The Economic Journal, **95** (*June* 1985), 358–382
Printed in Great Britain

MIGRATION AMONGST THE BATSWANA*

Robert E. B. Lucas

The hypothesis presented in Todaro (1969), that the likelihood of finding work in town influences the rate of rural–urban migration, now enjoys the status of a received doctrine. Assuming potential migrants indeed respond to this employment probability, the model of Harris and Todaro (1970) demonstrates that, in certain parametric ranges, urban job creation may actually exacerbate unemployment and even reduce national product. This result has had considerable influence on policy formulation in LDCs, by emphasising that, in the urban sector, the social opportunity cost of labour may not be insignificant despite burgeoning unemployment.

Yet neither the Todaro hypothesis nor prevalence of the Harris–Todaro parametric range has been adequately, empirically explored. Many estimates of macro migration equations do exist, normally relating the proportion of population migrating to average wages in differing locations and occasionally to average population characteristics. But in Lucas (1975), I show that the popular nonlinear specification of such macro functions may well display serious specification error bias; a nonlinear function of the aggregate variables is not simply the average of underlying micro migration decisions related to the disaggregated variables. Thus, although a few estimates of macro migration equations have also incorporated average unemployment rates, usually in developed country contexts and with mixed results, these analyses are at best very circumscribed tests of the Todaro and Harris–Todaro theories.

On the other hand, surprisingly few estimates of micro migration response functions have appeared. In part, this probably reflects the difficulty of dealing with unobserved variables. For example, suppose whether or not each person moves to town is to be related to his or her wage at home compared to the relevant wage in town. In a typical survey, for those individuals who remain in the village the wage they would command in town is not observed; for those who have moved to town, the prior rural wage is usually not reported. DaVanzo (1976) deals with this problem of unobserved variables, in a US context, by introducing a multivariate wage equation to predict potential wages in alternative locations for each person using their observed personal characteristics.

However, this broad approach suggested by DaVanzo has not previously been extended to incorporate unobserved employment probabilities in alternative locations. Thus, in surveying the literature in 1980, DaVanzo concludes:

* I am particularly grateful to Carol Kerven, John Case, David Cooper and John Harris for many discussions and for their tremendous effort in collecting and preparing the data. Funding for this study and permission to use the data were granted by the Government of Botswana, but results and interpretations are the responsibility of the author alone. Many helpful suggestions for redrafting were provided by two referees and an associate editor.

However, except for Hay's (1974) analysis of Tunisian migration, which used proxy variables in lieu of data on actual urban income and employment rates, to my knowledge there have been no tests of the Todaro model with microdata.[1]

The present study undertakes analysis of the effects of employment probabilities and wage alternatives on migration decisions in the context of Botswana, adopting micro level data both in estimating a multinomial migration function and to develop prediction equations for wages and likelihood of employment, thus extending DaVanzo's earlier work both to include employment probabilities and to an LDC context. Intra-rural movements, short- and long-term urban migration, and emigration to South Africa are each examined as possible outcomes. Part I outlines certain features of Botswana essential to interpretation of the ensuing results. In II, some issues pertaining to specification and estimation of the model are briefly outlined. The results on the estimated earnings and employment probability equations are discussed in Part III, and IV presents the estimated migration equations and in particular emphasises their relationship to the Todaro and Harris–Todaro hypotheses. This apparently represents the first systematic tests of these hypotheses, at least from micro data, though the implications of the Harris–Todaro model have already had considerable influence in shaping development policies.

I. THE BOTSWANA CONTEXT

Botswana is a land-locked, semi-arid country in the Southern African cone. About the size of Texas or France, it has a population of somewhat under 1 million – a population which is highly peripatetic rendering this an ideal context for a migration study. Since independence in 1966, much of this migration has been to the five urban centres of Botswana. These towns, with the exception of the Orapa mining compound in the north, are situated along the only railway line which runs north–south close to the eastern border. None of the towns is large. Gaborone, the capital, is the largest, having approximately 50,000 persons. But three of the towns – Gaborone, Selebi-Pikwe and Orapa – are new since independence so their growth rate has been substantial.

Botswana is juxtaposed to South Africa to the south and migration for employment in South Africa is very high. Currently, about 20,000 Batswana men are taken each year on contract to the South African mines, but given the sharp rise in mine wages since 1974 there exists a substantial excess supply of Batswana labour.[2] Migration into South Africa is restricted by the South Africans though some clandestine movement does occur.

The overwhelming majority of Botswana's population lives in nucleated villages. Some of the large villages exceed the smaller towns in population and the distinguishing features of villages are thus social structure, employment

[1] DaVanzo (1980), pp. 27–8. See also Hay (1980). Behrman and Wolfe (1985) do explore micro migration responses of Nicaraguan women in relation to predicted employment probabilities, in a study conducted concurrently and independently though with quite a different focus from the present study.

[2] See Lucas (1982, 1983). The Batswana are the peoples of Botswana.

activities and facilities available, rather than size. The large villages are usually also the tribal royal villages of the major tribal groups. Affiliated with these are other, typically smaller (or 'peripheral') villages. The sand veldt, more arid area occupying the western two-thirds of the country has quite small 'subperipheral' settlements.

A stylised land use pattern may be discerned around a 'typical' village. Beyond the nucleated village, though sometimes more than one day's walk away, lie scattered arable fields called lands. At greater distances, normally in the sand veldt portion, the village has associated cattle posts where at least the larger cattle herds are kept. Altogether, the village, its lands and cattleposts are referred to here as a village catchment area (v.c.a.). Schapera (1943) describes a pattern of households maintaining a home in the village, a dwelling at the lands and perhaps a third at a cattle post. In a traditional scenario, the family resides in its village dwelling, with some of the members moving to the lands when authorised by the chief at ploughing time. In fact, it seems more likely many households are too poor to maintain multiple dwellings. In addition, lands settlements are becoming more permanent as boreholes provide more disseminated water sources and control by the chiefs wanes. Nonetheless, substantial seasonal migration persists for those households having a separate lands dwelling. The cattle posts are more permanently occupied, by men and boys going from the village homes, but also by servant families – frequently Basarwa (bushmen) – relatively settled at the posts.

In the semi-arid climate, crop husbandry is both a very marginal and risky activity. Most of the time spent on crops is by women, though men do most of the ploughing. Even livestock tending, in this country where cattle outnumber people, does not provide large amounts of employment and much of this is done by herd boys for wages. For men, amidst extensive periods of non-employment, small wage jobs as herd boys, on freehold (commercial) farms or in village occupations are important. Among women there is more self employment, particularly in brewing beer from sorghum. Thus, in contrast to many LDCs, employment in agriculture on own account does not dominate time-use in the rural areas. In towns, much of the employment for men is in government service, mining (diamonds and copper), or parastatal organisations such as the Botswana Meat Commission, while amongst urban women jobs as domestic servants play a major role. Although wage employment has grown substantially since independence, the potential future decline in recruitment of men to the South African mines, and concern with finding an appropriate mechanism for more evenly dispersing the benefits of very rapid growth generated by rents from mineral extraction are focusing attention on further, substantial employment generation. [See Lipton (1978)]. But the relative merits of current alternatives – of rural employment creation through the Arable Lands Development Programme and Rural Industrial Areas Project, of jobs in the new diamond mine at Jwaneng and projected coal mines, of disputed Government minimum (urban) wage policy, and parastatal pay and hiring strategy – hinge in part upon their consequences for the composition and magnitude of migration patterns.

II. THE MODEL AND ESTIMATION ISSUES

A. Specification

The model is comprised of four groups of equations: for migration to South Africa, internal migration, employment and earnings. Each of these equations is estimated from micro data on individual persons. The data are taken from the National Migration Survey in Botswana, 1978–9 (NMS).

Micro migration equations are normally interpreted to be supply decisions of individuals or their families. However, whether or not individuals enter South Africa from Botswana is in fact a mixture of potential migrants' choices and South African selection. There exists a very obvious excess supply of Batswana wishing to find jobs in South Africa – particularly men seeking mine contracts – but restricted from moving by immigration laws and a well-policed border. For this reason a two-step procedure is adopted: the first group of equations examines the binomial outcome of migration to South Africa or not; the second group analyses the conditional probabilities of various internal choices of location given nonmovement to South Africa.

To understand the approach adopted to internal migration, two essential features of the survey must be mentioned. During the survey, four rounds of interviews were conducted, reinterviewing inhabitants of each selected dwelling every three months. At each interview, the respondents were asked who was absent from the dwelling and where each absentee was presently located. On this basis, each member of a rural home, other than those absent in South Africa, is placed in one of six categories:[1]

(*i*) Those persons present in the main home in all four rounds of the survey.

