Letters*, 18: 297-300.
- Sjaastad, L. (1962). "The Costs and Returns of Human Migration", *The American Economic Review*, 70: 80-93.
- Squire, L. (1981). *Employment Policy in Developing Countries*. New York: Oxford University Press.
- Stark, O. (1991). The Migration of Labour. Cambridge: Basil Blackwell.
- Stichter, S. (1985). Migrant Laborers. Cambridge: Cambridge University Press.
- Stiglitz, J.E. (1974). "Wage Determination and Unemployment in LDCs", Quarterly Journal of Economics, 88: 194-227.
- Stiglitz, J.E. (1982). "Alternative Theories of Wage Determination and Unemployment: the Efficiency Wage Model", in Gersovitz, M., E. Diaz-Alejandro, G. Ranis and M. Rosenzweig, eds., *The Theory and Experience* of Economic Development, pp. 78-106. London: George Allen and Unwin.

- Todaro (1969). "A Model of Labour, Migration and Urban Unemployment in Less Developed Countries", *American Economic Review*, 59: 138-48.
- Todaro (1977). Economic Development of the Third World. New York: Longman Publishing.
- Turnham, D. (1971). The Employment Problem in Less Developed Countries: A review of the evidence. Paris: OECD.
- United Nations (1980). Patterns of Urban and Rural Population Growth. New York: United Nations.
- United Nations (1990). World Population Monitoring 1989. New York: United Nations.
- United Nations (1991). World Population Prospects 1990. New York: United Nations.
- Weiss, A. (1980). "Job Queues and Layoffs in Labor Markets with Flexible Wages", Journal of Political Economy, 88: 526-38.
- Williamson, J.G. (1988). "Migration and Urbanization", in Chenery, H. and Srinivasan, T.N., ed., *Handbook of Development Economics*, Volume 1: 425-65. Amsterdam: North-Holland Publishing.
- World Bank (1990)., World Development Report. Washington, D.C.: The World Bank.
- van der Gaag and Vijverberg (1989). "Wage Determinants in Côte d'Ivoire: Experience, Credentials and Human Capital", *Economic Development and Cultural Change*, 37: 371-81.
- Vijverberg, W. (1989). "Labor Market Performance as a Determinant of Migration", Living Standards Measurement Study Working Paper No. 59, Washington: The World Bank.
- Vijverberg, W. and J. van der Gaag (1990). "Testing for Labour Market Duality: The Private Wage Sector in Cote d'Ivoire", *Living Standards Measurement Study Working Paper No. 66*, Washington: The World Bank.
- Yap, L. (1977). "The Attraction of Cities: A Review of the Migration Literature", Journal of Development Economics, 4: 105-127.

Zachariah, K. and J. Conde (1981). *Migration in West Africa: Demographic Aspects*. New York: Oxford University Press.

# CHAPTER 2.

# RURAL-URBAN MIGRATION WITH EFFICIENCY WAGES

#### 1. Introduction

The process of internal migration in developing countries has attracted a considerable amount of research effort. From the beginning, one of the central quests has been an explanation for a persistent rural-urban wage gap. Lewis (1954) interpreted the wage gap as simply the reflection of immobility and, concentrating on the rural sector, proposed income sharing as the explanation for the difference between rural wages and marginal productivity. The implication is that the market system tends to allocate too much labour to the rural sector, a view often challenged by later evidence indicating only seasonal labour surplus in most low income countries. Next, Harris and Todaro (1970) (HT) suggested that an institutionally fixed urban wage was responsible both for the wage gap and for high levels of rural-urban migration in the face of urban unemployment queues. This model generated an enormous literature by itself, and although severe criticisms, particularly concerning the notion of an institutionally fixed urban wage, are now generally levied against it, it did generate the critical idea that migrants attach some uncertainty to earnings.

More recently, the search for an explanation for the continued wage differential in the face of high migration rates has centred on the role of information. Several explanations have been advanced in the literature, including Pessino (1991), who, in the context of a dynamic learning model, argues for the role of imperfect information on the part of the migrant, and the Roy (1951) model which suggests that the observed wage differential (based on the experience of migrants) is overestimated for non-migrants.

We propose to model internal migration in the HT tradition, but with an expanded role for information. In explaining the persistence of rural-urban wage gaps in developing countries, the conventional HT model argues for an equilibrium gap, with industrial wages fixed by minimum wage legislation, urban unemployment and flexible rural wages accounting for the divergence. Potential migrants compare the expected value of the wage in urban (formal) employment to the wage in the rural area. Due to the presence of urban unemployment, the migrant is likely to be initially unemployed, as suggested by Harris and Todaro (1970), or working in the informal sector, as suggested initially by Hart (1973).<sup>65</sup>

The HT equilibrium condition, equation 8 on page 129 of Harris and Todaro (1970), is  $W_A = W_U^e = \overline{W}_M \frac{N_M}{N_U}$ , so that, in equilibrium and in the presence of free flow between sectors, the agricultural wage ( $W_A$ ) is equal to the expected urban wage( $W_U^e$ ). Given a random job selection process, the expected

<sup>&</sup>lt;sup>65</sup> Although the HT model has also generated a large number of extensions, as surveyed in Battacharya (1993), the flavour of these models remains largely unaltered.

urban wage is calculated by multiplying the fixed manufacturing wage ( $W_M$ ) by the probability of finding a manufacturing job  $\left(\frac{N_M}{N_U}\right)$ , where the numerator represents the number of manufacturing jobs and the denominator represents the total urban labour force. The inherent assumption is that the currently unemployed will, in future, have the same probability of employment as the currently employed – by, for example, acquiring a similar set of skills. In our notation,  $W_A = W_F \frac{L}{N}$ , where  $W_F$  is the formal sector wage. Defining unemployment as  $U = \frac{N-L}{N}$  and re-arranging, we arrive at the well known HT result

$$U = 1 - \frac{W_A}{W_F} \,. \tag{1}$$

Significantly, Harris and Todaro left the reason for downward rigidity in urban wages unspecified. We attempt to shed some light on this black box by endogenizing both urban wages and the probability of employment. Our proposed model is a straightforward extension of the monitoring model of efficiency wages, originally proposed by Shapiro and Stiglitz (1984) (SS).<sup>66</sup>

Initially, we specify a very simple model: one urban sector and a homogeneous labour force (some initially urban and some rural) with unrestricted flow of labour. In this model, (1) workers choose how much effort to put in

<sup>&</sup>lt;sup>66</sup> The earliest application of efficiency wages appears to have been in the economic development literature with Leibenstein (1957) originally linking higher wages and enhanced productivity through improved nutritional standards. Since then, the notion of efficiency wages has generated an extensive literature concentrating on explanations other than the nutrition-productivity relationship and with applications mainly to developed economies.

(binary variable: workers either shirk or not), (2) firms cannot directly observe the workers' choice,<sup>67</sup> (3) it is costly to check up on workers, and (4) there are two possible states for workers in the urban area, employment and unemployment. Subsequently, we follow Jones (1987) who extends the basic SS framework to two employment sectors, primary and secondary, and allows workers to be one of two types, high or low turnover.

We interpret the two employment sectors in Jones (1987) as the formal and informal urban sectors in a developing country, and we add a rural agricultural sector from which workers may move to the urban area, effectively incorporating the prospect of rural-urban migration into the monitoring model. Within the urban sector itself, we add the possibility of mobility from the informal sector and distinguish between effort in the formal and the informal sectors. As in Jones (1987), the urban labour force is composed of low turnover workers, interpreted here to be either original urban residents or permanent migrants, and high turnover workers, interpreted here to be temporary migrants. Although workers know their own probability to return to the rural area (either zero or positive), employers do not, and this asymmetry of information essentially gives rise to an urban-rural wage gap and an unemployment queue.

In our model, return migration is not linked to "migration failure". Instead, the probability to return is considered exogenous and not dependent on

<sup>&</sup>lt;sup>67</sup> Firms' inability to sort workers is consistent with a random probability of obtaining employment.

urban labour market outcomes. This is consistent with empirical evidence, for example Yap (1977, p. 256), suggesting that disappointed workers are a small fraction of the return flow. Additionally, conventional thinking now views the informal sector as a vibrant and productive part of the urban economy while recent survey findings suggest that some migrants, the temporary movers in our model, move explicitly to take advantage of income earnings opportunities in this sector.<sup>68</sup> Other workers, either urban residents or permanent migrants, strictly prefer unemployment over informal sector employment. We take such sectoral choices as signalling on the part of workers who essentially know their own characteristics. Employers, given their relative lack of information, find such signals to be of use in the design of an optimal hiring policy, and the equilibrium solution reached depends critically on whether workers completely sort themselves in terms of sectoral choices.

The proposed model suggests:

 a microeconomic foundation for a rural-urban wage gap as well as an intraurban formal-informal sectoral wage structure. Unlike previous models, for example Pessino (1991), our explanation is based on imperfect information on the part of the employer instead of the employee;

<sup>&</sup>lt;sup>68</sup> In this way, we diverge from the HT tradition of characterizing the informal sector as strictly a low-productivity "subsistence" sector with earnings typically lower than earnings in the rural sector and assuming that migrants are more likely to endure periods of unemployment. For a more complete exposition of our view of labour market operations, see the previous literature review chapter. For further corroborating evidence, and a view sympathetic to our assumptions, see Williamson (1988).

- ii. lower levels of unemployment and rural-urban migration than HT-type models, countering one of the main criticisms of HT as mentioned in the literature review;
- iii. formal sector employers prefer, and optimally discriminate in favour of, resident workers (a category including both urban residents and permanent migrants), and discriminate against temporary migrants;
- iv. the formal sector tends to be the dominion, or under complete screening,
   even the exclusive domain, of resident workers while the informal sector is
   restricted to migrant workers;
- v. under certain specified conditions, (a) informal sector wages may be higher than formal sector wages and (b) informal sector wages may be negative;
- vi. a non-human-capital based explanation for the incidence of luxury unemployment, or the idea that unemployed workers in developing countries optimally forgo informal sector employment;
- vii. migrants tend to be less likely to endure periods of unemployment. In general, it is only permanent migrants which are likely to find unemployment the optimal choice. This is in contrast to most HT-based models and in accordance with evidence cited in Yap (1977) and Williamson (1988).

We start by providing an introduction to the basic SS monitoring model of efficiency wages and then successively add a rural sector, an urban informal sector, and return migration. The last section concludes with a summary and suggestions for further research.

### 2. The Monitoring Model of Efficiency Wages

Using the methodology of asset equations, SS posit the following expected lifetime utility function<sup>69</sup> for workers deciding not to shirk:

$$rV_{F}^{N} = W_{F} - e + b(V_{U} - V_{F}^{N}), \qquad (2)$$

where the subscript F denotes formal sector employment, the subscript U indicates the employee is not in the formal sector (SS call this alternative state unemployment, an alternative may be to label it the informal sector), the superscript N denotes the not shirking state, and r represents the rate of time preference. The first term,  $W_F$  - e, wages less effort, represents the immediate payout, while the last term represents the expected capital gain associated with a state change (b is the probability the worker will leave the firm for exogenous reasons like relocation).

Analogously, those deciding to shirk have the following expected lifetime utility function:

$$rV_{F}^{S} = W_{F} + (b+q)(V_{U} - V_{F}^{S}), \qquad (3)$$

where the superscript S denotes the choice of shirking, and q represents the probability of being monitored (and fired).<sup>70</sup>

<sup>&</sup>lt;sup>69</sup> See Appendix A for a complete derivation.

<sup>&</sup>lt;sup>70</sup> Assuming that monitoring technologies are unlikely to be perfectly designed or implemented, false positives and negatives may occur. To simplify the analytical treatment, we have incorporated the false positives in the exogenous separation parameter and the false negatives in the probability of being fired. Thus, exogenous separation may come about because of worker relocation, product market changes or a false positive. On the other hand, a shirking worker (continued...)

Employees evaluate the utility of each of these streams and choose not to shirk if, and only if, the no shirking condition (NSC),  $V^N \ge V^S$ , is met.

Employees are assumed to be risk neutral, so that their objective can be taken to be the maximization of lifetime expected utility. If the value of the no shirking stream is larger, employees will then maximize expected discounted wages less effort costs. Substituting for the required utility functions, we can then re-write the NSC to yield the minimum wage, W<sup>\*</sup>, employers must pay to ensure employees choose to supply effort.

Re-arranging (2), we get  $V_F^N = \frac{1}{(r+b)} \left( W_F - e + b V_U \right)$ . Re-arranging (3), we get  $V_F^S = \frac{1}{(r+b+q)} \left[ W_F + (b+q) V_U \right]$ . Setting  $V_F^N \ge V_F^S$  and solving for  $W_F$ we get

$$W_F \ge rV_U + \left(r + b + q\right) \frac{e}{q} = W^* \quad . \tag{4}$$

The critical wage, W<sup>\*</sup>, is higher

(a) the higher the required effort, e. Because employees dislike putting forth
 effort, a higher level of required effort reduces the value of the no shirking
 discounted utility stream. To again induce workers to choose not to shirk,
 the employer must increase the wage offer.

(b) the higher the expected utility associated with the alternate state,  $V_U$ . A higher value associated with the alternate state reduces the cost, and therefore the value of the punishment of being fired.

<sup>&</sup>lt;sup>70</sup>(...continued)

gets the benefit of a false negative if monitored. This treatment does not alter the model's conclusions in any material way.

- (c) the lower the probability of being monitored, q. Similar to (b), this case implies a reduction in the cost of shirking.
- (d) the higher the discount rate, r. This means that employees attach relatively more weight to the short-run gains from shirking compared to the losses incurred when one is eventually caught.
- (e) the higher the exogenous quit rate, b. If the employee is going to leave the firm, he does not worry about the future losses from being caught, but tries instead to maximize his short-run gains.

Intuitively, if firms wish to elicit effort from their employees, they must set their wage above the value of the alternative state, unemployment in this case. The increment will then serve as an economic penalty for those workers (considering) making the transition to the alternative state. Equation (4) indicates that the increment must be at least equal to the net expected utility from shirking, e/q, times the "hazard rate" associated with shirking (where b+q is the instantaneous probability of job separation for a shirker and r is the weight attached to the short-run gain from shirking).<sup>71</sup>

If the wage is instead set at a level that just leaves the workers indifferent between employment and unemployment, workers will not feel compelled to expend more than the lowest possible level of effort, e=0 in our model. In this

<sup>&</sup>lt;sup>71</sup> Since q is a Poisson process, the expected duration is in fact  $\frac{1}{q}$ . The waiting time for the first event, or between consecutive events, in a Poisson process,  $\lambda$ , follows an exponential distribution with mean  $\theta = 1/\lambda$ .

case, unemployment is preferable to putting forth the firm's desired level of effort. The increment necessary to ensure workers work as diligently as the firm desires represents the surplus extracted by the employees and ultimately leads to a less than socially efficient level of employment. If firms could capture this rent, employment levels would move to the competitive level.

Analogous to (2) and (3), the asset equation for  $V_U$ , the lifetime utility of unemployment, is given by:

$$rV_U = a(V_E - V_U) \quad , \tag{5}$$

where the parameter a represents the job acquisition rate and  $V_E$  is the expected utility of employment. Notice that the above formulation assumes a zero dividend associated with being out of the formal sector. The interpretation is either that the alternative sector is indeed unemployment and there are no unemployment insurance benefits, or else that the prevailing wage in this sector is so low that it just equals the effort required (costless monitoring).

In equilibrium,  $V_E = V_F^N$ , so we can solve simultaneously for  $V_E$  and  $V_U$ . Re-arranging (5), we get  $V_U = \frac{a}{r+a} V_E$ , which, using  $V_E = V_F^N$ , yields  $V_U = \frac{a(W_F - e)}{r(r+b+a)}$ . (6)

Substituting this in the NSC, equation (4), we get

$$W_F \ge e + \frac{e(a+b+r)}{q} = W^* \quad . \tag{7}$$

As before, the critical wage is greater:

(a) the smaller the detection probability, q;

(b) the larger the effort, e;

(c) the higher the exogenous separation rate, b; and

(d) the higher the discount rate, r.

Additionally, the critical wage is greater the higher the job acquisition rate, i.e. the higher the flow out of unemployment, a. If an individual could immediately obtain employment after being fired, then the costs of shirking would not serve to deter much shirking.<sup>72</sup>

We can now eliminate a from the solution by using the equal flow condition:

$$bL = a(N-L) \quad , \tag{8}$$

which asserts that, in steady state, the flow out of unemployment, a(N-L), is equal to the flow into unemployment, bL, where N is the total urban labour force and L is the number of jobs in the formal sector.

Re-arranging the equal flow condition, equation (8), as

$$a = b \frac{L}{N-L} \quad , \tag{9}$$

and substituting into (7), yields  $W_F \ge e + \frac{e}{q} \left( \frac{bN}{(N-L)} + r \right) = W^*$ , which, defining the rate of unemployment as  $U = \frac{N-L}{N}$  and substituting, becomes

$$W_F \ge e + \frac{e}{q} \left( \frac{b}{U} + r \right) = W^*$$
 (10)

<sup>&</sup>lt;sup>72</sup> This implies that a higher number of formal sector jobs would necessitate a higher formal sector wage, as claimed by Bulow and Summers (1986).

From this equation, it is immediately obvious that  $W_F$  is inversely related to the rate of unemployment. That is, a higher unemployment rate or a larger penalty associated with being unemployed, means that a lower wage is required to induce non-shirking. Notice that full employment, L=N, means that a=∞, so that a shirking worker could expect to be immediately re-hired. Absent the penalty associated with shirking then, utility maximizing employees will always choose to shirk, so that full-employment is inconsistent with no shirking. That is, a stable equilibrium configuration requires a queue, or 'a worker discipline device', of potential job seekers.

To this point, we have closely followed SS in modelling the urban labour market. As we successively adapt the model to conditions in developing countries we will add a second urban sector, and find that, with a heterogeneous labour force, it may actually be profitable for some workers to stay unemployed and improve their chances of attaining the rents associated with the high wage sector. Significantly, whether this queue takes the form of unemployment or temporary employment in a different sector depends either on the possibility of signalling or on the degree of intersectoral mobility. The first modification to the SS model, however, will be the addition of a rural sector, as a way of introducing the possibility of rural-urban migration.

# 3. Adding a Rural Sector

Although the rural labour market in Africa is often thought to be dominated by self-employed and family labour, recent articles<sup>73</sup> have found an active exchange of labour, both on large commercial farms and also on smaller "peasant" farms. Outside the large farms, most rural job openings appear to be seasonal and most hired labour is apparently of the casual type. The bulk of the evidence, for example Adams (1991), points to wages below government specified minimum levels, and to largely local markets.

Given the low wage outcomes and the high likelihood of abundant labour market information arising out of local hiring practices, it seems plausible to assume a competitively clearing rural labour market. Consider then, the following lifetime utility function for participation in the rural sector:

$$rV_A = W_A \quad , \tag{11}$$

where  $W_A$  is the competitively determined agricultural, or rural, wage. Assuming free flow between the agricultural and the informal sector, i.e. with the size of the urban labour force, N, endogenized, we reach migration equilibrium at  $V_U = V_A$ . Together with the four valuation functions, this condition gives us 5 equations in 5 unknowns,  $W_F$ , U, a, L and N. Given the formal sector's production function,

<sup>&</sup>lt;sup>73</sup> See Adams (1991) for a recent characterization of the rural labour market in the Masvingo province of Zimbabwe. For India, Rosenzweig (1980) finds that rural labour market behaviour is reasonably captured by the competitive model, little evidence of wage rigidity and some evidence that labour is, to some degree, immobile between districts.

we can solve explicitly for  $W_F$ , U and a, we can solve implicitly for L, and given L, we can solve explicitly for N.

Substituting the migration equilibrium condition,  $rV_U = rV_A = W_A$ , in the NSC, equation (4), we get the following relationship between the urban and the agricultural wages:

$$W_F \geq W_A + \left(r + b + q\right) \frac{e}{q} = W^*.$$
(12)

This equation gives a condition that  $W_F$  must obey if there is to be a positive level of unemployment when migration equilibrium is reached. The intuition is similar to equation (4) with the alternative state being agriculture instead of unemployment. In order to ensure effort is provided, urban wages must exceed agricultural wages by the net expected utility from shirking (e/q) weighted by the employee's probability of exiting urban employment (b+q) and the discount rate (r). Significantly, this No Migrating Condition (NMC) implies an equilibrium divergence between observed rural and urban wages, not due to lack of information about employment opportunities on the part of workers,<sup>74</sup> but due to a lack of information on the part of employers about the expected tenure of employees.

Substituting  $rV_U = W_A$  in (6), the migration equilibrium condition becomes

$$\frac{a(W_F - e)}{r(r + b + a)} = \frac{W_A}{r} \quad . \tag{13}$$
  
Substituting  $U = \frac{N - L}{N}$  in equation (9) and re-arranging, we get  $a = b \frac{1 - U}{U}$ .

<sup>&</sup>lt;sup>74</sup> Pessino (1991, p. 68), for example, argues for the role of imperfect information on the part of migrants.

We can now substitute this in (13) to get<sup>75</sup>  $U = \frac{b(W_F - e - W_A)}{rW_A + b(W_F - e)}$ , which, substituting for W<sub>F</sub> from (12) yields

 $U = \frac{be}{W_A q + be} \qquad (14)$ Finally, substituting the last equation in  $a = b \frac{1-U}{U}$ , we can solve for a:

Finally, substituting the last equation in  $a = b \frac{1-U}{U}$ , we can solve for a:  $a = \frac{qW_A}{e}$ (15)

The rate of job acquisition from unemployment is greater:

- (a) the greater is the probability of being monitored or the greater is the amount of resources devoted to monitoring.
- (b) the higher is the rural wage. An increase in the agricultural wage leads to reverse migration as some migrants now find it worthwhile to return home. At least some of these returnees are likely to be currently unemployed, thus enhancing the probability of finding employment for those continuing to search.
- (c) the lower is the required level of effort. A lower e lowers W\*, which lowers the penalty associated with shirking and raises the number of formal sector jobs.

<sup>&</sup>lt;sup>75</sup> Notice that in this model, U=0 implies that e=0 (and  $W_F = rV_U = W_A$ ). Full employment is again fundamentally inconsistent with not shirking as it implies the absence of an economic penalty associated with shirking. According to this model then, we would generally expect to see a sectoral wage gap, even if the migration equilibrium condition is met. To see this, re-arrange the migration equilibrium condition, equation (13), to yield  $W_F = \frac{r+b+a}{a}W_A + e$ .

For the intuition behind equation (15), re-arrange to  $a\frac{e}{q} = W_A$ . That is, the agricultural wage,  $W_A$ , is equal to the (formal) job acquisition rate out of unemployment, a, times the expected utility gain from shirking in formal sector employment, e/q.

Assuming that hiring in the formal sector follows the marginal product rule,  $F'(L)=W_F$ , we can implicitly define L in terms of only exogenous variables:

$$F'(L) = W_A + (r + b + q) \frac{e}{q}$$
 (16)

We can now define N in terms of L and exogenous variables by

substituting (14) in 
$$U = \frac{N-L}{N}$$
 to yield  
 $N = L \frac{qW_A + be}{qW_A}$ 

1

or, re-arranging,

$$N = L\left(1 + \frac{e/q}{W_A/b}\right) \quad . \tag{17}$$

That is, equilibrium is indeed characterized by involuntary unemployment, as the size of the endogenously determined labour force, N, is greater than the number of available jobs, L. The size of the queue, or N - L, is directly related to the ratio of the net expected utility from shirking to the expected utility from voluntarily moving to the rural sector.

The immediate policy implication is that, with a fixed number of formal sector jobs, increasing the rural wage reduces the size of the queue necessary to ensure utility maximizing employees choose to provide effort. It's also true that, ceteris paribus, a higher value of b, or a higher probability of voluntarily moving back to the rural area, increases the required size of the queue, which is contrary

to the effect of diverting additional resources to monitoring, or increasing the probability of being monitored, q. Predictably, a higher level of effort must be balanced with a larger queue, or a higher penalty to shirking.

#### 3.1 Comparing the Two Models

To compare the level of unemployment generated by our model to that of the HT model, we assume a common equilibrium pair<sup>76</sup> ( $W_F$ ,  $W_A$ ) and posit that the level of unemployment generated by our model is lower than the HT model:

$$\frac{be}{W_A q + be} < \frac{W_F - W_A}{W_F}$$

Algebraic manipulation yields  $W_F > W_A + \frac{be}{q}$  as the necessary condition for our model to yield a lower level of unemployment. For the intuition behind this result, re-arrange to  $\frac{W_F - W_A}{b} > \frac{e}{q}$ . The interpretation is that the net loss in expected utility from voluntarily moving back to the rural sector (the lost wage differential divided by the exogenous probability of return) must be greater than the net expected utility from shirking. In other words, as long as the incentive to provide effort arising out of the prevailing wage structure is high enough, our characterization of the urban wage setting mechanism yields a shorter queue for urban formal employment. Further algebraic manipulation reveals that this condition is necessarily implied by the No Migrating Condition, equation (12), proving that, at least with migration free flow and an equilibrium configuration

 $<sup>^{76}</sup>$  The specific question we examine is this: setting a common agricultural wage, if b, e, q and r are such that the urban wage generated by our model is equal to the fixed urban wage in the HT model, which formulation yields a lower unemployment rate?

involving common wage levels, our model does indeed yield a lower level of urban unemployment.

We can also solve for N and similarly compare the size of the urban labour force generated by our model to that generated by the HT model. Solving the HT model for N and L yields

$$N = L \frac{W_F}{W_A} \tag{18}$$

To compare the size of the total urban labour force generated by the two models, set:

$$L \frac{q W_A + b e}{q W_A} < L \frac{W_F}{W_A}$$

Algebraic manipulation again yields  $W_F > W_A + \frac{be}{q}$  as the condition under which our model yields a smaller total urban labour force. The implication is that, because both models hold the urban based population constant, our model predicts a lower level of migration, which is inherently consistent with one key structural difference between the two models: while our model endogenizes the wage setting process, the HT model holds the formal sector wage exogenous. Because the formal sector wage in our model responds to the size of the queue for urban formal employment, as more people migrate from the rural areas and unemployment rises, not only does their likelihood of acquiring a formal job fall, but  $W_F$  also has a tendency to decrease. Both effects tend to reduce the attractiveness of migrating by lowering the return to migration. By contrast, in the HT model, as more people migrate to the urban areas, the only effect which lowers the return to migration is the falling formal job acquisition rate.<sup>77</sup>

# 3.2 Comparative Statics

Proceeding with the examination of our model we now turn to an analysis of the comparative statics.<sup>78</sup> First, we analyse the hypotheses generated by our formulation for U. Differentiating (14) with respect to  $W_A$  yields

$$\frac{\partial U}{\partial W_A} = \frac{-qbe}{(qW_A + be)^2}$$

The negative sign means that an increase in the agricultural wage is predicted to lead, through a narrowing of the wage gap, to lower migration rates, a smaller urban labour force and, consequently, to a lower unemployment rate.

In addition to the usual substitution of rural for urban employment implied by an increase in  $W_A$ , our model contains an additional effect associated with the endogeneity of  $W_F$ . The migration equilibrium condition of  $V_U = V_A$  implies that as  $V_A$  rises, through an increase in  $W_A$ , so, in equilibrium, must  $V_U$ . From the employee point of view, this means that the penalty associated with shirking declines, which in turn reduces the incentive to supply effort. For the urban

<sup>&</sup>lt;sup>77</sup> In other words, by assuming every job turns over every period and ascribing to every worker an equal probability of finding employment, the HT model (unrealistically) increases the attractiveness of migration for rural dwellers. Therefore, at every level of  $W_F$ , more potential migrants take the decision to try their luck at the urban employment lottery, generating relatively higher rates of unemployment in the urban area.

<sup>&</sup>lt;sup>78</sup> While our model endogenizes  $W_F$ , the HT model endogenizes  $W_A$ . The main impact is that partials with respect to these two wages obviously have different interpretations for each model. In this section, we only analyse our model.
employer, the implication is that the minimum wage required to elicit effort increases, which leads to a decrease in the number of available jobs and, consequently, to a higher unemployment rate.

The NSC in equation (12) shows that an increase in  $W_A$  is matched by a proportional increase in  $W_F$ . Because a potential migrant attaches some uncertainty (the probability of finding a formal job, for example) to  $W_F$ , a proportional increase in the levels of both wages leads to an overall lower level of migration: the (discounted) increase in  $W_F$  results in an effective narrowing of the wage gap. That is, in the end, the induced increase in  $W_F$  does not completely counter the increase in  $W_A$ , which allows the substitution effect to predominate and leads to a net decrease in the level of U. The initial suggestion is fairly orthodox: governments concerned about high levels of rural-urban migration should consider augmenting rural incomes. Whether such a policy increases overall welfare remains to be investigated.

With respect to the level of unemployment, our model yields the following additional partials:

$$\frac{\partial U}{\partial b} = \frac{q e W_A}{(q W_A + b e)^2} ,$$

$$\frac{\partial U}{\partial e} = \frac{b q W_A}{(q e W_A + b e)^2} , \text{ and}$$

$$\frac{\partial U}{\partial q} = \frac{-b e W_A}{(q W_A + b e)^2} .$$

The interpretation of these results is that either an increase in the quit rate,

b, or an increase in the required level of effort, e, necessitates a compensatory

increase in  $W_{F}$ . The subsequent rationing of formal sector jobs leads to a higher unemployment rate. On the other hand, an increase in the probability of being monitored and fired, q, results in a decrease in the effort eliciting critical wage, a subsequent increase in the number of formal sector jobs and, consequently, in a decrease in U.

We can similarly consider the partial derivatives of N and L. Partially differentiating equations (16) and (17) with respect to  $W_A$ , yields

 $\frac{\partial N}{\partial W_A} \bigg|_{\frac{L}{L}} = \frac{-Lbe}{qW_A^2} \text{ and } \frac{\partial L}{\partial W_A} = \frac{1}{F''(L)}.$  These partials are both negatively signed and the interpretation is fairly orthodox in each case. As the agricultural wage rises and the wage gap narrows, there should be less incentive to migrate, and the size of the urban labour force, N, should diminish. Similarly, an increase in the agricultural wage leads, in turn, to an increase in the urban wage and a decrease in the number of formal sector jobs.

Partially differentiating L with respect to the other exogenous variables again yields orthodox results: the number of formal sector jobs decreases with an increase in r, b or e through the necessity to raise  $W_F$  to offset those changes, but it increases with an increase in q due to the extra incentive to work independent of the wage effect.

Similarly differentiating N, holding L constant, also yields orthodox results: the size of the urban labour force increases, or migration is induced, with an increase in the quit rate, b, (due to the higher probability of potential migrants landing a formal sector job), and in the effort required, e, (due to the associated increase in  $W_F$ ). However, N increases with a decrease in the detection probability, q, which also raises the valuation by a potential migrant of a formal sector job.

In summary, our original model differs from the Harris-Todaro formulation by explicitly modelling the urban wage setting mechanism. The resulting model of efficiency wages induces a rigid and relatively high urban wage, not by some institutional assumption as in HT, but in response to an agency problem.<sup>79</sup> Under the assumption of free flow, rural dwellers migrate to the urban centre until the expected utility of being unemployed is just equal to the expected utility of remaining in the rural area. The informational structure generates an unemployment queue and an equilibrium urban-rural wage gap with lower levels of unemployment and migration than predicted by HT-type models.

The model can be made more relevant to current conditions in the developing world by the addition of an urban informal sector. Because owners in this sector are assumed to be able to elicit effort costlessly,<sup>80</sup> informal sector wages are competitively set and likely to be lower than formal sector wages. Rural dwellers wishing to migrate now have the choice of entering a low-wage

<sup>&</sup>lt;sup>79</sup> In terms of policy implications, our model is not subject to the HT finding that the first best solution could be reached by the "simple" abrogation of minimum wage legislation.

<sup>&</sup>lt;sup>80</sup> To simplify the analytic treatment, we normalize the cost of monitoring in the informal sector to zero and treat the intersectoral difference as the cost of monitoring in the formal sector. This normalization does not materially affect any of the model's predictions.

sector or scratching out a living in the unemployment/self-employment sector (the so-called survival informal sector). The advantage of the unemployment state is a higher probability of landing a formal sector job, due either to a more intensive search, or else to employers' hiring preferences.

# 4. Adding An Informal Sector

Adding an informal sector to the model gives us the following specification.

Formal sector:

$$rV_{F}^{N} = W_{F} - e + b(V_{U} - V_{F}^{N})$$
 (2)

$$r V_F^S = W_F^+ (b+q) (V_U^- V_F^S)$$
(3)

Unemployment sector:

$$rV_U = a_U(V_F - V_U) \tag{5'}$$

Agricultural sector:

$$rV_A = W_A \tag{11}$$

Informal sector:

$$rV_{I} = W_{I} - e_{I} + a_{I}(V_{F} - V_{I})$$
(19)

Additional notation is as follows:  $a_i = formal job acquisition rate from$ state i, where i = U (unemployment) or I (informal sector).

Free flow between the informal and unemployment sectors, means that, in equilibrium, the expected utility of being in each sector must be equal. In addition, free flow of migrants from the rural to the urban area means that the expected utility of being in the rural area must also be equal, in equilibrium, to the expected utility of being in either the informal or the unemployment sector. That is, the migration equilibrium condition is now  $V_U = V_A = V_I$ . The equal flow condition is  $a_I L_I + a_U (N - L_F - L_I) = bL_F$  and the unemployment rate<sup>81</sup> is defined as  $U = \frac{(N - L_F)}{N}$ . Endogenous variables are  $W_F$ ,  $W_I$ , N,  $L_F$ ,  $L_I$ , U and  $a_U$ , while exogenous variables are  $W_A$ , r, b, q, e, e<sub>I</sub> and  $a_I$ .

#### 4.1 Solving the Model

Equations (2) and (3) yield the familiar NSC arising out of the condition  $rV^{N} \ge rV^{S}$ . Specifically, we have  $W_{F} \ge rV_{U} + (r + b + q)\frac{e}{q} = W^{*}$ . Using the migration equilibrium condition,  $V_{U} = V_{A}$ , we solve for  $W_{F}$  as before:

$$W_F \geq W_A + \left(r + b + q\right) \frac{e}{q} = W^*.$$
(12)

This is our familiar No Migrating Constraint (NMC).

To solve for  $a_U$ , re-write (5') as  $(r + a_U) V_U = a_U V_F$ , and re-write (19) as  $(r + a_I) V_I = W_I - e_I + a_I V_F$ . Setting  $V_U = V_A$  and substituting for  $V_F = V_F^N$  and  $W_F$ , we arrive at

$$a_U = \frac{q W_A}{e}$$

which is equal to equation (15) in the original model and has a similar interpretation.

<sup>&</sup>lt;sup>81</sup> The proposed definition of unemployment conforms to the customary practice in (some) developing countries of including informal sector employees among the unemployed. This is largely due to difficulties in enumerating informal firms, many of which de Soto (1989) claims to benefit from avoiding government regulations.

To solve for  $W_1$  we set  $V_0 = V_1$ , the free flow condition between unemployment and the informal sector, and get

$$W_I - e_I = \frac{r(a_U - a_I)}{r + a_U} V_F$$

That is, the immediate reward of choosing informal sector employment,  $W_i - e_i$ , must be just equal to the benefits accruing from a more intensive formal sector job search usually associated with the state of unemployment. If the possibility of landing a formal sector job is the same from either sector, i.e. if  $a_U =$   $a_i$ , then  $W_i = e_i$ . In other words, this equilibrium condition ensures that the informal sector wage is always bid down to the point of indifference: with  $a_U =$   $a_i$ , the wage is just high enough to compensate for the effort expended; whereas with  $a_U > a_i$ , the wage is just high enough to compensate for the effort expended as well as the reduced probability of finding a formal sector job.

Substituting  $a_U$  in the last expression and re-arranging, we get

$$\frac{W_A}{r} = \frac{1}{r+a_I} \left( W_I - e_I + a_I V_F \right)$$

That is, the benefits of informal sector employment (the excess of the wage over the effort expended to earn it) plus the chance at a formal sector job must equal the costs (agricultural wage foregone because of migration).

Substituting for  $V_F$ <sup>82</sup> we get

<sup>82</sup> To express  $V_F = V_F^N$  in terms of exogenous terms only, substitute (2), (5') and  $a_U$  in  $V_F^N$ , to yield  $V_F = \frac{re + qW_A}{rq}$ .

$$W_I = e_I + W_A - a_I \frac{e}{q} \quad . \tag{20}$$

That is, while the informal sector employee must be compensated for his effort and his agricultural wage foregone,  $e_1$  and  $W_A$ , the informal sector employer is compensated for the chance at "striking gold" conferred by informal employment, which is the probability of transferring from informal to formal employment multiplied by the expected gain in utility from shirking in formal employment,  $a_I \frac{e}{q}$ . In this way, the current model provides an intuitively appealing reason for the low wages typically associated with informal sector employment. Notice further that if  $e_I < a_I \frac{e}{q}$ , the cost of working in the informal sector is less than the expected gain from shirking in formal employment, and  $W_1$  $< W_A$ . In this case, rural dwellers migrate in spite of moving from a (relatively) high to a low, at least initially, wage area.

Working in the informal sector can then be viewed as a lottery. Some employees hold the winning ticket, and in addition to earning the surplus associated with working in the formal sector, e/q, earn rents, ex-post, from their informal sector employment. The remainder of the employees share in the cost associated with informal employment granting them a chance at formal sector employment (in the form of reduced wages), but never reach the formal sector. Ex-post, informal sector wages for these unsuccessful employees are too low; the difference is earned by the successful transferees in the form of (ex-post) rents.

Migration from the rural area occurs until the gap between informal sector and agricultural wages is such that the equality in equation (20) holds. At that point, the payoffs are such that migrants become indifferent between the choices of migrating to informal sector employment or staying at home. Note that if  $a_i=a_U$ , we can substitute for  $a_U$  and show that  $W_I=e_I$ , as mentioned above. Although the assumption of costless monitoring would normally make employee behaviour solely a function of the firm's wage, the free flow assumption means that conditions in other sectors become pertinent in the wage determination process of the informal sector.

Comparative statics reveals that the wage prevailing in the informal sector is higher

- (a) the higher the effort required in the informal sector  $(e_1)$ ,
- (b) the higher the agricultural wage (W<sub>A</sub>), so that rural wage subsidies would likely also serve to increase formal and informal sector wages,
- (c) the lower the formal job acquisition rate from informal employment  $(a_i)$ ,
- (d) the lower the required effort in the formal sector (e), and
- (e) the higher the probability of being fired from a formal sector job (q), so that informal sector employers prefer not to see improvements in formal sector monitoring technology (or indeed in the quality of information available to formal sector employers).

Thus far, we have followed standard practice in assuming free flow from the rural area into both unemployment and the informal sector. In equilibrium then, the utility of participating in either of these sectors must be equal, in which case some, as yet unspecified, workers will choose unemployment over informal sector employment. That is, our equilibrium entails unemployment which is in some sense voluntary (informal sector jobs are always available at the going wage), although it remains an open question whether employment at relatively low wages serves to alleviate any of the social problems usually associated with unemployment. Notice in particular that, with a high enough degree of mobility between the informal and the formal sectors, the model admits the possibility of a negative informal sector wage.<sup>83</sup>

To find the condition regulating movement between the informal and the formal sectors substitute (20) in (12) to yield

$$W_F = (W_I - e_I) + (r + b + q) \frac{e}{q} + a_I \frac{e}{q}$$

Thus,  $W_F$  is higher with a larger payoff from informal sector employment, which can happen either in the form of higher net wages ( $W_I - e_I$ ) or, with a given informal wage, in the form of an increased probability of winning the lottery and moving to formal sector employment,  $a_I$ .<sup>84</sup> Notice that formal sector wages must be high enough to compensate not just for the foregone net informal sector wages, but also for the (short-run) cost of providing effort, the second term, and the

<sup>84</sup> By increasing the cost of access to the formal sector, negative informal sector wages should then serve to motivate the worker to reduce the probability of losing the benefits of formal sector employment. For the formal sector employer, this should translate into an attenuation of the monitoring problem.

<sup>&</sup>lt;sup>83</sup> To the extent that skills acquired in the informal sector are valuable assets in the formal sector, the informal sector may serve as a training ground with apprentices paying their masters for the (greatly improved) chance at employment in the high wage sector. This would then be similar to employees paying a bond upon joining the labour force. On the existence of negative informal sector wages in Africa, see ILO (1991).

expected utility gain from moving to the formal sector and shirking, the third term. The discount factor, r, is the weight attached to short-run gains or losses; b + q is the instantaneous probability of job separation<sup>85</sup> for an employee who chooses to shirk; and e/q is the expected utility gain from shirking, so that the second term on the right side of the equality can be interpreted as the foregone utility gain if the formal sector employee chooses not to provide effort, or the cost of shirking and exiting the formal sector. The third term is the expected payoff for the lottery ticket included in the compensation package of informal sector employees.<sup>86</sup>

The current model also admits the possibility that informal sector wages may be higher than formal sector wages. As mentioned in the literature review chapter, this is consistent with recent findings that (some) informal sector participants, typically managers, earn more than (some) formal sector employees. From the last equation,  $W_I > W_F$  requires  $|e_I| > |(r+b+q+a_I)e/q|$ . That is, a reversal in the usual sectoral ranking of wages requires the cost of working (and providing effort) in the informal sector to be greater than the sum of the expected

<sup>&</sup>lt;sup>85</sup> Two independent Poisson events, b and q, together occur at the rate b + q. For a reference, see Ross (1989, p. 353).

<sup>&</sup>lt;sup>86</sup> If we interpret our model as a continuous time Markov chain, a stochastic process with the conditional probabilities of being in a future state depending only on the present state, we can also interpret our transition probabilities, (b + q) and  $a_1$ , as, respectively, the long-run proportion of time spent out of the formal sector and the long-run proportion of time spent in the formal sector. In this case, b + q $+ a_1 = 1$ . One example of a continuous time Markov chain is in fact a Poisson process. For reference, see chapter 6 of Ross (1989), especially pp. 268-269.

utility gain of shirking in formal employment and the expected utility gain of transferring to the formal sector.

Assuming that formal sector hiring follows the marginal product rule,  $F'(L_F) = W_F$ , we can implicitly define the level of formal sector employment,  $L_F$ , as follows:

$$F'(L_F) = W_A + (r+b+q)\frac{e}{q}$$

Similarly, we can implicitly define the level of employment in the informal sector:

$$G'(L_I) = W_I = e_I + W_A - a_I \frac{e}{q}$$

To solve for N we substitute  $a_U$ ,  $L_F$  and  $L_I$  in the equal flow condition. Subsequently substituting N and  $L_F$  in the unemployment definition yields the equation for U.

In summary, the current model improves on our original specification by the addition of the urban informal sector. The importance of this sector in the context of developing countries is derived not only from its relatively large share of economic production, but also from the sector's central role in expanding the opportunity set available to migrants, and possibly urban residents. This model structure allows us to model not only rural-urban movement but also, within the urban area, informal-formal movement, and to explain formal-informal wage patterns documented in the development literature. In particular, the model allows, under certain conditions, negative wages in the informal sector, wages to be higher in the informal than in the formal sector, wages to be higher in agriculture than in the informal sector, but restricts urban formal sector wages to be higher than agricultural wages.

Relative to the first model presented, behaviour in the formal sector is unchanged (other than the possibility of a larger unemployment pool leading to a lower level of efficiency wages), although formal sector wages are now responsive to conditions in the informal sector. It is then economic conditions across the different sectors that determine how, or if, labour movements take place, and given that informal sector jobs are readily available at the going wage, the current model is implicitly consistent with any level of rural-urban migration. Given that a temporary absence from the labour force does not penalize a worker willing to accept a job in the informal sector, the current model is also implicitly consistent with return migration. In the next section, we attempt to explicitly account for the impacts on the urban labour market from the possibility that some rural-urban migrants may actually engage in return, or temporary, migration.

# 5. Adding Return Migration

In the models presented so far all workers are assumed to have similar characteristics and to be equally productive, the usual homogeneity assumption. In this section, we introduce heterogeneity into the urban labour force by adding the possibility of permanent as well as temporary migration. As mentioned in both the literature review and the chapter on migration patterns in Malaysia, this type of migration behaviour conforms with extensive documentation on circular

107

migration patterns in developing countries, particularly in Africa and Southeast Asia.

Our earlier assumptions that informal sector jobs are readily available, and that informal sector employers are able to elicit effort (relatively) costlessly, imply that, for an otherwise equally productive labour force, the possibility of return migration on the part of some workers does not affect behaviour in the informal sector. Consequently, the informal sector continues to be characterized as a perfectly competitive labour market where jobs at the going wage are always available and can be reclaimed by return migrants. For the formal sector, however, the differential impact of low and high turnover employees may be substantial and we investigate the implications by adding the possibility of return migration from the formal sector back to the rural areas. Our model specification is now as follows.

Formal sector:

$$r V_F^{Ni} = W_F - e + b (V_{NF}^i - V_F^{Ni}) + \pi^i (V_A - V_F^{Ni})$$
(2')

$$r V_F^{Si} = W_F + (b+q) (V_{NF}^i - V_F^{Si}) + \pi^i (V_A - V_F^{Si})$$
(3')

Unemployment sector:

$$rV_{U}^{i} = a_{U}(V_{F}^{i} - V_{U}^{i})$$
(5")

Agricultural sector:

$$rV_A = W_A \tag{11}$$

Informal sector:

$$r V_I^i = W_I - e_I + a_I (V_F^i - V_I^i)$$
(19)

Additional notation is as follows:  $\pi$  = probability of return which equals zero for 'urban residents and permanent migrants' but is strictly positive for 'temporary migrants';<sup>87</sup> superscript 1 denotes the type of worker, where I = 1 for 'resident' or I = 2 for 'migrant'; and subscript NF denotes a state other than the formal sector. The model structure is outlined in Appendix B.

### 5.1 Solving the Model

While waiting for a formal job, migrants should have a higher propensity to participate in the informal sector (and accept a lower current wage) instead of remaining unemployed. Theoretical support for such a hypothesis derives from the predictions of the theory of intertemporal substitution.<sup>88</sup> Empirical support derives from evidence presented by, for example, Yap (1977) indicating that migrants, upon arrival in urban areas, tend to have shorter unemployment spells and a higher participation rate in the informal sector. In terms of our model, such rationale leads to migrants placing a higher value on informal sector work than on idle unemployment,  $V_1^2 > V_0^2$ , while the opposite should hold true for residents,

<sup>&</sup>lt;sup>87</sup> From here on, 'migrant workers' refers only to temporary migrants ( $\pi > 0$ ), while 'resident workers' refers to the group of urban residents and permanent migrants.

<sup>&</sup>lt;sup>88</sup> As workers move from a low to a high wage region, their rate of substitution of income for leisure increases along with the wage rate. If these workers also retain a positive probability to return, intertemporal substitution of labour implies that these workers should (1) strive to increase their supply of labour while in the high wage region, and (2) enjoy more leisure upon their return to the low-wage area. In other words, during their stay in the urban area and given a positive likelihood of returning to the low-wage area, migrant workers value leisure less than resident workers.

 $V_U^{1} > V_1^{1}$ . Whether this means that only migrants participate in the informal sector has a critical impact on the equilibrium configuration.<sup>89</sup>

In solving this model, we examine two types of equilibrium:<sup>90</sup> (1) complete screening, under which firms are able to effectively sort workers, and (2) incomplete screening, under which workers from either group have the same probability of acquiring a formal sector job. Without complete information on  $\pi$  (observationally indistinguishable), firms can achieve complete screening only with status dependent hiring (allowing firms to carry out a discriminatory hiring policy favouring low risk or resident, i.e. cheaper, workers).<sup>91</sup> The key requirement turns out to be that workers sort themselves out in terms of choosing either unemployment or the informal sector while searching for a formal sector job. In turn, employers have an incentive to discriminate against migrants and, given the opportunity, must decide between complete or partial exclusion.

<sup>&</sup>lt;sup>89</sup> Notice that the usual justification for labelling informal employees as underemployed does not apply in this model because residents and migrants are not identical economic agents. That is, migrants employed in the informal sector would, under perfect information, not be able to earn a higher wage in the formal sector.

<sup>&</sup>lt;sup>90</sup> This model is somewhat similar to Jones (1987). To the Jones framework, our model adds an agricultural sector, allows mobility from the informal sector and distinguishes between effort in the formal (or 'primary') and the informal (or 'secondary') sectors.

<sup>&</sup>lt;sup>91</sup> Residents, or type 1 workers, are cheaper because their lower exogenous turnover probability is translated into a lower effort eliciting minimum wage. The incomplete information (on migrants) is a result of the inability to distinguish between those who intend to stay in the urban area permanently ( $\pi$ =0) and those who are on a temporary stay ( $\pi > 0$ ).

Given any market transaction with an outcome dependent on participant characteristics, the presence of imperfect information on those characteristics will result in some effort being diverted to screening, or the gathering of information on individuals.<sup>92</sup> Because firms operating in a competitive market are generally unable to capture the returns to screening, such costs must then be borne by the participants being screened: job applicants in the labour market.

In our model, there is imperfect information on only one side of the market: employers are uncertain about employees' effort level. Ex-ante, employers do not know whether the job applicant is a low or a high risk employee, while ex-post, monitoring is not costless. Although this asymmetry of information does raise the possibility of shirking and, in turn, does eliminate any prospect of an equilibrium with no screening, the competitive environment necessitates the passing of screening costs on to the job applicants. In turn, the willingness of these individuals to bear these costs depends both on their level of risk aversion, and on the state of their information about their own characteristics. In our model, because workers are risk neutral and know their own type (either high risk,  $\pi > 0$ , or low risk,  $\pi = 0$ ), they are willing to assume those costs. In particular, the low risk individuals have an incentive to reveal themselves in order to capitalize on the higher value of a formal job, a position for which they happen to be preferred.

111

<sup>&</sup>lt;sup>92</sup> For a concise discussion see Stiglitz (1974).

In the first case examined, the low risk workers shoulder the screening costs by choosing to remain unemployed while searching for a formal job. If the prevailing conditions result in the high risk workers simultaneously choosing the informal sector over unemployment, screening is complete, and only low risk workers are hired by formal sector firms. Compared to the usual form of the luxury unemployment hypothesis<sup>93</sup> proposed in the migration literature, this signalling explanation for involuntary unemployment retains the characteristic that the unemployment and informal sectors are each composed of distinct types of workers, but differs in one important aspect. In our model workers are equally productive (once they decide to supply effort).

In the second case examined, conditions are such that some high risk workers also choose unemployment. Screening is consequently incomplete, and a proportion of formal sector jobs is filled by high risk workers (at least temporarily). In equilibrium, the probabilities of job acquisition remain constant and stationarity involves a fixed proportion of jobs filled by each of two worker types.

# 5.2 Complete Screening Equilibrium

The lower probability of separation ( $\pi^1 = 0$ ) associated with urban residents and permanent migrants leads to a lower effort-eliciting critical wage. These workers, in effect allowing themselves fewer external alternatives, have

<sup>&</sup>lt;sup>93</sup> See, for example, Sahn and Alderman (1990) who in a study of Pakistan find two distinct groups of unemployed workers – young, more educated men waiting for a formal sector-type job and passing over informal sector typeopportunities and older, less educated men with relatively substantial family support.

more to lose by shirking, and so choose to supply the required level of effort at a lower wage. Equations (2') and (3') yield the familiar NSC arising out of the condition  $rV^{N_i} \ge rV^{S_i}$ . In this case, the NSC takes the form:

$$W_F^i \ge V_{NF}^i(r+\pi^i) - \pi^i V_A + (r+b+q+\pi^i) \frac{e}{q} = W^*.$$
 (21)

As before, a higher exogenous quit rate raises the critical wage as it lowers the value of maintaining the job. That is, workers with  $\pi = 0$  (permanent migrants and residents) tend to have a lower effort-eliciting critical wage because staying on the job is worth more to them than to workers with  $\pi > 0$  (temporary migrants). We detail first the behaviour of residents, for whom the critical wage is equal to:

$$W_{F}^{1} = r V_{U}^{1} + (r + b + q) \frac{e}{q} \quad .$$
 (22)

On the other side of the labour exchange, firms would increase profits by offering  $W_F^1$ , which is lower than the alternative  $W_F^2$ , provided effort is elicited from their labour force. With equally productive workers (at least once they choose to supply effort), the monitoring model suggests that formal firms might then engage in discriminatory hiring. However, the employers' inability to identify workers by type implies that the rate of formal job acquisition from any one of the alternative sectors must be equal for both groups of workers, and a complete screening solution then requires that employers use a (perfectly discriminating) signal to screen applicants. In this model, the signal turns out to be the selection of sectors (unemployment or informal) made by workers.

Specifically, the requirement<sup>94</sup> is that unemployment be the choice of only the low risk workers, while high risk workers choose to work in the informal sector, in which case,  $a_U^1 = a_U^2 = a_U$  and  $a_I^1 = a_I^2 = 0$ , and the formal sector wage is solely determined by the behaviour of (low risk) resident workers.<sup>95</sup> Under complete screening then, formal sector employees choose not to shirk and their value of formal employment,  $V_F^1$ , is equal to the no-shirking valuation,  $V_F^{N1}$ , so that  $rV_U^1 = \frac{a_U}{r+b+a_U} (W_F - e)$ . Substituting  $rV_U^1$  in equation (22), the NSC becomes

$$W_F^1 = \left(r+b+q+a_U\right)\frac{e}{q} \quad , \tag{23}$$

or equivalently, using  $a_U(N^1-L_F) = bL_F$ ,  $W_F^1 = \left(r+q+\frac{bN^1}{N^1-L_F}\right)\frac{e}{q}$ .

The condition for residents to choose unemployment is that  $V_U^1 > V_I^1$ , which, along with the NSC, yields the following sectoral allocation condition for residents

<sup>&</sup>lt;sup>94</sup> An additional proviso is that the supply of resident labour not be a binding constraint. That is, there should be enough resident workers to fill all the available positions and the pool of unemployed workers should be large enough that formal sector wages are not driven up.

<sup>&</sup>lt;sup>95</sup> Notice that employers must choose between either a low wage (enough to induce effort from the low risk workers) and a high wage (enough to induce effort from both high and low risk workers). Any wage level between these extremes serves only as an extra transfer of wealth from the employer to the employees. For employers, the question then centres on whether the rents earned by low risk employees under a high wage regime are greater or lesser than the costs of offering a low wage and increasing the level of employee turnover. Given some level of screening, the low wage policy must be preferred as in Jones (1987, p. 195).

$$W_I - e_I < a_U \frac{e}{q} \qquad (24)$$

In words, resident workers prefer unemployment over informal sector employment if compensation in the informal sector is less than the expected utility gain of acquiring a formal sector job from unemployment and shirking. Formal sector employers will now minimize formal sector compensation subject to meeting this minimum threshold, the implications of which will become evident as we discuss migrant behaviour and profit maximization below.

Simultaneously, the condition for migrants to choose informal employment over unemployment is that  $V_1^2 > V_0^2$ , which, with  $a_i = 0$  and  $rV_1^2 = W_1 - e_1$ , yields<sup>96</sup>

$$W_I - e_I > \frac{a_U W_F^1}{(r+b+q+\pi+a_U)}$$

In words, as long as compensation in the informal sector obeys the above condition, migrants prefer the immediate rewards of informal employment (and are thus willing to sacrifice their chance at a formal job) to the prospects provided by choosing unemployment. For an intuitive interpretation, re-arrange to

$$(r+b+q+\pi)(W_{I}-e_{j}) > a_{U}[W_{F}^{1}-(W_{I}-e_{j})].$$

That is, informal employment is chosen if the probability of shirking and exiting the formal sector (and entering the informal sector) times the net informal wage is

<sup>&</sup>lt;sup>96</sup> To solve for  $V_U^2$ , set the formal sector wage at  $W_F^1$ , migrants' value of formal employment at  $V_F^{S2}$  and simultaneously solve for  $V_U^2$  and  $V_F^2$ .

greater than the probability of getting a formal job from unemployment times the increase in compensation. Re-arranging once more and using (24), we get

$$(W_I - e_I) \left[ 1 + \frac{\pi}{r + b + q + a_U} \right] > a_U \frac{e}{q} > (W_I - e_I) , \qquad (25)$$

which imposes bounds on a<sub>U</sub>.

Given unrestricted migration and informal sector compensation consistent with (25), the rural-urban flow continues until  $V_I = V_A$ . Using this migration equilibrium condition, and with  $a_I = 0$ , we arrive at

$$W_I = W_A + e_I \quad , \tag{26}$$

which, together with the NSC shown previously, pins down the wages in both the formal sector, for a given level of  $a_U$ , and the informal sector. The relationship postulated implies that even temporary migrants do not experience a decrease in wages upon taking an informal sector job. This conforms with Williamson (1988) who notes that, unlike HT-type assumptions, migration is not just a lottery system for the best-paid jobs, and migrants do not (necessarily) earn less in cities than where they originate from.

Combining equations (26) and (24) shows that the probability of formal job acquisition is higher (for those still eligible) with heterogeneity,  $a_U > W_A \frac{q}{e}$ , than without,  $a_U = W_A \frac{q}{e}$ , which is consistent with restricting hiring to a segment of the population. In other words, the existence of the informal sector, by raising the equilibrium value of  $a_U$  and allowing employers an opportunity to restrict hiring, results in an increase in  $W_F^*$ .

Combining the migration equilibrium condition,  $V_1 = V_A$ , and the sectoral allocation condition for migrants,  $V_1^2 > V_U^2$ , we obtain  $V_A > V_U^2$ . In other words, (temporary) migrants are attracted to the urban sector, not by the income opportunities in the formal sector, which are only accessible through a spell of unemployment, but by informal sector earnings which, as shown by (26), are strictly greater than agricultural wages. Using the NSC, equation (23), informal wages, equation (26), and the sector allocation condition for residents, equation (24), we can also show unequivocally that  $W_F > W_A$  as  $W_F > W_A + (r+b+q)\frac{e}{q}$ . This condition serves to provide permanent migrants with enough incentive to endure the necessary spell of unemployment.

In turn, the relationship between formal and informal sector wages depends on relative working conditions, or on the relative levels of effort required. Although under the present solution we cannot rule out  $W_F < W_I$ , nonshirkers do value the short-term compensation in the formal sector,  $W_F - e$ , more than the short-term compensation in the informal sector,  $W_I - e_I$ , as the sectoral allocation condition for residents reduces to  $(W_F - e) > (W_I - e_I) \frac{r + b + a_U}{a_U}$ . We can also show for  $W_F > W_I$  it is sufficient, although not necessary,

that  $e + (r+b)\frac{e}{q} > e_1$ , a condition consistent with the usual assumption about poorer working conditions in the informal sector, whether due to more labour intensive tasks or closer monitoring.<sup>97</sup> Under the complete screening equilibrium

<sup>&</sup>lt;sup>97</sup> This condition raises the possibility of expanding the model by having output in the informal sector dependent on the labour of both managers and (continued...)

then, a perfectly discriminatory hiring policy  $(a_1 = 0)$  ensures that wages in each of the urban sectors are strictly greater than agricultural wages, but does not rule out informal sector wages being higher than formal sector wages.

On the other side of the labour market, we can show that formal sector employers are behaving as profit maximizers in instituting discriminatory hiring practices. To elicit effort from their workers, these employers have a choice of offering either  $W_F^1$  or  $W_F^2$ . If  $W_F^2 > W_F^1$ , migrants would shirk at  $W_F^1$ , and the latter would certainly be the profit maximizing wage provided the work force was composed entirely of residents. To solve for the formal sector wage necessary to elicit effort from migrants, use equation (21) and  $V_I^2 = V_A$ , to arrive at

 $W_F^2 = W_A + (r+b+q+\pi)\frac{e}{q}$ . To show that indeed  $W_F^{2>}W_F^{-1}$ , use (26) to arrive at  $W_I - e_I < \frac{e}{q}(r+b+q) + a_U \frac{e}{q}$ , which clearly holds given (24). We can then conclude that, with equally productive workers, employers do indeed maximize profits by restricting hiring to resident workers.

Assuming a marginal productivity hiring rule, we can pin down the levels of employment in both sectors, which arise from the following conditions:  $F'(L_E) = W_E$ , and

$$G'(L_I) = W_I$$
.

Given the competitive nature of the informal sector, it must be true that  $N^2 = L_1$ , in

<sup>&</sup>lt;sup>97</sup>(...continued)

employees and allowing for different effort levels. Such a framework would then allow managers' wages to be higher than  $W_F$  (due to higher effort levels) and employees' wages to be lower than  $W_F$ . This would be consistent with some informal sector participants earning more than formal sector employees as alluded to by Kahnert (1987) and de Soto (1989).

which case it must also be true that, combining (23), (24), (25) and the marginal productivity rule,

$$a_U \frac{e}{q} + e_I > G'(N^2) > e_I + a_U \frac{e}{q} \left[ \frac{r + b + q + a_U}{r + b + q + \pi + a_U} \right]$$
 (27)

We can also solve for the remaining endogenous variables  $a_U$ ,  $N^2$  and U. At the complete screening equilibrium, the equal flow condition for resident workers reduces to  $a_U(N^1 - L_F) = bL_F$ . We can then solve for  $a_U$  by using the above implicit equation for  $L_F$ . Finally, substituting both  $L_F$  and  $N^2$  into  $U = (N^1 - L_F) / N$ , yields an implicit function for U.<sup>98</sup>

We have now characterized equilibrium behaviour on the part of both employees and employers in both the formal and the informal sectors. Workers voluntarily sort themselves with resident workers choosing unemployment as a signal of their low risk characteristic, and migrants choosing the short-term compensation of informal sector employment over the possibility of earning the rents associated with a formal job. On the other side of the labour market, formal firms use status dependent hiring (inferring worker type from the observed worker choice) to achieve perfect discrimination.

These discriminatory hiring patterns are not eroded by market forces because the difference in the efficiency wages for different groups of workers

<sup>&</sup>lt;sup>98</sup> Note that the standard definition of the unemployment rate leads, in this solution, to the total exclusion of migrant labourers from the ranks of the employed, so that migration comes to be seen as just adding to the unemployment rolls. Alternatively, we could define  $U=(N-L_F-L_I) / N$ , in which case the only workers considered as unemployed would be those resident workers currently without a formal job. In this case, migration would be seen as the necessary pre-requisite to the survival of the informal sector.

makes such non human-capital based discrimination optimal for employers. In formal sector employment and with equal wages, workers with a shorter expected job duration incur a smaller penalty for (shirking and) being dismissed. In our model, migrants, who have a relatively high risk of separation, supply effort in formal sector employment only at a wage higher than residents. On the other side, formal sector employers, given their monitoring technology, maximize profits by restricting hiring to the equally productive, but less likely to separate, group of resident workers.

The proposed sectoral segregation of workers is efficient in the sense that the firms generally believed to have the least cost effective monitoring technology, formal sector enterprises, end up with the workers less likely to shirk, residents, while the firms with (relatively) costless monitoring technology, informal sector enterprises, end up with the workers more likely to shirk, migrants. Formal sector employers discriminate against migrants solely on the basis of profit maximizing considerations as migrants (are perceived to) have lower expected employment durations than residents.<sup>99</sup> Moreover, the model

<sup>&</sup>lt;sup>99</sup> Clearly, testing the validity of the assumption of differential quit rates would be an important, albeit difficult, empirical exercise. Separating the effects of employment history and migration status on current behaviour would necessitate a complete history of both work and migration. Although the key difference is in quit rates, differences in the utility of shirking between migrants and residents would suggest that all layoffs should be treated as voluntary departures. Additionally, notice that the difference in penalties between quitting a formal versus an informal job combined with the self sorting by workers implies that the appropriate question centres on differences in expected job tenure in formal employment only.

provides an explanation for the concept of luxury unemployment which is not based on differences in human capital, so that even after controlling for productivity differences, workers in unemployment are systematically different from workers who optimally choose to work in the informal sector instead of remaining unemployed.

The presence of heterogeneity, unlike the simple addition of the informal sector, has actually changed the wage setting behaviour in the formal sector by raising the minimum effort eliciting wage. This occurs in spite of the fact that the firm is able to hire only low risk workers. Restricting the hiring pool means that the probability of acquiring a formal job has increased for resident workers ( $a_U > W_A \frac{q}{e}$  versus  $a_U = W_A \frac{q}{e}$ ) but disappeared for migrants. In return, formal firms must raise their wage offer to ensure that their employees do not shirk. Similarly, informal wages have also increased as the compensation package,  $W_I - e_I$ , is now greater than, as opposed to equal to,  $a_U \frac{e}{q}$ . Equilibrium is characterized by perfect sectoral sorting of workers by type and the realization of a zero level of shirking at the relatively lower  $W_F^{-1}$ . The resulting pattern of occupational segregation will not be eroded by market forces.

A cursory examination of (27) reveals that wages in the informal sector may be either so low that some migrants actually choose unemployment, or so high that some resident workers choose to dabble in the informal sector. Assuming that formal sector firms hire strictly from the rolls of the unemployed, the complete screening solution survives the case of high informal sector wages because it is still true that only residents choose unemployment. However, the case of low informal sector wages results in some migrants also choosing unemployment, and this does invalidate complete screening because some migrants will be hired into the formal sector. We analyse this situation in the next section.

### 5.3 Incomplete Screening Equilibrium

If compensation patterns do not obey the complete screening sectoral allocation conditions, equation (27), the labour market becomes characterized by less than complete screening, and the undisclosed nature of worker heterogeneity leads to all unemployed workers, migrant and resident alike, having an equal probability of acquiring a formal sector job.<sup>100</sup> Under such conditions, employers still find it beneficial to hire strictly from unemployment because this remains the only sector with low turnover workers. In this instance, of course, screening is not complete, and given that probabilities of job acquisition remain constant, stationary equilibrium now involves a constant proportion of jobs filled by each of the two worker types. However, given that employers continue to offer a wage just high enough to induce effort from the low turnover employees, equilibrium

<sup>&</sup>lt;sup>100</sup> Compounding this effect, informal sector workers may be able to pass themselves off as unemployed: although informal sector-based search is likely to be less intensive than unemployment-based search ( $a_1 < a_U$ ), it may well happen that  $a_1$  is greater than zero through worker initiative rather than employer design. Setting  $a_1 = 0$  assumes that employers have the ability to distinguish between unemployed workers and informal sector workers, so that the source of the incomplete screening arises strictly because it is now optimal for (some) migrant workers to choose unemployment over informal sector employment and employers are unable to distinguish between residents and migrants.

must now be characterized by a positive number of firings as well as the voluntary<sup>101</sup> unemployment of migrants and the involuntary<sup>102</sup> unemployment of residents.

From the right side of inequality (27) on page 119, we know that  $(N^2-L_1)$  migrants will choose unemployment if

$$G'(L_I) = W_I = e_I + a_U \frac{e}{q} \left[ \frac{r+b+q+a_U}{r+b+q+\pi+a_U} \right],$$
 (28)

where  $L_1 < N^2$  and  $a_U$  is the equilibrium formal job acquisition rate common to both migrants and residents. The left side of the inequality still holds as residents still prefer to remain unemployed over working in the informal sector.

For resident workers, equation (23) is still the applicable efficiency wage: the behaviour of these workers is unchanged from the complete screening solution. However, relative to the perfect discrimination case under complete screening, the larger pool of unemployed workers leads to a lower acquisition rate of formal sector jobs, such that now  $a_U^1 = \frac{bL_F^1}{N^1 - L_F^1}$ . Substituting this expression in the NSC clearly leads to a lower efficiency wage for these low turnover

<sup>&</sup>lt;sup>101</sup> Voluntarily unemployed workers are individuals who turn down an employment opportunity otherwise filled by observationally identical workers. In our model, migrants who choose unemployment over informal sector employment are considered to be voluntarily unemployed because such jobs are assumed to be always available and such positions are filled by workers with similar characteristics.