(*ii*) Persons sometimes or always absent, whose most frequent location of absence across rounds is the household's own lands or cattle post.[2]

(*iii*) Absentees with most frequent location elsewhere in rural Botswana.

(*iv–v*) Those having urban Botswana reported as most frequent location of absence are subdivided into two categories. Anyone who is never present in the main home in any round, or who is sometimes present but reports an average duration of absence of twelve months or more is termed a long term urban migrant. All others are considered short term.

[1] To study migration through reported absentees has at least two inherent weaknesses. Whole households which have departed are naturally missed, though this is probably comparatively rare in Botswana. Moreover, there is probably some selectivity in recalling absent members or in being defined as absent – a problem which is discussed subsequently in the text. An alternative approach, and probably a preferable one, is to ask each individual about their migration history, though experience in Botswana suggests difficulties in gleaning particulars about the household left behind. Given the absentee basis of the NMS migration information some care has to be taken over the treatment of multiple dwelling households described. Each household tends to consider one of its dwellings as the main home. Thus, for example, with the exception of those Basarwa households settled at the cattle posts, no one would normally be thought of as absent from the cattle post. The present study therefore focuses only on village dwellings (including their members absent at the lands, cattle post, town or elsewhere), plus those non-village rural dwellings which are today becoming more permanent and now reported as main homes. Given the rapid urban growth and establishment of new towns, virtually all urban inhabitants actually come from such rural main homes, and the vast majority came in the twelve year period between independence and the survey.

[2] Matching locational data across rounds shows a very remarkable mobility of absentees – hence the use of most frequent location.

(*vi*) A number of persons are not mentioned in one or more rounds of the survey, but once information is reported on them they are never subsequently absent. Such persons are referred to here as new arrivals, though they have very similar characteristics to those always present and indeed some may well be household members simply forgotten in earlier rounds.[1]

The dependent variables for the South Africa and internal migration equations respectively are the binomial observation on being in South Africa or not and the six-fold polytomous internal location measure. Binomial and multinomial logit equations relate these dependent variables to various explanatory variables, which in both the internal and South Africa migration equations include three broad classes of measures:

(*a*) personal characteristics – sex, age, years of schooling completed, if head of household, if 'married';[2]

(*b*) household characteristics – cattle owned (very highly correlated with overall wealth in Botswana households), and number of persons in the household (including those absent);

(*c*) locational factors – if the household's main dwelling is in a nonvillage (lands or cattle post) location, region (north, south or west) and village catchment area type (large, peripheral or subperipheral).[3]

All of these explanatory variables are entered linearly except region and village type which are used to define separate regressions. Except in the south of Botswana, very few women are reported in South Africa, so the South Africa migration equations are estimated for males only, with this exception.[4]

In the internal migration equations, five explanatory variables relating to employment are also explored: daily earnings in the home village and in town; probabilities of being in employment at home and in town; and an indicator of involvement in own account agriculture.[5] Since these employment alternatives

[1] Some of the new arrivals are reassigned to categories (*i*) to (*v*) by checking their reported movements in the month prior to first being listed in the household. Persons described by the respondents as temporary visitors are omitted from the analysis. It should be emphasised that by the standards of the developed countries, quite who is and is not a member of the household is a remarkably difficult matter to define in Botswana. Considerable care had to be taken with this in the survey design.

[2] Formal marriage is comparatively rare in Botswana (partly owing to the long history of male migration to South African mines). 'Married' therefore is allowed to embrace actually married, engaged or having some other partner.

[3] North is defined by the North East and Central Districts plus Maun village. South comprises Kweneng, Kgatleng, Southern and South East Districts and West is represented in the sample by Kgalagadi plus Ngamiland (except Maun).

[4] Four separate regression equations are defined for internal migration: for large villages in the north; large and peripheral villages in the south; subperipheral villages in the north plus south; and the west region altogether. The distance to town variable commonly included is omitted here since it is well proxied by this regional disaggregation. In the case of the binomial South Africa migration equation, to preserve degrees of freedom the west and subperipheral areas are combined in the one equation, with a dummy to represent the subperipheral north and south. A further dummy is inserted in the South Africa equations to discern peripheral as opposed to large village catchment areas within the north and south. This latter variable is excluded from the internal migration equation, having an insignificant effect there.

[5] These five employment measures are excluded from the separate South Africa migration equations since in the context of obvious excess supply, local opportunities are of little relevance. (This is confirmed by estimates including these measures, not reported here.) However, selection of male recruits for the South African mines is reputed to be affected by prior employment in the mines. A dummy

are central to the Todaro and Harris–Todaro hypotheses, their measurement warrants closer consideration.

The problem of unobserved values for earnings and employment alternatives is, as outlined in the introduction, to be tackled here by developing prediction equations. These equations are constrained to contain only explanatory variables known for each person present or absent in order to be able to use these values for prediction. For example, occupation of rural dwellers if they were to be in town is not observed. Accordingly, an earnings equation is estimated, of a fairly conventional human capital type, expressing the logarithm of daily wages in cash and kind plus net self-employment earnings as dependent upon years of schooling, experience (age minus schooling minus 5), and experience squared.[1] The regression equations for earnings are estimated from data on earnings of persons actually present in the various locations within the NMS sample.[2] A set of binomial logistic equations is also estimated relating whether or not a person is employed (for wages or in nonfarm self-employment) to schooling, age, age squared, sex, if head of household, and if 'married'.[3]

In estimating the internal migration decision functions, these earnings and employment equations are used to predict four work related explanatory

---

variable is therefore added to the South Africa migration equation indicating whether or not each male would be a novice miner if recruited. Those males not currently in South Africa are defined as novices unless they report a prior history of employment in the South African mines. Current miners are defined as novice unless their total period of engagement to date exceeds nine months (the duration of one normal contract or tour of duty).

[1] For monetisation of wages in kind see Lucas (1981*a*), appendix B. During pretesting it became apparent self-employed respondents could report their gross income (beer brewed, fruit sold, etc.) reasonably well, but a typical interviewer could not disentangle costs and net income. Separate information on typical costs per unit in each observed self-employment activity was therefore compiled from more detailed interviews and the averages applied to convert gross to net earnings. (I am most grateful to Steve Haggbladde of the Ministry of Commerce and Industry, Government of Botswana, and to Jane Scanlan, Scott Stuart and David Kunin of Dartmouth University for compiling these ratios. Details are given in Lucas (1981*a*), appendix C.) Separate earnings equations are estimated for males and for females, in each of five location types: Gaborone, other Botswana towns, large village catchment areas, peripheral, and subperipheral vca. In the urban equations two pairs of further variables are also included: the first pair measures number of years since the person arrived in town and this number squared; the second are dummies for two of the three towns (Lobatse and Francistown; Selebi-Pikwe is therefore the reference) in the group of other towns. Finally, a dummy for whether the person was solely self-employed is also explored.

[2] DaVanzo (1980) notes the selectivity bias inherent in such estimates, since observed wages depend on migration outcomes which in turn depend upon anticipated wages. No attempt is made to correct for this here. The additional problem as to whether actual urban migrants earn the same as indigenous urban residents is not an issue in Botswana since essentially everyone is a migrant and differing stages of arrival are explicitly recognised. However, a further assumption must be maintained: since only today's earnings are observed and used in estimating earnings equations, insertion of the predicted values in the migration function assumes current magnitudes are pertinent. For those who migrated some time ago this is not obvious: either one must maintain the dispersion of opportunities has remained relatively stable or that current magnitudes are those relevant in persuading past migrants not to return (even though transition costs ought to apply in reverse). In Botswana this is less of a problem than it might be elsewhere, since the vast majority of urban dwellers have arrived in the quite recent past, given the emergence of new towns.

[3] Separate employment equations are estimated for Gaborone and Lobatse, Francistown and Selebi-Pikwe, the large, peripheral, and subperipheral vca respectively. In the urban cases, time in town and its square are also inserted, plus a dummy for the second town in each class. The grouping of towns here reflects adjacent towns in north and south. (In the earnings equations, Lobatse is grouped with the northern towns, otherwise the number of earnings observations in this class is comparatively small. Estimates in each town separately suggest similar functional earnings relations for these three towns.)

variables for each individual in the migration regression, namely the daily earnings and chance of being employed both in town and in the home village.[1] In the case of absentees in town, the earnings and employment probability predicted by the equations for that town are adopted. For those not in town, a weighted average of predicted probabilities and wages for the different towns is taken.[2] In predicting, values are adopted for new arrivals in town.