<sup>&</sup>lt;sup>102</sup> Involuntarily unemployed workers are those individuals who cannot find employment under the same terms as apparently identical employees. In our model, unemployed residents are considered to be involuntarily unemployed because they are observationally identical to employed residents and are willing to supply effort at the same wage, but are unable to find such a job.

workers: the lower  $a_U$  increases the cost of transition into unemployment which, in turn, induces workers to supply effort at a lower wage.

With the prevailing wage in the formal sector determined by the behaviour of resident workers, migrants hired into formal employment from unemployment will shirk,<sup>103</sup> so that  $V_F^2 = V_F^{S2}$ . Substituting  $W_F^*$  in  $V_F^{S2}$ , we get

$$V_F^2 = \frac{W_F^* + (b+q+\pi)V_A}{r+b+q+\pi} , \qquad (29)$$

which, substituted in  $V_U^2$  and using the migration equilibrium  $V_U^2 = V_A$ , yields

$$W_{A} = \frac{a_{U}}{r+b+q+\pi+a_{U}}W_{F}^{*} \quad . \tag{30}$$

This relationship is exactly as in the complete screening perfect discrimination case and again establishes that  $W_F^* > W_A$ , a necessary condition for a positive level of rural-urban migration in a model which ascribes the motivation for migration to a pursuit of utility or income.

With respect to other sectoral wages, the results and the rankings also follow the complete screening perfect discrimination case. To verify that informal sector wages are strictly higher than agricultural wages, set  $V_1^2=V_A$  to arrive at  $W_1 = e_1 + W_A$ . Owing to the fact that migrants are just as likely to choose unemployment as informal employment, informal sector entrepreneurs have now lost the ability to extract compensation for the lottery ticket of access to formal employment. Formal sector employers continue to discriminate against migrants.

<sup>&</sup>lt;sup>103</sup> In this way, the proposed model predicts that disguised unemployment (if shirkers were replaced by non-shirkers the total output would increase) exists not in the informal sector as usually assumed, but in the formal sector.

Whether equilibrium formal sector wages are in turn higher than informal sector wages depends on the level of effort required in the informal sector. It is certainly straightforward to show that formal sector wages are higher than the immediate benefits in the informal sector. Substituting  $W_I - e_I = W_A$  in (30), yields

$$W_I - e_I = \frac{a_U}{r + b + q + \pi + a_U} W_F^*$$

This is the appropriate utility comparison for shirkers, and clearly, net of effort, informal sector wages are lower than formal sector wages. For formal sector workers who do supply effort, we can also show that, even net of effort, compensation in the formal sector is still higher than the immediate compensation in the informal sector, as the left inequality in the sectoral allocation condition, equation (27) on page 119 can be re-written as

$$(W_F - e) \left[ \frac{a_U}{r + b + a_U} \right] > W_I - e_I$$

It also clear that current labour market conditions, as exemplified by the previous inequality, admit the possibility that informal sector wages may be higher than formal sector wages, with the key determining factor being the relationship between the required level of effort in the formal and in the informal sectors.

Re-arranging (30) gives a quadratic expression which can be solved to yield  $a_{ty}$ . The solution

$$a_{U} = \frac{\left[(r+b+q)\frac{e}{q} - W_{A}\right] \pm \sqrt{\left[(r+b+q)\frac{e}{q} - W_{A}\right]^{2} + 4\frac{e}{q}(r+b+q+\pi)W_{A}}}{2\frac{e}{q}}$$

gives the equilibrium access rate to formal employment for unemployed workers as a function of the agricultural wage and underlying model parameters.

Following Jones (1987, p. 195) and solving the equal flow conditions with  $L_F^2 = L_F - L_F^{-1}$ , yields

$$\frac{bL_F^1}{N^1 - L_F^1} = \frac{(b + q + \pi)(L_F - L_F^1)}{N^2 - (L_F - L_F^1) - L_I}$$

which has the solution

$$L_F^1 = \frac{-A \pm \sqrt{A^2 - 4(q + \pi)(b + q + \pi)N^1L_F}}{-2(q + \pi)}$$

where  $A = b(N^2 - L_F - L_I) + (N^1 + L_F)(b + q + \pi)$ . With the formal sector wage set at  $W_F^1$ , hiring in the formal sector obeys

$$F^{(L_F^1)} \frac{\partial L_F^1}{\partial L_F} = W_F^*$$

indicating that the level of hiring in the formal sector is related to the composition of that sector's labour force.

Substituting  $a_U$  in the equal flow condition for resident workers yields the equilibrium  $L_F^{-1}$ , which, substituted into the formal jobs aggregation condition, yields the equilibrium  $L_F^{-2}$ . Additionally, we can express  $W_F$  in terms of only exogenous variables by substituting  $a_U$  in the NSC.

To solve for N<sup>2</sup>, use the equal flow conditions for migrants and residents and set  $a_{U}^{2} = a_{U}^{-1}$  to yield:

$$\frac{(b+q+\pi)L_F^2}{N^2-L_F^2-L_I} = \frac{bL_F^1}{N^1-L_F^1}$$

Solving for  $N^2$ , the size of the migrant labour force in the urban area, we get the following expression in terms of  $N^1$ , the size of the resident labour force,  $L_F$  and

L<sub>1</sub>, employment levels in the formal and informal sectors, as well as some turnover parameters:

$$N^{2} = L_{F}^{2}\left[(b+q+\pi)N^{1}-\frac{(q+\pi)}{b}\right]+L_{J}$$

Note that the level of migration is positively related to the number of formal sector jobs "allocated" to migrants,  $L_{F}^{2}$ , as well as to the size of the informal sector. Finally, to implicitly solve for  $L_F$  and  $L_I$ , use the marginal productivity hiring rule.

The last expression also predicts that the level of migration for any urban centre is positively related to the size of the resident labour force as the partial effect of  $N^1$  on  $N^2$  is positive.<sup>104</sup> This result may help to explain why the process of urbanization has led to the emergence of megacities throughout the developing world.

In summary, the incomplete screening equilibrium with  $a_1 = 0$  differs from the complete screening perfect discrimination case by the employment of migrants in the formal sector,  $L_F^2$ , the positive number of firings in equilibrium, at

<sup>104</sup> 
$$\frac{\partial N^2}{\partial N^1} = \frac{\partial L_F^2}{\partial N^1} \left[ (b+q+\pi)N^1 - \frac{(q+\pi)}{b} \right] + L_F^2(b+q+\pi)$$
 which is positive  
with  $\frac{\partial L_F^2}{\partial N^1} > 0$ . In words, provided that a larger pool of resident workers, N<sup>1</sup>,

 $\partial N^1$ 

leads to an increase in the number of formal sector jobs filled by migrants,  $L_{F}^{2}$ , larger population centres will attract more migrants. A larger N<sup>1</sup>, or a larger formal sector queue, leads to a reduction in the effort eliciting wage which leads to an increase in the level of formal sector employment. Given that employers are hiring both residents and migrants, part of the increase in the number of jobs should indeed go to migrants.

the rate  $qL_F^2$ , the positive number of migrants returning to the agricultural region from the formal sector, at the rate  $\pi L_F^2$ , the distinct reasons for the choice of unemployment by residents and migrants and the aforementioned connection between the level of hiring to the makeup of the labour force.

The sectoral ranking of wages remains as before, although the reduced probability of job acquisition by residents ensures that  $W_F^*$  is lower, so that the low wages in the informal sector, translate not just into a lower job acquisition probability for residents, but also into a reduced compensation package once such a job is found. However, the inability of employers to observe migration status and optimally discriminate leads to some lost output as some of the hires choose to shirk. For employers then, the lower wage rate serves as (partial) compensation for the costs brought about by the higher rate of turnover in equilibrium.

In conclusion, the presence of workers with different expected job tenures in formal sector employment makes it profitable for (formal sector) employers to discriminate against the high turnover workers. The resulting pattern of occupational segregation arises not out of human capital considerations, but out of profitmaximizing behaviour, and will, therefore, not be eroded by market forces. In the likely scenario that information on quit behaviour is known solely to the worker, it maybe be possible for employers to infer behaviour from signalling on the part of workers. In particular, if workers perfectly sort themselves, with low turnover workers choosing to remain unemployed and high turnover workers choosing to work in the informal sector, occupational segregation will be complete as formal sector employers optimally restrict hiring to the unemployed, or the low turnover workers. In this way, the model provides explanations both for discrimination on the part of the employer and the luxury unemployment hypothesis which are not based on differences in human capital. The necessary sectoral sorting on the part of workers requires wages in the informal sector to be high enough that high turnover workers optimally choose informal employment over unemployment.

In the case that informal sector wages are so low as to cause some high turnover workers to actually choose unemployment, formal employers still find it optimal to restrict hiring to the unemployed, but now some high turnover workers actually land formal jobs. With formal employers offering a wage just high enough to elicit effort from low turnover employees, high turnover employees maximize their utility by shirking and, consequently, a positive level of layoffs is experienced in equilibrium. The level of employment in the formal sector will then be partly determined by the composition of the labour force.

### 6. Summary

In the absence of institutional mechanisms, as assumed by HT, and in the presence of unemployment queues, increasing urbanization and continued migration flows, it must be that formal sector firms are maximizing their profits by maintaining their relatively high wages. We argue that the link is through the effect of high wages on productivity. Additionally, the presence of urban dualism with respect to firm organization and the ability of efficiency wage models to generate a dualistic structure as in Jones (1987) and Esfahani and Salehi-Isfahani (1989), led us to extend the monitoring model of efficiency wages initially proposed by Shapiro and Stiglitz (1984) and expanded on by Jones (1987) to the prospect of rural-urban migration.

The current model can be extended in a number of important ways. First, the evidence on migration with pre-arranged jobs, for example Banerjee (1991), along with the ready availability of informal jobs (which blunts the importance of pre-arranged jobs, at least for temporary migrants) leads us to believe that some portion of the permanent migrants in our model are likely to move to a pre-arranged formal sector job. This is likely to occur because of the financial costs of long searches from unemployment and suggests that we ought to incorporate the possibility of rural-based search.

Second, the incomplete screening solution suggests that once a migrant has been hired into a formal sector job, his 'value' has apparently increased: given their formal sector experience, other formal firms are likely to ascribe them a higher probability of being a low turnover worker. Since the wage offered is bound to be less than the required efficiency wage (i.e. migrants will shirk), this increased "capitalized value" is the real prize in being hired into a formal sector job. Optimally then, migrants will quit and return home before they are monitored and fired. If they can do this, and assuming other formal sector firms are unable to completely discern between quits for exogenous reasons, q, and quits to return home,  $\pi$ , migrants retain their higher capitalized value which is translated into a higher chance of being re-hired into a formal sector job.

In effect, having a formal job on their "resume"<sup>105</sup> and having subsequently returned home, migrants have a higher probability (compared to their original migration decision) to re-migrate to the urban area because their valuation of urban opportunities has increased. The observed pattern would then be several moves over a lifetime, at least for "successful" migrants. In terms of the information flow in the labour market, migrants would like to compound the initial error of being hired by switching from formal job to formal job with a home interlude between assignments.

The result of this plan on the equilibrium solution would be a lower number of dismissals (reducing them to zero in the extreme case of perfect execution). Additionally, employers are likely to end up with an overestimate of b, i.e. they will ascribe at least some of the quits to return home to quits of the exogenous type. With the present model, there is nothing to trigger the "optimal" time for this return home: the transition variables are Poisson probabilities (i.e. constant over time) so that it is just as "optimal" to leave immediately upon being hired as to wait twelve years. We may then need to add some type of cost recovery function. This scheme may also serve to explain why families of migrants would

<sup>&</sup>lt;sup>105</sup> It seems reasonable to expect employers facing a problem of adverse selection to appeal to any fragment of employment history at their disposal. It also seems reasonable to assume that ex-employers would be unable to distinguish between an exogenous quit and a "temporary" quit to return home for a (short) while.
be willing to fund fairly long searches for a formal job when  $\pi > 0$ : in all likelihood the migrant will continue to participate in family life (ie. remit), if only to ensure there is a home to return to.

Note, however, that this type of reputation effect, where the employee would not want to be fired may also allow the employer to use it as an (implicit) threat in order to elicit effort, as originally mentioned by Shapiro and Stiglitz (1984, p. 443). Of course, for this to be true there must some penalty associated with being fired over and above the immediate loss of employment. If there is indeed a cost to the loss of reputation, migrants may be induced to supply (more) effort at the current wage, at least prior to their last planned foray into the urban labour market.

Finally, it is likely that workers endowed with more human capital are motivated to differentiate themselves from workers with less. Similarly, "senior" migrants, those with at least one successful urban "tour of duty" (where success means landing a formal sector job and then quitting before being monitored and fired), have a desire to differentiate themselves from the "novice" migrants. Within our framework, this may be interpreted to mean that "seniors" only migrate with a pre-arranged job. That is, once you have a reputation you would not want to sully it by being unemployed or working in the informal sector.

In general, the proposed model appears to conform more closely with stylized facts prevailing in developing countries, and should then be better suited for policy analysis. Compared to HT-type models, the predicted levels of unemployment and migration are lower, migrants to cities do not necessarily have lower incomes than they earned before migrating, migrants are not more likely to experience unemployment nor is migration solely a lottery system for formal sector jobs. In addition, the model (1) provides non human-capital based explanations for systemic discrimination by formal sector employers and the existence of luxury unemployment, or the idea that the unemployed optimally bypass informal sector opportunities; and (2) argues that it is imperfect information on the part of the employer, instead of imperfect information on the part of the worker as predicted by, for example, Pessino (1991), that generates the crucial rural-urban wage gap. The structure of the model also appears to have wide support from the literature as we detail in the first chapter.

Although formal policy analysis based on the proposed model remains for future work, it is clear that the HT first-best solution of removing the institutional impediment (e.g. repealing minimum wage legislation) is not applicable. In addition, it is not a priori clear that the proposed HT combination of a subsidy to formal employment combined with a physical restriction on labour migration would enhance overall welfare. While it is true that formal jobs create rents, raising the revenue necessary for such a subsidy is likely to be problematic. In fact, preliminary investigation indicates that a balanced budget scheme which would tax formal employment and subsidize rural employment would increase utility and reduce the level of migration compared to the laissez faire solution.

## APPENDIX A

Starting with a utility function defined over wages (equated to consumption at any point in time) and effort, U(w,e), and assuming separability, risk neutrality and disutility of effort, we can write the instantaneous utility as U = w - e. Setting the model in continuous time, workers choose effort to maximize expected discounted utility:

$$U = E \int_0^\infty (W_t - e_t) \exp(-rt) dt$$

To derive the continuous time valuation equations, consider a short time interval [0,t], and re-write the utility of being employed as the instantaneous (dividend) benefit plus the benefit of changing employment status (transition or capital gain). For a worker supplying effort, we then arrive at:

$$V_E = wt - et + [btV_U + (1-bt)V_E] e^{-rt}, \qquad (A.1)$$

where  $V_E$  = expected utility of an agent currently in employment,  $V_U$  = expected utility of an agent currently in unemployment and b = exogenous quit rate.

Substituting  $e^{-rt} \approx (1 - rt)$  and re-arranging, we get

$$V_E = \frac{wt - et + (1 - rt)bt V_U}{1 - (1 - rt)(1 - bt)} .$$

We now wish to take limits as  $t \rightarrow 0$ . Because the limit for both numerator and denominator is zero and because we do have two one-sided limits, we can use L'Hôpital's rule to arrive at:

$$V_E = \frac{w - e - r(btV_U) + bV_U(1 - rt)}{r(1 - bt) + b(1 - rt)}$$

$$135$$

$$(r+b) V_E - 2rbt V_E = w - e + bV_U - 2rbt V_U$$

$$r V_E = w - e + b(V_U - V_E) + 2rbt(V_E - V_U)$$

Now, taking limits as  $t \rightarrow 0$ :

$$rV_E = w - e + b(V_U - V_E)$$

This is then designated as the expected lifetime utility of a non-shirker,  $V^N$ , equation (2) in Shapiro-Stiglitz.

We can similarly derive  $V^s$ , the expected lifetime utility of a worker who (optimally) chooses to supply no effort. To equation (A.1) above, add the probability of being monitored and fired, q, and drop the effort variable, e, to get

$$V_E = wt + [(b+q)tV_U + (1-(b+q)t)V_E] e^{-rt} .$$

Again substituting  $e^{-rt} = (1 - rt)$  and re-arranging, we get

$$V_E = \frac{wt + (1 - rt)(b + q)tV_U}{1 - (1 - rt)[1 - (b + q)t]}$$

Again using L'Hôpital's rule:

$$V_E = \frac{w - r(b+q)tV_U + (b+q)V_U(1-rt)}{r[1-(b+q)t] + (b+q)(1-rt)}$$

Re-arranging:

$$(r+b+q)V_E - 2(b+q)rtV_E = w + (b+q)V_U - 2(b+q)rtV_U$$

$$rV_E = w + (b+q)(V_U - V_E) + 2(b+q)rt(V_E - V_U)$$

Taking limits at  $t \rightarrow 0$ :

$$rV_E = w + (b+q)(V_U - V_E)$$

We have then arrived at the "fundamental asset equation for a shirker", equation

(1) in Shapiro-Stiglitz.

#### APPENDIX B

The model in section five introduces labour force heterogeneity by ascribing different probabilities of leaving formal sector employment for resident and migrant workers. This appendix details both the structure and the solution procedure for each of the two possible equilibrium configurations. First, the complete screening equilibrium has the following structure.

Migration equilibrium condition:(migrants) $V_I = V_A$ .Sectoral allocation condition:(migrants) $V_I^2 > V_U^2$ ; and(residents) $V_U^1 > V_I^1$ .

Informal labour market clearing condition:  $N^2 = L_I$ .

Equal flow condition: (residents)  $\mathbf{a}_{U} (\mathbf{N}^{1}-\mathbf{L}_{F}) = \mathbf{b}\mathbf{L}_{F}$ . Unemployment rate:  $U = \frac{(N^{1}-L_{F})}{N^{1}}$ . Aggregation condition: (urban labour force)  $\mathbf{N} = \mathbf{N}^{1} + \mathbf{N}^{2}$ .

The level of hiring in each sector is set by the marginal productivity rule.

Endogenous variables are  $W_F$ ,  $W_I$ , N,  $N^2$ ,  $L_F$ ,  $L_I$ , U and  $a_U$ .

Exogenous variables are  $W_A$ ,  $N^1$ , r, b, q, e,  $e_I$ ,  $a_I$  and  $\pi$ . The flowchart below details the solution process for this model.



-----

Under the incomplete screening equilibrium, we have the following structure.

Migration equilibrium condition:(migrants)
$$V_I = V_A = V_U$$
.Equal flow condition: (residents) $\mathbf{a}_U (\mathbf{N}^1 - \mathbf{L}_F^{-1}) = \mathbf{b} \mathbf{L}_F^{-1}$ ; and(migrants) $\mathbf{a}_U (\mathbf{N}^2 - \mathbf{L}_F^{-2} - \mathbf{L}_I) + \mathbf{a}_I \mathbf{L}_I = (\mathbf{b} + \mathbf{q} + \pi) \mathbf{L}_F^{-2}$ .Aggregation conditions:(urban labour force) $\mathbf{N} = \mathbf{N}^1 + \mathbf{N}^2$ ; and(formal sector jobs) $\mathbf{L}_F = \mathbf{L}_F^{-1} + \mathbf{L}_F^{-2}$ .

The level of hiring in each sector is set by the marginal productivity rule. Endogenous variables are  $W_F$ ,  $W_I$ , N,  $N^2$ ,  $L_F$ ,  $L_F^1$ ,  $L_F^2$ ,  $L_I$ , U and  $a_U$ . Exogenous variables are  $W_A$ ,  $N^1$ , r, b, q, e,  $e_I$ ,  $\pi$  and  $a_I$ . The flowchart below details the solution process for this model.



Figure B-2. Incomplete Screening

## 140

## **BIBLIOGRAPHY**

- Adams, J. (1991). "The Rural Labour Market in Zimbabwe", Development and Change, 22: 297-320.
- Banerjee, Biswajit (1991). "The Determinants of Migrating With a Pre-arranged Job and of the Initial Duration of Urban Unemployment", *Journal of Economic Development*, 36: 337-351.
- Bhattacharya, P. C. (1993). "Rural-Urban Migration in Economic Development", Journal of Economic Surveys, 7(3): 243-281.
- Bulow, J.I. and L. Summers (1986). "A Theory of Dual Labor Markets with Application to Industrial Policy, Discrimination and Keynesian Unemployment". *Journal of Labor Economics*, 4(3): 376-414.
- de Soto, Hernando (1989). The Other Path: The Invisible Revolution in the Third World. New York: Harper & Row.
- Esfahani, H.S. and D. Salehi-Isfahani (1989). "Effort Observability and Worker Productivity: (1989) "Effort Observability and Worker Productivity: Towards an Explanation of Economic Dualism". *The Economic Journal*, 99: 818-36.
- Harris, J.E. and M. Todaro (1970). "Migration, Unemployment and Development, a Two-Sector Model". *American Economic Review*, 60: 126-42.
- Hart, Keith (1973). "Informal Income Opportunities and Urban Employment in Ghana", *The Journal of Modern African Studies*, 11: 61-89.
- International Labour Organization (1991). "The Urban Informal Sector in Africa in Retrospect and Prospect: An Annotated Bibliography". *International Labour Bibliography No. 10.* Geneva: International Labour Office.
- Jones, S.R.G. (1985). "Dual Labour Markets, Productivity and Unemployment". Discussion Paper 85-22, April (Department of Economics, University of British Columbia, Vancouver).
- Jones, S.R.G. (1987). "Screening Unemployment in a Dual Labor Market". *Economics Letters*, 25: 191-5.

- Kahnert, Friedrich (1987). "Improving Urban Employment and Labor Productivity", World Bank Discussion Papers No. 10. Washington, D.C.: The World Bank.
- Leibenstein, H. (1957). Economic Backwardness and Economic Growth. New York: Wiley.
- Mazumdar, D. (1990). "Government Intervention and Urban Labour Markets in LDCs", in Scott, M. and Lal, D., ed., Public Policy and Economic Development: Essays in Honour of Ian Little, pp. 345-360. Oxford: Clarendon Press.
- Ross, Sheldon (1989). Introduction to Probability Models, 4<sup>th</sup> Edition. New York: Academic Press.
- Roy, A. D. 1951. "Some Thoughts on the Distribution of Earnings", Oxford Economic Papers (New Series), 3: 135-146.
- Sahn, D. and H. Alderman (1990). "Formal and Informal Sector Wage Determination in Urban Low-Income Neighborhoods in Pakistan". Living Standards Measurement Study Working Paper No. 65. Washington, D.C.: The World Bank.
- Shapiro, C. and J.E. Stiglitz (1984). "Equilibrium Unemployment as a Worker Discipline Device". American Economic Review, 74: 433-44.
- Stiglitz, J.E. (1975). "Information and Economic Analysis", in Parkin, M. and Nobay, A.R., ed., Current Economic Problems, pp. 27-52. Cambridge: Cambridge University Press.
- Williamson, J.G. (1988). "Migration and Urbanization", in Chenery, H. and Srinivasan, T.N., ed., *Handbook of Development Economics*, Volume 1: 425-65. Amsterdam: North-Holland Publishing.

## CHAPTER 3.

## MIGRATION PATTERNS IN MALAYSIA

### 1. Introduction

Our analysis of lifetime internal migration in Malaysia uses the Malaysian Family Life Survey-2 (MFLS-2) conducted by the Rand Corporation and the National Population and Family Development Board of Malaysia in Peninsular Malaysia. The primary objective of this chapter is to provide an introduction to the data set through a series of descriptive statistics and bivariate analysis. Additionally, we compare MFLS-2 data with established patterns and highlight some of the issues fundamental to the process of internal migration in the Malaysian context. Statistical inference is taken up in a companion chapter.

Given the overwhelming importance of the rural-urban dichotomy and the ethnic split in both the economic and political spheres, the discussion is arranged largely along those two dimensions. Following a brief introduction to Malaysia and to MFLS-2, we discuss demographics, migration patterns, migration spells, education selectivity, age selectivity, and life cycle effects. The final section concludes with a summary of our findings and some suggestions for further research.

142

## 2. Malaysia

The country of Malaysia is located in Southeast Asia and is composed of Peninsular Malaysia, which lies in the southern part of the Malay Peninsula, and the northern half of the island of Borneo, which lies across the South China Sea from Peninsular Malaysia. These two regions are typically considered to differ in a structural way, and previous analyses have either focussed on Peninsular Malaysia, where most of the population lives, or, at least, conducted separate studies. As MFLS-2 is a representative survey only of Peninsular Malaysia, our study retains a similar focus. Appendix A contains a map of Peninsular Malaysia showing the eleven states and one federal territory, and three tables of development indicators gathered mainly from World Bank (1994) with comparative figures for Malaysia, South Korea, Philippines and China.

Malaysia is a multiracial society with 58% of the population of Peninsular Malaysia in 1988 composed of native Malays, traditionally peasant farmers, 32% Chinese, many of whom have been in Malaysia for more than one generation and originally came to work in the tin mines and branched out in commerce, and 10% Indians, many of whom work in rubber estates or in public services. Ethnic considerations, although generally well managed by policy makers, have had profound impacts in politics, with the constitution traditionally attempting to balance the political power of the Malays and the economic power of the Chinese, but also in economic policy, especially since the race riots of 1969.<sup>106</sup>

At the time of independence from Britain in 1957, Malaysia fit the typical characterization of a single-crop economy.<sup>107</sup> Since then, an emphasis on export diversification and a willingness to plan for new industries has led to significant flows of foreign investment and an uninterrupted and quite spectacular increase in national income. Since 1957, the only recession of significance was the result of global economic conditions in the mid-1980s but lasted only two years (1985 with a decrease in GDP of 1.0% and 1986 with an increase of 1.2%). Throughout, Malaysia has enjoyed annual increases in GDP in the order of 5-8%, as shown in the first table. In per-capita terms, GDP increased, on average, 3.5% between 1960 and 1970, 5.1% between 1970 and 1982 (Spinanger, 1986) and 3.2% between 1980 and 1992 (World Bank, 1994).

<sup>106</sup> See Jomo (1990) for an account of Malaysia's pre- and post-independence history.

<sup>107</sup> According to Jenkins and Lai (1990), at the time of independence from Britain in 1957, agriculture and mining were the dominant economic sectors, with the main exports being rubber (over half of export earnings and a quarter of gross domestic product) and tin (between 10 and 20 percent of total export revenues). Unlike most single-crop economies, however, Malaysia has been credited with having a reasonably good and widely accessible educational and health system as well as better infrastructure and civil service in place at the time of independence. For a discussion of these issues see Spinnanger (1986).

1	956-60	1961-65	1966-70	1971-75	1976-80	1981-5	1986-8
Average	41	5.0	54	73	86	51	4.8

Table 1. Gross Domestic Product Annual Growth Rates in Percentage (constant prices)

Source: Jomo, K.S. (1990), Table 3.1, page 39.

Despite the generally impressive macroeconomic record portrayed above, Jomo (1990) reports on page 6 that the almost singular pursuit of income growth by Malaysia's government was accompanied in the 1960s by "a growing gap between town and country and growing inequality within all the major ethnic groups" in the distribution of income. Evidence provided by Young, Bussink and Hasan (1980) shows that, historically, the northern region, comprising the states of Perlis, Kedah, Kelantan and Terengganu, had the lowest level of income per capita and the highest incidence of poverty, while the federal territory of Kuala Lumpur had the highest level of income per capita and the lowest rate of poverty. Similarly, rural areas and Malays tend to have the lowest level of income and the highest rate of poverty. These patterns continue to the present day as discussed in Jomo (1990) and Spinanger (1986) and form the basis of regional groupings used in Spinanger (1986) and in this paper.

The growing inequality in the income distribution in all ethnic groups, but especially among the Malays,<sup>108</sup> fuelled ethnic-based interpretations of a deteriorating political climate and led to political upheavals, including racial riots in 1969 and the replacement of the Prime Minister after that year's elections. The

<sup>&</sup>lt;sup>108</sup> See Jomo (1990, p. 143) for a brief review and interpretation.

subsequent rise to power of elements favourable to the concerns expressed by the Malay majority led the government to devote more resources to the pursuit of both employment creation and a more equal distribution of income. Also not lost on the government was the simultaneous need for continued high rates of growth in order to allow the new fiscal responsibilities to be funded in a climate of social harmony.

New policy measures included a renewed emphasis on widening the accessibility of basic services, and a more active role in enhancing the access of native Malays (*bimuputras*) to public sector jobs through an expansion of the public service and to the modern private sector by setting up a 30% quota for native Malays for most jobs. These policies were officially outlined in the New Economic Policy (NEP) introduced in 1970. Although subsequently tempered as the 1980s brought to light both growing government deficits and a domestic savings shortage, these policies did serve to restore a sense of political harmony and also appear to have had a significant impact on the incidence and the distribution of poverty<sup>109</sup>.

Although the NEP may have heralded a re-alignment of some of the goals of economic development, Malaysia's policy makers remained true to their longterm goals of bringing about structural change in a stable environment. By main-

<sup>&</sup>lt;sup>109</sup> For a view sympathetic to government achievements which accepts government figures as solid evidence see Spinanger (1986) and Demery and Demery (1992). For a more critical view which questions the credibility of some of the claims made by government data see Jomo (1990).

taining the high degree of openness which characterized the economy prior to independence, economic policy has continued to follow the principle of comparative advantage, and also continued to stress export diversification and a harmonious relationship with the private sector, in the pursuit of rapid economic growth. The establishment of export processing zones in the early 1970s attracted American and Japanese electronic manufacturers and made Malaysia the second largest exporter of electronic components after Japan. Although rising costs and changing technology would later reduce the number of unskilled workers, the electronic industry drew most of its initial recruits from the rural areas of Malaysia, thus fulfilling some of the government promises of increased industrial jobs for the mainly Malay rural workers. The attractions for the foreign companies were low wages, motivated workers and generous tax exemptions.

Providentially, the 1970s also brought about large increases in the prices of primary products and Malaysia, as mainly a primary producer and exporter, took great advantage. In addition to the traditional exports of tin and rubber (including synthetic rubber in which Malaysia had earlier invested), palm oil, timber, cacao and pepper also became export crops over the 1960s and 1970s, and petroleum production came online just as oil prices soared after 1973. Through all the good fortune the government remained concerned with widening the industrial base, and consequently, Jomo (1990, p. 43) reports that the share of manufacturing increased from approximately 8% of gross domestic product at the time of independence to over 13% in 1970 and to 20% in 1980. At the same time, the share of agriculture fell from 40% at the time of independence to just under 31% by 1970 and to just under 23% by 1980.

The decade of the 1980s opened with an unexpected decrease in commodity prices which led the government to implement a vigorous spending program, including the purchase of foreign-held corporations and an expansion of the public sector, that lasted approximately four years. This was viewed as a Keynesiantype smoothing action and signalled the government's intent to maintain the high rates of economic growth, especially in light of the NEP. On the natural resource front, the 1980s witnessed the start of petroleum gas production, almost exclusively for export to Japan, providing yet another source of revenue for Malaysia. By 1988, the share of manufacturing stood at 23.9% of gross domestic product, while the share of agriculture was down to 21.2%.

A two-year long recession in the mid-1980s along with rising government budget deficits and persistent current account deficits led to a tighter fiscal policy, and a new emphasis on more skilled manufacturing (especially in the face of rates of unemployment in the order of 3% and rising wages) and heavier industry. The persistence of large deficits in the services account have led to the encouragement of native freight and insurance companies and the introduction of investment inducements aimed at reducing the outflow of repatriated profits. Rising concern about the sustainability of export-led growth has also led the government to introduce a new population policy with the aim of increasing the size of the domestic market by raising Malaysia's population from 18.7 million in 1992 to 70 million

148

by the year 2100. The government now recommends families have five children in order to achieve this demographic goal.<sup>110</sup>

In summary, it appears that the success story of Malaysia has been built on macroeconomic and political stability. Although there have been instances of discord, such as the riots of 1969, and regress, such as the recession of the mid-1980s, it appears that the pursuit of high rates of economic growth has been, at least partly, consistent with income distribution and regional and employment policies which aim to reduce traditional inequalities. The consistent gains of the 1960s were augmented in the 1970s by the discovery of oil, favourable external conditions in the form of rising commodity prices, and the birth of the electronics industry. The first half of the 1980s maintained the higher rate of growth began in the 1970s, partially due to an expansionary fiscal policy. The regress came about with the recession of 1985 and 1986, mainly as a result of world economic conditions, but the high rates of growth did resume in 1987 and the momentum has carried forward into the first half of the 1990s, with improved commodity prices and strong growth in the exports of manufactured goods.

#### 3. Malaysian Family Life Survey-2

There have been two waves of the Malaysian Family Life Survey (MFLS). The first wave, MFLS-1, was carried out in 1976-77 and surveyed a nationally representative sample of 1,262 households in peninsular Malaysia each containing

<sup>&</sup>lt;sup>110</sup> For a discussion, see Jomo (1990, section 9.2).

an ever-married woman under age 50. As a follow-up to MFLS-1, field work for MFLS-2 was carried out between August 1988 and January 1989. Both surveys provide household-level retrospective and current data on women and their husbands, including data on employment, migration and training, as well as basic demographic and educational information. MFLS-2 also includes detailed information on each household's wealth, earned income and intergenerational transfers in the year preceding the interview.

MFLS-2 consists of four samples: (1) the "Panel" sample, made up of the 1,262 households interviewed in MFLS-1, of which 889, or 72%, were re-interviewed; (2) the "Children" sample, for which were eligible children 18 or older of the original 1,262 MFLS-1 female primary respondents, with a final sample size of 1,096, of whom 499 were living in Panel households; (3) the "New" sample, for which were eligible women aged 18-49 or ever-married women under age 18, with 2,184 respondents, only 6 of whom were under 18; and (4) the "Senior" sample, made up of 1,357 respondents, of whom 671 were male, aged 50 or older.

The data was collected with eight survey instruments and we focus on MF23, Male Life History. This questionnaire was administered to all male primary respondents in the Children sample and to current husbands of all female primary respondents in the Panel, Children and New samples, for a total of 1,550 men for the Panel and Children samples and 1,513 men for the New sample. It excludes some husbands aged 50 and over of women in the New sample, who were instead administered the Senior questionnaire. For the migration history questionnaire, respondents were asked for a complete listing of their inter-district moves from the age of 15 to the survey date in 1988.

Approximately half of the interviewing staff were selected from the regular staff of the National Population and Family Development Board (LPPKN), with the other half hired temporarily for the MFLS-2 project. Selection of interviewers was in part conditioned by the desire to have households contacted by interviewers of the same ethnic group. All interviewers underwent 22 days of training sessions which included survey overviews, basic interviewing skills and practice interviews.<sup>111</sup> Field work was carried out in teams led by a senior field supervisor with experience in household surveys. Interviewers were responsible for determining eligibility of respondents for each of the samples, and for collecting information using the printed instruments as well as revisiting the household if necessary. Procedures for collecting, processing and storing the data followed LPPKN practices and were also reviewed and approved by RAND. The median interview length per household was 64 minutes while the median for the MF23 guestionnaire was 17 minutes. Consistency checks, both computerized and manual, were undertaken before, during and after data entry and complemented considerable supervision of data entry personnel.

The data for our study is drawn from the Panel and Children samples. For the Panel sample, 889 of the 1,262 women interviewed for MFLS-1 were success-

<sup>&</sup>lt;sup>111</sup> The Interviewer's Instruction Manual is part of the documentation available with MFLS-2. The present discussion is based largely on section 6 of Haaga et al (1993).

fully re-interviewed. Although no follow-up to MFLS-1 was planned so that detailed location information was not originally collected, interviewers learned that 31 of the women died in the intervening period while 2 had moved either to East Malaysia or to another country, leaving 1,229 women eligible to be re-interviewed. Among the women successfully located, 13 refused to participate in MFLS-2 while 21 were unable to participate for other reasons. The other 306 women were not located.

Of the 889 women in the Panel sample, 768 reported that they were married at the time of the interview, and for 717 of these cases, the Panel respondent's spouse completed MF23. Of the 51 other cases, 8 husbands refused to be interviewed, 10 were never home, 1 was ill, 30 were not living in the household and 2 did not complete the questionnaire for other reasons. Additionally, 11 husbands completed MF23 while their spouses did not complete the female questionnaire.

Geographically, the response rates for the Panel sample were consistent with all states having response rates exceeding 60% and all but 3 of the 37 districts represented in MFLS-1 having response rates of at least 50%. The next table reproduces the results in Table 14 in Haaga et al (1993) and reveals that only for Chinese respondents living in metropolitan towns was the follow-up rate below 60%.

Ethnic Group Metropolitar		Smaller Cities	Rural
Malays	63.8%	77.2%	87.0%
Chinese	45.3%	65.5%	66.8%
Indians	80.6%	67.7%	65.4%
Total	54.9% (244)	69.8% (272)	79.4% (713)

 Table 2. Follow-up Rates for Potential Panel Sample Members (MFLS-1 Women)

 by Ethnic Group and Urban/Rural Residence in 1976

The working sample size of 1,456 men with a total number of 2,289 recorded migration moves was arrived at by omitting observations with intra-district moves and moves before the age of 15, and omitting individuals with moves to either a military base or unknown type of destination for a total of 426 omitted migration moves and 94 individuals. For purposes of the survey, a migration move for males was defined to involve an individual crossing a district boundary so that intra-district moves were not routinely recorded.<sup>112</sup> Similarly, the survey took 15 to be the age at which an individual becomes economically independent, and again, moves before the age of 15 were only infrequently recorded. Moves to a military base were not included because of the potential for movement associated with leaves of absence or involuntary transfer to different stations, while moves with unknown destination were dropped primarily because of the potential interest in differentiating between types of destination.

<sup>&</sup>lt;sup>112</sup> For females, all moves, including intra-district residence changes, were recorded.

The quality of data collected from the Female Life History questionnaire (MF22) in the New Sample is assessed by Sine and Peterson (1993). Data on marital status, fertility, infant and fetal mortality, birth weights, contraception, breastfeeding, and education are compared to known or expected patterns and to other data sources. In general, the report finds that the degree of consistency is high and concludes that the data are generally representative of the Peninsular Malaysian population. In the next section, we provide a description of the demographic patterns in MFLS-2 and compare them with 1980 Census data.

#### 4. Demographics

Although Malaysia ranks as one of the faster developing LDCs in terms of GNP growth over the last three decades, increases in the urbanization rate have been relatively restrained. The 1982 World Development Report (World Bank, 1982, p. 149) indicates that, from 1960 to 1980, the proportion of Malaysia's population in urban areas increased from 25% to 29%. Over the same time period, the same proportion for countries in the same broad income group, the 'middle-income economies', jumped 12%, from 33% in 1960 to 45% in 1980. Part of the explanation, however, is definitional as Malaysia has defined an urban area to consist of a minimum of 10,000 people, a comparatively stringent classification.<sup>113</sup>

(continued...)

<sup>&</sup>lt;sup>113</sup> The definition of urban areas usually includes more than a quantitative threshold. Usually included are references to forms of local government,

In our sample, the urban share was approximately 26.7% with the

respondents spatially distributed as shown in Table 2.

		Place of Residence at Age 15							
		R	URAL		URBAN				
Migrant:	Kam- pung	Estate	Land Scheme	New Village	Small Town	Large Town	City	Total	%
NO	393	55	2	53	113	35	14	665	45.7
YES	461	55	6	41	137	70	21	791	54.3
Total	854	110	8	94	250	105	35	1456	100.0
%	58.6	7.6	0.5	6.5	17.2	7.2	2.4	100.0	

Table 3. Rural-Urban Composition in Panel Sample of MFLS-2

Note: Kampung is a Malay word meaning village. Estate refers to a rubber plantation or large agricultural operation. Land scheme refers to settlement schemes supported by the Federal Land Development Authority which cleared large tracts of virgin land and subdivided them into 10-acre family holdings. The classification of new villages as rural and small towns as urban follows Census department records as reported in Peterson et al (1993, p. 426). According to the 1980 Population and Housing Census of Malaysia, as referenced in Schätzl (1988, p. 62), Kuala Lumpur had a population of 937,817 with the next largest city, lpoh, having a population of 300,325. There were 9 other settlements with a population between 135,000 and 250,000, and 2 others with a population between 75,000 and 88,000. Peterson et al (1993, p. 696-97) reports that in 1988 there were 82 towns with a population of 10,000-74,999.

## 4.1 Ethnic Composition

As mentioned previously, native Malays make up approximately 58% of

the population of Peninsular Malaysia and are, traditionally, rural-based. The

 $^{113}$ (...continued)

<sup>&</sup>quot;gazetted areas", or lists of characteristics such as population density or percentage of population engaged in non-agricultural work. With respect to the threshold, United Nations (1989) indicates that most countries use 5,000 or less as the criterion. In Southeast Asia, India, Pakistan, Korea and Bangladesh all use 5,000 while Indonesia and Thailand do not specify quantitative criteria in their definitions.

Chinese and Indian communities form the only significant minorities, with the Chinese being traditionally urban based and dominating commerce, and the Indians being traditionally either agricultural estate workers or public service employees. Table 4 and the associated tests of independence<sup>114</sup> provide some support for these generalizations.

Residence						
At Age 15	Malay	Chinese	Indian	Other	Total	
Urban	128	181	79	2	390	
	(14.88 %)	(46.77 %)	(40.31 %)	(15.38%)	(26.79 %)	
Rural	732	206	117	11	1,066	
	(85.12 %)	(53.23 %)	(59.69%)	(84.62 %)	(73.21 %)	
Total	860	387	196	13	1,456	
	(100.0 %)	(100.0 %)	( 100.0 %)	( 100.0 %)	(100.0 %)	
Pearson chi2(3) = 160.0643 Pr = 0.000 likelihood-ratio chi2(3) = 158.2863 Pr = 0.000						

Table 4. Ethnic Distribution by Rural/Urban

We find that a significantly larger proportion of Malays are indeed ruralbased, followed by Indians and Chinese, in that order. Given that the majority of population are Malay, it is also true that most of the rural residents are Malay (ap-

<sup>&</sup>lt;sup>114</sup> The null hypothesis for both the Pearson  $\chi^2$  and the likelihood-ratio tests is that a particular outcome measure (for example, ethnicity) is independent of a particular characteristic used to define different samples (for example, residence). If  $n_{ij}$  (*i*=1,...,I, *j*=1,...,J) is the number of observations in the *i*<sup>th</sup> row and *j*<sup>th</sup> column, the Pearson  $\chi^2$  statistic is defined as  $\chi^2 = \sum_i \sum_j (n_{ij} - m_{ij})^2 / m_{ij}$ , where  $m_{ij} = [(\text{row } i \text{ total})(\text{column } j \text{ total})] / \text{sample size}$ . The likelihood-ratio  $\chi^2$  statistic is defined as  $L^2 = 2\sum_i \sum_j n_{ij} \ln(n_{ij}/m_{ij})$ . For both tests, the degrees of freedom are (I-1)(J-1). For reference, see Stata Corporation (1993, pp. 212-13).

proximately 69% based on 732 Malays out of a total of 1,066 rural residents).

From the above table, the ethnic distribution in our sample is 59.1% Malay,

26.6% Chinese and 13.5% Indian. As mentioned previously, the population in

Peninsular Malaysia in 1988 was 58% Malay, 32% Chinese and 10% Indian. The

differences are consistent with our data reflecting a lower urbanization rate,

26.7%, than the official rate, 29%.

# 4.2 General Area of Origin

These ethnic groups also differ markedly in their general area of origin as shown by Table 5.

Table 5. Lume Distribution by Region of Residence at Age 15								
General		Ethnic Composition						
Area	Malay	Chinese	Indian	Other	Total			
Northern	423 ( 49.19 % )	63 (16.28 %)	10 (5.10%)	10 ( 76.92 % )	506 (34.75 %)			
Southern	164 (19.07 %)	78 (20.16 %)	10 (5.10%)	0	252 (17.31 %)			
Western	267 (31.05 %)	214 (55.30%)	164 ( 83.67 % )	3 (23.08 %)	648 (44.51 %)			
Foreign	6 (0.70%)	32 ( 8.27 % )	12 (6.12 %)	0	50 (3.43 %)			
Total	860 (100.0 %)	387 (100.0 %)	196 (100.0 %)	13 (100.0%)	1,456 (100.0 %)			
	Pears	Pearson chi2(9) = 331.9660 Pr = 0.000						

Table 5. Ethnic Distribution by Region of Residence at Age 15

Note: the Northern region encompasses the states of Trengganu, Kelantan, Perlis and Kedah; the Southern region encompasses the states of Pahang, Melaka and Johor; the Western region encompasses the states of Perak, Selangor, Negeri Sembilan, Penang, and the federal territory of Kuala Lumpur. These divisions are due to Spinanger (1986).

The patterns evident in Table 5 are consistent with a double income gap, the customary rural-urban difference overlayed by a Malay-non Malay disparity. The table shows a higher concentration of Malays in the less developed, and more rural, northern area, and a higher concentration of non-Malays in the relatively more developed, and more urban, Western states. Such patterns highlight the interaction of geography and ethnicity in Malaysia are widely cited in the literature, including Schätzl (1988).

### 5. Migration Patterns

Consistent with a free market economy described by significant regional income variation, the Malaysian labour market is characterized by significant levels of internal migration. Out of a total of 1,456 survey respondents, 791 report having made at least one move over their lifetime. In this section, we first assess the consistency of the migration data in the Panel and Children samples of MFLS-2 and then examine the impact of spatial and ethnic considerations on prevailing migration patterns.

### 5.1 Rural-Urban Distribution

We start our discussion of migration patterns by scrutinizing directional tendencies for the first two migration moves recorded by each individual migrant.

After disaggregating the data along the lines of ethnicity and area of origin, we examine the impact of each of these two factors, separately, on the incidence (participation rate) and on the frequency (total number of moves) of migration. We conclude the discussion on spatial considerations with a summary of individual lifetime migration patterns (up to the third move recorded by respondents).

The spatial distribution of the first migration move recorded by respondents is shown in Table 6 below.

	Destin		
Origin	Rural	Urban	Total
Rural	268	295	563
	(33.9 %)	( 37.3 % )	(71.2%)
Urban	49	179	228
	(6.2 %)	( 22.6 % )	(28.8%)
Total	317	474	791
	(40.1 %)	( 59.9 % )	(100.0 %)

Table 6. Destination and Origin of First Migration Moves

Two important conclusions follow from the above table: first, the rural areas supply the bulk of the migrants, both to other rural areas as well as to the urban areas; and, second, the level of urban to urban migration is quite significant. This is largely consistent with the descriptive literature on the Malaysian economy, for example Mazumdar (1981).

	Desti		
Origin	Rural	Rural Urban	
Rural	149	54	203
	(25.1 %)	(9.1%)	(34.2 %)
Urban	159	232	391
	( 26.8 % )	( 39.0 % )	(65.8%)
Total	308	286	594
	(51.9%)	(48.1 %)	(100.0%)

Table 7 presents the spatial distribution for the second migration move.

Table 7. Destination and Origin of Second Migration Moves

Comparison of Tables 6 and 7 highlights the reversal in the relative importance of migration flows by origin. While for the first migration a larger percentage of the migration flows originates from an rural area, the opposite is true for the second migration. The next two moves, details of which are not shown, extend the urban-bias with 58% of moves in the third migration, and 67% of moves in the fourth migration, originating in an urban area.

Two possible explanations for this pattern reversal are (1) that return migration is important in the Malaysian context, or (2) that workers originally urban-based (making mainly urban to urban moves) migrate more often. To pursue the first hypothesis, the next table shows the proportion of second or higher migration moves that represented a return to the district of residence at age 15.

	Return to district of 1			
Residence at age 15	YES	NO	Total	
Rural	359	655	1,014	
	( 24.0 % )	( 43.7 % )	(67.7%)	
Urban	119	365	484	
	( 7.9 % )	( 24.4 % )	(32.3 %)	
Total 478		1,020	1,498	
(31.9%)		(68.1%)	(100.0%)	
Pearson chi2 (1) = 17.6452 Pr = 0.000 likelihood chi2(1) = 18.1181 Pr = 0.000				

Table 8. Incidence of Return Migration (Second or Higher Migration Moves)

Return migration, mainly by workers residing in a rural area at age 15, does appear to be important with 35% (359 out of 1,014) of migrations by such workers representing a return move. Overall, almost 32% of second of higher moves are to the district of residence at age 15. Table 9, by restricting attention to the second migration move, highlights the return pattern more clearly.

	Return to district				
Residence at age 15:	YES NO		Total		
Rural	214 ( 36.0 % )	214 (36.0%)	428 (72.1 %)		
Urban	66 (11.1%)	100 (16.8 %)	166 (27.9%)		
Total	Total         280 (47.1 %)         314 (52.9 %)				
Pearson chi2(1) = 5.0342 Pr = 0.025 likelihood-ratio chi2(1) = 5.0662 Pr = 0.024					

Table 9. Incidence of Return Migration (Second Migration Moves Only)

Fully 50% of the second moves by rural-origin workers are return moves, compared to 40% for urban-origin workers, and 47% overall. The  $\chi^2$  statistics indicate that rural workers are significantly more disposed to return to their district of origin than their urban counterparts, with fully 50% of second moves by migrants from a rural area representing a return home.

To examine the explanation that urban-based workers tend to migrate more often, I compare both the participation rates and the average number of moves by migrants along rural-urban lines. First, I find urban workers to be marginally more likely to engage in migration, as shown in Table B.1 on page 195. Second, I find that urban migrants also tend to make marginally more migration moves (t-statistic equal to 1.94), as shown in Table B.2 on page 195. A t-test on the equality of means over the entire population, i.e. including each group's nonmigrants in the sample, indicates the average difference of 0.5 migration moves to be statistically significant at any level greater than 0.58%.<sup>115, 116</sup> This result represents the cumulative effect of a marginally higher participation rate for urban workers and marginally more moves for those urban workers who do choose to migrate.

In a more general test, the  $\chi^2$  statistics for Table 10 below (marginally) reject residence at age 15 as an important determinant of the number of migration moves for migrants. This test covers approximately 73% of migrants (575 of a total of 791 workers with at least 1 move).

	Total number of migration moves:				
	1	2	3	Total	
Urban	62	54	36	152	
	( 40.79 % )	(35.53 %)	(23.68 %)	(100.00 %)	
Rural	135	195	93	423	
	(31.91%)	(46.10%)	(21.99%)	(100.00%)	
Total	197	249	129	575	
	( 34.26 % )	(43.30%)	(22.44 %)	(100.0 %)	
Pearson chi2(2) = 5.6008 Pr = 0.061 likelihood-ratio chi2(2) = 5.6294 Pr = 0.060					

Table 10. Number of Migrations by Rural/Urban

<sup>115</sup> Interestingly, conducting a similar test along the lines of birth place yields an insignificant difference (Pr | t | = 0.2979). This may indicate that family moves early in an individual's life are an important determinant of migration status.

<sup>116</sup> Including non-migrants in the number of observations is more relevant if (some of) the zeroes are generated not by people who would never migrate but, instead, by people whose choice, under the current circumstances, is not to migrate. In other words, the zeroes should be included if there is not a separate process, such as some level of aversion to migration, generating at least some of the (current) absences from migration. While it appears that area of origin plays only a marginal role in the determination of both participation rates and the number of moves undertaken by a migrant, significant compositional differences are revealed by Table 10. In particular, we note that the distribution of the total number of moves has a peak at 1 for urban residents and a peak at 2 for rural residents. We return to this topic in the next section when we further disaggregate the data to consider the impact of ethnicity on the migration process.

Table 11 gives the spatial distribution of all migration moves undertaken by MFLS-2 respondents past the age of 15. Although the migration literature has traditionally concentrated much of its efforts in explaining rural to urban migration, it appears that, at least in Malaysia, rural-rural and urban-urban routes are even more widely travelled. In this case, concentrating on rural-urban migration would overlook approximately 4 out of every 5 migration moves.

	Destir		
Origin	Rural	Urban	Total
Rural	593	483	1,076
	(25.9%)	(21.1 %)	(47.0%)
Urban	384	829	1,213
	(16.8 %)	( 36.2 % )	(53.0 %)
Total	977	1,312	2,289
	(42.7 %)	( 57.3 % )	(100.0%)

Table 11. Spatial Distribution of Migrations by Area of Origin

## 5.2 Ethnic Composition

In this section, we consider the impact of ethnicity on the direction of migration, on the participation rate and the frequency of migration, and on return migration. Table 12 decomposes the proportion of moves undertaken by workers of different ethnic groups into the four different combinations of rural and urban areas, and reveals some significant differences among ethnic groups with respect to the direction of their migration moves.

ORIGIN:	Rural		Ur		
DESTINATION:	Rural	Urban	Rural	Urban	Total
Malay	49.85 %	17.44 %	11.92 %	20.79 %	100.00 %
Chinese	31.83 %	14.11 %	11.29 %	42.77 %	100.00 %
Indian	33.71 %	15.46 %	8.38 %	42.45 %	100.00 %

Table 12. Rural/Urban Distribution by Ethnic Group

Apparent from the previous table are the relative preferences of Malays for a rural, and of non-Malays for an urban, destination. These two trends are consistent with the previously noted (1) predominance of rural-rural and urbanurban paths, (2) the relative concentration of Malays in rural areas and non-Malays in urban areas, and (3) the relative importance of return migration.

Table 13 shows the ethnic distribution for migrants and non-migrants, as well as the results of two  $\chi^2$  tests for the independence of the rows and columns.

	At least 1 move beyond age 15:			
Ethnic Group	NO	YES	Total	
Malay	361	499	860	
	( 41.98 % )	( 58.02 % )	(100.00 %)	
Chinese	216	171	387	
	( 55.81 % )	( 44.19 % )	(100.00 %)	
Indian	80	116	196	
	( 40.82 % )	(59.18%)	(100.00 %)	
Other	8	5	13	
	(61.54%)	( 38.46 % )	(100.00%)	
Total	665	791	1,456	
	( 45.67 % )	( 54.33 % )	( 100.00 % )	
Pearson chi2(3) = 23.9569 Pr = 0.000 likelihood-ratio chi2(3) = 23.9156 Pr = 0.000				

 Table 13. Migration Participation Rates by Ethnic Group

The  $\chi^2$  statistics associated with this table reveal the observed differences to be significant. Alternatively, defining a migrant as a worker with at least two migration moves or as a worker with up to three migration moves yields results similar to those presented above. In essence, Chinese workers are, on average, less likely to migrate than their counterparts. Once the decision is made to migrate, however, there is no significant difference in the total number of moves between Chinese and non-Chinese, or between Malays and non-Malays, as shown below in Table B.3, on page 196. A more general tabulation not only confirms the lack of importance of ethnicity in determining the total number of migration moves made by migrant workers, but also reveals little difference in the ethnic

167

composition of those with 1, 2 or 3 moves. This is shown in Table B.4 on page 196.

Next, we examine the relationship between place of origin and ethnicity. Differentiating the sample of non-Malay workers along rural-urban lines and performing the usual tests of independence on the decision to migrate yields a  $Pr(Pearson \chi^2) = 0.548$ . Statistically then, it appears that spatial considerations have relatively little impact on the decision to migrate for non-Malay workers. Table 14, however, reveals significant differences between rural and urban Malays with regards to willingness to move.

	At least 1 migration move after age 15:			
Residence at age 15	YES	NO	Total	
Rural	403	329	732	
	(55.1 %)	( 44.9 % )	(100.0 %)	
Urban	96	32	128	
	( 75.0 % )	( 25.0 % )	(100.0 %)	
Total	499	361	860	
	(58.0 % )	( 42.0 % )	(100.0 %)	
Pearson chi2(1) = 17.7949 Pr = 0.000 likelihood-ratio chi2(1) = 18.7413 Pr = 0.000				

Table 14. Participation Rates Among Malays

Additionally, while urban and rural non-Malays tend to make approximately the same number of migration moves<sup>117</sup>, urban Malay workers tend to

<sup>&</sup>lt;sup>117</sup> A mean comparison test does not reject the null hypothesis of equality as the t-stat=0.42 with 290 degrees of freedom has a Pr > |t| = 0.6720.
make a significantly higher number of moves than their rural counterparts as shown in Table 15.

Residence	Number of Inter-District Migrations for Malay Workers			
At Age 15:	Observations	Mean	Std. Dev.	
Urban	96	3.3	2.4	
Rural	403	2.8	2.0	
combined	499	2.9	2.1	
H <sub>0</sub> : mean(x) = mean(y) t = 2.35 with 497 d.f. Pr > $ t $ = 0.0192				

Table 15. Number of Migrations for Malay Workers

The earlier conclusion that area of origin plays only a marginal role in the migration process is not substantiated when we disaggregate the data by ethnic group. In fact, we have just shown that, for Malays, both the participation rate and the number of migrations are higher for urban than rural-based workers.

Comparing urban workers across ethnic groups, I find that urban Malay workers have a higher participation rate in migration than other urban workers, indicating that the earlier finding of a marginally higher participation for urban workers is mainly driven by the sub-sample of urban Malay workers.

At least 1 Move After					
Age of 15:	Malay	Chinese	Indian	Total	
NO	32	99	30	161	
	(25.0%)	(54.7 %)	(38.0 %)	(41.5%)	
YES	96	82	49	227	
	( 75.0 % )	( 45.3 % )	(62.0%)	(58.5 %)	
Total	128	181	79	388	
	(100.0 %)	(100.0 %)	(100.0%)	(100.0 %)	
Pearson chi2(2) = 27.7422 Pr = 0.000 likelihood-ratio chi2(2) = 28.4201 Pr = 0.000					

 Table 16. Participation Rates by Ethnic Group for Urban Workers

By contrast, urban Malay workers are found not to make significantly more moves than urban non-Malay workers<sup>118</sup>, an indication that the earlier finding that urban workers tend to make marginally more moves is not entirely driven by the sub-sample of urban Malay workers.

The next table reveals that the impact of ethnic considerations on return migration is statistically marginal.

<sup>&</sup>lt;sup>118</sup> A mean comparison test does not reject the null hypothesis of equality as the t-stat=1.14 with 226 degrees of freedom for a Pr |t| = 0.2562.

Return to District of Res-					
idence At Age 15:	Malay	Chinese	Indian	Total	
NO	643	210	165	1,018	
	(68.6 %)	(64.0%)	(73.3 %)	(68.3 %)	
YES	295	118	49	473	
	(31.4%)	(36.0%)	( 26.7 % )	(31.7%)	
Total	938	328	225	1,491	
	(100.0 %)	(100.0 %)	(100.0%)	(100.0 %)	
Pearson chi2(2) = 5.4267 Pr = 0.066 likelihood-ratio chi2(2) = 5.4519 Pr = 0.065					

Table 17. Return Migration Participation Rates by Ethnic Group

In summary, the interaction of space and ethnicity yields two significant differences: (1) Malay workers resident in an urban area at age 15 are more likely to engage in migration and then tend to make more moves than Malay workers resident in a rural area at age 15, and (2) urban Malay workers are more likely to engage in migration, but tend to make a similar number of moves, as other urban workers. It would then appear that while the dominant effect for participation in migration is ethnic background, the key determinant for the number of moves by a migrant is the rural-urban split for residence at age 15. Given the previously noted significance of the rural-urban split with regards to participation in return migration, Table 17 indicates that a similar discrepancy does not hold across ethnic boundaries. The relatively low percentage of Indian workers that do return to their district of residence at age 15 is consistent with their more recent arrival in Malaysia.

#### 6. Migration Spells

The nature of our data allows us to examine migration spells. Continuing our focus on the effects of residence and ethnicity, we can seek answers for questions such as: once a migration move is undertaken, do rural workers tend to have longer or shorter spells? Given the dearth of migration duration data for developing countries, this is an area of migration literature without well established empirical facts.

#### 6.1 Rural-Urban Distribution

While mean comparison tests reveal no statistical difference for the duration of the first migration move between rural and urban workers, similar tests for all subsequent moves indicate a significant difference at any level greater than 2.7% (see Appendix C for details). Table 18, for example, shows that, for the second migration move, migrants originally from a rural area tend to have significantly longer stays than migrants not originally from a rural area.

Residence	Migration Spells in Years for the Second Move After Age			
At Age 15:	Observations	Mean	Std. Dev.	
Urban	166	5.9	8.7	
Rural	428	8.6	10.5	
combined	594	7.9	10.1	
H <sub>0</sub> : mean(x) = mean(y) t = $-2.94$ with 592 d.f. Pr >  t  = $0.0034$				

Table 18. Migration Spells by Area of Origin

In addressing this question, we need to consider not only the usual dimensions of space and ethnicity, but also the effects of censoring. We start by noting the lack of a statistical difference between rural and urban workers with respect to "moving" age for the first four moves, as detailed in Appendix D. Combined with our earlier finding that urban migrants make (marginally) more moves, the equality in the moving ages would lead to an initial hypothesis that rural migrants will tend to settle down sooner thereby generating longer censored spells than their urban counterparts. This explanation would be consistent with both equality in moving ages and yet longer spells for rural origin workers.

In fact, while any difference in the average length of **non**-censored spells by rural and urban workers is statistically insignificant (consistent with similar moving ages), the next table indicates the opposite for censored spells (consistent with rural migrants settling down sooner and making fewer moves, on average).

Residence	Censored Migration Spells in Years			
At Age 15:	Observations	Mean	Std. Dev.	
Urban	228	10.7	10.9	
Rural	561	13.0	11.5	
combined	789	12.4	11.3	
H <sub>0</sub> : mean(x) = mean(y) t = -2.69 with 787 d.f. Pr > $ t $ = 0.0073				

Table 19. Censored Migration Spells by Rural/Urban Origin

As of the survey date, rural migrants have, on average, been settled down for almost 3 years longer than urban migrants. Combined with the finding that rural-origin workers are, on average, 2 years older than urban workers (as discussed in the next section on Age Distribution), this implies that rural migrants tend to make their last recorded move at an earlier age than urban migrants. The next table shows this difference in the age at the last move to be significantly different.<sup>119</sup>

Residence	Age At the Last Recorded Move			
At Age 15:	Observations	Mean	Std. Dev.	
Urban	228	28.6	10.3	
Rural	563	27.1	8.9	
combined	791	27.6	9.3	
H <sub>0</sub> : mean(x) = mean(y) t = 2.03 with 789 d.f. Pr >  t  = 0.0428				

Table 20. Age at Last Move by Rural/Urban Origin

#### 6.2 Ethnic Composition

Ethnicity is judged not to affect the length of migration spells as mean comparison tests over all migration moves, as well as over single moves, consistently reject any difference in the average length of spells, or the age at the last move, of Malays compared to non-Malays or Chinese compared to Malays. A

<sup>&</sup>lt;sup>119</sup> There is no significant difference between Malays and non-Malays for the age at the last recorded move.

difference does, however, exist between rural Malays who leave for a rural destination instead of the city. Specifically, I find that rural Malays whose first migration move is to an urban centre tend to leave, on average, more than 2 years earlier than rural Malays who first move to another rural area.<sup>120</sup>

#### 7. Age Distribution

The principal implication of the human capital theory of migration is that the propensity to migrate should be highest for younger workers who would generally have the longest horizon over which to recuperate their initial investment. In this section, we examine the impact of age on migration status.

As MFLS-2 was a representative survey, all age groups are represented. The main benefit for our study is that the working data set contains a considerable amount of older respondents for whom we would expect migration careers to be completed, an important consideration for meaningful analysis of the number of lifetime moves. Graph 1 shows the age distribution for our sample.

<sup>&</sup>lt;sup>120</sup> The lack of difference in the age at the start of a move may be considered evidence against significant levels of migration for education, at least after the age of 15, from the rural areas. Under the premise that migration for education is likely to be associated with an urban destination, however, the finding that ruralurban migration starts, on average, two years earlier than rural-rural migration may be an indication that there is some migration for education. A more convincing test would compare spells of unemployment upon arrival in the city combined with listing school/training as the "reason for not working".



U

Disaggregating, we find that Indian workers are younger than either of their Ma-

lay or Chinese counterparts as shown in Table 21.

Table 21. Average Age by Ethnic Group

Race	Observations	Mean	St'd. Dev
Malay	860	39.09	14.19
Chinese	387	40.55	14.67
Indian	196	37.37	14.25

Similarly, rural-origin workers are, on average, 2 years older than urban workers, although the difference is marginally insignificant (t-statistic=1.93). The source of this age difference can be traced to the group of urban Malays who are, on the average, 4 years younger than either their Malay or urban counterparts, as shown in the Tables E.1 and E.2 on page 200.

That is, dividing the sample both spatially (rural and urban) and ethnically (Malay and non-Malay), the group of 128 urban Malays are, at an average age of

35, approximately four years younger than each of the other three groups. No such difference is evident for the rural-urban split in either the Chinese or the Indian ethnic groups.

Dividing the sample spatially and then into migrants and non-migrants, I find that the sub-sample of urban non-migrants are, on the average, 3.5 years younger than each of the other segments, a statistically significant difference. The sample statistics for each of these segments are as follows in Table 22.

RURAL	Observations	Mean	Std. Dev.
Migrant	563	40.3	13.5
Non-Migrant	503	38.9	15.4
URBAN			
Migrant	228	39.1	13.9
Non-Migrant	162	36.1	14.0

Table 22. Average Age by Migration Status

The sub-sample of urban non-migrants are significantly younger than any of the other sub-samples (for each pair-wise comparison, the t-stat is greater than 2.0). Overall, people who had migrated at least once in their lifetime were, at 40 years of age, an average of 2 years older than non-migrants (Pr > |t| = 0.0136). The source of this difference in age is evident from Table 23.

At Least 1 Migra-	Non-Malay Respondents' Current Age			
tion Move:	Observations	Mean	Std. Dev.	
NO	304	37.7	14.5	
YES	292	41.2	14.4	
combined	596	39.4	14.5	
Ho: mean(x) = mean(y) t = -2.98 with 594 d.f. Pr >  t  = 0.0030				

Table 23. Average Age for Non-Malays by Migration Status

That is, non-Malay migrants tend to be significantly older than non-Malay non-migrants, although this is not true for Malays. Specifically, it turns out that this difference is mainly driven by the Indian sub-sample as shown by Table 24. This pattern is consistent with Indians being the most recent arrivals in Malaysia. Immigrants would necessarily have at least one recorded migration move upon arrival in Malaysia while their offspring would have been in Malaysia from the start and would only record a move if they migrated internally subsequent to their fifteenth birthday.

At Least 1 Migra-	Indian Respondents' Current Age			
tion Move:	Observations	Mean	Std. Dev.	
NO	80	32.7	14.0	
YES	116	40.6	13.6	
combined	196	37.4	14.3	
Ho: mean(x) = mean(y) t = -3.97 with 194 d.f. Pr >  t  = 0.0001				

Table 24. Average Age for Indians by Migration Status

In accordance with their relatively earlier arrival, no such difference is evident for the sub-sample of Chinese workers. In the next section, we consider the effects of education on the migration process.