The fifth employment measure reflects involvement in own account agriculture. In the survey, each person present was asked, task-by-task, whether they helped with the household's crops or livestock, and of each absentee it was asked whether they would have been involved if present. From these answers a single dummy variable is generated and included as an explanatory variable, indicating whether or not a person did or would have performed some agricultural task if present. The average versus marginal product opportunity cost quandary is thus circumvented here by including neither measure. This may either be interpreted that the included, going, rural earnings variable for each individual is a reasonable proxy for opportunity cost for those persons most likely to perform own account agricultural activities – not an implausible notion for wages of herd boys, freehold farm workers, or earnings in beer brewing – or that opportunity cost varies little from person-to-person given their involvement.

Given the central role performed by employment as a potential determinant of internal migration and the restriction of migration to South Africa for employment only, analysis is confined throughout to adults of working age who are not in school at present.[3]

B. Notes on Estimation Issues

Three particular issues with respect to estimation must be discussed before proceeding to the results. These may be emphasised by outlining a discrete choice model of the type popularised by McFadden (1974). Consider a simple utility function, in which tastes have two components: the representative consumer part wherein tastes vary systematically with some vector of measured attributes of persons; and a stochastic disturbance element, reflecting individuals' idiosyncracies. Thus:

$$u^{\alpha} = u(w^{\alpha}, e^{\alpha}, \mathbf{p}^{\alpha}, \epsilon^{\alpha}), \tag{1}$$

where u^{α}, w^{α}, e^{α} and $\mathbf{p}^{\alpha}$ are utility derived, wage, whether or not employed, and a vector of personal attributes respectively for person α and ϵ^{α} is a stochastic disturbance term. α's problem is to pick from a set of locations offering various

[1] No attempt is made here to correct earnings for locational differences in cost of living. The typical baskets of goods consumed are so different between urban and rural areas – for example in housing – as to encounter a severe index number problem. Moreover, only household expenditure (rather than consumption) data exist to establish weights in Botswana and application of such weights to deflate even cash income entails some fairly restrictive assumptions.

[2] Weights are taken according to approximate migration outcomes, differentiating between north, south and west. Since no observations exist for Orapa it is assumed similar to Selebi-Pikwe (another mining town).

[3] Working age is defined to be 15 to 54, the upper bound being imposed by precoded age category data in those instances where precise age is unknown.

wages and employment probabilities represented by vectors $\mathbf{W}^\alpha$ and $\mathbf{E}^\alpha$ for individual α. The probability of any person α, drawn randomly from the population, choosing location i is then some function:

$$m_i^\alpha = m(\mathbf{W}^\alpha, \mathbf{E}^\alpha, \mathbf{p}^\alpha, e^\alpha). \tag{2}$$

The internal migration portion of our model essentially consists of three sets of equations – the estimated version of (2), the earnings and employment equations:

$$m_i^\alpha = m(\hat{\mathbf{W}}^\alpha, \hat{\mathbf{E}}^\alpha, \mathbf{p}^\alpha), \tag{3.i}$$

$$w_i^\alpha = w(\boldsymbol{\pi}^\alpha), \tag{3.ii}$$

$$e_i^\alpha = e(\boldsymbol{\rho}^\alpha), \tag{3.iii}$$

where $\hat{\mathbf{W}}^\alpha$ and $\hat{\mathbf{E}}^\alpha$ are vectors of locational earnings and employment probabilities for person α predicted from (3.*ii*) and (3.*iii*), w_i^α and e_i^α are earnings and an employment dummy for person α present in location i, $\boldsymbol{\pi}^\alpha$ and $\boldsymbol{\rho}^\alpha$ are vectors of information about person α, and stochastic disturbances are omitted for simplicity.

Given this stylised statement of the model, an identification problem now becomes apparent. If w and e were linear functions, and if $\mathbf{p}^\alpha$, $\boldsymbol{\pi}^\alpha$ and $\boldsymbol{\rho}^\alpha$ contained the same elements, perfect collinearity would appear between $\hat{\mathbf{W}}^\alpha$, $\hat{\mathbf{E}}^\alpha$ and $\mathbf{p}^\alpha$ and the separate effects of these variables in (3.*i*) could not be identified. The identifying restrictions imposed here reside mostly in the nonlinear specifications of (3.*ii*) and (3.*iii*). In particular, the common human capital specification adopted for (3.*ii*) implies a semilogarithmic form and (3.*iii*) is appropriately estimated as a binomial logistic function. In addition, such elements as marital status and being head of household are excluded *a priori* from the wage equations but included in the employment determinants serving further to distinguish between $\mathbf{W}^\alpha$ and $\mathbf{E}^\alpha$.

The second issue arises from the use of predicted values as proxies for their unobserved counterparts, which introduces errors-in-variables. This is so, both in the sense that true, available opportunities may differ from our estimates and in that individuals may form their expectations of wage and employment availability 'irrationally'. However, given the amount of movement back and forth in Botswana, expectations are probably quite realistically formed. Provided the instrumental variables $\boldsymbol{\pi}^\alpha$ and $\boldsymbol{\rho}^\alpha$ are not then correlated with the disturbance term in the migration equations, no serious inconsistency should ensue. Certainly the use of a locational average wage as a proxy instead would normally result in biased estimates, for the deviation between average and true wage would almost certainly be correlated with the personal characteristics, $\mathbf{p}^\alpha$, appearing in the migration function. Indeed, this is precisely one of the difficulties associated with justifying the popular macro migration functions, in addition to the previously mentioned argument that taking expectations throughout a nonlinear variant of (2) would not provide an aggregate equivalent expressed in average values alone.

The remaining problem in estimation is one of sample selectivity. Section *A* mentioned an inherent problem in using absentee data on migrants: households

are likely to be selective in which migrants they will report as absentees. In turn, it is well-known that such sample selectivity can bias estimates. To correct for this potential bias a hazard rate variable is added to the explanatory variables in the migration equations.[1] In fact the hazard rate applied has three components. The first uses censal data to calculate the probability of being reported as an absentee, given that one is a migrant, for a number of population cells delineated by sex, schooling, age and location.[2] Multiplying this probability by the second component – probability of sample selection resulting from the stratification schema – a hazard rate is derived and inserted in the South Africa migration equation. Given the obvious excess supply of labour to South Africa from Botswana at present, the probability of being in the sample for estimation of the internal migration equations is then perceived as conditional upon not having migrated to South Africa. The hazard rate for internal migration is therefore defined by the product of the previous two elements and the probability of not migrating to South Africa as predicted by the binomial equation for this latter event.

III. SOME FINDINGS ON EARNINGS AND EMPLOYMENT

(*A*) *Earnings.* Space limitations preclude a full discussion of the results on earnings and employment, however the salient features may briefly be outlined. For both sexes, mean earnings are much higher in urban than rural areas – monthly wage earnings of males, for example, are about 68% higher in urban areas. Within the rural areas, earnings further decline as one proceeds from large to peripheral to subperipheral village catchment areas. Everywhere women earn less on average than men, the mean monthly wage of women in Gaborone and the large vca being about 65 and 77% respectively of males' wages.

However, after controlling for differences in schooling and experience and their effects, the intercepts in the regressions on daily earnings reveal a less clear cut picture. Some selected estimates are presented in Table 1, with standard errors of estimated coefficients in parentheses.[3] Using the overall estimates to predict earnings (in Pula per day) in the various localities for certain school, age and sex categories, one obtains the results shown in Table 2.[4]

Whereas males preserve the pattern of higher urban earnings noted for the average earnings, this is not necessarily true for women – especially those with less schooling (reflecting the low wages amongst urban domestic servants though an estimate for lodging and meals is included). Even amongst males, for those with no schooling, earnings in Gaborone and the large vca hardly differ,

[1] Heckman (1976) examines the properties of this simple estimator for the case of continuous dependent variables with normal distribution. Subsequently, Manski and McFadden (1981), have explored several maximum likelihood esimators strictly more appropriate for the case of discrete choice analysis. The hazard rate is the inverse of Mill's ratio and is given by $n/(1-N)$ where n and N are, respectively, values of the density and distribution function of the standard normal distribution.

[2] For a more exact description, see Lucas (1981*a*), appendix A.

[3] The full results are tabulated in Lucas (1981*a*), appendix D.

[4] In the urban contexts, time in town is set equal to zero as is the self-employment dummy. At the time of the survey, the Pula was worth approximately $1.20.

Table 1

Selected Earnings Equations

Dependent variable: logarithm of daily wage and self-employment earnings (Pula)

	Males		Females	
	Gaborone	Subperipheral vca.	Gaborone	Subperipheral vca.
Intercept	−0·567	−1·524	−1·400	−1·144
	(0·183)	(0·364)	(0·237)	(0·431)
Education (years)	0·143	0·145	0·205	0·150
	(0·011)	(0·029)	(0·016)	(0·036)
Experience (years)	0·060	0·079	0·026	0·054
	(0·013)	(0·024)	(0·020)	(0·029)
Experience^2	−0·0007	−0·0009	−0·00004	−0·0007
	(0·0002)	(0·00036)	(0·00043)	(0·0005)
Time in town (years)	0·058	—	0·101	—
	(0·020)	—	(0·039)	—
Time in town^2	−0·0020	—	−0·003	—
	(0·0008)	—	(0·002)	—
Self-employment dummy	−0·939	—	—	—
	(0·184)	—	—	—
R^2	0·54	0·18	0·50	0·11
F-statistic	41·69	11·57	43·10	6·10
Number of observations	220	166	222	151

Table 2

Predicted Earnings

	Predicted earnings (Pula per day)							
	Male				Female			
	Age 15 years Education (years)		Age 30 years Education (years)		Age 15 years Education (years)		Age 30 years Education (years)	
	0	10	0	10	0	10	0	10
Gaborone	0·96	2·37	1·64	4·98	0·32	1·92	0·46	2·80
Selebi-Pikwe	1·11	1·90	1·96	4·27	0·38	1·25	0·56	2·65
Large vca.	0·92	1·48	1·69	3·43	0·74	1·59	0·91	2·21
Peripheral vca.	0·76	1·86	1·14	3·34	0·69	1·53	0·98	2·69
Subperipheral vca.	0·44	0·93	0·89	2·48	0·51	1·43	0·79	2·74

though this is not true for those with more schooling, nor for those without schooling if one compares the large vca with Selebi-Pikwe (where higher wage, unskilled jobs exist in this mining town). Earnings for 15-year-old women with form 3 completed (10 years of schooling) are higher in the large vca than in Selebi-Pikwe, but this is not true of their older counterparts. A similar age effect is seen amongst educated males in comparing the peripheral villages and Selebi-Pikwe. Thus, the dangers in the common practice of adopting simple averages for urban and rural wage alternatives are well illustrated: relative wages in alternative locations depend critically upon characteristics of the individual and may even reverse at differing characteristic levels.

With the exception of the subperipheral vca, the predicted daily earnings of women do, however, remain below those of men. Thus, the sex-differential observed in the simple means is not entirely accounted for by differences in age-education structure of the male and female workforces alone. Whether the residual difference is attributable to discrimination has been the subject of considerable debate in other contexts, though certainly the tasks performed for pay by women and men are quite different in the Botswana context (Knight and Sabot (1982)).

The earnings equation estimates exhibit the usual properties. Earnings rise with schooling and with experience, eventually peaking and tending to decline as one ages. However, a couple of additional features are worth emphasising.

Earnings are higher amongst those who have been in town longer, with this time effect tapering off, as may be seen for the Gaborone case in table 1.[1] From this, one cannot discern which of three plausible effects is at work: the longer one stays in town the better the job obtained and new arrivals can emulate this advantage; the earlier arrivals have already cornered the better jobs – a vintage effect; those who fail to obtain higher wages do not stay in town for long and the observed correlation results from reversed causality. But it should be emphasised, no matter which effect underlies the observed pattern, it is the new arrivals in town who receive the lowest pay, given their age and educational background.

Secondly, our measure of earnings adds together wage and net self-employment earnings. Estimating separate equations for these two types of earnings reveals that even self-employment earnings tend to rise with schooling and with age, though the number of degrees of freedom for such estimates is small. To allow for a difference in level of earnings a dummy for self-employment was initially included in all of the earnings regressions. The coefficient on this dummy proved indistinguishable from zero except in the case of urban males, and the dummy was subsequently suppressed with this exception. This is consistent with formal sector jobs – largely male, urban jobs with government or the parastatals – paying above the self-employment alternative, while jobs in herding, as agricultural labourers or in domestic service pay close to the opportunity cost in self-employment. From the result in Table 1, wage earnings in Gaborone are estimated to be more than 2·5 times higher than in self-employment for comparable males.

(*B*) *Employment.* The probability of being either in wage work or self-employed in urban Botswana varies for males (ages 15 to 54 and not in school) from 0·74 in Lobatse and Francistown to 0·78 and 0·80 in Selebi-Pikwe and Gaborone respectively. In contrast, the likelihood of employment outside of own-account agriculture in rural areas is 0·35, 0·21 and 0·33 in large, peripheral and subperipheral village catchment areas respectively. (The rural employment rates may, however, have been particularly low owing to prevailing drought conditions during the survey.) That the likelihood of employment in the subperipheral is as high as in large vca reflects the comparative prevalence of low

[1] The implied turning point in earnings occurs at a date of arrival in town prior to establishment of the new towns and is therefore irrelevant.

paying, herding, wage jobs in subperipheral areas. In neither urban nor rural areas does reported self-employment outside of agriculture reach more than 3 to 5% among such males. Employment rates of women are lower than of men in urban areas (the former ranging from 35% in Francistown to 64% in Gaborone). The reverse is true in rural areas other than the subperiphery. However, half of female employment in large vca and two-thirds in peripheral and subperipheral areas is in nonfarm self-employment – predominantly in brewing beer and distilling khadi (often sold with cooked foods), but also in some crafts and retailing. Despite the lower incidence of self-employment by women in towns, overall urban employment rates are generally higher than rural rates for women – the latter being 0·39 in the large vca and 0·27 elsewhere, excluding own-account agricultural activities.

This brief, descriptive outline has focused on the employment rate rather than unemployment. Certainly not all of the non-employed are actively seeking work. For example, many men, after a number of contracts on the South African mines, effectively retire though still comparatively young. While not actively seeking work, some of these men would no doubt work if a suitable job presented itself. Nor are all of those not employed in the rural area active in own-account agriculture – there is a very considerable level of rural inactivity, particularly amongst men (Lipton (1978), Mueller (1985)). Unemployment is a very elusive concept in a society such as Botswana – hence the focus on the employment rate.

Some determinants of this rate are explored in Table 3, which reports the estimated, maximum likelihood, employment coefficients of the binomial logistic employment-nonemployment equations, where 'employment' includes wage work and nonfarm self-employment.[1] It must be emphasised for purposes of interpreting these results that 'employment' is certainly not synonomous with productive activity in this context. Own account agriculture activities are excluded, for example, in part because they are to be represented separately in the migration equations as discussed previously, but other productive activities such as housework are also excluded unless the product or service is meant for sale. The chi-squared statistics reported in Table 3 are for a likelihood ratio test of the null hypothesis that all coefficients except the intercept are zero. In each regression this null hypothesis may be rejected with less than 1% significance. Asymptotic standard errors of coefficients are presented in parentheses.

With respect to at least sex, education and age, one cannot distinguish effects on labour force entry as opposed to success in finding employment from these results. The probability of employment rises with schooling at a rate equal to approximately two percentage points for each additional year of schooling in

[1] The co-efficients are reported such that the non-employment co-efficients are equal and opposite in sign, rather than being normalised. The multinomial logit package used also calculates the partial derivative of the category probabilities with respect to each variable, evaluated at the sample means of the independent variables and constrained, of course, to sum to zero across categories. All statements in the text with respect to marginal effects of continuous variables refer to these values. The program also provides estimates of the standard errors of such marginal effects and, unless otherwise noted, marginal effects mentioned may be assumed significantly different from zero at the 5% significance level at least. I am most grateful to Randy Brown of Mathematica for permission to use this program developed by him.