#### 8. Education

Education has been repeatedly identified as a key contributor to the process of development, particularly since Schultz (1961) conceptualized investment in education as an investment in human capital. Today, education is universally acknowledged as a key component of human resource development, and educational expenditure and accessibility are two of the most cited measures of a developing country's performance. In the migration literature, it is well established<sup>121</sup> that education has an important effect on migration. Yap (1977) and

<sup>&</sup>lt;sup>121</sup> See Gould (1993, chapter 7) for a brief discussion on the relationship between education and both internal and international migration, and Sabot (1979, (continued...)

others have particularly noted that migrants tend to be better educated than workers at their place of origin. In this section, we examine the impact of education on both the propensity to migrate and the direction of migration, and we also compare educational levels of migrants and non-migrants.

### 8.1 Educational Selectivity

Stratifying the sample by the level of educational achievement, I find a clear difference in the willingness to migrate among different educational groups. This is one form of educational selectivity in migration, a characteristic commonly noted in the literature<sup>122</sup> and evident in Table 25.

<sup>121</sup>(...continued)

p. 105) for a discussion in the African context.

<sup>&</sup>lt;sup>122</sup> See, for example, Sabot (1979), p. 105.

Highest Level of	At Least 1 Migration Move?		
Schooling	NO	YES	Total
No School	58	63	121
	( 47.93 % )	( 52.07 % )	(100.00%)
Primary	325	332	657
	( 49.47 % )	( 50.53 % )	(100.00%)
Secondary	265	283	548
	( 48.36 % )	( 51.64 % )	(100.00 %)
Post-secondary	17	113	130
	(13.08 %)	( 86.92 % )	(100.00 %)
Total	665	791	1,456
	( 45.67 % )	(54.33 %)	(100.00%)
Pearson chi2(3) = 61.3201 Pr = 0.000 likelihood-ratio chi2(3) = 69.3333 Pr = 0.000			

Table 25. Propensity to Migrate by Educational Group

The strong effect of the level of education on the incidence of migration becomes slightly more pronounced when we stratify the sample by the type of location at age 15. As shown in Appendix F on page 201, among workers with a post-secondary education who reside in a rural location at age 15, fully 92% engage in migration. As we do not (and may ultimately be unable to) differentiate between migration for education and work, these figures may capture some migration for education effect, which is compatible with the usual charac-terization of educational opportunities as urban-biased.

By focussing on migrants originating in rural areas and dividing the sample into those headed for an urban or a rural area, we can shed further light on the prevailing pattern of migration by educational group. As shown in Table 26, there is a definite preference by more educated workers for the urban setting with a corresponding preference for the countryside by those with less education. Although some migration for education effects may be included, this distribution is also compatible with the traditional occupational opportunity sets at each type of location.

Highest Level of	First Migration Move:		
Schooling	Rural-Rural	Rural-Urban	Total
No School	36	16	52
	( 69.23 % )	( 30.77 % )	(100.00%)
Primary	151	114	265
	( 56.98 % )	( 43.02 % )	(100.00 %)
Secondary	72	119	191
	( 37.70 % )	(62.30%)	(100.00 %)
Post-secondary	9	46	55
	(16.36%)	( 83.64 % )	(100.00%)
Total	268	295	563
	( 47.60 % )	( 52.40 % )	(100.00 %)
Pearson chi2(3) = 48.1305 Pr = 0.000 likelihood-ratio chi2(3) = 50.6914 Pr = 0.000			

Table 26. Direction of First Migration by Educational Group for Rural Respondents

Although not often mentioned in the migration literature, it is also possible that education, quite apart from the effect it has on the propensity to migrate, also has an impact on the number of moves over a migrant's lifetime, and this effect need not be in the same direction. In fact, we might expect that better educated workers possess more accurate labour market information, and might then require fewer moves over their lifetime to find a compatible employment match. If this were true, education would tend to increase the propensity of migration but decrease the total number of moves. Stratifying the data on the lifetime number of moves by educational strata, we find that the only statistically significant difference is that post-secondary workers tend to make 1 to 1.5 more moves than other groups. It would then appear that education has similar effects on the propensity and on the frequency of migration. That is, better educated workers tend to have both a higher participation in, and a higher frequency of, migration.

One prominent stylized fact previously mentioned is that migrants tend to be more educated than non-migrants. We start our analysis by comparing the level of education of workers resident in rural versus urban areas at age 15. Given the likelihood of easier access to the education system for urban dwellers, we should expect these workers to be, on average, better educated. This is confirmed by mean comparison tests, both for residence at birth and at age 15 as shown in Appendix G on page 202.

Similarly, migrants tend to have significantly more education than workers who stay behind. Mean comparison tests confirm the education gap when we compare all migrants to all non-migrants, Malay migrants to Malay non-migrants, and all non-Malay migrants to all non-Malay non-migrants. Appendix H on pages 203 and 204 details the results. Similar tests comparing migrants with at least two moves to all others reveal similar results and are shown in Appendix I on page 205.

### 9. Life Cycle Effects

The previously mentioned lack of difference in moving age does not hold if we divide the sample into censored and non-censored observations. For the first move after the age of 15, migrants with censored spells (those whose first move is also their last move) are, on average, 3 years older at the time of the move, with a similar pattern holding for subsequent moves. This result, unchanged by spatial or ethnic considerations, leads to the conclusion that workers who start their migration career later in life tend to have fewer moves.

For participation in migration, I find evidence of a common start age. With the average age for the first recorded move being approximately 21 years of age, more than three quarters of those engaging in migration have made their first move by the age of 23, and only 9% of migrants make their first move at 30 years of age or older. As can be ascertained from Graph 2 and the underlying data presented in Appendix J on page 206, 287, there is also some evidence of a discontinuity in the propensity to start migrating between the ages of 16 and 20, with a marked increase in the percentage of migrants starting their migration career within this age span. Such a clear life cycle pattern would be reason to model the



decision to participate in migration separately from the process generating the number of moves.

Graph 2 also plots the cumulative frequency for the age at the last recorded move with the average age being approximately 27.5 years. Relative to the first move, the distribution for the last move is considerably more uniform, indicating that the evidence for a common stop age is relatively less convincing or, at least, indicating relatively weaker life cycle patterns. In fact, while the range between 15 and 20 years of age contains approximately 62% of respondents in the case of the start process, the next table shows that the range would have to be considerably wider to contain similar percentages in the case of the stop process. The underlying data are shown in Appendix K on pages 207 and 208.

Age Range	Percentage
25 - 30	25
25 - 35	37
20 - 30	56
20 - 35	68

Table 27. Percentage of Migrants Making Their Last Move Within an Age Range

The feasibility of effectively modelling at least the start process is enhanced by the presence of a variety of information on family background like inheritance of physical assets from parents, and other family endowment factors such as parents' education, number of older siblings and parents' occupation.

There is also a possibility that the migration decision ought to be modelled separately by area of residence. Certainly, typical lifetime migration patterns appear to be different for urban versus rural workers. Appendix L on pages 210 and 211 contains a summary of the lifetime migration patterns reported by MFLS-2 respondents up to three moves. For respondents who reside in a rural area at age 15, the listing reveals the relative importance of return migration for those who do seek an urban job (of the 295 rural-urban migrants, 138 have by the survey date returned to a rural area). Among rural-urban migrants who make a second move

to an urban area (there are a total of 113 such rural-urban-urban migrants), the majority (57) make a further move to an urban area. This may be illustrative of the type of learn-by-doing sequential migration modelled by Pessino (1991). The third relatively important pattern among rural residents, involves two or three moves all from rural areas to rural areas.

For respondents who start their exposure to migration in an urban area, the overwhelmingly popular pattern is to move between urban areas only. The listing reveals that of the 228 migrants who resided in an urban area at age 15, 119 made two moves to other urban areas while 70 of these actually made a third move to another urban area while 39 stopped after their second move. These findings serve to reinforce the fact that internal migration patterns in developing countries is not at all dominated by the standard rural-urban moves.

#### 10. Conclusion

This chapter has served to highlight some of the issues fundamental to the process of internal migration in Malaysia by highlighting some of the patterns in the Panel sample of the MFLS-2 migration data. The analysis has confirmed some stylized facts common in the development literature, namely that education plays an important role in migration decision and that migrants tend to start their migration career at a relatively young age. Specifically, we confirmed that migrants tend to be more educated than those who choose to stay behind, with the more educated among the migrants preferring urban rather than rural destinations. We also confirmed the idea generated by human capital theory that potential migrants have an incentive to start their migration career earlier rather than later in life. Accordingly, we found that while three quarters of migrants make their first move by the age of 23, only 9% of make their first move past the age of 30.

The analysis has also confirmed some stylized facts common in the literature on Malaysia, namely that spatial and ethnic considerations interact and that rural-urban moves are not the most common type of migration. In general, we found that Malay workers are more likely to engage in migration than non-Malays, with urban Malay workers having a higher participation rate in migration and also making a significantly higher number of moves over their lifetime. We also found that the standard rural-urban move makes up only 21% of the total number of moves recorded in the Panel sample, with the most common type of move being urban-urban. This is consistent with findings by Mazumdar (1981).

Given the retrospective nature of the data, we were also able to discern some patterns not usually accessible to researchers. We found that education appears to have a positive impact both on the incidence and on the frequency of migration as workers with post-secondary education have both a higher participation rate and, on average, a higher number of moves. We were also able to evaluate the importance of return migration and we found that indeed return migration is an important component of the migratory patterns in Malaysia with workers residing in a rural area at age 15 being significantly more disposed to return to their original district. Additionally, using the migration spell data, we also found that,

187

by making their last move at an earlier age, rural migrants tend to settle down sooner than urban migrants.

Although this section on descriptive statistics and bivariate analysis has served to highlight what theory seeks to explain, it does not allow us an explanation of the underlying process. As economists and policy makers we are interested in the marginal impact of determining factors, for which we must resort to statistical inference and actually estimate parameters from the data gathered in a multivariate context. Such an approach may yield results at odds with findings garnered from a bivariate analysis, where all else is not held constant, because of the impact of correlation patterns among the determining factors. In a companion chapter, we pursue an event count approach to modelling lifetime internal migration in Malaysia and we indeed find some results for which the econometric model gives importantly different results than would be apparent from cross tabulation procedures. Specifically, we find that some factors associated in this data chapter with an increased number of migrations, namely residing originally in the poorer Northern region and comparing rural and urban Malays, turn out to be insignificant determinants in a multivariate context. Additionally, while cross tabulations show that mother's (father's) education has a significantly positive (insignificant) effect on the number of migrations, the econometric model yields the reverse pattern with mother's education being insignificant and father's education being positively significant. Most significantly, two variables, the effect of working at a young age and the effect of an urban family moving between the respondent's birth and fifteenth birthday, give strictly opposite results. In cross-tabulations, these two factors have, respectively, significantly negative and positive impacts on the number of moves. In the econometric model, however, the effects are still statistically significant but the sign is reversed with working at a young age leading to a higher number of moves, and early family movements leading to fewer moves for urban families.

In addition to our own work on the determinants of lifetime migration, the comprehensive nature of the data set is certainly conducive to a variety of behavioural studies. For example, the presence of retrospective data on the various life courses, such as marriage and work, allows simultaneous estimation of different trajectories, while the data on migration allows the use of duration modelling to study the determinants of migration spells. These are subjects about which development economists know relatively little, mainly because this type of data has hitherto been unavailable for developing countries.

To conclude, we should note that our findings on the existence of a common start age for the migration process lead us to believe that the start process should be modelled separately, so that in modelling the number of lifetime migration moves, for example, it would appear appropriate to use a double hurdle type of model, where a choice of zero moves could arise either from the participation or from the frequency decision. The data also suggest that migration behaviour differs by region of residence with the differences between rural/urban workers noted above. Table 26 on the direction of the first move by educational stratum certainly indicates that there are compositional differences between different streams. It would then seem interesting to split the population into a rural and an urban sample and estimate separate equations for each sample. Additionally, the methodology of multinomial and nested logits could then be useful to investigate differences in the structure of the migration decision by using the type of lifetime patterns shown in Appendix L. Specifically, this would allow us to test whether migration decisions are sequential (consistent with the idea of learn-by-doing).

### APPENDIX A

Figure A-1. Malaysian Regional and State Divisions



	South	Korea	Phili	ppines	Ch	ina	Mala	aysia
Capital	Se	oul	Ma	anila	Bei	jing	Kuala I	Lumpur
Area	98,50	0 km <sup>2</sup>	300,0	000 km <sup>2</sup>	9,600,0	000 km <sup>2</sup>	330,00	)0 km <sup>2</sup>
<b>Population</b> (millions, mid-1992)	43	3.7	64.3		1,162.2		18	1.6
GNP per capita (US\$, 1992)	6,7	790	7	770		70	2,7	'90
<b>Population average annual growth</b> (%, 1980-1992)	1	.1	2	2.4	1	.4	2.	5
GNP per capita, average annual growth (%, 1980-1992)	8	.5	-	1.0	7	.6	3.	.2
Life expectancy at birth (1992)	7	'1		65 69		7	'1	
<b>Daily calorie supply per capita (1988)</b> <sup>1</sup>	2,8	378	2,	255	2,0	532	2,6	586
Primary school enrollment rate (1991, total and female)	107	109	110	111	123	118	93	93
Illiteracy rate (age 15+, %, 1990)	4		10		27		22	
Human Development Index (1990) <sup>1</sup>	0.87	(high) 0.6 (medium)		medium) 0.61 (medium)		0.79 (m	nedium)	

Table A-1. Comparative Geographic, Social and Economic Indicators I

Source: World Bank (1994) (WB1994), Tables 1, 25 and 28. <sup>1</sup> Todaro (1994).

	South	Korea	Philip	opines	Ch	nina	Mal	aysia
Structure of Production (% of GDP): <sup>1</sup>	1970	1992	1970	1992	1970	1992	1970	1992
Agriculture	26	8	30	22	••	27	29	
Industry	29	45	32	33		34	25	
Manufacturing	21	26	25	24			12	
Services, etc.	45	47	39	45		38	46	
Demography and Fertility: <sup>2</sup>								
Crude birth rate (per 1,000 population)	30	16	38	32	33	19	36	28
Crude death rate (per 1,000 population)	9	6	11	7	8	8	10	5
Married women of childbearing age using contraception (%, 1988-1993)		17	40		83		56	
				<u> </u>				
Urbanization: <sup>3</sup>								
Urban population as % of total (1970 and 1992)	41	74	33	44	18	27	27	45
Average annual growth rate (%) of urban popu- lation (1970-80 and 1980-92)	5.3	3.4	3.8	3.8	2.7	4.3	5.0	4.8
Pop. in capital city 1990 (% of urban and total)	36	26	32	14	4	1	22	10

Table A-2. Comparative Geographic, Social and Economic Indicators II

Source: <sup>1</sup> Table 3 in WB1994; <sup>2</sup> Table 26 in WB1994; <sup>3</sup> Table 31 in WB1994.

	South	Korea	Philip	ppines	Ch	ina	Mala	aysia
Health and nutrition: <sup>1</sup>	1970	1990	1970	1990	1970	1990	1970	1990
Population per physician	2,220	1,070	9,270	8,120	1,500		4,310	2,590
Population per nursing person	1,190	510	2,690		2,500		1,270	380
Low birthweight babies (%, 1990)		9	1	15		9	9	)
Infant mortality rate (per 1,000 live births)	51	13	66	40	69	31	45	14
Income distribution: <sup>2</sup>	1990		1988		1990		1989	
Lowest 20 percent	7.4		6.5		6.4		4.6	
Highest 20 percent	42	2.2	47.8		41.8		53.7	
Highest 10 percent	2	7.6	32.1		24.6		37.9	
Infrastructure: <sup>3</sup>	1980	1990	1980	1990	1980	1990	1980	1990
Access to safe water (% urban population)	86	100	65	93	••	87	90	96
Access to safe water (% rural population)	61	76	43	72		68	49	66
Households with electricity (% total, 1984)	100		46				6	4

Table A-3. Comparative Geographic, Social and Economic Indicators III

Source: <sup>1</sup> Table 27 in WB 1994; <sup>2</sup> Table 30 in WB 1994; <sup>3</sup> Tables A-2 and 32 in WB1994.

### APPENDIX B

Note: Appendices B-L display data from MFLS-2.

	Mig					
Residence at age15:	YES	NO	Total			
Rural	563 (52.8%)	503 ( 47.2 % )	1,066 ( 100.0 % )			
Urban	228 (58.5%)	162 (41.5 %)	390 (100.0%)			
Total	1,456 (100.0 %)					
Pearson chi2(1) = 3.6700 Pr = 0.055 likelihood-ratio chi2(1) = 3.6846 Pr = 0.055						

Table B-1. Migration Participation Rates by Rural/Urban

Table B-2. Number of Migrations by Rural/Urban

Residence At	Number of Inter-District Migrations					
Age 15:	Observations	Mean	Std. Dev.			
Rural	563	2.80	2.06			
Urban	228	3.12	2.27			
combined	791	2.90	2.09			
H <sub>0</sub> : mean(x) = mean(y) t = 1.94 with 789 d.f. Pr > $ t  = 0.0528$						

	Number of Inter-District Migrations					
Malay:	Observations	Mean	Std. Dev.			
NO	292 2.89		2.03			
YES	499	2.88	2.04			
combined	791	2.88	2.04			
H <sub>0</sub> : mean(x) = mean(y) t = 0.03 with 789 d.f. Pr > $ t  = 0.9722$						

Table B-3. Number of Migrations by Malay/Non-Malay

Table B-4. Distribution of the Number of Migrations by Ethnic Group

	Total number of migration moves:					
Ethnic Group	1	2	3	Total		
Malay	123	161	78	362		
	( 62.44 % )	( 64.66 % )	( 60.47 % )	(62.96%)		
Chinese	45	51	28	124		
	( 22.84 % )	( 20.48 % )	(21.71 %)	(21.57%)		
Indian	28	35	22	85		
	(14.21 %)	(14.06 %)	( 17.05 % )	(14.78%)		
Other	1	2	1	4		
	(0.51%)	(0.80%)	( 0.78 % )	(0.70%)		
Total	197	249	129	575		
	(100.0%)	(100.0%)	(100.0%)	(100.0%)		
likelil	Pearson chi2(6) = 1.2717 Pr = 0.973 likelihood-ratio chi2(6) = 1.2605 Pr = 0.974					

### APPENDIX C

Mean Comparison tests on migration spells

(a) First migration move only:

Ru Ag	ral At e 15?		)bs	Mean	Std. De	ev.
	NO YES	2   5	28 63	6.40 6.87	8.94 9.26	
cc	ombined	7	91	6.73	9.17	
	H <sub>o</sub> :	mean(x)	= mear	n(y) (	assumin	g equal

variances) t = -0.66 with 789 d.f. Pr > |t| = 0.5068

(b) Second and subsequent moves:

Rural At | Age 15? | Obs Mean Std. Dev. NO | 712 5.95 8.11 YES | 1579 6.83 9.13 combined | 2291 6.56 8.83 H<sub>0</sub>: mean(x) = mean(y) (assuming equal variances) t = -2.22 with 2289 d.f. Pr > |t| = 0.0267

### APPENDIX D

Mean Comparison Tests on Age at the Start of a Move

(a) First migration move only:

Rural At Age 15?	(	Obs	Mean S	td. Dev.	
NO YES		228 563	21.0 21.3	6.30 5.90	
combined	7	791	21.2	6.02	
H <sub>0</sub> : variances)	mean(x)	= mean	n(y) (a	ssuming ea	qual
	t Pr >  t	t = -0.   = 0.5	67 with 059	789 d.f.	

# (b) Second migration move only:

Rural At Age 15?	Obs	Mean	Std. Dev	
NO YES	166   428	24.04 24.21	7.37 7.24	
combined	594	24.16	7.27	
H <sub>0</sub> : variances)	mean(x) =	mean(y)	(assuming	equal
	t = Pr >  t  =	-0.24 wit 0.8102	th 592 d.f	

(c) Third migration move only:

Rural At Mean Std. Dev. Age 15? | Obs \_\_\_\_\_ 26.44 112 NO | 7.85 233 26.16 8.32 YES | \_\_\_\_\_ 26.25 8.16 combined | 345  $H_0$ : mean(x) = mean(y) (assuming equal variances) t = 0.30 with 343 d.f. Pr > |t| = 0.7669

(d) Fourth migration move only:

Rural At Age 15? | Obs Mean Std. Dev. \_\_\_\_\_ \_\_\_\_\_ 76 27.67 6.00 NO | 27.29 8.36 YES 140 \_\_\_\_\_+\_\_\_+ combined | 216 27.42 7.60  $H_0$ : mean(x) = mean(y) (assuming equal variances) t = 0.36 with 214 d.f. Pr > |t| = 0.7229

.

# APPENDIX E

Ethnic	Urban Respondents' Current Age					
Group:	Observations	Mean	Std. Dev.			
Non-Malay	262	39.23	14.73			
Malay	128	35.54	12.13			
combined	390	38.02	14.02			
H <sub>0</sub> : mean(x) = mean(y) t = 2.46 with 388 d.f. Pr > $ t  = 0.0144$						

Table E-1. Average Age for Urban Workers by Ethnic Group

Table E-2	Average	Age for	Malays h	Rural/I	Irhan Origin
Table E-2.	Average	Age IOI	Ivialays Uy	ruial/c	Juan Ongin

Residence	Malay Respondents' Current Age			
At Age 15	Observations	Mean	Std. Dev.	
Urban	128	35.54	12.13	
Rural	732	39.72	14.44	
combined	860	39.09	14.19	
H <sub>0</sub> : mean(x) = mean(y) t = $-3.09$ with 858 d.f. Pr >  t  = $0.0021$				

### APPENDIX F

	Residence At Age 15					
	URBAN RURAL		RURAL			
Highest Level of Schooling	At Least 1 Mo	Migration ve?		At Least 1 Mo	Migration ve?	
	NO	YES	Total	NO	YES	Total
No School	5	11	16	53	52	105
	(31.25%)	( 68.75 % )	(100.00 %)	(50.48%)	( 49.52 % )	(100.00 %)
Primary	59	67	126	266	265	531
	( 46.83 % )	(53.17%)	(100.00 %)	( 50.09 % )	( 49.91 % )	(100.00 %)
Secondary	86	92	178	179	191	370
	(48.31 %)	( 51.69 % )	(100.00 %)	( 48.38 % )	(51.62 %)	(100.00 %)
Post-secondary	12	58	70	5	55	60
	(17.14%)	( 82.86 % )	(100.00 %)	(8.33%)	(91.67%)	(100.00 %)
Total	162	228	390	503	563	1,066
	(41.54 %)	(58.46 %)	(100.00 %)	(47.19%)	(52.81 %)	(100.00 %)
Pearson chi2 likelihood-ratio chi2	Pearson chi2(3) = 22.6687 Pr = 0.000 Pearson chi2(3) = 38.8131 Pr lihood-ratio chi2(3) = 24.6938 Pr = 0.000 likelihood-ratio chi2(3) = 45.7791 Pr			Pr = 0.000 Pr = 0.000		

Table F-1. Education Effects by Rural/Urban

## APPENDIX G

Residence	Years of Education			
At Birth	Observations	Mean	Std. Dev.	
Urban	298	8.86	4.14	
Rural	1,145	6.71	3.95	
	1.440		4.00	
Combined	1,443	7.16	4.08	
H <sub>0</sub> : mean(x) = mean(y) t = 8.27 with 1141 d.f. Pr > $ t $ = 0.0000				

Table G-1. Comparison of Years of Education by Residence at Birth

Table G-2.	Comparison of Years of Education by Residence at Age 15
1 abie 0-2.	comparison of Tears of Education by Residence at Age 15

Residence	Years of Education			
At Age 15	Observations	Mean	Std. Dev.	
Urban	389	8.68	4.32	
Rural	1,054	6.59	3.85	
Combined	1,443	7.16	4.08	
H <sub>0</sub> : mean(x) = mean(y) t = $8.82$ with 1141 d.f. Pr >  t  = $0.0000$				

### APPENDIX H

At Least 1 Move	Years of Education			
After Age 15	Observations	Mean	Std. Dev.	
NO	660	6.47	3.58	
YES	783	7.73	4.38	
Combined	1,443	7.16	4.08	
H <sub>0</sub> : mean(x) = mean(y) t = $-5.88$ with 1441 d.f. Pr >  t  = $0.0000$				

Table H-1.	Comparison of	of Years of	Education	by Migration	Status

Table H-2	Comparison of	Vears of Edu	cation by Mi	oration Status	for Malays
	Companson or	I Cais UI Luu	ication ov ivi	ikialivii Slalus	IUI IVIAIAVS

At Least 1 Move	Years of Education for Malay Workers				
After Age 15	Observations	Mean	Std. Dev.		
NO	360	6.31	3.72		
YES	494	7.65	4.48		
Combined	854	7.08	4.22		
H <sub>0</sub> : mean(x) = mean(y) t = -4.62 with 852 d.f. Pr > $ t $ = 0.0000					
At Least 1 Move	Years of Education for Non-Malay Workers				
--	--	------	-----------		
After Age 15	Observations	Mean	Std. Dev.		
NO	300	6.67	3.40		
YES	289	7.87	4.23		
Cambinal	590	7.26	2.07		
Combined	589	7.20	3.8/		
H <sub>0</sub> : mean(x) = mean(y) t = $-3.80$ with 587 d.f. Pr >  t  = $0.0002$					

Table H-3. Comparison of Years of Education by Migration Status for Non-Malays

# APPENDIX I

Mean Comparison Tests on Years of Education by Migration Status.

(a) All migrants to all non-migrants:

At Least 2 Moves?	Obs	Mean S	Std. Dev.
NO YES	855   588	6.57 8.00	3.67 4.49
combined	1443	7.16	4.08
H <sub>0</sub> :	<pre>mean(x) =     t =     Pr &gt;  t  =</pre>	mean(y) = -6.64 with = 0.0000	1441 d.f.

#### (b) Malay migrants to Malay non-migrants:

At Least 2 Moves? Obs Mean Std. Dev. -----482 6.41 NO ł 3.82 372 7.95 4.55 YES ------7.08 4.22 combined | 854

> H<sub>0</sub>: mean(x) = mean(y) t = -5.34 with 852 d.f. Pr > |t| = 0.0000

## (c) Non-Malay migrants to non-Malay non-migrants:

#### At Least

2 Moves? | Obs Mean Std. Dev. -------+----------373 NO | 6.77 3.46 YES | 216 8.10 4.39 combined | 589 7.26 3.87  $H_0: mean(x) = mean(y)$ t = -4.06 with 587 d.f. Pr > |t| = 0.0000

# APPENDIX J

Listing of Age at First Move:

Age	Frequency	Percent	Cumulative
15	21	2.65	2.65
16	92	11.63	14.29
17	98	12.39	26.68
18	112	14.16	40.83
19	86	10.87	51.71
20	79	9.99	61.69
21	51	6.45	68.14
22	45	5.69	73.83
23	29	3.67	77.50
24	22	2.78	80.28
25	26	3.29	83.57
26	18	2.28	85.84
27	15	1.90	87.74
28	11	1.39	89.13
29	13	1.64	90.77
30	5	0.63	91.40
31	14	1.77	93.17
32	6	0.76	93.93
33	4	0.51	94.44
34	8	1.01	95.45
35	6	0.76	96.21
36	8	1.01	97.22
37	2	0.25	97.47
38	1	0.13	97.60
40	4	0.51	98.10
41	1	0.13	98.23
42	3	0.38	98.61
43	1	0.13	98.74
44	4	0.51	99.24
46	1	0.13	99.37
4 /	1	0.13	99.49
49	1	0.13	99.62
52	1	0.13	99.75
	1	0.13	99.8/ 100.00
00	1	0.13	100.00
Total	791	100.00	

# APPENDIX K

Listing of Age at Last Recorded Move:

Age	Frequency	Percent	Cumulative
15	2	0.25	0.25
16	19	2.40	2.65
17 1	29	3.67	6.32
18	38	4.80	11.13
19	33	4.17	15.30
20	54	6.83	22.12
21	49	6.19	28.32
22	54	6.83	35.15
23	50	6.32	41.47
24	41	5.18	46.65
25	55	6.95	53.60
26	28	3.54	57.14
27	30	3.79	60.94
28	28	3.54	64.48
29	34	4.30	68.77
30	23	2.91	71.68
31	25	3.16	74.84
32	20	2.53	77.37
33	20	2.53	79.90
34	15	1.90	81.80
35	14	1.77	83.57
36	14	1.77	85.34
37	14	1.77	87.10
38	7	0.88	87.99
39	13	1.64	89.63
40	11	1.39	91.02
41	3	0.38	91.40
42	4	0.51	91.91
43	6	0.76	92.67
44	7	0.88	93.55
45	2	0.25	93.81
46	4	0.51	94.31
47	5	0.63	94.94
48	1	0.13	95.07
49	6	0.76	95.83
50	2	0.25	96.08
52	2	0.25	96.33

<u>Listing of</u>	Age	<u>at Last Recor</u>	ded Move	(cont'd):
Age	1	Frequency	Percent	Cumulative
	-+			
53	1	3	0.38	96.71
54	1	5	0.63	97.35
55	1	7	0.88	98.23
56	1	3	0.38	98.61
57	1	3	0.38	98.99
58	1	1	0.13	99.12
59	1	2	0.25	99.37
60	1	3	0.38	99.75
62	I	1	0.13	99.87
63	1	1	0.13	100.00
	-+			
Total	1	791	100.00	

# APPENDIX L

# Listing of Migration Patterns

Legend: R = rural node U = urban node D = Did not move (any further)







Figure L-2. Lifetime Migration Patterns for Urban-At-Age-15 Workers

# BIBLIOGRAPHY

- Demery, David and Lionel Demery (1992). "Adjustment and Equity in Malaysia", Development Centre Studies. Paris: OECD.
- Gould, W.T.S. (1993). *People and Educaton in the Third World*. Essex: Longman Group.
- Haaga, John G. et al (1993). The Second Malaysian Family Life Survey: Overview and Technical Report. Santa Monica: RAND.
- Jenkins, Glenn P. and Andrew Lai (1990). "Trade, Exchange Rate, and Agricultural Policies in Malaysia", World Bank Comparative Studies: The Political Economy of Agricultural Pricing Policy. New York: Oxford University Press.
- Jomo, K.S. (1990). Growth and Structural Change in the Malaysian Economy. London: MacMillan Press Ltd.
- Mazumdar, Dipak (1981). The Urban Labor Market and Income Distribution: A Study of Malaysia. New York: Oxford University Press.
- Peterson, Christine E., Jeffrey Sine and Deborah Wesley (1993). The Second Malaysian Family Life Survey: Codebook. Santa Monica: RAND.
- Sabot (1979). *Economic Development and Urban Migration*. Oxford: Oxford University Press.
- Schätzl, Ludwig H. (ed.) (1988). "Growth and Spatial Equity in West Malaysia", Institute of Southeast Asian Studies: Research Notes and Discussions Paper No. 63. Singapore: General Printing and Publishing Services Pte Ltd.
- Schultz, Theodore W. (1961). "Investment in Human Capital". American Economic Review, 51: 1-17.
- Sine, Jeffrey and Christine E. Peterson (1993). The Second Malaysian Family Life Survey: Quality of Retrospective Data for the New Sample. Santa Monica: RAND.
- Spinanger, Dean (1986). Industrialization Policies and Regional Economic Development in Malaysia. New York: Oxford University Press.

- Stata Corporation (1993). *Stata Reference Manual: Release 3.1*, vol. 3, 6<sup>th</sup> ed., 212-13. College Station, Texas.
- Todaro, Michael P. (1994). *Economic Development, Fifth Edition*. London: Longman Group Ltd.
- United Nations (1989). Prospects of World Urbanization, 1988. New York: United Nations.
- Yap, Lorene Y. (1977). "The Attraction of Cities: A Review of the Migration Literature", *Journal of Development Economics*, 4: 239-264.
- Young, Bussink and Hasan (eds.) (1980). *Malaysia: Growth and Equity in a Multiracial Society*. Baltimore: The Johns Hopkins University Press.
- World Bank (1982). World Development Report 1982. New York: Oxford University Press.

.

World Bank (1994). World Development Report 1982. New York: Oxford University Press.

# CHAPTER 4.

# LIFETIME MIGRATION IN MALAYSIA: AN EVENT COUNT APPROACH

# 1. Introduction

Labour mobility is an important public policy issue in both developed and developing countries and has spawned a voluminous literature, especially in the economic and demography domains. In the economic field, empirical studies of labour migration can be divided into three types: descriptive studies, those that use aggregate data and those that use individual data.<sup>123</sup> While the descriptive studies attempt to present an overall picture of the migration process and its participants, the other two types estimate parametric models usually drawing upon the human capital approach to migration pioneered by Sjaastad (1962).

The studies employing aggregate data use the gravity model framework in attempting to test the theory proposed by Sjaastad. As adapted by economists and economic geographers from physics, the gravity model proposes that the migration flows between different areas is a function of the "mass" of each location,

<sup>&</sup>lt;sup>123</sup> Yap (1977) and Greenwood (1975) survey studies using aggregate data, both gross and net migration flows. In addition, Yap (1977) discusses findings by descriptive studies and gives reference to relevant surveys.

typically represented by population and income, and the distance between them. Often, measures of economic conditions, such as the unemployment rate, are also included. The single most important conclusion arising from this research effort has been the finding that areas with a high rate of in-migration tend to also be characterized by a relatively high rate of out-migration. Additional evidence that the propensity to migrate is higher for those with a previous move led Vanderkamp (1971) to conclude that migrants are not a homogeneous group and that, in particular, migration behaviour differs between new and return migrants, the latter tending to be less responsive to considerations of incomes and distance (a key explanatory variable in gravity models). The data used in these studies is generally census type data on either gross or net movement<sup>124</sup> between provinces or metropolitan areas.

The studies employing individual-level data are of more recent vintage and focus on the relationship between an individual's characteristics and the migration decision, typically with the intention of measuring the wage gain thought to be associated with migration. These studies generally employ cross-section data and follow the framework of Robinson and Tomes (1982). Central to this approach has been the presence of a non-random selection mechanism and how to econometrically account for its effects on the wages of different groups. The treatment follows the initial application to occupational choice by Roy (1951).