Table 3

*Determinants of Likelihood of Employment**

	Gaborone-Lobatse	Francistown-Selebi-Pikwe	Large vca	Peripheral vca	Sub-peripheral vca
Intercept	−1·659	−0·765	−1·946	−2·179	−1·597
	(0·492)	(0·622)	(0·284)	(0·213)	(0·331)
Female	−0·452	−0·679	0·110	0·226	0·005
	(0·107)	(0·134)	(0·059)	(0·048)	(0·075)
Education (years)	0·065	0·059	0·033	0·060	0·046
	(0·013)	(0·019)	(0·009)	(0·007)	(0·012)
Age (years)	0·108	0·063	0·085	0·070	0·064
	(0·031)	(0·038)	(0·017)	(0·013)	(0·020)
Age²/100	−0·128	−0·087	−0·112	−0·084	−0·074
	(0·045)	(0·055)	(0·024)	(0·018)	(0·030)
Head of household	0·682	0·712	0·487	0·365	0·184
	(0·128)	(0·180)	(0·088)	(0·067)	(0·116)
'Married'	−0·113	−0·208	−0·021	−0·003	−0·175
	(0·114)	(0·128)	(0·067)	(0·048)	(0·085)
Time in town (years)	0·009	0·008	—	—	—
	(0·023)	(0·030)	—	—	—
Time in town²/100	−0·051	−0·076	—	—	—
	(0·096)	(0·120)	—	—	—
Lobatse/Selebi-Pikwe	0·250	0·381	—	—	—
	(0·134)	(0·129)	—	—	—
No. observations	697	469	1,477	2,981	940
Chi-squared	149·70	122·87	100·29	201·96	41·05
Deg. of freedom, χ^2	9	9	6	6	6

* Nonlinear equations, in which the binomial dependent variables are logistic functions of the explanatory variables. The dependent variable takes value one if employed for wages or in non-farm self-employment, zero otherwise.

every location's regression. No matter whether this represents a participation or unemployment effect, at least the monetary returns to schooling are then underestimated by the co-efficient in the wage equation alone. As one ages, the likelihood of being employed also tends to rise though at a diminishing rate, probably reflecting both initial unemployment among school leavers with only primary education and the greater ease with which older women can finance self employment activities. Comparing men and women of equal age and schooling, the women are significantly less likely to be employed if in town, but somewhat more likely to be in employment in the rural areas (though insignificantly so in the subperipheral areas). For example, the odds of a woman being in employment in Francistown–Selebi-Pikwe are estimated to be 0·51 of the odds for a man of comparable schooling, age and so forth, but in the peripheral areas this ratio is 1·25.

On the other hand, being head of household and marital status are more likely to affect labour force entry than success in job seeking. Heads of household – with greater responsibility for financial support of the household, at least prior

to retirement and dependence on their children – are much more likely to be in employment, whereas 'married' persons, at least in the subperipheral areas and northern towns, display a tendency to less employment, other things being equal.

The fact that migrants come to town despite existing unemployment, is now commonly modelled as a process of migration in order to search for a job. The results in Table 3 do suggest a rise in the rate of employment with duration in town – a rise which tapers off – but the effect is both very tiny and statistically insignificant. In part this may stem from measuring time in town in years, rather than examining the process within the first year, a weakness imposed by the data which report only calendar year of arrival in town. From this result a model of search successfully completed within the first months in town cannot be rejected. However, noting the effect of time in town in the earnings equations' estimates, it may also be argued that new arrivals in town cannot afford to be long unemployed and consequently soon accept a job – albeit a low paying one. The continued process of search is then for a better paying job, perhaps in the formal sector. At least after the first year in town, the advantage of a longer period of residence is in the expectation of higher wages, not in a higher probability of being employed.

IV. ESTIMATES OF THE MIGRATION EQUATIONS

A. Migration to South Africa

The estimated binomial logistic functions on the outcome of migration to South Africa in four population classes are presented in Table 4. The reported coefficients refer to the category of migrants. The null-hypothesis of no explanatory power in the approach is again rejected with more than 1 % significance for each of the four separate equations estimated.

The effect of being a novice miner on the likelihood of a man being recruited to work in South Africa is very large and statistically significant. Since 1974, the South African mines have reduced novice recruitment in general – aiming to reduce turnover of experienced workers – but have reduced foreign novice intake in particular and real mine wages have more than trebled. It is clear that most of the observed novice effect in Table 4 is today an inability of novices to be recruited rather than perpetuation of a choice not to go.

Amongst men, it is the less educated who migrate to South Africa. In part this no doubt reflects individuals' choices. Under the South African apartheid system, skilled jobs are the domain of whites, and the premium to being educated in an unskilled job is relatively small. Before 1974, the majority of Batswana migrants to South Africa's mines were under 30, yet the decline in likelihood of being a migrant with respect to age in Table 4 is small and generally insignificant for men (though this may result from excessive linearity in the specification). To some extent this no doubt reflects the much longer job retention now common as high wages are paid but only to experienced miners. Heads of household are less likely to be migrants to South Africa, though the standard error is relatively large for the west and subperipheral

Table 4

*Determinants of Migration to South Africa**

	Males			Females
	South	North	West and Subperipheral vca	South
Intercept	1·107	0·720	0·776	−1·380
	(0·242)	(0·292)	(0·449)	(0·373)
Education (years)	−0·041	−0·052	−0·054	−0·005
	(0·015)	(0·016)	(0·032)	(0·024)
Age (years)	−0·005	−0·0130	−0·012	0·020
	(0·006)	(0·0065)	(0·010)	(0·008)
Head of household	−0·543	−0·683	−0·267	−0·074
	(0·173)	(0·208)	(0·287)	(0·311)
'Married'	−0·157	0·127	−0·098	−0·221
	(0·123)	(0·127)	(0·206)	(0·178)
Household size	−0·008	−0·014	−0·003	0·016
	(0·009)	(0·009)	(0·011)	(0·015)
Number cattle owned	0·0018	−0·0015	−0·0070	0·0001
	(0·0011)	(0·0009)	(0·0038)	(0·0016)
Nonvillage home	−0·112	−0·122	0·198	−0·906
	(0·101)	(0·110)	(0·189)	(0·277)
Peripheral vca	0·163	0·212	—	0·480
	(0·090)	(0·139)	—	(0·169)
Subperipheral vca	—	—	−0·288	—
	—	—	(0·191)	—
Novice	−0·764	−0·879	−0·561	—
	(0·088)	(0·113)	(0·181)	—
Hazard rate	−48·8	−22·1	−41·1	−60·4
	(7·74)	(6·40)	(11·8)	(11·6)
Number of observations	834	1,275	553	858
Chi-squared	167.18	115·89	49·78	50·12
Degrees of freedom, χ^2	10	10	10	9

* Nonlinear equations, in which the binomial dependent variables are logistic functions of the explanatory variables. The dependent variable takes value one if in South Africa, zero otherwise.

regression. An element of simultaneity may, however, enter here, for there is no doubt some tendency to delay forming one's own household until the sojourn in South Africa ends. From the north and south, men are more likely to go to South Africa from peripheral rather than large village catchment areas. The former villages are generally lower on the poverty scale and men from these villages established their experience records before the rapid mine wage escalation and have subsequently benefitted. The lack of clear correlation of male migration to South Africa with cattle owned by their household would have been surprising before 1974, for it was the relatively poor who went, but dramatically higher wages have enabled substantial cattle accumulation out of migrants' earnings (Lucas (1981*c*, 1983)).

The pattern of female migration to South Africa is less clear, though certainly only negligible numbers of women absentees are reported in South Africa from the north and west. In contrast to men, it is the older women who go to South Africa from the south, significantly more coming from the peripheral vca and with a far lower propensity to emigrate from homes outside of the village centre (labelled non-village home). These female migrants may well include significant numbers visiting friends or relatives or shopping, as compared to men whose migration is mostly structured recruitment. Clandestine female migration to South Africa as domestic servants or farm workers is certainly under-reported.

(B) Internal migration

But the primary objective of this paper is to explore the role of earnings and employment alternatives in influencing internal migration and to these core results this section is now ready to turn. Tables 5 and 6 show two of the four estimated internal migration equations.[1] It should be emphasised that, in contrast to previous tables, 5 and 6 contain only one, multinomial logistic, regression each. The columns of these tables represent all six categories of the polytomous, dependent, migration variable.

The higher is the person's earnings if in town the more likely is that person to move to town. The higher the chance of being in employment in town the greater the probability of moving to town. But the converse is also found to be true. Namely: the higher the wage for a person in the home village and the greater the chance of being employed if at home, the lower is the propensity to migrate to town. These statements are based on the long term urban migration co-efficients normalised on any one of the categories always present at the main home, absent at the household's own lands or cattle post, or absent elsewhere in rural Botswana. Almost without exception this pattern holds up in each pair-wise normalisation for all four of the regressions and the normalised differences are almost everywhere significantly different from zero at at least the 5% significance level.[2]

This probably represents the first systematic, micro evidence on the effects of employment as well as earnings alternatives on the decision to migrate to town or not. The evidence lends considerable support to a Todaro type model though the present analysis adds a rural employment probability effect and eschews the more specific expected wage formulation.[3]

What are the magnitudes of these effects? Computing the estimated partial

[1] The two not tabulated are for the west and the large and peripheral vca in the south.

[2] The only notable exceptions occur in the effect of local wage in the south, which has the wrong sign but is indistinguishable from zero when normalised on any of the three rural categories, and none of the effects differs from zero in the west when normalised on own lands or cattle post. In the west, however, the village, lands, cattle post settlement pattern is far less well defined than in the east. The category of new arrivals is generally similar to those always present at home.

[3] Internal migration equations were also estimated inserting expected wages (the product of predicted earnings and employment probability) for both town and home village, while retaining the four separate components as above. In no instance did the partial derivative of the probability of long term migration to town with respect to employment or earnings (evaluated at the sample mean from the linear plus interactive terms) significantly reverse sign with this insertion. Indeed, in almost all cases the sign of these derivatives remained consistent with the discussion in the text.

Table 5

*Internal Migration from Large Village Catchment Area Homes in the North**

$N = 593$ $\chi^2(65) = 372{\cdot}30$	Always present	Own lands or cattle post	Other rural areas	Botswana towns: short term	Botswana towns: long term	New arrivals
Intercept	0·861	1·164	2·358	−0·308	−5·429	1·353
	(1·228)	(1·455)	(1·092)	(1·515)	(1·393)	(1·386)
Female	−1·825	−1·434	−1·200	0·384	5·359	−1·284
	(0·962)	(1·217)	(0·854)	(1·285)	(1·061)	(1·095)
Education (years)	0·203	0·029	0·081	0·001	−0·375	0·060
	(0·078)	(0·090)	(0·067)	(0·101)	(0·088)	(0·086)
Age (years)	0·049	0·023	0·011	−0·034	−0·056	0·007
	(0·016)	(0·019)	(0·016)	(0·029)	(0·023)	(0·019)
Head of household	−0·075	0·317	0·738	0·981	−2·131	0·171
	(0·566)	(0·688)	(0·528)	(0·928)	(0·866)	(0·658)
'Married'	−0·426	0·007	−0·132	0·132	0·650	−0·231
	(0·292)	(0·335)	(0·274)	(0·443)	(0·363)	(0·339)
Local employment	7·073	−1·258	−0·053	−2·225	−8·639	5·103
	(3·104)	(3·928)	(2·82)	(4·809)	(3·848)	(3·619)
Local earnings	1·526	2·376	1·591	−2·664	−3·890	1·061
	(0·818)	(1·119)	(0·745)	(0·955)	(0·855)	(0·935)
Urban employment	−5·627	−3·248	−2·357	1·825	13·521	−4·113
	(2·524)	(3·068)	(2·274)	(3·233)	(2·722)	(2·900)
Urban earnings	−1·746	−1·929	−1·279	1·994	4·035	−1·075
	(0·739)	(1·004)	(0·665)	(0·833)	(0·762)	(0·835)
Own account agric.	−0·941	1·982	−0·290	−0·024	0·324	−1·051
	(0·216)	(0·281)	(0·189)	(0·296)	(0·257)	(0·255)
Household size	−0·022	0·005	−0·007	0·023	0·007	−0·007
	(0·016)	(0·018)	(0·014)	(0·020)	(0·018)	(0·017)
Number cattle owned	−0·0004	0·0029	0·0013	−0·0029	−0·0003	−0·0005
	(0·0010)	(0·0010)	(0·0008)	(0·0018)	(0·0014)	(0·0012)
Hazard rate	27·14	−32·03	−20·49	41·04	−36·74	21·08
	(8·49)	(11·65)	(8·45)	(12·36)	(12·47)	(9·38)
Fraction of observations	0·231	0·157	0·239	0·083	0·143	0·147

* A single, nonlinear equation in which the polytomous dependent variable is a logistic function of the explanatory variables. The dependent variable takes discrete values one through six according to the person's absentee status.

derivatives of the probability of long-term migration to town with respect to each of the four employment variables (evaluated at the sample means), reveals that a null-hypothesis of equal but opposite effects of the urban compared to the home area pairs of employment variables cannot be rejected.[1] In other words, a small rise in urban earnings has approximately the same effect on promoting long-term urban migration as does an equal absolute decline in rural earnings, and a similar statement holds for employment probabilities. Using these partial derivatives to approximate the effect of a 5% increase in urban

[1] The solitary exception is for wages in the south, where local wage has the wrong sign but is statistically indistinguishable from zero.

Table 6

*Internal Migration from Subperipheral vca Homes in the North and South**

$N = 708$ $\chi^2(70) = 359{\cdot}13$	Always present	Own lands or cattle post	Other rural areas	Botswana towns: short term	Botswana towns: long term	New arrivals
Intercept	8·182 (2·015)	8·766 (2·659)	4·361 (1·484)	−12·830 (3·054)	−14·063 (2·934)	5·583 (1·838)
Female	−8·872 (1·861)	−2·930 (2·498)	−3·495 (1·321)	8·384 (2·493)	11·679 (2·439)	−4·766 (1·670)
Education (years)	0·144 (0·105)	0·241 (0·124)	0·071 (0·083)	−0·270 (0·204)	−0·149 (0·196)	−0·037 (0·100)
Age (years)	0·020 (0·033)	0·052 (0·046)	0·016 (0·030)	−0·044 (0·084)	−0·036 (0·084)	−0·008 (0·035)
Head of household	3·141 (0·712)	2·382 (0·994)	0·493 (0·537)	−4·236 (1·397)	−2·365 (1·102)	0·584 (0·642)
'Married'	0·365 (0·427)	−0·246 (0·611)	0·101 (0·369)	0·165 (0·968)	−0·973 (0·926)	0·588 (0·430)
Local employment	13·363 (7·639)	3·855 (11·889)	9·210 (6·011)	−7·952 (13·409)	−23·290 (11·215)	4·813 (7·829)
Local earnings	5·330 (1·677)	−2·919 (2·577)	1·055 (1·074)	−3·152 (1·963)	−4·197 (2·050)	3·882 (1·360)
Urban employment	−13·449 (3·949)	−11·082 (6·141)	−8·343 (2·969)	16·334 (5·284)	23·043 (4·723)	−6·502 (3·860)
Urban earnings	−5·575 (1·153)	0·477 (1·478)	−0·874 (0·714)	3·406 (1·211)	5·223 (1·363)	−2·658 (0·907)
Own account agric.	0·019 (0·200)	−0·003 (0·288)	−0·038 (0·174)	0·858 (0·405)	0·191 (0·401)	−1·028 (0·210)
Household size	0·005 (0·016)	−0·081 (0·026)	0·015 (0·013)	−0·006 (0·034)	0·075 (0·027)	−0·008 (0·016)
Number cattle owned	0·0010 (0·0012)	0·0049 (0·0014)	0·0006 (0·0010)	0·0012 (0·0020)	−0·0104 (0·0035)	0·0027 (0·0011)
Non-village home	−0·531 (0·225)	−0·464 (0·306)	0·784 (0·206)	−0·200 (0·456)	0·186 (0·444)	0·225 (0·229)
Hazard rate	101·64 (22·55)	−67·58 (23·10)	−13·62 (15·24)	45·56 (42·22)	−71·35 (34·59)	5·35 (18·36)
Fraction of observations	0·264	0·072	0·383	0·034	0·040	0·208

* A single, nonlinear equation in which the polytomous variable is a logistic function of the explanatory variables. The dependent variable takes discrete values one through six according to the person's absentee status.

earnings from its sample mean, would imply a rise in propensity to long term urban migration ranging from about 5 percentage points from large vca in the north to about 0·4 percentage points from the west and subperipheral east. Given the relative population bases of rural and urban Botswana (a little over 15% urban), even if 1% of rural inhabitants moved to town, the consequences for the towns would be quite substantial.

A 1 percentage point increase in the chances of being employed in town is estimated to raise urban migration propensities by amounts ranging from 1·6

percentage points from the large vca of the north to 0·1 and 0·3 percentage points in the west and subperipheral areas respectively.[1] In other words, the creation of each additional urban job is estimated to attract more than one additional adult to town – the essential point with regard to urban shadow wage in the Harris–Todaro framework – though one cannot discern how many of the extra adults come as dependents as opposed to active job seekers.[2] To countervail this effect through rural job creation at present pay requires several rural jobs for each urban job – approximately in proportion to relative adult population, given the roughly equal and opposite effects, if no migration is to be induced and all new urban jobs are thus to be filled by prior, non-employed urban dwellers.