<sup>&</sup>lt;sup>124</sup> Although the theoretical implications are slightly different, models using gross and net migration data have the gravity model as a common theoretical foundation.

By using lifetime migration data on individuals, this paper is in line with recent work on the analysis of the individual's decision to migrate. We go further in that we model not just the decision to migrate (participation), but also the number of moves (frequency) undertaken over a lifetime by using an event count model to take specific advantage of the discrete nature of the data. Our analysis uses the Malaysian Family Life Survey-2 (MFLS-2), a survey which provides household-level retrospective and current data on women and their husbands, including data on employment, migration and training, as well as basic demographic and educational information. A companion data chapter provides a descriptive introduction to the data set.

Our results indicate that the conventional Poisson model inadequately captures the migration process. Instead, we find that migration should be modelled as a two step decision, with individuals making both a participation and a frequency decision. The proposed model allows non-migration to be a result of either decision, so that stayers could be either potential migrants who choose to stay or non-migrants, workers who would "never" migrate.

The nature of the data set employed relieves some significant problems with previous studies, for example, the initial conditions problem, and the availability of information about different facets of the individual's background allows us to control for factors overlooked in most studies. Specifically, we are able to disentangle the effects of parent's and own education and we find that both effects induce both higher participation rates and a greater number of moves. Studies which do not control for parent's education levels would then be likely to overestimate the effect of own education. We also find some support for the effects of location specific capital. Specifically, early family moves lead both urban and rural workers themselves to have a higher participation rate in migration, but lead urban-based workers (no significant effect on rural workers) to make fewer moves. The inheritance of physical capital induces a lower participation rate and fewer moves for rural-based workers, but more moves for urban-based workers. These results are consistent with the notion of rural capital being (relatively) location specific if we accept the notion that early family moves lead to less location specific capital being accumulated.

Additionally, we find that most geographical and ethnic effects, so prominent in the accompanying data chapter, lose their significance in a multivariate context. Our results show that, ceteris paribus, the only significant effect is that urban Chinese workers tend to make significantly fewer moves than other workers. Finally, we find that inferences on the effects of both working at an early age and of early urban family moves on the number of lifetime moves by the respondent are reversed from the predictions of cross tabulation exercises such as those presented in the data chapter. Such reversals serve to confirm the importance of assessing the impact of determinants in a multivariate, rather than a bivariate, context.

This chapter is organized as follows. The second section reviews the key results from the literature on micro-based migration studies and discusses some of

the limitations of these studies. The third section discusses methodology and proposes a statistical model for lifetime migration based on a Poisson data generation process. In addition to the usual Poisson and Negative Binomial specifications, we consider a double hurdle model proposed by Greene (1994). The fourth section discusses theoretical implications and estimation results. The fifth section proposes further research and the last section summarizes.

# 2. Literature Review

Micro-based studies on the propensity to migrate have focused on the effect of either regional economic differences or migrants' characteristics, or both, on the rural-urban migration decision. Although no unified theory of migration exists, numerous determinants of migration have been identified. In her review of the migration literature, Yap (1977, p. 239) documents the generally agreed description of the typical rural-urban migrant as "young, better educated than the average rural resident, and predominantly male in Africa and South Asia". Yap (1977, p. 239) also notes that migrants tend "to move to places with higher income and employment levels than their origin locations, and friends and relatives in the destination increase the attraction". Although these findings are largely consistent with the human capital theory of migration, originally formulated by Sjaastad (1962), the static approach generally employed does not

explain the high incidence of subsequent migration noted by, for example, DaVanzo (1983).

Harris and Todaro (1970) extend the basic model by introducing the idea of uncertainty with respect to the destination wage in order to capture continued high rural-urban migration rates in the face of apparently high rates of urban (under)unemployment. Pessino's (1991) refinement incorporates a learning-bymoving process in which younger workers are especially inclined to gather information by moving. The obvious implication from both these extensions is that most return migration would tend to consist of disappointed workers who overestimate either the destination wage or the probability of finding employment. In this way, Pessino hypothesizes, and finds some empirical support for the notion, that the motivation to migrate in Peru differs by the periodicity of migration.<sup>125</sup> This is consistent with Vanderkamp (1971)'s assertion about the different motivations of original and subsequent migrants.

A key feature of empirical, as well as theoretical, work on migration to date has been the lack of distinction between factors influencing the incidence and the frequency of migration.<sup>126</sup> Tangentially relevant, however, has been the

<sup>&</sup>lt;sup>125</sup> As discussed in a later section, Pessino (1991) approximates the distinction between original and subsequent moves, elsewhere referred to as the periodicity of migration, by the origin of the move studied. This is necessary because of the lack of a complete migration history but is clearly a less than optimal treatment.

<sup>&</sup>lt;sup>126</sup> Empirically, this is reflected in the predominance of surveys conducted in urban areas with a sampling universe restricted to migrants. Clearly, inference in (continued...)

discussion on return migration. First, Yap (1977, p. 257) does mention that ruralurban migrants returning to their rural residence "had relatively high status in the rural communities. They were more likely to be from land-owning families and had slightly more education than either their rural non-migrant counterparts or the migrants still in Bogota". The implication is that these migrants had the financial ability to migrate and conduct a destination-based search but, given their (considerable) opportunities at home, also had a high reservation wage. For a number of reasons mentioned in the literature<sup>127</sup> and independent of their family wealth or status, a higher level of education would also be expected to raise the probability of migration so that both effects are working in the same direction.

Additionally, DaVanzo (1983) introduces the idea of location specific capital to differentiate between migrants who move back and migrants who move on. The central idea is that ties to a specific area, like home ownership or relatives living close by, should increase the probability of the migrant returning to the location of origin. Given that a higher number of moves is likely to result in the acquisition of less location-specific capital, there should be a positive

(...continued)

these cases is quite limited, e.g., Banerjee (1991).

<sup>&</sup>lt;sup>127</sup> These include the existence of a positive relationship between education and information about conditions farther away (wider horizons), the likelihood of a national, rather than regional, market for more skilled jobs, and the possibility that investment in education may be a sign of willingness to improve on present conditions, including a higher tendency to migrate in pursuit of higher earnings. Additionally, education may affect tastes and preferences in the direction of increasing the likelihood of migration.

relationship between the propensity to move and the number of previous moves. The implications of DaVanzo's location specific capital are similar to the previous notion that the availability and quality of local income earning opportunities have a significant impact on the incidence of migration.

In our case, there are two variables which might be considered as indicative of the level of location specific capital. First, if the family migrated between the respondent's birth and fifteenth birthday, they have likely accumulated a relatively lower level of location specific capital by the start of the respondent's exposure to migration. We would then hypothesize that these more frequent moves might increase the subsequent incidence of migration for the respondent. Second, inheritance of some form of capital asset, like a house, land or a business, is likely to increase both local income earning opportunities and also the probability of staying, or at least returning, home.

#### 2.1 Limitations of Previous Studies

In the gravity-type migration studies, the level of aggregation, typically interstate or interprovincial migration, tends to obscure differential individual and subgroup responses and to produce a narrow definition of a migration move. As documented by Yap (1977, p. 243): "typically, most of the moves within a country are within and not between states". Problems of aggregation bias are discussed by Vanderkamp (1971) who finds that different subgroups tend to respond differently to similar levels of prices and incomes, while the use of one single measure as the estimate for area income in attracting all types of migrants points to a potentially serious problem of errors in variables.

The micro-based literature is enormous but more sophisticated econometric analysis started with Robinson and Tomes (1982). The central feature of this approach is the incorporation of a Heckman-style selectivity correction mechanism to control for the non-random nature of the migration participation decision. While the theoretical framework draws on the human capital approach pioneered by Sjaastad (1962), particular difficulties with these studies relate to the lack of direction with respect to selecting variables that strongly affect the chances for participation but not the outcome under study, typically wages. Additionally, there exists the possibility that the imposed distributional assumptions result in inconsistent estimates.

Aside from methodological considerations, the main drawback of these studies centres on the cross-section nature of the surveys used. By providing only a snapshot of what is a lifetime mechanism, cross-section surveys are unlikely to be fully representative of the migration process. First, by classifying people into just two types, migrant and non-migrant, these studies ignore the potentially differential aspect of subsequent, particularly return, migration. In the MFLS-2 data, three quarters of migrants have multiple moves, with one third eventually returning to the original district of residence. As noted (with foresight!) by Yap (1977, p. 241) "multiple moves are not unusual". Tunali (1986) did employ a double selection framework to try to capture the effects of both migration and re-

migration on the earnings process but the model fit was poor, likely a result of the noted sensitivity to mispecification of these models.

Second, cross section type surveys will tend to understate the importance of temporary or short-term moves, and may also misclassify a frequent migrant's original residence. Yap (1977, p. 241) notes that "many migrants whose previous residence was an urban one had rural origins" and also that the "proportion of the urban migrant population who were actually born in other urban locations is relatively high", with Santiago, for example, at 70% in 1960 and Seoul at 30% in 1970. Although some surveys do ask about the place of birth as well as current and last location, such as those employed by DaVanzo (1983) and Pessino (1991), none of these surveys collects a complete migration history. Without further information, not only is it difficult to locate individuals at the start of their migration exposure, which may or may coincide with their place of birth, but it is also impossible to ascertain an individual's pattern or number of moves. The latter is potentially crucial, especially in light of DaVanzo's findings about the positive relationship between the propensity to move and the number of previous moves.

The lack of a complete migration history led Pessino (1991) to approximate the distinction between original and return moves by designating moves from a rural area as original and moves from an urban area as return moves. Hence, from estimates of regionally disaggregated migration decision functions, Pessino states that motivations to migrate also differ, presumably, by

the periodicity of migration. Given the significant proportion of original migrants out of urban areas, this is clearly a procedure which introduces some noise into the analysis, and which would benefit from having available a complete migration history.

Third, these surveys are typically undertaken in an urban area, see Banerjee (1991) for example, and so ignore urban-rural or rural-rural migration, which together represent approximately 43% of recorded migration moves in MFLS-2. Additionally, studies based on surveys which sample migrants only at the destination tend to exaggerate the private benefits to migration, particularly if migration disappointments are a significant part of migration flows.<sup>128</sup>

In contrast to previous studies, we have available a complete migration history for every individual in the survey and this allows us not only to model the lifetime nature of the process but also to precisely locate individuals at the outset of their period of exposure to migration. If motivations to migrate are indeed different for original and subsequent migrants, as proposed by Vanderkamp (1971), then an important contribution of this study may indeed be the relaxation

<sup>&</sup>lt;sup>128</sup> Although Yap (1977, p.256) agrees that "disappointed migrants do tend to move on or return home", she also concludes that "they are not a large fraction of the return population in rural areas or of the out-migration from a city". In the rest of the literature, there is some discord with regards to the relative importance of disappointed migrants in subsequent migration flows. Certainly, the evidence available is quite limited while the bulk of theory work to date, e.g., Harris and Todaro (1970), Yezer and Thurston (1976), Allen (1979) and DaVanzo (1983), imply that return migrants will tend to be disappointed workers. Clearly, the overestimate of the private benefits to migration is larger the larger is the percentage of disappointed migrants who respond by making further moves (either to return home or else to continue their migration career elsewhere).

of the initial conditions problem associated with cross-section estimation of the migration decision.

By using lifetime migration data on individuals, this paper is in line with recent work on the analysis of the individual's decision to migrate. The paper, however, goes further in that we model not just the decision to migrate, but also the number of moves undertaken over a lifetime. The use of lifetime migration data, advocated by Schultz (1982),<sup>129</sup> is supported by the potential relevance of life cycle effects and the documented high rates of return migration, both of which serve to highlight the lifetime nature of migration careers. Unlike most of the empirical work on migration, particularly in developing countries, our sample is not restricted by geographical location or migration status. Banerjee (1991), for example, uses data from a survey of migrants living in Delhi. This type of survey not only suffers from sampling bias owing to non-random selection but, more substantively, is likely to under represent return or short-term migrants. The resulting inferences are difficult to generalize and are certainly conditional on an individual having already chosen to migrate.

<sup>&</sup>lt;sup>129</sup> "A seemingly neglected objective in migration research is to develop satisfactory stochastic specifications of longitudinal migration models that can be estimated from time series. Progress along these lines could represent a major improvement over cross-sectional analyses" (Schultz (1982), p. 95).

# 3. Methodology

Data based on events come in two forms: event counts and durations between events. Event counts are defined to take on strictly non-negative integer values and record such occurrences as the number of trips to a recreation site in one weekend, the number of patents per firm, or the number of migrations over a lifetime. This type of data allows inferences about the factors associated with a higher or lower number of occurrences. Alternatively, the use of duration data allows inferences about the probability, or the risk, of a change in state. In this case, the dependent variable, defined to take on any non-negative real number, would measure time spent between events or the duration of a particular state. Examples of duration data include the length of each stay at recreation sites, time between patent applications or the duration of migration moves.

Given the structure of MFLS-2, we can perform our analysis in terms of either event counts or durations. For each individual respondent we have a complete migration life history consisting of the sequence of time changes ('start' and 'end' time of each spell) and of the discrete states occupied (origin and destination district for each move undertaken). As usual, theory offers precious little guidance in the selection of a statistical model, but given the retrospective nature of the data set together with the possibility of recall bias, we choose to analyse event counts rather than durations. The total number of moves should be easier to recall and relatively free of any recall bias whereas the duration of every

single move is more likely to be recalled with some error, possibly systematic in nature.

We also choose to aggregate the data to one observation per individual lifetime rather than building a panel-type data set with potentially several observations per individual. Although the aggregate method does impose a loss in terms of the efficiency in the use of the data, a panel type data set would be more likely to be affected by recall bias and be characterized by some type of data 'heaping'. The use of aggregate lifetime count data is then likely to be a more robust, if less efficient, procedure.

The next section discusses the methodology of event count models. First, we present the base Poisson model and discuss its restrictive, albeit econometrically convenient, assumptions. We then present models based on compound Poisson distributions that allow us to relax the Poisson assumption of equal mean and variance.

### 3.1 Event Count Models

The common starting point in analysing event count data is to posit a Poisson data generation process.<sup>130</sup> For stationarity, this type of stochastic

<sup>&</sup>lt;sup>130</sup> The use of ordinary least squares (OLS) to analyze event count data leads to a series of well-documented problems including non-sensical negative predictions, heteroskedastic error terms and inefficiencies. If the dependent variable is dis-tributed Poisson then, aside from employing the incorrect functional form, OLS is also biased and inconsistent. Aside from shortcomings on technical statistical criteria, King (1986, 1987) presents evidence that the use of OLS can seriously affect inference. One way to solve these as well as other problems introduced by ad-hoc correction procedures is to employ methods of analysis that explicitly involve stochastic processes and take into account the explicitly discrete form of the data, such as the Poisson regression model (see, for (continued...)

process requires (1) events occurring in non-overlapping intervals of time to be independent of each other, and (2) the probability of one event occurring in any given short time period to be the same and the probability of more than one event occurring to be small. Although we eventually do generalize our model to take into account some specific departures from the Poisson assumptions, commonly referred to as "non-Poissonness", we start by considering the standard Poisson regression model and then progress to the conventional second choice, Negative Binomial models, and eventually to the double hurdle framework. Such a progression allows us to assess the usefulness of the added complexity of double hurdle models.

# 3.1.1 Poisson Regression Model

Let  $\mathbf{Y}_i^t$  be a random variable recording, for individual i, the number of events occurring in the interval of time  $(\mathbf{t},\mathbf{t}+\boldsymbol{\delta}]$  where  $\boldsymbol{\delta} > 0$ . For  $\mathbf{Y}_i^t$  to follow a Poisson distribution, two assumptions must be satisfied. The first assumption follows from the fact that the Poisson is an exponential distribution, and states that, for each unit of observation, i.e., for each  $\mathbf{Y}_i^t$ , there is a constant rate<sup>131</sup> of occurrence,  $\lambda_{i}^t$ , so that  $\lambda_i^t = \lambda_i^{t+\delta}$  for all t and  $\boldsymbol{\delta} > 0$ . Note that (1) by including covariates as explanatory variables, we do allow  $\lambda$  to vary across individuals, so

(...continued)

example, Cox and Isham, 1980, and King, 1989a).

<sup>&</sup>lt;sup>131</sup> In terms of duration modeling, this is equivalent to specifying a constant hazard rate, implying that the hazard function is independent of elapsed duration, a type of model that follows the application of the exponential distribution to the random time of waiting for an event.

that, in particular,  $\lambda_i^t$  need not equal  $\lambda_j^t$ ; and (2) given the Poisson characteristic of equality between the mean and the variance of a process, the application of OLS to data characterized by different  $\lambda_s$  results in heteroskedastic error terms.

Essentially, this homogeneity assumption ensures that we have repeated drawings from the same probability distribution for each individual in our data set, a conjecture which is generally problematic for economic phenomena.<sup>132</sup> In particular, the presence of life cycle effects would lead us to hypothesize that  $\lambda$  is not constant over an individual's lifetime but is instead dependent on the individual's age. To control for observed systematic differences over an individual's lifetime we include, as explanatory variables, the individual's age linearly, squared and cubed. We also allow for the presence of unobserved heterogeneity, which shows up in the form of overdispersion, by changing the initial distributional assumptions and estimating a Negative Binomial model. We expand on this point in section 3.1.1.2.

The second assumption required for a stochastic process to be Poisson follows from the memoryless property of the exponential distribution and states that the probability of an event occurring during an observation period be

<sup>&</sup>lt;sup>132</sup> There are two important arguments that may allow us to relax the constant rate assumption within the context of the Poisson regression model. First, the aggregation of "a large number of independent and uniformly sparse variables of any type is approximately a Poisson process" as noted by Amburgey and Carrol (1984, pp. 41-2) as referenced in King (1988, p. 840). Second, the Poisson distribution is the limiting form of "a very large number of other distributions and naturally occurring situations" as noted by King (1988, p. 840), including the Negative Binomial.

independent of all previous history, including the length of time since the previous event and the number of previous occurrences. This assumption ensures that each trial can be characterized as time independent and implies that experiencing the event in the past does not render individuals more, or less, likely to experience the event in the future. Violation of the assumption would suggest either true or spurious state dependence,<sup>133</sup> and would imply a loss of information for methods of estimation which do not allow for some form of time dependence. In the context of event count modelling, the loss of information could be alleviated by allowing  $\lambda$  to vary with respect to time, as in the panel type estimator used by Hausman, Hall and Griliches (1984), or else by allowing for the time varying unobserved heterogeneity as accommodated by the Negative Binomial regression.

If the event generating process is indeed Poisson,<sup>134</sup> the unconditional probability distribution for  $Y_i$  is given by

<sup>&</sup>lt;sup>133</sup> See Heckman (1981) for an elaboration on the distinction between true and spurious state dependence. The influence of location specific capital on the migration decision would be an example of true state dependence. A mover might find it relatively difficult to make the first move, but once done, that move is likely to lead to a loss in location capital and a change in preferences, prices or constraints making further moves more likely. Spurious state dependence, on the other hand, occurs when an apparent relationship between previous occurrences and the probability of experiencing the outcome arises as a result of unobserved individual heterogeneity.

<sup>&</sup>lt;sup>134</sup> Note that although researchers typically observe only the total count for each individual over the period of observation, the Poisson assumptions relate to the (unobserved) underlying process generating the observed counts.

$$Prob[Y_i | t_i] = \frac{e^{-(t_i \lambda)}(t_i \lambda)^{Y_i}}{Y_i!}, \qquad Y_i = 0, 1, 2, ...$$

where  $\lambda$  is the mean occurrence rate per time unit and  $t_i$  is the length of the interval over which  $Y_i$  is observed. A key property of the Poisson distribution is the equality between the mean and variance of the underlying process, commonly designated as  $\lambda_i$ .

Following standard practice, and for expository purposes, we eliminate  $t_i$ from the model by setting it equal to 1, a normalization which does not entail any loss of generality in the case of an equal observation period for all individuals. (We reconsider this assumption in the next section.) If we further assume independence of observations, it follows that the joint distribution<sup>135</sup> is given by

$$L = \prod_{i=1}^{n} \frac{e^{-\lambda_i} (\lambda_i)^{Y_i}}{Y_i!} ,$$

and the log-likelihood is given by

$$\log L = \sum_{i=1}^{n} \left[ y_i \log(\lambda_i) - \lambda_i - \log(y_i!) \right] \quad . \tag{1}$$

This is the common form of the Poisson regression model.

To implement this estimator, it is now only necessary to specify the functional form so that we may introduce covariates into the analysis. Following standard practice, we use the exponential function:

$$E(y_i \mid X) = \lambda_i = \exp(x_i' \beta)$$
(2)

This functional form is appropriate not only because  $\lambda_i$  is restricted to be nonnegative but, as noted by King (1988), because it also explicitly deals with the

<sup>&</sup>lt;sup>135</sup> For a discussion about likelihood theory in the context of the Poisson and the Negative Binomial models, see Gourieux, Monfort and Trognon (1984) or King (1989b).

heteroskedastic nature of event count data. The effect on  $\mathbf{y}_i$  of a one unit change in  $\mathbf{x}_i$  is now  $\beta \lambda_i$ , and a fixed change in  $\mathbf{x}_i$  would have a greater effect on  $\mathbf{y}_i$  the larger is the expected value,  $\lambda_i$ . In other words, the exponential specification ensures that the first migration, moving  $\mathbf{y}_i$  from 0 to 1, takes relatively more effort, or more of a change in  $\mathbf{x}_i$ , than the fifteenth migration which would move  $\mathbf{y}_i$ from 14 to 15. This premise is consistent with (1) a migration process characterized, to a significant degree, by temporary, or return, migration; and (2) DaVanzo (1983)'s empirical finding of a positive relationship between the propensity to move and the number of previous moves.<sup>136</sup>

Substituting equation (2) into equation (1), and after dropping the last term in equation (1) as it does not vary with the parameters to be estimated, the loglikelihood is given by

$$\log L = \sum_{i=1}^{n} \left[ y^{i}(x_{i}^{\prime}\beta) - e^{x_{i}^{\prime}\beta} \right]$$

Implementation of this estimator is now straightforward.



<sup>136</sup> Graphing the exponential function with  $\mathbf{E}(\mathbf{y}_i \mid \mathbf{X}) = \lambda_i$  on the Y-axis and  $\mathbf{x}_i$  on the X-axis, the exponential curve bends upwards and becomes steeper as we move away from the origin and along the X-axis, ensuring that as  $\lambda_i$  increases, a fixed change in  $\mathbf{x}_i$  has a larger effect on  $\mathbf{y}_i$ . In other words, as more migratory moves are undertaken, it becomes relatively easier to make one more migration.

While assuming equal periods of observation proves to be econometrically and expositionally convenient, the age distribution of our sample makes it difficult to sustain the argument, and we reconsider the equal exposure assumption in the next section.

#### 3.1.1.1 Unequal Periods of Observation

In analysing the number of lifetime migrations, we consider each individual's migration exposure to start at the age of 15 and continue until 1988, the time of the survey. Ideally, we would like to have a sample consisting of individuals who have completed their migration career. Instead, we have a sample characterized by censoring in the sense that we have a wide range in age, or the length of exposure to migration. Allowing t to vary over individuals then, it follows that we should control for lifetimes of differing length, as the length of the observation period is likely to influence the expected count. The time independence assumption allows convenient accumulation of the underlying Poisson process, with parameter  $\lambda_i^t$  and taking place over each of the observation periods t+ $\delta$ , to an aggregate Poisson process, with parameter  $\lambda_i = \sum_{t=1}^{T_i} \lambda_i^t$ . This is strictly a restatement of the well known result that the sum of T identical Poisson variables has a Poisson distribution. Given a constant rate,  $\lambda_i^t$ , for each individual, we can simplify to  $\lambda_i = T_i \lambda_i^t$ , where  $T_i$  is the individual's total exposure to migration (age at 1988 less 15, in our case), and  $\lambda_i$  is the expected count over the entire period of exposure. The observed counts are therefore:

$$\lambda_i = T_i e^{\beta_0 + \beta_i x_i} = (age_i - 15)e^{\beta_0 + \beta_i x_i}$$

For estimation purposes,<sup>137</sup> the last equation would take the following form:

$$\lambda_i = e^{\ln(age_i - 15) + \beta_0 + \beta_l x_i}$$

with the coefficient on  $\ln(age_i - 15)$  constrained to 1.0 as the independence assumption dictates that the expected count increases proportionately with the period of exposure.

The need to account for unequal periods of observation makes identification of life cycle and cohort effects problematic, especially given the high level of aggregation of our initial analysis. To control for different periods of exposure, we do include the logarithm of age as an explanatory variable but we do not restrict the coefficient to be 1.0. To control for life-cycle and cohort effects we also include the square and cubic of **ln(age)**, dummy variables for people reaching the age of 15 before 1950, a period of disruption due to the Second World War and its aftermath, and after 1970, when a major change in policy regime occurred, as well as interactions between the **ln(age)** variables and the period dummies.

We interpret the difference between the estimated coefficient on ln(age) and 1.0 to be the cumulative effect of life-cycle and cohort effects. As we have not articulated a full stochastic model, it is not possible (and indeed may not be

$$\ln L(\hat{\beta}, y) = \sum_{i=1}^{n} \left\{ y_i(x_i \hat{\beta}) - t_i e^{x_i \hat{\beta}} \right\}.$$

<sup>&</sup>lt;sup>137</sup> Section 5.8 of King (1988) formally derives the log-likelihood function in the case of different periods of observation at equation 5.20:

possible in any case) to separately identify these effects. Nevertheless, we proceed with inference on the remaining variables on the assumption that we have properly controlled for both life-cycle and cohort effects.

#### 3.1.1.2 Poisson Restriction

A property of the Poisson distribution mentioned above is that the mean of the underlying process,  $\lambda$ , is equal to its variance. This restriction is often violated in empirical applications of the Poisson regression model, resulting either in underdispersion, if the estimated mean turns out to be greater than the variance or, more commonly, overdispersion, if the estimated mean turns out to smaller than the variance. The reason overdispersion is relatively more common, especially in social science applications, derives from the implausibility that  $\lambda$  is completely determined by the covariates on which data has been gathered. In the presence of unobserved systematic differences, and in contrast to a completely determined specification of  $\lambda$ , the prediction error is bound to increase as the unobserved variable affects the observed, but not the estimated, outcome. In the end, the variance of the observed count will be greater than the variance estimated by the Poisson regression model. In fact, the only way for underdispersion to result is for the underlying process to be somehow regulated as would occur with negative duration dependence.

Aside from unobserved individual heterogeneity, examined in the context of time series cross section data by Hausman, Hall and Griliches (1984), other sources of overdispersion include a violation of either the homogeneity or the independence assumption, and the presence of a separate non-Poisson mechanism in the underlying data generation process. Breaching the homogeneity assumption implies some type of unobserved heterogeneity, time-varying or otherwise, and always results in overdispersion, as argued above.

Violating the independence assumption results in overdispersion in the case of positive state duration dependence (the equivalent of an increasing hazard in duration modelling) and underdispersion in the case of negative state duration dependence (the equivalent to a decreasing hazard in duration modelling). In the first instance, the results are again similar to the case of unobserved heterogeneity as the effect of the elapsed duration increases the observed, but not the predicted, outcomes. In the second instance, however, the negative nature of the state dependence constrains the series of observed outcomes while the estimated count remains relatively unaffected. The result is a larger variance in the estimated counts than in the observed counts, or underdispersion.

Finally, neglecting some element of non-Poissonness in the underlying process affects the predictions from the estimated model in a manner similar to unobserved heterogeneity. In this case, we tend to observe an overabundance of specific values. In the presence of some significant level of aversion to migrating, for example, we might observe relatively more zeros than if the process had only a Poisson component. If we neglect to model the non-Poisson process, and as we constrain the estimated mean to be equal to the observed mean by the inclusion of a constant in the regression model, we observe that the estimated outcomes will

be packed closer around the mean<sup>138</sup> than the observed outcomes. That is, the variance in the observed outcomes is greater than the variance in the estimated counts resulting in overdispersion.

In the migration case, the consistent finding<sup>139</sup> that most of the recorded moves are subsequent moves has led to the conclusion that the migration decision is path dependent. The common explanation, formalized by Pessino (1991), is that a learning process initiated by the first move increases subsequent probabilities of a move. Dependence of the rate of migration on previous occurrences constitutes a breach of the independence assumption in the Poisson statistical model and, similar to the case of an omitted variable, leads to overdispersion.

Additionally, the selective nature of the migration process, which is well documented in both the theoretical and the empirical literature,<sup>140</sup> is likely to lead to the presence of at least two processes: one determining the decision to migrate or not migrate, and a second, separate, process determining the frequency of migration. In the extreme, some individuals are likely to never migrate,

<sup>&</sup>lt;sup>138</sup> To compensate for the relative lack of zeros and the relative abundance of outcomes between zero and the mean, the model also takes some mass away from the higher end of the distribution.

<sup>&</sup>lt;sup>139</sup> See references in Morrison (1971) and Vanderkamp (1971).

<sup>&</sup>lt;sup>140</sup> See, for example, Allen (1979) and Robinson and Tomes (1982).

regardless of relative prices and income,<sup>141</sup> while others base their decision to migrate on the (perceived) distribution of prices and income. In the presence of such processes, the base Poisson model, which does not allow for the presence of a separate participation process, should significantly underpredict the number of zeros in the sample data and exhibit overdispersion. Partially to account for this possibility, we discuss the prospect of relaxing the strict assumptions in the base Poisson model in the next section.

## 3.2 Relaxing the Poisson Assumptions

After estimating a base Poisson model and finding evidence of overdispersion, we consider two important generalizations. Although it has become standard practice to proceed to a hypothesis that the data generation process is governed by a less restrictive distribution, commonly the negative binomial, modelling the extra variation by the estimation of explanatory equations may prove effective given our central aim of inference, as noted by Dean (1992). We start by pursuing this modelling approach with our specification of a double hurdle model, also known as the zero altered Poisson model. We then consider the Negative Binomial regression model as well as the zero altered Negative Binomial.

<sup>&</sup>lt;sup>141</sup> Theoretical support for this notion may be provided by the Roy model of self selection and earnings with non-movers characterized as having infinite costs of moving due either to a (perceived) large amount of location specific capital or else a capital constraint. Although the studies on geographical location choice and its effect on wages fall under the general framework of the Roy model, the preceding argument has not yet been specifically layed out in the literature.

# 3.2.1 The Zero Altered Poisson Model

Using the Poisson model to explain the total number of migrations is akin to postulating the following decision process:



Figure 10. Decision Process for Poisson Model

Alternatively, under the Zero Altered Poisson model (ZAP), the decision process follows the scheme:



Figure 11. Decision Process for ZAP Model

The source of the extra-Poisson variation is thus posited to be an excess number of zeros arising out of the separate participation process depicted in Figure 2. A clear indication of the suitability of this approach would be the presence of a substantially greater number of zeros than otherwise predicted by the base Poisson model.

The double hurdle model relaxes the restriction of equal mean and variance by introducing a behaviour splitting variate, **q**<sub>i</sub>, a parameter between 0 and 1. The aim, as expressed by the second figure above, is to capture the fact that survey respondents with zero moves are likely to be of two types: type I would never migrate, regardless of relative prices and income, while type II do base their decisions on relative prices and income, but may simply not have found conditions to be favourable enough to engage in migration. By considering a model which allows zeros to arise either from the binary part of the model or the count part of the model, Greene (1994) allows non-participation to result from either an aversion to participation or as an outcome of a standard corner solution. Following Jones (1989), this is a double hurdle model with independence assumed between the participation and the consumption decision, and without first hurdle dominance.<sup>142</sup> We use the logistic distribution for the participation hurdle.

As specified by Greene (1994), the mean is

$$\mathbf{E}\left(\mathbf{y}_{i}\right)=\left(1-\mathbf{q}_{i}\right)\boldsymbol{\lambda}_{i}$$

<sup>&</sup>lt;sup>142</sup> First hurdle dominance is consistent with an interpretation of zero moves as a discrete choice rather than a marginal adjustment. In this case, (potential) movers, i.e. those passing the first hurdle, would not have the choice of zero moves, but would instead always choose one or more moves. This would leave the zero moves choice to be credited only to non-movers, i.e. those not passing the first hurdle.
and the variance is

$$\operatorname{Var}(\mathbf{y}_{i}) = \lambda_{i} (1 - q_{i}) [1 + \lambda_{i} q_{i}] .$$

A smaller value for  $\mathbf{q}_i$  corresponds to a lower percentage of Type I agents, as indicated by the presence of fewer zeros in the data. To see that the splitting mechanism does indeed introduce the possibility of overdispersion, consider

$$\frac{Var(y_i)}{E(y_i)} = 1 + \lambda_i q_i = 1 + \left[\frac{q_i}{1 - q_i}\right] E(y_i)$$

A larger estimate for  $\mathbf{q}_i$ , indicating an increase in the excess number of zeros, leads to an increase in the  $\frac{q_i}{1-q_i}$  term and a greater level of overdispersion. This serves to explain how excess zeros can show up as overdispersion.