On the other hand, a marginally greater likelihood of being involved in own account agriculture if in the rural home area does nothing to deter movement to town either on a short or long term basis. Those who would be active in own account agriculture if in the home area are slightly more likely to go to town on a short term rather than a long term basis – perhaps reflecting seasonal returns of the potentially more useful members – but this is untrue in the north and is nowhere statistically significant. In this semi-arid country, with recurrent crop failures and cheap grain imports from South Africa within the Southern African Customs Union, crop husbandry is simply not a very viable proposition, and minding one's own livestock takes little time for the typical household. Yet own account agriculture does have an effect on intrarural movements.[3] As one should expect, given the disposition of lands and cattleposts described earlier, commitment to family farming necessitates movement from the main home to own lands-cum-cattle-post. The data confirm a very strong and statistically significant movement in this regard but only in the somewhat bigger villages in the east (the north and south regions). In the west and subperipheral east, many fields are closer to the main home so that actual relocation for involvement in own account agriculture is not necessary and is statistically indistinguishable from zero. Those who remain at one of the household's

[1] If the marginal estimated effects are used to evaluate the elasticities of long term urban migration with respect to local and urban earnings and employment, in all but two cases these elasticities prove greater than one. Thus, the weighted averages of these estimated elasticities across regions are:

	Employment	Earnings
Local	−1·61	−1·72
Urban	3·81	2·46

[2] To illustrate, consider a rough calculation for Botswana. Suppose the *de facto*, adult population of rural and urban Botswana are initially 450 and 75 thousands respectively and the employment rate in town is 0·6. If the marginal co-efficient on urban employment for the average person is 0·66 (approximately the weighted average of the locational marginal co-efficients), a rise in probability of urban employment by 1 percentage point would attract about 3,000 new migrants. To raise the employment rate to 0·61 requires 2,580 extra jobs – 750 for the initial residents, 1,830 for the new migrants. Thus 2,580 extra jobs attract 3,000 adults to town, a ratio of nearly 1·2 to 1, leaving 420 more adults non-employed in town than before the new jobs were created. In contrast to the Harris–Todaro model, it is, however, dubious in the Botswana context just how productively employed these 420 people would have been in the rural sector, given the low village employment rates and evidence on low marginal product in own account agriculture. (See Harris and Todaro (1970) and Lucas (1981 *b*)).

[3] The apparent effect of own account agriculture on new arrivals is an artifact. This independent variable, as described in part II, is constructed from cumulative data across all four rounds of the survey. By definition, new arrivals report no such information in earlier rounds and hence appear less likely to have a responsibility in the household's agricultural tasks.

dwellings (either main home, lands or cattle post) throughout the year, rather than moving to other rural areas, are more likely to be committed to own account agriculture though the effect is statistically significant only in the north and south cases. Thus, involvement in or potential responsibility for family agricultural tasks tends to encourage intra-rural mobility amongst the household's dwellings at least in the east, and positively discourages other forms of intra-rural mobility.

The direct role played by personal characteristics in influencing migration – as described in connection with $\mathbf{p}^{\alpha}$ in (2) – may now be briefly considered. In the sample, slightly less than half of adults reported absent in urban Botswana are female, though more than half of the total sample is female (after subtracting absentees in South Africa.)[1] Moreover, more than half of the long term absentees in town are male. Using the regression estimates to control for differences in employment opportunities and so forth, these simple correlations are, however, reversed. *Ceteris paribus*, being female significantly and substantially raises the chance of migrating to town on a long-term basis, rather than falling into any one of the three rural categories. The same tends to be true of short-term urban migration, but this latter effect is uniformly significant at the 5% level only for subperipheral areas. The particular attraction of towns to the younger women is probably partly linked to the relative dearth of eligible male partners in rural Botswana attributable to the joint effects of South African and urban migration by men.[2] Except in the west it is, moreover, estimated that being female raises the propensity to move elsewhere in the rural sector other than to the household's own dwellings, again reversing the simple correlation. It seems that, given employment alternatives, commitment to agricultural tasks, age, education and so forth, women tend to be inherently more peripatetic in Botswana.

The most highly educated group on average in the sample are the long-term migrants, followed by the short-term urban migrants. This is consistent with the findings in most LDCs and is usually interpreted as education causing urban migration. However, there are really two components – one working through the effects of education on relative job opportunities, the other is a direct attitudinal factor. In fact, for the average individual, it is estimated that a marginal increase in education would significantly reduce the probability of being a long-term urban migrant from both the north and south and has no effect elsewhere, given employment and wage opportunities. That education inculcates attitudes only conducive to urban life is rejected. Education does, however, tend to promote urban migration, not by changing tastes, but by relative improvements in the expected earnings in town.

Age, on the other hand, retains a partial effect in a similar direction to its simple correlation – the young tend to migrate to town, the older folk remain always present at the main home. Though consistent with the 'bright lights' of town attracting the young, a more likely underlying cause is that the young

[1] Women tend to be under-reported as absentees compared to men, an element recognised in computing the hazard rate.

[2] On migration and the role of women in Botswana, see Izzard (1981).

have a longer life horizon over which to reap the returns to moving into town (Sjaastad (1962)).

Heads of household might be anticipated to be more committed to staying at home. Comparing the effect of being household head on long term urban migration relative to being always present at the main home and to being absent at one of the household's other dwellings (at the lands or cattle post), this portion of the hypothesis is everywhere maintained.[1] The difference in co-efficients is less than zero in every instance at the 10% significance level and usually at the 5% level. But this does not necessarily mean household heads tend to stay at one of the household's rural dwellings. For example, in the south being head raises the probability of short-term movement to town or elsewhere in the rural area, and this tends also to be true in the north though confidence levels are lower there. The subperiphery acts in the opposite direction in this latter respect. It seems heads of rural households tend not to move to town permanently, but may use their privileged position to take shorter journeys from home.[2]

Adults from larger households – as measured by number of members present or reported absent – are more likely to migrate to town (particularly on a long-term basis) and also more likely to move elsewhere within the rural sector. Thus, one can be at least 90% confident a marginal rise in household size increases the probability of long-term migration to town from the south, subperiphery and west and increases the likelihood of movement within the rural sector in the last two cases also. Other instances are generally in the same direction though with lower levels of confidence. These findings are consistent with the notion that the opportunity cost of leaving a rural home diminishes as household membership expands, though amounts of land, agricultural equipment and capital associated with self-employment are not maintained constant in the regression.[3] But the magnitude of these effects is not very large: an extra member of the household raises the probability of any adult moving to another rural area or to town by only one percentage point or less in each of the regions.

In so far as migration is perceived as an investment, greater wealth might be expected to promote out migration by lowering the costs of financing. On the other hand, enhanced rural assets may detract from departure rates when incomplete rental markets exist or the need for personal supervision prevails – which certainly seems to be the rule for cattle in Botswana. Together with any effects of wealth on risk aversion and on consumption of psychic benefits associated with any particular milieu, this leaves uncertain whether wealth accumulation should be hypothesised to swell the migration flow. *Ceteris paribus*, our results show greater wealth – as proxied by number of cattle owned – tends on the margin to discourage migration to town or to rural areas other than the

[1] No co-efficient on head of household in the west is estimated, owing to perfect categorisation in the data.

[2] The use of the term rural households is to emphasise that one cannot rule out any tendency for heads to move to town accompanied by their household, thus establishing a truly urban household and disappearing from our sample.

[3] In Lucas (1981*b*), I showed, with data from an earlier survey, that both land and crop equipment are positively associated with number of members present in the household in Botswana.

household's own dwellings, though the magnitude of the effect is certainly very small.[1] It should, however, be remembered that wealth may also act indirectly on migration. For example, the wealthy are more likely to educate their children and, as shown in Table 3, education does more to raise the probability of employment in town than in the villages, thus promoting urban migration amongst the wealthy.[2]

Finally, we may examine certain locational effects. Not surprisingly, whether the main home of the individual is a village or nonvillage (lands-cum-cattle-post) home does significantly affect intrarural migration. In each case, adults from nonvillage homes are more likely to move elsewhere in the rural sector at some point in the year, as compared to persons from village homes.[3] Whether this is necessitated by employment availability, some other form of social intercourse in the village or seasonal lack of water cannot be discerned. Moreover, persons from nonvillage homes are significantly less likely to move to another household dwelling during the year rather than to remain at the main home – which is not unexpected when the main home is already outside of the village – though no difference is observed in the subperiphery where village – non-village distinctions are anyway blurred.

The effects of the regions themselves, as distinct from variations in employment opportunities and personal attributes of their inhabitants, are reflected in the intercept terms. Of the four regions distinguished, the west and subperiphery are more remote from the towns along the railway line and represent much smaller settlements than do the north and south regions. In pair-wise comparisons of intercepts it is readily seen that short-term movements into town are less frequent from the more remote regions, presumably reflecting the deterrent effect of distance. Long term migration to town is also less frequent, yet originating from the west is less a deterrent to long-term urban migration than is coming from the subperiphery even though the west is further removed from the towns. Moreover, the difference in intercepts is less for long-term than for short-term movements to town, suggesting that distance is much more a barrier to short term mobility into town than to long-term relocation. Lastly, intrarural migration beyond the household's own dwellings is more likely in the west and subperiphery than in the north and south, (though the difference is insignificant in comparing the north with the subperiphery.) When distance deters visits to town, intrarural movement, probably to larger settlements, may thus be a substitute, perhaps also leading to step migration into town in the long run.

[1] An exception is in the north, where possession of more cattle is associated with marginally higher intrarural migration.