We have now introduced the Poisson regression model and dealt with overdispersion by extending the basic model to allow for the possibility of one likely source of extra-Poisson variation. If our main interest is inference concerning regression parameters, and if the situation is one in which overdispersion routinely occurs, then accounting for this via estimating equations is likely to be a viable approach. It is also possible to pursue a different approach and specify an alternative distribution that may be more flexible. This approach would become more appropriate if the estimation of tail probabilities is of particular importance. The usual candidate is the Negative Binomial distribution. 3.2.2 The Negative Binomial Model

The advantage of positing that the data generation process is governed by the Negative Binomial instead of the Poisson distribution is that we relax the restriction that the mean of the process be equal to the variance. The Poisson is a limiting form for the Negative Binomial, a distribution which has become the conventional choice in likely cases of overdispersion, especially if the independence assumption is unlikely to be met.<sup>143</sup>

The Negative Binomial also arises by assuming that  $\lambda$  varies according to a Gamma distribution, a flexible distribution that produces compound Poisson distributions that are computationally tractable. The Gamma distribution has mean and variance, respectively,

 $E(y) = \phi$  and  $Var(y) = \phi(\sigma^2 - 1)$ ,

where  $\phi$  is the Poisson equivalent to the mean rate of occurrence and  $\sigma^2$  is a shape parameter. As  $\sigma^2$  approaches 1, the Gamma distribution collapses to a spike over the mean of the process, leaving  $\lambda$  constant, in which case we recover the base Poisson model.<sup>144</sup>

In deriving the Negative Binomial distribution, then, we start with the Poisson model and drop the homogeneity assumption by letting  $\lambda$  vary within an observation according to the Gamma distribution. Using compounding distributions, we eventually arrive at the Negative Binomial distribution

<sup>&</sup>lt;sup>143</sup> See, for example, Johnston, Kotz and Kemp (1992, pp. 223-4). Essentially, the Negative Binomial distribution has more mass in the tails than the Poisson distribution.

<sup>&</sup>lt;sup>144</sup> Although the usual derivation of the Negative Binomial distribution, the Greenwood-Yule version, which follows from relaxing the constant rate assumption, it is also possible to derive this compound distribution by relaxing the independence assumption. Johnson et al. (1992) refer to Foster (1952) and Thompson (1954), two derivations which follow from a stochastic process characterized by some form of state dependence.

### $f_{nb}(y_{_{1}}|\lambda,\sigma^{2})$

with  $\lambda > 0$  and  $\sigma^2 > 1$  for non-trivial cases. The parameter  $\lambda$  is still the expected number of events. The restriction of  $\sigma^2 > 1$  means that (1) we escape the equal mean and variance restriction as well as the independence assumption, from which the restriction flows;<sup>145</sup> and (2) the Negative Binomial model allows only over-dispersion, not underdispersion.

The extra parameter in the Negative Binomial distribution, commonly labelled as  $\alpha$ , is referred to as the overdispersion parameter and is also most commonly interpreted as representing unobserved heterogeneity.<sup>146</sup> A test of significance on  $\alpha$  is then a natural test of overdispersion. In practice, the estimates arising out of this model tend to differ very little from the Poisson estimates, although the standard errors are usually larger, indicating that the Poisson estimates are still consistent but inefficient, as noted by King (1988).

We have now introduced two distinct ways of dealing with extra-Poisson variation. The double hurdle Poisson model attempts to model specific non-Poisson mechanisms included in the data generation process. Conversely, the Negative Binomial model, attempts to capture any overdispersion present in the

<sup>&</sup>lt;sup>145</sup> See King (1989a) and Johnson et al. (1992).

<sup>&</sup>lt;sup>146</sup> The size of the parameter  $\sigma^2$  is related to the structure of the number of events within each observation. Ceteris paribus,  $\sigma^2$  will be larger if these events are either positively related or heterogeneous, that is, in the presence of overdispersion. Conversely, a smaller value for  $\sigma^2$  means that the Negative Binomial distribution is closer to the Poisson. This parameter is then the counterpart to the overdispersion term in the hurdle Poisson model.

data without specifying the source or the form. Given the likelihood of overdispersion arising from different sources, combining a separate process for the generation of zeros with a more flexible distribution is more likely to be empirically appealing.

#### 3.2.3 The Zero Altered Negative Binomial Model

As noted by Greene (1994), the advantage of the Zero Altered Negative Binomial model (ZANBR) is that it allows a distinction to be made between the overdispersion arising from a non-Poisson mechanism and the overdispersion arising from individual heterogeneity. In this model, the mean is the same as the ZAP model, but the variance becomes

$$\operatorname{Var}(\mathbf{y}_{i}) = (1 - q_{i}) \lambda_{i} [1 + (q_{i} + \alpha)\lambda_{i}] .$$

To see the differential impact of these two sources of overdispersion, consider

$$\frac{Var(y_i)}{E(y_i)} = 1 + \lambda_i q_i = 1 + \left[\frac{q_i + \alpha}{1 - q_i}\right] E(y_i)$$

Notice that the introduction of the separate mechanism into the Negative Binomial model has a cumulative effect on the overdispersion term. In other words, the term in brackets is greater than  $\alpha$  for all positive  $q_i$ .

It is also possible to use covariates to parameterize the separate participation process for both the ZAP and the ZANBR models. This would allow us to test whether some factors affect the participation and the frequency of migration differently. For example, it might be the case that higher education, noted in the literature to be associated with higher migration rates, leads to a higher level of participation but, perhaps because of better information and an enhanced ability to find a good match quickly, it reduces the number of moves over a lifetime. To distinguish among competing specifications, we designate the models holding constant the extra zero altered parameter,  $\gamma$ , as ZAP(g) and ZANBR(g) while the models which parameterize  $\gamma$  with covariates are labelled as ZAP and ZANBR.

#### 3.3 Specification Testing

Given the different models to be estimated it is essential that we be able to compare them and assess the relevance of more complicated specifications. Specifically, we must evaluate the significance of the overdispersion arising from each of two processes, individual heterogeneity and a separate participation process, as well as the significance of the additional parameterization in the ZAP and ZANBR specifications. First, we check for the presence of overdispersion, and choose between the base Negative Binomial and Poisson specifications, with two regression based tests proposed by Cameron and Trivedi (1990) and a more general conditional moments test proposed by Pagan and Vella (1989). Under the regression tests, the null hypothesis is

$$H_0$$
: var(y) =  $\lambda_1$ 

and the alternative hypothesis is

H<sub>1</sub>: var(y<sub>1</sub>) = 
$$\lambda_1 + \alpha g(\lambda_1)$$

where  $g(\lambda_i) = \lambda_i$  and  $g(\lambda_i) = \lambda_i^2$  are suggested by Cameron and Trivedi as the ideal test specifications. As mentioned by Greene (1993) on page 679, the test is

carried out by testing the significance of the single coefficient in the ordinary least squares regression of

$$z_{i} = \frac{\left[(y_{i} - \mu_{i}^{2})^{2} - y_{i}\right]}{\mu_{i}\sqrt{2}}$$

on a constant term, for  $g(\lambda_i) = \lambda_i$ , or  $\lambda_i$  without a constant term, for  $g(\lambda_i) = \lambda_i^2$ .

Although convenient, the Cameron and Trivedi tests require the form of overdispersion serving as the alternative hypothesis to be specified. It is possible to formulate a more general test by using conditional moment (CM) restrictions. One such test is proposed by Pagan and Vella (1989) and is also discussed by Greene (1993, section 22.3.4d). For the Poisson model, the moment conditions to be tested are

$$E\left[z_{i}\left[(y_{i}-\lambda_{i})^{2}-\lambda_{i}\right]\right]=0$$

where the  $z_i$  vector does not include the constant term in  $x_i$ . For this test, the null hypothesis is again  $Var[y_i] = \lambda_i$  but the form of overdispersion is not specified.

Additional tests for the relevance of overdispersion in the base Poisson model are available. If the underlying Poisson process is characterized by significant overdispersion, the  $\alpha$  term in the Negative Binomial model should be significantly different from zero. Alternatively, as the Poisson model follows from the Negative Binomial model by restricting  $\alpha$  to zero, the standard likelihood ratio (LR) test is applicable. This test is available to compare common specifications of Negative Binomial and Poisson models, such as the base models, ZANBR to ZAP and ZANBR(g) to ZAP(g). Additionally, we compare the predicted outcome probabilities for the various models proposed to the sample probabilities as a measure of fit.

Second, we compare the zero altered to the unaltered models in order to assess the relevance of the separate participation decision, or the clustering at zero in the sample data, as one possible source of any overdispersion initially found by the initial set of tests. Unfortunately, the ZANBR and the ZAP models do not nest the base Negative Binomial and the Poisson models and, consequently, the respective zero altered and unaltered log-likelihoods are not directly comparable. This raises the need to consider a test statistic for non-nested models, and we follow Greene (1994)'s suggestion in using a test originally proposed by Vuong (1989). This moment-based test requires estimation of both models as well as the computation of the sample of predicted probabilities, and essentially computes a t-statistic for a zero mean of the variable

$$m_{i} = \log \left[ \frac{Prob(y_{i} \mid model \# 1)}{Prob(y_{i} \mid model \# 2)} \right]$$

Asymptotically, the test statistic  $V = \frac{\sqrt{Nm}}{s_m}$ , where N is the sample size and s is the standard deviation of the variable  $\mathbf{m}_i$ , has a standard Normal distribution. As noted by Vuong (1989) and Greene (1994), the test statistic is directional in that large positive values favour model #1 while large negative values favour model #2. An absolute value less than the critical value favours neither model.

Third, we compare the more generally parameterized zero altered versions of each of the Negative Binomial and Poisson specifications, ZAP and ZANBR,

to the single parameter ZAP(g) and ZANBR(g), respectively, in order to evaluate the significance of parameterizing the extra  $\gamma$  parameter with a set of covariates instead of a constant term. Unfortunately, ZAP does not nest ZAP(g) and, by the same token, ZANBR does not nest ZANBR(g), unless the participation variables are the same as the frequency variables, in which case, and as noted by Greene (1994), the restriction is that the participation parameters all be the same proportion,  $\gamma$ , of the frequency parameters. At present, we assess the relevance of the added parameterization by noting whether some of the participation parameters have a significant t-statistic.

Finally, we need to assess the relevance of the overdispersion left in the data once we have accounted for the separate participation process. Conveniently, the ZANBR specification does nest the ZAP specification, as ZANBR(g) nests ZAP(g), given the single restriction that the negative binomial parameter,  $\alpha$ , equal 0, and we can then make use of the LR statistic for this test.

#### 4. Econometric Model

We analyse the reduced form determinants of lifetime internal migration. The dependent variable is the number of migrations over an individual's lifetime. The regressors are standard demographic, geographic and family background variables at age 15, the start of the exposure to migration, as well as the number of years of schooling at the survey date. A complete list of variables along with descriptive statistics are provided in Appendix A.

#### 4.1 Theoretical Considerations

Following the approach pioneered by Sjaastad (1962), the theoretical framework universally employed in empirical studies of migration claims that migration is an individual investment decision in human capital. The main implications of this theory of migration are well known and have been outlined elsewhere, including Pessino (1991). For our application, the human capital theory of migration yields two empirically verifiable implications. First, the theory predicts a negative relationship between participation in migration and age. Younger workers have a longer horizon over which to enjoy the longer-term (net) benefits of migration and we should then expect they should be relatively more willing to incur the immediate costs. This simple relationship, while generally validated by empirical studies, is not able to account for the life cycle type effects documented in the previous data chapter and also by Schultz (1982).

One explanation proposed by Schultz (1982) for the relatively high incidence of migration among younger workers essentially boils down to the presence of a separate participation process or what Schultz refers to as "cohort selectivity effects". Within any birth cohort, those whose characteristics and skills are best suited for the uncertainties associated with migration will, according to the human capital theory of migration, leave and try their luck at a relatively young age. Therefore, over time in a cohort, the composition of the origin population changes and migration cannot be regarded as being generated by a constant stochastic process. It is not that (some of) the older non-migrants

choose not to migrate because they are no longer younger; it is, instead, that they choose not to migrate because of their preferences, prices or constraints. To capture this type of process accurately, Schultz mentions that longitudinal data, as available in MFLS-2, and a stochastic specification of longitudinal migration models, such as the stochastic point processes we propose, are needed. In the zero altered specifications we estimate, a coefficient on **ln(age)** significantly different from one would suggest the presence of life cycle and period effects, net of the impact of the participation process.

Second, human capital theory predicts that education should increase mobility. Several explanations for this effect have been widely circulated in the literature including the fact that more educated workers participate in a national, rather than a regional, labour market; schooling enhances a worker's skills, a proxy for the ability to deal with a changing environment, and also enhances a worker's information about opportunities farther away; and, given the larger investment in productive skills, educated workers tend to be more responsive to regional differences in economic incentives. Given the dearth of previous work on longitudinal migration models, the literature is silent with respect to the possibility that education affects the participation and frequency of migration differently. The general version of the zero-altered models, ZAP and ZANBR, will provide some empirical evidence on this important question as we parameterize the separate participation process with a set of covariates including years of education. If we find that education enhances participation in migration among the general population while reducing the frequency of migration among migrants, it is likely that the primary role of education is an indicator of the quality of information possessed by potential migrants. In this case, we would hypothesize that while the enhanced quality of information reduces the uncertainties of migration, it also allows a good match to be found more quickly.

The migration literature provides a further explanation for individual heterogeneity with respect to migration rates, the idea of physical or psychic location specific capital. Although previously hinted at in previous papers, DaVanzo (1983) provides the first empirical evidence that the documented positive relationship between the number of previous moves and propensity to move may be related to factors such as home ownership. We incorporate the idea of location specific capital in two ways. First, we include a dummy variable which equals one if the respondent's family moved between the respondent being born and reaching the age of fifteen. In line with DaVanzo's reasoning, families with more frequent moves are likely to develop less attachment to a specific area and family members are, therefore, more likely to be mobile. It is also possible that people with more location specific capital are more likely to return home, if they do migrate. We might then find location specific capital having opposite effects in the participation equation and in the frequency equation. Second, we distinguish people who have inherited physical assets from those who have not. The claim is that if physical capital is indeed location specific, that is, in the presence of less than perfect capital markets, heirs may be induced to migrate

less often than otherwise. A related question centres on the potential difference between rural and urban inheritances. With generally thinner markets in the rural areas, it is more likely that rural inheritances fall more squarely under the rubric of location specific capital, so that we would expect rural inheritances to restrain migration more than urban inheritances.

#### 4.2 Parameterizing the Participation Decision

While the migration literature, especially the human capital approach, does propose some variables to explain the propensity to migrate, there is a paucity of guidance on variables which might affect only one of the participation or the frequency of migration. Econometrically, this would leaves us without any exclusion restrictions on which to base parameter identification, forcing us instead to rely on functional form. As a result, we pursue an iterative modelling cycle in parameterizing the participation process in the zero-altered models. The first stage involves fitting a Probit model to arrive at a set of significant covariates for the participation decision, while the second stage involves estimation of the stipulated zero altered model. Although less than optimal in terms of probability theory, this type of specification search should allow us to both generate the necessary exclusion restrictions as well as make an initial contribution to the identification of variables which affect the participation and the frequency of migration differently.<sup>147</sup> Among the hypotheses of interest would be the potentially different effects of education and family inheritance.

#### 4.3 Estimation Results

This section presents estimates for the models previously discussed. We first report estimates for the base Poisson model and test for the presence of overdispersion. Subsequently, we consider the Negative Binomial specification as well as the zero altered versions of the Poisson and the Negative Binomial models. The results for the base Poisson model are shown in Appendix B. Test results also shown in appendix B uniformly suggest the presence of overdispersion. Both formulations of the Cameron and Trivedi (1990) (CT) tests have significant t-ratios while the conditional moments test (PV) has a significant  $\chi^2$  statistic. The histogram shown in Appendix C shows that while the sample data contain 665 zero outcomes, the Poisson model predicts just 408 such outcomes. As mentioned in section 3.1.1.2, such a result implies that the overdispersion suggested by both the CT and PV tests may be at least partly due to an excess number of zeros, suggesting a separate process for participation in

<sup>&</sup>lt;sup>147</sup> Technically, the lack of availability of the necessary extraneous information either from theory or previous studies leaves us with only two options: making some (possibly incorrect) assumption about the correct set of independent variables, or (possibly) incurring some level of pre-test bias. Although the selection of pre-testing could possibly be justified by the search for the increased efficiency to be garnered from incorporating extra information in our estimation procedure, our choice of modeling strategy was made not just because of the opportunity to identify some important empirical relationships but also because of the ready availability of a second sample in MFLS-2 which will be used as additional evidence in future work.

migration. We explore this possibility after considering the standard alternative to the Poisson model, the Negative Binomial model.

The more general specification based on the Negative Binomial distribution introduces an additional source of variation by allowing  $\lambda$  to vary according to a Gamma distribution. The estimates for this model are shown in Appendix D and the extra variable, labelled as Alpha, is highly significant, in line with our earlier finding of overdispersion. Similarly, the LR statistic comparing the log-likelihood from the Negative Binomial and the Poisson models categorically rejects the latter model in favour of the more general approach. Given our overriding concern with inference, it is encouraging to note that the estimates change very little between these two models, although the standard errors are (predictably) larger under the negative binomial specification. This is an indication that although the Poisson estimates are inefficient, they are still consistent. The number of predicted zeros is 607, a substantial improvement over the base Poisson model. Next, we pursue the zero altered approach and the explanation of a separate participation process.

#### 4.3.1 Parameterizing the Participation Decision

A zero altered model attempts to augment the probability of a zero outcome by modelling a separate participation process. To identify variables significant for the migration equation, Appendix E presents the maximum likelihood estimates for a Probit model. In general, only variables with significant t-ratios are included in the participation equation for the zero-altered

models. This gives us some exclusion restrictions and helps to ensure we get parameter identification from other than the functional form. However, we do include an indicator variable for individuals whose first job was by age fifteen, WRK YONG, in both equations because of the interest in testing the possibly differential impacts of an early job on the participation and the frequency of migration. Given that individuals in our sample have unequal exposure periods to migration, it follows that both the frequency and the participation of migration should be a function of the length of exposure. As a result, we also include the natural logarithm of the individual's age less fifteen, LNAGE, and cohort dummies for those individuals reaching the age of fifteen before 1950, BEFOR 50, as well as those reaching the age of fifteen after 1970, AFTER 70, along with interaction terms in both equations.<sup>148</sup> The parent education variables (parents with only primary education, PAR PRED, is significant in the Probit) are included in both the participation and the frequency equation (disaggregated into mother and father in the latter) while the parent occupation variables are included only in the frequency equation. Interaction terms between the respondent's area of origin and the presence of younger sisters, R YNGSI and U YNGSI are included only in the participation equation while an indicator

<sup>&</sup>lt;sup>148</sup> The possibility of life cycle effects suggests that higher powers, normally a square and cubic, of LNAGE should be included in both the participation and the frequency equations. As mentioned earlier, we did try including such terms in our specifications, but in all cases, there were clear indications of overparameterization as both the coefficient estimates and the standard errors for the constant term, the cohort dummies as well as the interaction terms became very large.

variable picking up individuals with younger brothers, AN YNGBR, is included only in the frequency equation.

#### 4.3.2 Zero Altered Models

Estimation results for both ZAP and ZANBR are presented in Appendix F. In the case of the participation equation, the models estimate the probability of not moving so that a negative sign actually indicates an increase in the incidence of migration. For both models, the participation equation does yield a substantial proportion of significant variables with some changes in the numerical estimates, particularly BOARD and U MV 15, although there are no sign changes. The diagnostic test allowing a comparison to the respective unaltered models, the Vuong statistic, favours the zero altered model in each case. The predicted number of zeros rises significantly to 663 for the ZAP and 662 for the ZANBR. The Negative Binomial model, by contrast, predicts only 607 zeros. The chart in appendix G confirms that, on the basis of in-sample predictions, the zero altered models appear to fit the data better than either the base Negative Binomial or Poisson models.

The ZANBR model allows a simultaneous assessment of the significance of the separate participation process and other sources of extra-Poisson variation. The extra variable, again labelled as Alpha, is highly significant indicating that some level of overdispersion is left even after we model the participation decision and favouring the more general specification. On the basis of the LR test, ZANBR is to be preferred to ZAP as the test statistic of  $2 \times (2,392.228$ - 2,349.304) = 85.776 is significant at the 1% level. The chart in appendix G, though, reveals that over the most frequent outcomes, zero, one and two, the ZAP outperforms the ZANBR for in-sample predictions. Consistent with its more flexible distribution, the ZANBR appears to do a better job than the ZAP in the right hand tail of the distribution, but at the expense of the rest of the distribution, leading us to prefer the relatively parsimonious specification of the ZAP model.

To assess the relevance of parameterizing the zero altered parameter with covariates, consider not only the number of significant variables in the participation equations shown in appendix F, but also the chart shown in appendix H. Clearly, the added generality sharpens the performances of the zero altered model. In summary, it appears that the model best suited to explain lifetime migration is a double hurdle Poisson model, or ZAP, with the participation equation parameterized by covariates. Given the overriding concern with inference, however, it is again noteworthy that in choosing between the ZAP and the ZANBR model, the estimated coefficients in the frequency equation change very little, while the two changes noted above in the participation equation do not involve a change of sign.

#### 4.3.3 Zero Altered Poisson Model: The Participation Equation

The results for the participation equation reveal that the incidence of migration is positively related to the respondent's level of education, YRSEDUC, and residence status while in school, BOARD. This confirms the universal finding that migrants tend to be more educated than non-migrants. Participation

in migration is also affected by the level of parental education. Compared to the reference group of parents with no education, parents with only primary education, PAR PRED, positively increase the probability of their children migrating. Interestingly, the same effect is not evident for children of parents with secondary education, PAR SEC. Given the often noted correlation of education with income level and family assets, it may be that children of highly educated parents have enhanced earning opportunities locally and are consequently less inclined to migrate.

The finding that parent's education complements the own education effect on migration is noteworthy. Parent's education is part of the family endowment of the next generation. If, in general, it is not observed (data on parents are often not collected, or if collected, not used by analysts) and is positively correlated with an observed variable, namely education level of the respondent whose behaviour is being modelled, then quantitative analyses that do not control for the endowment effect will overestimate the effects of own education. As detailed previously, there is abundant evidence, confirmed by our results, that own education is strongly associated with migration. However, our results also indicate that part of the education effect up to now unambiguously assigned to own education may reflect the indirect influence of family endowment.

Geography and ethnic background have been previously highlighted as playing key roles in migration in Malaysia. Consistent with these earlier accounts, we find that urban Chinese tend to engage in migration significantly less than other groups. Previous studies, such as Mazumdar (1981), have found that the Chinese, particularly those in an urban area, are at or near the top of the earnings distribution in Malaysia, and further, that Chinese tend to dominate commerce. For purposes of migration, this should translate into relatively more local opportunities and (perhaps) more location specific capital and, consequently, less incentive to migrate.

We also find strong support for the hypothesis that location specific capital should inhibit migration. Specifically, we find that if the family moves around between the respondent's birth and fifteenth birthday, the respondent is more likely to engage in migration. This holds for both rural and urban dwellers and highlights the importance of gathering complete migration history data, as in MFLS-2. Additionally, rural respondents inheriting either land or house from the parents are significantly less likely to engage in migration although no such relationship was found for urban respondents. We elaborate on this point in the next section.

With respect to the age responsiveness of the incidence of migration, our results indicate that, for the base group, the probability to migrate is unrelated to the respondent's age. Although contrary to the standard cross-section based findings of a negative relationship between age and the probability to migrate, this result is consistent with the combination of (1) the human capital theory of migration, and the prediction that a migrant-type will tend to migrate earlier, rather than later, in life; and (2) the type of data we employ which involves, for

the most part, workers who are beyond the recognized threshold<sup>149</sup> age for a first move of 18-20. Given that the coefficient on LNAGE essentially estimates the relationship between the probability of having ever migrated and age among a relatively old sample, we would naturally expect that other characteristics, those determining whether a worker is a mover or a stayer, but not age, are significantly related to the probability of being a migrant. For the post 1970 group, however, we do find a significantly positive increase in the age responsiveness of migration as the coefficient on the interaction term, LNAG 70, is significantly different from zero. This is likely a result of the age censoring in this group as some of the post 1970 people have in fact still to come to the standard threshold age for a first move.

Finally, our results do not support the notion that the major changes in policy in 1970 have significantly increased the incentives for labour mobility. We find that our cohort dummy for people reaching the age of 15 on or after 1970, AFTER 70, is significantly positive or that, ceteris paribus, the probability of migration is significantly lower for this group than the base group. As the 1970 changes targeted Malays, we did split the post 1970 group into Malays and others but found the coefficient estimates and standard errors were virtually identical. This may be an indication that the major effect of policy changes in the decade of the 1970s came through the export promotion and industrialization policies, and the associated rate of increase in economic activity, rather than the affirmative

<sup>&</sup>lt;sup>149</sup> See Schultz (1982) and the accompanying data chapter.

action hiring programs begun in 1970 in the public sector. We also find no significant differences for the pre-1950 group either in the mean probability of migrating or in the age responsiveness of the incidence of migration.

#### 4.3.4 Zero Altered Poisson Model: The Frequency Equation

The results for the frequency equation confirm that the level of education, quite apart from increasing the probability of migration, also increases the frequency of migration. This is a novel result given the lack of previous work on the determinants of lifetime migration. The positive impact rules out the possibility that more educated workers, although more likely to migrate, might engage in fewer moves as a result of better information.

Interestingly, the WRK YONG coefficient indicates that, ceteris paribus, there are two types of workers who tend to be recurrent movers, those with a higher level of education and those who start their working career early. The result on WRK YONG is also noteworthy because in bivariate analysis the effect on the number of migrations is negative as reported in Table J-1 on page 206, 287.

Similar to the results in the participation equation, the number of migration moves is positively related not just to the level of own education, but also to the level of parental education. The significance of the coefficient on parents' education suggests the presence of positive inter-generational effects, even after controlling for the respondent's level of education, again highlighting that part of the effect commonly attributed to own education is likely due to

parent's education. It's also noteworthy that the father's level of education appears relatively more important than the mother's, a ranking which is reversed in bivariate analysis.

The coefficient on LNAGE is significantly less than 1.0 indicating that the rate of migration falls with age. This is consistent with the predictions of human capital theory which suggests that older people, given a shorter horizon over which to defer the initial costs and enjoy the longer-term benefits, have smaller gains from moving; and is also consistent with the negative relationship between age and the probability to migrate found in cross-section studies and noted by Schultz (1982). We do note, however, that the decrease in the rate of migration with respect to age is much less pronounced for the post 1970 group of workers which, again, is indicative of the positive effect of the changes occurring in the 1970s. The explanation for the negative coefficient on the post 1970 dummy, AFTER 70, is again likely to revolve around the issue of age censoring.

The impact of ethnic and geographical considerations alluded to in the initial data analysis is by and large not supported by the insignificance of the general area dummy variables nor by the fact that among the interaction terms of geography and ethnicity, only urban Chinese differ markedly from the reference group of urban Malays. The results do, however, confirm that, ceteris paribus, urban Chinese are the least mobile group in Malaysia, both in terms of participation rates and also in terms of frequency of moves.

Interestingly, the effects of family inheritance differ by area of residence. For those in a rural area, inheriting land or house tends to have a negative impact on the total number of migrations undertaken (RUR LORH is significantly negative). However, the opposite holds for those in an urban area as URB LORH is significantly positive. The implication is that the notion of location-specific capital works in the rural areas but not in the urban areas. This is consistent with the standard arguments in the development literature that (1) capital asset markets in developing countries, particularly those in rural areas, are thin; or (2) capital assets in rural areas are more likely to be productive-type assets with a substantial amount of plot-specific experience translating into an inability to sell such assets for a market value close to the potential returns of using them personally.

The impact of family moves prior to the respondent reaching the age of 15 also differs by area of residence, with no effect estimated for rural dwellers, R MV15, but a negative effect estimated for urban dwellers, U MV15, implying that early moves by urban families tend to actually dampen the number of moves by the respondent. Similar to WRK YONG, this result is also noteworthy because cross tabulations imply that U MV15 would have the opposite effect, as reported in Table J-2 on page 206, 287.

In summary, we have found evidence in favour of a separate participation process in migration. This is consistent with the large literature on the selectivity of migration and suggests the use of a double hurdle model. We estimated a double hurdle-type Poisson model which allows for zeros at each stage. Notably, we found that own education, as well as parental education, enhances both the probability, and the frequency, of migration. Our finding on the intergenerational effects of education on the migration process appears to be unique in the literature. We also found that while family inheritance in rural areas dampens both the probability and the frequency of migration, family inheritance in urban areas has no effect on the probability of migration, but actually increases the frequency of migration. Early family moves tend to decrease the incidence of migration in both rural and urban areas, but while they decrease the frequency of moves in urban areas, they have no apparent impact on the frequency of moves in rural areas. Consistent with the standard descriptions of the income distribution in Malaysia, we found that urban Chinese tend to, ceteris paribus, be less mobile, with a lower level of both participation and frequency, than other regionally disaggregated ethnic groups. Disaggregation of the sample into subsamples of (1) natives of Malaysia, (2) urban residents, and (3) rural Malays revealed results broadly consistent with the entire sample.

#### 5. Further Research

This exploratory work has served to establish both migration patterns and a set of factors which help to determine an individual's migration career. As further evidence of the variables identified as separately affecting the participation and the frequency decisions, the second sample available in MFLS-2 should be used to estimate the zero altered models proposed by this study. To specifically account for the importance of return migration as well as the fact that migration flows are not unidirectional would require data on district characteristics, now available for MFLS-2, and a multinomial Logit estimator. In this way, we would be able to distinguish between the migration process of urban and rural dwellers as well as the migration process of individuals who prefer urban or rural locations as their destinations.

Further work should also recognize that migration decisions are but one facet of the typical individual's life trajectory. Bartel (1979) proposes that a second life domain which is likely to be closely related to migration decisions and patterns is the individual's occupational career. We should then expect to find that work experience and other occupational variables, such as the number of occupation changes, are significant covariates in a (sequential) migration model. This would suggest a simultaneous equations approach would be appropriate and we could set up a SURE type Poisson estimator, along the lines proposed by King (1989a). It is also clear that some of the migration decisions are taken in the context of the family. The availability of female migration history as well as marriage and fertility history should allow some formulation of family migration functions.

To consider the possibility that the timing, as well as the occurrence, of events affects future behaviour we could set up either a panel type estimator along the lines of Hausman, Hall and Griliches (1984), or else estimate a hazard rate model. Either of these approaches would effectively increase the information set available to the decision maker by incorporating time-varying covariates.

It is well known that the specification search we conducted in parameterizing the zero altered process has deleterious effects on the quality of inference. One possible solution, is to follow Veall (1992) and bootstrap the process of model selection, or else, to make use of the additional sample (the New Sample) available in MFLS-2, perhaps for out of sample forecast evaluation.

Normally, it is not possible to distinguish the cause of overdispersion. However, if  $\lambda$  depends on the number of previous events (a violation of the independence assumption), we can use the expected utility framework to formulate a test for these departures from the Poisson assumptions. First, we posit that each possible choice of the number of migrations for an individual is an alternative with its own expected utility level. Second, we note the direct correspondence between assuming that  $\lambda$  is independent of previous events and assuming that the Independence of Irrelevant Alternatives (IIA) holds between the different alternatives. For example, the choice between the utility streams associated with one and two moves should be independent of the presence, or absence, of the stream associated with making zero moves. Finally, we formally test for IIA, and hence for the associated compliance with the Poisson assumptions, by estimating a sequence of multinomial and nested Logit models.<sup>150</sup>

<sup>&</sup>lt;sup>150</sup> Two notes of caution should be sounded. First, the power of the test is undetermined, and second, the test does not distinguish between a departure from (continued...)

#### 6. Summary

This paper pursues an event count approach in modelling lifetime migration in Malaysia. Of particular importance is our finding that migration should be modelled as a two step decision. First, the individual decides whether to migrate or not; this is the participation or incidence decision. Subsequently, the individual decides how many moves should be made; this is the frequency decision. Also noteworthy is that the complete migration history available in MFLS-2 alleviates the well known initial conditions problem associated with cross-section estimation. With respect to estimation results, our findings of positive intergenerational effects of parent's education on both participation and frequency, positive effects of own education on frequency and negative effects of location specific capital for rural workers seem most noteworthy.

Initially, we estimate a Poisson model and find overwhelming evidence that the underlying process is characterized by overdispersion with a clear underestimation of the number of zeros present in the data, 408 predicted versus 665 actual. We then turn to a standard extension, the Negative Binomial model, which although not identifying the source of overdispersion, adjusts for its presence by the introduction of a stochastic term. The estimated coefficients are found to change very little, which is encouraging given our emphasis on inference, and the overdispersion parameter is found to be statistically significant.

<sup>(...</sup>continued)

the homogeneity assumption and a departure from the independence assumption.

The added variation also serves to ameliorate the underestimation of zeros, with the predicted number rising to 607.

Next, we concentrate on the presence of extra zeros as a source of overdispersion and estimate a double hurdle model which allows zeros as the result of both the standard corner solution and a separate discrete choice. Empirically, this is a two step model with the first process having a binary outcome (the participation decision) in modelling the possibility that some agents are unlikely to ever participate in the event. The second process generates the count variable, and includes the possibility that potential participants may consciously choose zero to be their optimal level of participation. A natural application would appear to be the migration process, where some agents might have a non-trivial level of aversion to migrating while other individuals base their decision not to migrate on more standard economic variables like relative prices and income. We find ample support for our parameterization of both equations and conclude that migration should be modelled as a two stage decision. In terms of economic theory, this implies a non-random selection mechanism consistent with the Roy model of selfselection.

### APPENDIX A

268

-

Regression Variables:	
LHS	
TOT MIG = Total number of	f inter-district migrations reported by respondent
RHS	
(1) Demographics:	
AG15 RUR	0/1 = 1 if resided in a rural area at age 15
AFTER 70	0/1 = 1 if reached age 15 after 1970 (born after 1970)
BEFOR 50	0/1 = 1 if reached age 15 before 1950 (born before 1950)
LNAGE	natural log of respondent's age at 1988 minus 15
LNAG 50, LNAG 70	LNAGE × BEFOR 50, LNAGE × AFTER 70
YRSEDUC	total years of education
BOARD	0/1 = 1 if room/board during secondary school or college
	(versus living with relatives, non-relatives or other)
SOUTHERN	0/1 = 1 if residing in the southern region at age 15
WESTERN	0/1 = 1 if residing in the western region at age 15
FOREIGN	0/1 = 1 if residing in a foreign country at age 15
(omitted category is	NORTHERN)
URB CHIN, URB IND	0/1 = 1 if urban and Chinese, Indian
RUR MAL, RUR CHIN, RU (omitted category is	JR IND $0/1 = 1$ if rural and Malay, Chinese, Indian urban Malay)
(	
(2) Family Background:	
AN YNGBR, AN YNGSI	0/1 = 1 if any younger brothers, sister
R YNGSI, U YNGSI	0/1 AN YNGSI×AG15 RUR,AN YNGSI×(1-AG15 RUR)
PAR LORH	0/1 = 1 if inherited land or house from mother or father
RUR LORH, URB LORH	0/1 PAR LORH×AG15 RUR,PAR LORH×(1-AG15 RUR)
F SECED, M SECED	0/1 = 1 if father's/mother's highest education is secondary
PAR PRED	0/1 = 1 if father's/mother's highest education is primary
(omitted: father/mo	ther no education)
FATH WHI, MOTH WHI	0/1 = 1 if father's/mother's occupation while respondent
CATU DDD MOTU DDD	was age 3-15 was write collar 0/1 = 1 if fother's mother's accuration while more on test
FAIR PKD, MUTH PKD	was age 5-15 was production
(omitted: father/mo	ther died before the age of 5 and father/mother in agriculture)

(2) Family	/ Background	(cont'd):
------------	--------------	-----------

.....

MV B 15	0/1 = 1 if moved between birth and age 15
R MV15	0/1 = 1 if AG15 RUR=1 and MV B 15=1
U MV15	0/1 = 1 if AG15 RUR=0 and MV B 15=1
(3) Other:	
WRK YONG	0/1 = 1 if began first job at age 15
DIST NET	District net-migration rate (1970-80) for residence at age 15

### **Descriptive Statistics**

Note: the increase in the current sample size from that employed in the data chapter (1483 versus 1456) arises as a result of cases with unknown destination or unknown timing for some migration moves. The level of aggregation in this chapter allows us to retain these cases whereas the type of analysis in the data chapter necessitated such cases be dropped.

Variable	Mean	<u>Std. Dev.</u>	Skew.	Kurt.	Min.	Max.	<u>Cases</u>
BOB MIC	1 (10	0.010	2 0	0.0	0	15 00	1400
TOT MIG	1.649	2.218	2.0	8.2	0.	15.00	1483
AG15 RUR	0.7296	0.4443	-1.0	2.1	Ο.	1.00	1483
AFTER 70	0.3817	0.4860	0.5	1.2	Ο.	1.00	1483
BEFOR 50	0.1942	0.3957	1.5	3.4	0.	1.00	1483
LNAGE	2.9570	0.7663	-0.8	2.9	0.69	4.26	1483
LNAG 50	0.7419	1.513	1.5	3.4	Ο.	4.26	1483
LNAG 70	0.8187	1.097	0.7	1.8	0.	2.83	1483
YRSEDUC	7.2030	4.126	0.2	2.7	0.	21.00	1483
BOARD	0.1005	0.3007	2.7	8.1	0.	1.00	1483
NORTHERN							
SOUTHERN	0.1773	0.3821	1.7	3.9	0.	1.00	1483
WESTERN	0.4639	0.4989	0.1	1.0	Ο.	1.00	1483
FOREIGN	0.0161	0.1288	7.5	57.3	0.	1.00	1483
MALAY	0.5907	0.4919	-0.4	1.1	0.	1.00	1483
CHINESE	0.2670	0.4426	1.1	2.1	0.	1.00	1483
INDIAN	0.1335	0.3402	2.2	5.6	0.	1.00	1483
RUR MAL	0.5017	0.5002	0.0	1.0	Ο.	1.00	1483
RUR CHI	0.1416	0.3488	2.1	5.2	Ο.	1.00	1483
RUR IND	0.0781	0.2697	3.1	10.8	Ο.	1.00	1483
URB MAL	0.0890	0.2849	2.9	9.3	0.	1.00	1483
URB CHI	0.1254	0.3313	2.3	6.1	Ο.	1.00	1483
URB IND	0.0546	0.2273	3.9	16.4	0.	1.00	1483
AN YNGBR	0.6709	0.4700	-0.7	1.5	0.	1.00	1483
AN YNGSI	0.6763	0.4680	-0.8	1.6	0.	1.00	1483
PAR LORH	0.1510	0.3582	1.9	4.8	Ο.	1.00	1483
RUR LORH	0.1200	0.3251	2.3	6.5	Ο.	1.00	1483
URB LORH	0.0310	0.1734	5.4	30.3	0.	1.00	1483

	Descr:	<b>iptive</b>	Statistics	(cont'd)
--	--------	---------------	------------	----------

<u>Variable</u>	<u>Mean</u>	Std. Dev.	<u>Skew.</u>	Kurt.	Min.	Max.	<u>Cases</u>
F SECED	0.0600	0.2376	3.7	14.7	0.	1.00	1483
M SECED	0.0243	0.1540	6.2	39.2	0.	1.00	1483
PAR SEC	0.0701	0.2554	3.4	12.3	0.	1.00	1483
F PR ED	0.4612	0.4987	0.2	1.0	0.	1.00	1483
M PR ED	0.2765	0.4474	1.0	2.0	0.	1.00	1483
PAR PRED	0.5233	0.4996	-0.1	1.0	0.	1.00	1483
FATH WHI	0.2097	0.4072	1.4	3.0	0.	1.00	1483
MOTH WHI	0.0641	0.2449	3.6	13.7	0.	1.00	1483
FATH PRD	0.2481	0.4321	1.2	2.4	Ο.	1.00	1483
MOTH PRD	0.0708	0.2566	3.3	12.2	0.	1.00	1483
FATH AG	0.5266	0.4995	-0.1	1.0	Ο.	1.00	1483
MOTH AG	0.4080	0.4916	0.4	1.1	0.	1.00	1483
MV B 15	0.2313	0.4218	1.3	2.6	0.	1.00	1483
R MV15	0.1274	0.3336	2.2	6.0	Ο.	1.00	1483
U MV15	0.1038	0.3052	2.6	7.7	0.	1.00	1483
WRK YONG	0.3783	0.4851	0.5	1.3	0.	1.00	1483
DIST NET	-0.6468	12.65	0.6	3.4	-27.61	40.47	1458

\_

## APPENDIX B

### Poisson Regression Model

Log-likelihood	=	-2,686.890	LR Statistic≈	921.3
Restricted Log-L	=	-3,147.557		

<u>Variable</u>	Coefficient	Std. Error	<u>t-ratio</u>	Prob.
Constant	-0.697	0.4740	-1.472	0.1411
AN_YNGBR	0.098	0.0457	2.153	0.0313
F_SECED	0.358	0.0875	4.098	0.0000
M_SECED	-0.125	0.1380	-0.911	0.3621
PAR_PRED	0.223	0.0456	4.905	0.0000
FATH_WHI	-0.087	0.0592	-1.475	0.1403
MOTH_WHI	0.242	0.0867	2.795	0.0052
FATH_PRD	0.097	0.0538	1.818	0.0691
MOTH_PRD	0.067	0.0802	0.838	0.4020
YRSEDUC	0.081	0.0074	11.033	0.0000
BOARD	0.428	0.0692	6.194	0.0000
SOUTHERN	0.045	0.0589	0.774	0.4389
WESTERN	-0.152	0.0550	-2.784	0.0054
FOREIGN	0.518	0.1341	3.865	0.0001
URB_CHI	-0.539	0.0876	-6.162	0.0000
URB_IND	-0.184	0.1007	-1.829	0.0674
RUR_MAL	-0.094	0.0769	-1.232	0.2178
RUR_CHI	-0.186	0.0946	-1.973	0.0485
RUR_IND	-0.042	0.1087	-0.392	0.6949
R_MV15	0.192	0.0617	3.118	0.0018
U_MV15	0.003	0.0760	0.051	0.9595
WRK_YONG	0.083	0.0491	1.702	0.0888
RUR_LORH	-0.427	0.0776	-5.518	0.0000
URB_LORH	0.202	0.1001	2.025	0.0428
BEFOR_50	2.973	1.4960	1.990	0.0466
AFTER_70	-2.493	0.4954	-5.029	0.0000
LNAGE	0.170	0.1373	1.240	0.2150
LNAG_50	-0.682	0.3978	-1.716	0.0862
LNAG_70	0.881	0.1637	5.386	0.0000

.

-

#### Overdispersion Tests (Base Poisson Model)

In all cases, the null hypothesis is  $Var[y_i]=E[y_i]$ . The first two tests, due to Cameron and Trivedi (1990), formulate the alternative hypothesis as  $Var[y_i]=E[y_i]+g(E[y_i])$  where the specific form of g(.) is given below in (1) and (2). The conditional moments test, due to Pagan and Vella (1989), is more general in that the form of overdispersion is not specified. The null hypothesis of no overdispersion is uniformly rejected.

(1) g	$(\lambda_i) = \lambda_i$			
<u>Variable</u>	Coefficient	Std. Error	<u>t-ratio</u>	Prob.
TEST1	1.2287	0.1474	8.335	0.0000
(2) g	$(\lambda_i) = \lambda_i^2$			
<u>Variable</u>	Coefficient	Std. Error	<u>t-ratio</u>	Prob.
TEST2	0.50690	0.07439	6.814	0.0000

(3) Conditional Moments Test

Test statistic = 222.8072  $\stackrel{d}{\rightarrow} \chi^2(dof=29)$ 

5% critical value = 42.56

Note: the predicted number for each outcome, w, is calculated as the sample sum of the estimates of  $Prob[y_i = w]$ .



### APPENDIX D

# Negative Binomial Regression Model

NBR Log-lik	elihood		=	-2,430.677
Restricted	(Poisson)	Log-likelihood	=	-2,686.890

LR Statistic= 256.2

<u>Variable</u>	<u>Coefficient</u>	<u>Std. Error</u>	<u>t-ratio</u>	Prob.
Constant	-0.380	0.7852	-0.485	0.6277
AN_YNGBR	0.147	0.0773	1.903	0.0570
F_SECED	0.313	0.1634	1.918	0.0551
M_SECED	-0.130	0.2740	-0.475	0.6345
PAR_PRED	0.221	0.0753	2.943	0.0032
FATH_WHI	-0.111	0.1043	-1.066	0.2862
MOTH_WHI	0.247	0.1471	1.686	0.0918
FATH_PRD	0.112	0.0931	1.206	0.2280
MOTH_PRD	0.044	0.1313	0.341	0.7331
YRSEDUC	0.077	0.0141	5.518	0.0000
BOARD	0.479	0.1696	2.826	0.0047
SOUTHERN	0.059	0.1040	0.567	0.5704
WESTERN	-0.120	0.0910	-1.325	0.1851
FOREIGN	0.578	0.4843	1.195	0.2320
URB_CHI	-0.610	0.1560	-3.911	0.0001
URB_IND	-0.183	0.1978	-0.926	0.3543
RUR_MAL	-0.136	0.1401	-0.975	0.3296
RUR_CHI	-0.231	0.1614	-1.433	0.1520
RUR_IND	-0.122	0.1994	-0.615	0.5388
R_MV15	0.280	0.1203	2.330	0.0198
U_MV15	0.067	0.1532	0.440	0.6600
WRK_YONG	0.095	0.0764	1.253	0.2104
RUR_LORH	-0.427	0.1141	-3.746	0.0002
URB_LORH	0.179	0.2189	0.820	0.4125
BEFOR_50	3.125	2.6150	1.195	0.2320
AFTER_70	-2.942	0.7878	-3.736	0.0002
LNAGE	0.069	0.2256	0.306	0.7593
LNAG_50	-0.709	0.6981	-1.017	0.3092
LNAG_70	1.050	0.2552	4.118	0.0000
Alpha	0.828	0.0688	12.047	0.0000
# APPENDIX E

## Binomial Probit Model

Log-Likel:	hood	- 86	2.15	
Restricted	d (Slopes=0) L	og-L1,02	0.00	
Chi-Square	ed (32)	31	5.75	
<u>Variable</u>	Coefficient_	Std. Error	<u>t-ratio</u>	Prob.
Constant	-0.19223	0.8772	-0.219	0.8265
AN_YNGBR	0.12956	0.0803	1.613	0.1067
AN_OLDBR	0.08197	0.0738	1.111	0.2667
F_SECED	0.17577	0.1825	0.963	0.3356
M_SECED	0.20605	0.2626	0.785	0.4326
PAR_PRED	0.27193	0.0795	3.421	0.0006
FATH_WHI	-0.19045	0.1070	-1.780	0.0751
MOTH_WHI	0.18808	0.1509	1.247	0.2125
FATH_PRD	-0.05117	0.0941	-0.539	0.5898
MOTH_PRD	-0.05757	0.1412	-0.408	0.6836
YRSEDUC	0.06024	0.0138	4.359	0.0000
SOUTHERN	0.03502	0.1082	0.324	0.7461
WESTERN	-0.13590	0.0913	-1.489	0.1366
FOREIGN	4.97940	34.32	0.145	0.8846
AG15_RUR	-0.31558	0.4471	-0.706	0.4803
URB_CHI	-0.57679	0.1700	-3.393	0.0007
URB_IND	-0.11209	0.2122	-0.528	0.5974
RUR_MAL	0.36012	0.4016	0.897	0.3699
RUR_CHI	0.10369	0.4153	0.250	0.8028
RUR_IND	0.39359	0.4237	0.929	0.3530
BOARD	0.84697	0.1673	5.062	0.0000
R_MV15	0.49063	0.1117	4.391	0.0000
U_MV15	0.56814	0.1552	3.662	0.0003
WRK_YONG	-0.03033	0.0798	-0.380	0.7040
RUR_LORH	-0.35783	0.1150	-3.112	0.0019
URB_LORH	0.19660	0.2479	0.793	0.4277
BEFOR_50	2.29820	2.8630	0.803	0.4222
AFTER_70	-2.62700	0.8426	-3.118	0.0018
R_YNGSI	-0.08683	0.0935	-0.929	0.3529
U_YNGSI	0.02317	0.1584	0.146	0.8837
LNAGE	-0.03843	0.2476	-0.155	0.8766
LNAG_50	-0.53662	0.7598	-0.706	0.4800
LNAG_70	0.91883	0.2701	3.401	0.0007

# Binomial Probit Model (cont'd)

Frequencies of actual & predicted outcomes Predicted outcome has maximum probability.

	Pred	dicted		
Actual	0	1	TOTAL	
0 1	396 209	269 609	665 818	
Total	605	878	1,483	

## APPENDIX F

### Zero Altered Poisson Regression Model

Comparison of estimated models:

	Number	of zeros	
	Actual	Predicted	Log-likelihood
Poisson	665	408	-2,686.890
ZAP(g)	665	631	-2,452.969
ZAP	665	663	-2,392.228

Vuong statistic for testing ZAP vs. unaltered model (ZAP) is 15.867 Distributed as standard normal. A value greater than +1.96 favours the ZAP model. A value less than -1.96 rejects the ZAP model.

<ul> <li>Variable</li> </ul>	<u>Coefficient</u>	<u>Std. Error</u>	<u>t-ratio</u>	Prob.
Constant	-0.24955	0.3965	-0.629	0.5291
AN_YNGBR	0.08291	0.0429	1.933	0.0532
FATH_WHI	0.06415	0.0509	1.259	0.2079
MOTH_WHI	0.09950	0.0719	1.384	0.1665
FATH_PRD	0.10631	0.0492	2.163	0.0305
MOTH_PRD	0.08874	0.0757	1.173	0.2408
YRSEDUC	0.05718	0.0073	7.858	0.0000
BOARD	0.25496	0.0686	3.718	0.0002
SOUTHERN	0.03852	0.0541	0.712	0.4766
WESTERN	-0.09889	0.0513	-1.928	0.0538
FOREIGN	0.03317	0.1525	0.218	0.8278
URB_CHI	-0.29223	0.0857	-3.411	0.0006
URB_IND	-0.09586	0.0935	-1.025	0.3054
RUR_MAL	-0.10674	0.0690	-1.547	0.1219
RUR_CHI	-0.03356	0.0805	-0.417	0.6769
RUR_IND	-0.11714	0.0969	-1.209	0.2266
R_MV15	-0.06313	0.0639	-0.987	0.3235
U_MV15	-0.24749	0.7480	-3.309	0.0009
WRK_YONG	0.14972	0.4482	3.341	0.0008
RUR_LORH	-0.27595	0.7870	-3.506	0.0005
URB_LORH	0.17621	0.8702	2.025	0.0429
BEFOR_50	2.07220	1.2380	1.674	0.0942
AFTER_70	-1.13130	0.4612	-2.453	0.0142
LNAGE	0.25043	0.1126	2.225	0.0261
LNAG_50	-0.47639	0.3295	-1.446	0.1483
LNAG 70	0.40753	0.1573	2.591	0.0096

# Zero Altered Poisson Regression Model (cont'd)

# **Participation** Decision

Variable	Coefficient	Std. Error	t-ratio	Prob.
Constant	-1.5320	1.7000	-0.901	0.3676
R_YNGSI	0.2113	0.1598	1.322	0.1860
U_YNGSI	0.1068	0.2593	0.412	0.6803
PAR_PRED	-0.4891	0.1486	-3.291	0.0010
PAR_SEC	-0.1013	0.3545	-0.286	0.7751
YRSEDUC	-0.0811	0.0271	-2.991	0.0028
BOARD	-1.6977	0.4395	-3.862	0.0001
URB_CHI	0.9667	0.2765	3.497	0.0005
R_MV15	-1.0209	0.2454	-4.160	0.0000
U_MV15	-1.8529	0.4351	-4.259	0.0000
WRK_YONG	0.1088	0.1507	0.722	0.4705
RUR_LORH	0.4511	0.2099	2.149	0.0317
BEFOR_50	2.9682	4.7780	0.621	0.5345
AFTER_70	5.0308	1.7120	2.939	0.0033
LNAGE	0.4979	0.4927	1.011	0.3122
LNAG_50	-0.9146	1.2720	-0.719	0.4723
LNAG_70	-1.7354	0.5490	-3.161	0.0016

#### Zero Altered Negative Binomial Regression Model

Comparison of estimated models:

	Number	of zeros	
	Actual	Predicted	Log-likelihood
Poisson	665	408	-2,686.890
Neg. Bin.	665	607	-2,430.677
ZANBR (g)	665	648	-2,392.156
ZANBR	665	662	-2,349.304

Vuong statistic for testing ZANBR vs. unaltered model (NBR) is 8.7662 Distributed as standard normal. A value greater than +1.96 favours the ZANBR model. A value less than -1.96 rejects the ZANBR model.

<u>Variable</u>	Coefficient	<u>Std. Error</u>	<u>t-ratio</u>	Prob.
Constant	-0.52258	0.6561	-0.796	0.4258
AN_YNGBR	0.10062	0.0661	1.522	0.1281
FATH_WHI	0.01780	0.0820	0.217	0.8282
MOTH_WHI	0.12154	0.1176	1.034	0.3013
FATH_PRD	0.09886	0.0784	1.261	0.2074
MOTH_PRD	0.09002	0.1225	0.735	0.4625
YRSEDUC	0.06393	0.0118	5.432	0.0000
BOARD	0.25512	0.1143	2.231	0.0257
SOUTHERN	0.03380	0.0877	0.385	0.6999
WESTERN	-0.12291	0.0793	-1.550	0.1211
FOREIGN	0.15258	0.2535	0.602	0.5472
URB_CHI	-0.30631	0.1358	-2.255	0.0241
URB_IND	-0.10424	0.1431	-0.728	0.4663
RUR_MAL	-0.11576	0.1143	-1.013	0.3112
RUR_CHI	-0.08620	0.1324	-0.651	0.5150
RUR_IND	-0.10709	0.1589	-0.674	0.5003
R_MV15	-0.67916	0.1044	-0.651	0.5153
U_MV15	-0.28368	0.1184	-2.396	0.0166
WRK_YONG	0.16518	0.0722	2.288	0.0222
RUR_LORH	-0.35603	0.1158	-3.074	0.0021
URB_LORH	0.20647	0.1469	1.405	0.1599
BEFOR_50	2.78220	2.1250	1.310	0.1904
AFTER_70	-1.16570	0.6996	-1.666	0.0956
LNAGE	0.29808	0.1859	1.603	0.1089
LNAG_50	-0.66605	0.5648	-1.179	0.2383
LNAG_70	0.43920	0.2329	1.885	0.0594
Alpha	0.24134	0.0441	5.467	0.0000

# Zero Altered Negative Binomial Regression Model (cont'd)

# **Participation Decision**

Variable	Coefficient	Std. Error	t-ratio	Prob.
Constant	-2.7633	2.1710	-1.273	0.2031
R_YNGSI	0.2798	0.1932	1.448	0.1475
U_YNGSI	0.0716	0.3356	0.213	0.8311
PAR_PRED	-0.6336	0.1860	-3.406	0.0007
PAR_SEC	0.0402	0.4399	0.091	0.9272
YRSEDUC	-0.0845	0.0332	-2.547	0.0109
BOARD	-3.1379	1.9110	-1.642	0.1006
URB_CHI	1.1963	0.3494	3.424	0.0006
R_MV15	-1.4364	0.4213	-3.409	0.0007
U_MV15	-3.2546	1.5040	-2.165	0.0304
WRK_YONG	0.1985	0.1875	1.059	0.2895
RUR_LORH	0.3205	0.2695	1.189	0.2344
BEFOR_50	5.0517	5.8080	0.870	0.3844
AFTER_70	6.2053	2.1810	2.845	0.0044
LNAGE	0.7975	0.6253	1.275	0.2022
LNAG_50	-1.5075	1.5500	-0.972	0.3309
LNAG_70	-2.0892	0.6930	-3.015	0.0026





Figure G-1



## APPENDIX I

### Zero Altered Poisson Regression Model (g)

Comparison of estimated models:

-

	Number	of zeros	
	Actual	Predicted	Log-likelihood
Poisson	665	408	-2,686.890
ZAP(g)	665	631	-2,452.969

Vuong statistic for testing ZAP vs. unaltered model is 14.3374 Distributed as standard normal. A value greater than +1.96 favours the ZAP model. A value less than -1.96 rejects the ZAP model.

Variable	Coefficient	Std. Error	t-ratio	Prob.
Constant	-0.061	0.3610	-0.170	0.8651
AN_YNGBR	0.094	0.03831	2.475	0.0133
FATH_WHI	0.025	0.04693	0.549	0.5829
MOTH_WHI	0.115	0.06585	1.759	0.0786
FATH_PRD	0.098	0.04445	2.207	0.0273
MOTH_PRD	0.041	0.06647	0.628	0.5297
YRSEDUC	0.063	0.00659	9.631	0.0000
BOARD	0.304	0.06426	4.742	0.0000
SOUTHERN	0.060	0.04906	1.235	0.2167
WESTERN	-0.068	0.04618	-1.479	0.1392
FOREIGN	0.217	0.1569	1.385	0.1661
URB_CHI	-0.393	0.07483	-5.265	0.0000
URB_IND	-0.106	0.08948	-1.193	0.2327
RUR_MAL	-0.089	0.06335	-1.408	0.1592
RUR_CHI	-0.101	0.07446	-1.363	0.1728
RUR_IND	-0.086	0.09035	-0.960	0.3369
R_MV15	0.093	0.05757	1.618	0.1056
U_MV15	-0.048	0.06984	-0.699	0.4845
WRK_YONG	0.094	0.03909	2.408	0.0160
RUR_LORH	-0.330	0.06413	-5.158	0.0000
URB_LORH	0.138	0.08363	1.652	0.0985
BEFOR_50	1.983	1.1460	1.732	0.0833
AFTER_70	-1.703	0.3822	-4.449	0.0000
LNAGE	0.151	0.1030	1.471	0.1413
LNAG_50	-0.437	0.3053	-1.434	0.1516
LNAG_70	0.608	0.1272	4.784	0.0000
Gamma	-0.893	0.08571	-10.431	0.0000

Comparison of estimated models:

	Number	of zeros	
	Actual	Predicted	Log-likelihood
Poisson	665	408	-2,686.890
Neg. Bin.	665	607	-2,430.677
ZANBR (g)	665	648	-2,392.156

Vuong statistic for testing ZANBR vs. unaltered model is 7.9950 Distributed as standard normal. A value greater than +1.96 favours the ZANBR model. A value less than -1.96 rejects the ZANBR model.

<u>Variable</u>	<u>Coefficient</u>	<u>Std. Error</u>	<u>t-ratio</u>	Prob.
Constant	0.12448	0.5122	0.243	0.8080
AN_YNGBR	0.09878	0.0506	1.954	0.0507
FATH_WHI	-0.03070	0.0653	-0.470	0.6384
MOTH_WHI	0.12859	0.0921	1.396	0.1628
FATH_PRD	0.06775	0.0605	1.120	0.2625
MOTH_PRD	0.00133	0.0853	0.016	0.9876
YRSEDUC	0.06166	0.0090	6.900	0.0000
BOARD	0.40843	0.1053	3.878	0.0001
SOUTHERN	0.06371	0.0670	0.951	0.3416
WESTERN	-0.05703	0.0594	-0.961	0.3366
FOREIGN	0.47899	0.3001	1.596	0.1105
URB_CHI	-0.41562	0.1008	-4.122	0.0000
URB_IND	-0.09291	0.1282	-0.725	0.4685
RUR_MAL	-0.07760	0.0904	-0.858	0.3909
RUR_CHI	-0.15375	0.1035	-1.485	0.1376
RUR_IND	-0.07406	0.1280	-0.578	0.5630
R_MV15	0.20422	0.0785	2.603	0.0092
U_MV15	0.08828	0.0988	0.894	0.3714
WRK_YONG	0.04975	0.0502	0.991	0.3215
RUR_LORH	-0.30762	0.0764	-4.026	0.0001
URB_LORH	0.15067	0.1384	1.089	0.2762
BEFOR_50	2.11470	1.6390	1.290	0.1969
AFTER_70	-1.78680	0.5199	-3.437	0.0006
LNAGE	0.05400	0.1468	0.368	0.7129
LNAG_50	-0.47838	0.4370	-1.095	0.2736
LNAG_70	0.62951	0.1691	3.724	0.0002
Alpha	0.34629	0.0573	6.042	0.0000
Gamma	-1.72750	0.2776	-6.223	0.0000

	Number of Inter-District Migrations				
Work At Age 15:	Observations	Mean	Std. Dev.		
No	903	1.69	2.22		
Yes	553	1.38	1.90		
combined	1456	1.57	2.11		
H <sub>0</sub> : mean(x) = mean(y) t = 2.74 with 1454 d.f. Pr > $ t  = 0.0061$					

Table J-1. Number of Migrations by 'WRK YONG'

Table J-2. Number of Migrations by 'U MV15'

	Number of Inter-District Migrations				
U MV15 =	Observations	Mean	Std. Dev.		
0	1305	1.52	2.09		
1	151	2.07	2.27		
combined	1456	1.57	2.11		
H <sub>0</sub> : mean(x) = mean(y) t = $-3.08$ with 1454 d.f. Pr >  t  = $0.0021$					

٠

#### BIBLIOGRAPHY

- Allen, Jeremiah. 1979. "Information and Subsequent Migration: Further Analysis and Additional Evidence", *Southern Economic Journal*, 45: 1274-1284.
- Amburgey, T. L. and G. R. Carroll. 1984. "Time Series Models for Event Counts", Social Science Research, 13, 1: 38-54.
- Amemiya, Takeshi and Keiko Shimono. 19??. "An Application of Nested Logit Models to the Labor Supply of the Elderly". Unpublished manuscript.
- Banerjee, Biswajit. 1991. "The Determinants of Migrating with a Pre-Arranged Job and of the Initial Duration of Urban Unemployment", *Journal of Development Economics*, 36: 337-51.
- Barron, David N. 1992. "The Analysis of Count Data: Overdispersion and Autocorrelation" in Marsden, Peter V. (ed.), Sociological Methodology 1992, volume 22. Oxford: Basil Blackwell.
- Bartel, Ann P. 1979. "The Migration Decision: What Role Does Job Mobility Play?", American Economic Review, 69: 775-786.
- Cameron, A. Colin and Pravin K. Trivedi. 1986. "Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests", *Journal of Applied Econometrics* 1, 29-53.
- Cameron, A. Colin and Pravin K. Trivedi. 1990. "Regression-based Tests for Overdispersion in the Poisson Model", *Journal of Econometrics*, 46: 347-64.
- Cox, David R. and Valeris Isham. 1980. *Point Processes*. London: Chapman and Hall.
- DaVanzo, Julie. 1983. "Repeat Migration In the United States: Who Moves Back and Who Moves on?", *Review of Economic Studies*, 65, 4: 552-559.
- Dean, C. B. 1992. "Testing for Overdispersion in Poisson and Binomial Regression Models", Journal of the American Statistical Association, 87: 451-457.
- Foster, F. G. 1952. "A Markov Chain Derivation of Discrete Distributions", Annals of Mathematical Statistics, 23: 624-627.

- Gourieroux, C.; A. Monfort; and A. Trognon. 1984. "Pseudo Maximum Likelihood Methods: Applications to Poisson Models", *Econometrica* 52: 701-720.
- Greene, William H. 1992. Limdep 6.0: User's Manual and Reference Guide. New York: Econometric Software Inc.
- Greene, William H. 1993. Econometric Analysis. New York: Macmillan.
- Greene, William H. 1994. "Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models", Working Paper EC-94-10, Department of Economics, New York University.
- Greenwood, Michael J. 1975. "Research on Internal Migration in the United States: A Survey", *Journal of Economic Literature*, 8: 397-433.
- Harris and Todaro, "Migration, Unemployment and Development: A Two-Sector Analysis", *American Economic Review*, 60, 1:126-142.
- Hausman, Jerry; Bronwyn H. Hall; and Zvi Griliches. 1984. "Econometrics Models for Count Data with An Application to the Patents-R&D Relationship", *Econometrica*, 52, 4: 909-938.
- Heckman, James J. 1981. "Statistical Analysis of Discrete Data with Econometric Applications" in Manski, C.F. (ed.), *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge: The MIT Press.
- Johnston, N. L.; S. Kotz; and J. Kemp. 1992. Univariate Discrete Distributions, 2<sup>nd</sup> Edition. Toronto: John Wiley & Sons Inc.
- Jones, Andrew. 1989. "A Double-Hurdle Model of Cigarette Consumption", Journal of Applied Econometrics, 4: 23-29.
- King, Gary. 1988. "Statistical Models for Political Science Event Counts Bias in Conventional Procedures and Evidence for the Exponential Poisson Regression Model", American Journal of Political Science, 32, 3: 838-863.
- King, Gary. 1987. "Presidential Appointments to the Supreme Court: Adding Systematic Explanation to Probabilistic Description", *American Politics Quarterly*, 15, 3: 373-386.

- King, Gary. 1988. "Statistical Models for Political Science Event Counts: Bias in Conventional Procedures and Evidence for The Exponential Poisson Regression Model", American Journal of Political Science, 32, 3: 838-863.
- King, Gary. 1989a. Unifying Political Methodology: The Likelihood Theory of Statistical Inference. New York: Cambridge University Press.
- King, Gary. 1989b. "Event Count Models for International Relations: Generalizations and Applications", *International Studies Quarterly*, 33, 2: 123--148.
- Maddala, G. S. 1983. Limited Dependent and Qualitative Variables in Econometrics. New York: Cambridge University Press.
- Mazumdar, Dipak. 1981. The Urban Labour Market and Income Distribution: A Study of Malaysia. New York: Oxford University Press.
- Morrison, Peter A. 1971. "Chronic Movers and the Future Redistribution of Population: A Longitudinal Analysis", *Demography*, 8: 171-184.
- Nakosteen, Robert A. and M. Zimmer. 1980. "Migration and Income: The Question of Self-Selection", *Southern Economic Journal*, 46, 3: 840-851.
- Pagan, A. and F. Vella. 1989. "Diagnostic Tests for Models Based on Individual Data: A Survey", *Journal of Applied Econometrics*, 4: S29-S59.
- Pessino, Carola. 1991. "Sequential Migration Theory and Evidence from Peru", Journal of Development Economics, 36, 1: 55-87.
- Robinson, C. and N. Tomes. 1982. "Self-Selection and Interprovincial Migration in Canada", *Canadian Journal of Economics*, 15: 474-502.
- Roy, A. D. 1951. "Some Thoughts on the Distribution of Earnings", Oxford Economic Papers (New Series), 3: 135-146.
- Schultz, T. Paul. 1982. "Notes on the Estimation of Migration Decision Functions" in Sabot, Richard H. (ed), Migration and the Labour Market in Developing Countries. Boulder, Colorado: Westview Press.
- Sjaastad, L. 1962. "The Costs and Returns of Human Migration", American Economic Review, 70: 80-93.

Thompson, H. R. 1954. "A Note on Contagious Distributions", *Biometrika*, 41: 268-71.

- Tunali, Insan. 1986. "A General Structure for Models of Double-Selection and an Application to a Joint Migration/Earnings Process with Re-Migration" in Ehrenberg, Ronald G. (ed.), *Research in Labor Economics*, 8, Part B: 235-282.
- Vanderkamp, John. 1971. "Migration Flows, Their Determinants, and the Effects of Return Migration", *Journal of Political Economy*, 79: 1012-31.
- Veall, Michael R. 1992. "Bootstrapping the Process of Model Selection An Econometric Example", *Journal of Applied Economics*, 7, 1: 93-99.
- Vuong, Q. 1989. "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses", *Econometrica*, 57: 307-334.
- Yap, Lorene Y. 1977. "The Attraction of Cities: A Review of the Migration Literature", *Journal of Development Economics*, 4: 239-264.
- Yezer, Anthony M.J. and Lawrence Thurston. 1976. "Migration Patterns and Income Change: Implications for the Human Capital Approach to Migration", *Southern Economic Journal*, 45: 694-702.