[2] The raw data exhibit no clear pattern between cattle owned and whether the person is a migrant or remains at one of the household's dwellings, across the four regions. It should be mentioned though that there may be a selectivity problem in this regard in the Botswana data. Certainly some heads of urban households are well-known to possess very large herds of cattle. But these men may well be omitted from our sample since this group in particular is not likely to be considered absent from a rural home though many are first generation urban dwellers.

[3] The non-village dummy is omitted in the north, owing to perfect categorisation in the data.

V. CONCLUSIONS

Some determinants of whether rural-based Batswana adults remain at home, migrate to town on a short- or long-term basis, move intrarurally either to another household dwelling or elsewhere, or emigrate to South Africa have been examined. Both the effects of employment alternatives on the rate of internal migration and the selectivity of movements with respect to personal, household and locational factors were analysed.

It has been shown that the more likely a person is to be employed if in town, and the higher is the wage if employed, the greater is the chance that person will migrate to the urban sector on a long term basis. Moreover, at present rates of pay, the magnitude of the employment effect is sufficiently strong as to draw more than one adult to town for each extra urban job created, leaving more non-employed adults in town than before the job creation. Both the Todaro and Harris–Todaro hypotheses are thus supported. It is interesting to speculate on the reasons for this over-reaction to urban job creation and particularly to relate it to the standard job search model. To do so, it will be useful to recapitulate certain patterns discerned in the dispersion of earnings and employment among adults.

Earnings are found to rise significantly the longer one has been in town, though one cannot disentangle whether this reflects future prospects for new migrants, a vintage effect in reserving the better jobs, or tenacity with respect to urban residence selective on wage. The probability of being in employment also tends to rise with duration of stay in town (measured in years), but the effect is very small and statistically insignificant. Thus, at least after the first year in town, the advantage of a longer period of residence is in the expectation of higher wages, rather than in a higher probability of being employed. The corollary signifies that it is the new arrivals in town who receive the lowest pay, given their age and educational background.

Given these patterns, at least three possible interpretations of the over-reaction effect may then be suggested. The first is that extra jobs indeed attract additional job seekers. But success in finding and retaining jobs is not cumulative with time in town – those more likely to be unemployed continue to move in and out of jobs. The rise in earnings with duration of urban residence could still represent a future prospect available to the average newcomer, but presumably pay is less likely to rise amongst those continuing to oscillate between work and unemployment. The second possibility is that success in job search is cumulative but hidden by our data which report time in town only in years. Creation of extra jobs may then attract new searchers who find work within a few months of arrival, albeit at low pay, then continue to search for better paying jobs while employed. In a third potential scenario the additional non-employed persons may not be seeking jobs at all, but come with the job holder as dependent adults. Nonetheless, whatever mixture of these mechanisms is correct, the essential question for the shadow price of urban labour remains the rural opportunity cost of all those drawn to town per job created, though

appropriate definition of this opportunity cost has many well-known difficulties, particularly when rural labour productivity is highly seasonal.

Given the paucity of current crop cultivation opportunities in Botswana's semi-arid environment, wage and nonagricultural self employment form relatively more important alternatives in rural Botswana than in most LDCs. It was found that the greater are the chances of being in such wage work or self employment in the home village area, and the better is the pay for these tasks, the less likely a person is to move into town. But the pay differential between urban and rural sectors is so substantial, that for each extra job generated in town between five and six additional jobs are required in the rural sector if no fresh migration is to be induced. On the other hand, the risks of foot and mouth disease plus drought and availability of cheap grain imports within the Southern African Customs Union, render involvement in own account agriculture sufficiently unattractive as to fail to deter urban migration at all. Responsibility for tasks in the household's crop and animal husbandry does, however, significantly affect intrarural mobility, both promoting much greater movement between the various dwellings possessed by a household and simultaneously detracting from the propensity to migrate elsewhere within the rural sector.

As to selectivity of migration, it has been shown that after controlling for differences in employment opportunities amongst individuals, the simple correlations between propensity to migrate and certain personal characteristics may even be reversed. For example, educated persons are more likely to move into town. However, this well-known tendency is estimated to result from improvement in urban employment openings relative to rural opportunities as education increases, rather than from a change in taste for migration as such. Thus, although rural education does promote urban migration, this would change if better rural employment opportunities for the educated were available or, perhaps more realistically, if less emphasis were placed on creating particularly attractive urban jobs for the better educated. Selectivity with respect to age is not, however, reversed by the multivariate approach. Younger adults from villages nearer to the towns tend to move into town irrespective of current employment opportunities, perhaps reflecting the longer horizon over which the young may expect to reap any given level of benefits.

Thus, the technique of predicting employment and earnings alternatives from estimated micro equations has been shown to provide quite plausible results on migration responsiveness to locational work opportunities. Moreover, the identifying restrictions imposed have permitted a rich array of estimates on the separate effects of personal attributes. The natural extension of this type of approach, besides exploring different specifications and identifying restrictions, might be to the selection of individual towns or more specific rural areas by migrants. However, the data requirements, both for estimating the auxiliary employment and earning equations in each locale and to obtain sufficient observations within each discrete migration class, rapidly become quite massive.

Boston University

Date of receipt of final typescript: August 1984

References

Behrman, Jere R. and Wolfe, Barbara L. (1985). 'Micro determinants of female migration in a developing country: labor market, demographic marriage market and economic marriage market incentives.' In *Research in Population Economics*, vol. 5 (ed. T. P. Schultz and K. I. Wolpin). Greenwich: JAI Press.

DaVanzo, Julie. (1976). *Why Families Move: A Model of the Geographic Mobility of Married Couples.* Santa Monica: The Rand Corporation, R-1972-DOL.

—— (1980). 'Microeconomic approaches to studying migration decisions.' Rand Note no. N-1201-NICHD, The Rand Corporation (May).

Harris, John R. and Todaro, Michael (1970). 'Migration, unemployment and development: a two-sector analysis.' *American Economic Review*, vol. 60 (March), pp. 126–42.

Hay, Michael, J. (1974). 'An econometric analysis of rural-urban migration in Tunisia.' Ph.D. dissertation, University of Minnesota.

—— (1980). 'A structural equations model of migration in Tunisia.' *Economic Development and Cultural Change*, vol. 28 (January), pp. 345–58.

Heckman, James J. (1976). 'The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models.' *Annals of Economic and Social Measurement* vol. 5 (Fall), 475–92.

Izzard, Wendy. (1981). 'The impact of migration on the roles of women.' Paper presented at the National Migration Study Conference, Gaborone, Botswana, (December).

Knight, J. B. and Sabot, Richard H. (1982). 'Labor market discrimination in a poor urban economy.' *Journal of Development Studies* vol. 19 (October), pp. 67–87.

Lipton, Michael. (1978). *Employment and Labour Use in Botswana.* Ministry of Finance and Development Planning, Gaborone.

Lucas, Robert E. B. (1975). 'The supply-of-immigrants function and taxation of immigrants' incomes.' *Journal of Development Economics*, vol. 2 (September), pp. 289–308.

—— (1981 *a*). 'Determinants of migration decisions in Botswana.' Paper presented at the National Migration Study Conference, Gaborone, Botswana (December).

—— (1981 *b*). 'The distribution and efficiency of crop production in tribal areas of Botswana.' Discussion Paper, Population and Human Resource Division, World Bank.

—— (1981 *c*). 'Out migration, remittances and investment in rural Botswana.' Paper presented at the National Migration Study Conference, Gaborone, Botswana (December).

—— (1982). 'Mines and migrants in South Africa.' Unpublished paper. Boston: Boston University.

—— (1983). 'Emigration, employment and accumulation: the miners of Southern Africa.' Discussion Paper no. 4, Migration and Development Program, Harvard University (December).

Manski, Charles F. and McFadden, Daniel (1981). 'Alternative estimators and sample designs for discrete choice analysis.' In *Structural Analysis of Discrete Data*, (ed. Charles F. Manski and Daniel McFadden). Cambridge: MIT Press.

McFadden, Daniel. (1974). 'Conditional logit analysis of qualitative choice behavior.' In *Frontiers in Econometrics*, (ed. Paul Zarembka). New York: Academic Press.

Mueller, Eva. (1985). 'The value and allocation of time in rural Botswana.' *Journal of Development Economics.*

Schapera, I. (1943). *Native Land Tenure in the Bechuanaland Protectorate.* The Lovedale Press.

Sjaastad, Larry A. (1962). 'The costs and returns of human migration.' *Journal of Political Economy*, vol. 70 (October), pp. 580–93.

Todaro, Michael P. (1969). 'A model of labor migration and urban unemployment in less developed countries.' *American Economic Review*, vol. 59 (March), pp. 138–48.