A Stochastic Learning Model of Migration

John G. Cross

CENTER FOR RESEARCH ON ECONOMIC DEVELOPMENT
The University of Michigan
Ann Arbor, Michigan 48108

Discussion Paper No. 65

February 1977

ABSTRACT

A Stochastic Learning Model of Migration

While the comparative-static model of labor force location is straightforward enough, the theory of the dynamic process of migration is incomplete in that it fails to explain how it is possible for relatively low rates of migration to persist over long periods of time in the presence of a divergence between wage rates (corrected for unemployment) in various localities. Most migration theorists seem to assume that workers take rational account of unemployment rates when they compare the income potentials of alternative locations, but that the great majority of workers irrationally fail to act upon these comparisons once they are made. In this note, a migration theory is developed which relies entirely upon a theory of learning and information transmission, rather than comparative static optimization procedures. It is shown that this leads to the formulation of a migration model which is quite similar to the more intuitively based migration equations which are found in the current literature.

* * *

Lors que le modèle statique comparatif de l'emplacement de la force de travail est assez simple, la théorie du processus dynamique migratoire est incomplète en ce sens qu'elle échoue à expliquer comment des taux relativement bas de migration peuvent persister pendant de longues périodes, à un moment où existe une divergence entre les taux salariaux (rectifiés pour le chômage) dans diverses localités. La plupart des théoriciens de la migration semblent assumer que les travailleurs tiennent rationnellement compte des taux de chômage lorsqu'ils comparent les potentiels de revenu des emplacements alternatifs, mais que leur grande majorité échoue irrationnellement à agir d'après ces comparaisons une fois qu'elles ont été faites. Dans cet article a été développée une théorie migratoire basée entièrement sur une théorie de l'étude et de la transmission des renseignements plutôt que sur des procédés d'optimisation statique comparatifs. Il est démontré que ceci entraine la formulation d'un modèle de migration tout à fait semblable aux équations migratoires plus basées sur l'intuition et que l'on trouve dans la documentation courante.

A STOCHASTIC LEARNING MODEL OF MIGRATION*

John G. Cross

In the growing literature concerning behavioral, adaptive, or otherwise non-traditional theories of consumer and firm behavior, a great deal of stress is placed upon the need for models which are as effective in describing markets which are out of equilibrium as the traditional optimization theory is in describing comparative static equilibria.[1] Although most of the models which are developed in this vein are indeed well-defined under conditions of incomplete market adjustment, examples of the specific contributions which these might make to particular economic problem areas are not plentiful. The purpose of this note is to make one such application, using a simple stochastic learning model to support a theory of rural-urban labor force migration, particularly as it occurs in less-developed countries.

The comparative-static theory of labor-force location is straightforward. In most papers (such as Todaro [7], Harris-Todaro [3], Stiglitz [6]), it is maintained that the urban labor supply in less-developed countries is determined not by the prevailing wage rate alone, but by a composite of the urban wage and the (typically high) urban

*Although the theory described here is quite different, this note was originally stimulated by a suggestion by O. Onyemelukwe to the effect that a learning process may underlie village attempts to receive income by exporting labor to urban areas. I would also like to thank Richard C. Porter for comments on an earlier draft.

unemployment rate. Individual workers are assumed to compare the income which they can earn from employment in rural agriculture to that which could be received in the city, where the expected urban income is computed from the higher urban wage and various assumptions regarding the probability of finding employment there.

In the cases of the Todaro and Harris-Todaro papers, however, a good deal of stress is placed upon the importance of the dynamic process of migration itself. Todaro, for example, presents as his main hypothesis the relation

$$\frac{\dot{S}(t)}{S(t)} = F\left[\frac{V_u(t) - V_r(t)}{V_r(t)}\right] \tag{1}$$

where S is the size of the urban labor force and $\dot{S}$ its time derivative, $V_u(t)$ is the discounted present value of expected urban income, and $V_r(t)$ is the discounted present value of expected rural income. The Harris-Todaro paper similarly draws attention to migration as an ongoing phenomenon with the introductory observation that "migration not only continues to exist, but, indeed, appears to be accelerating" ([3], p. 126).

Despite these suggestions that the central problem of migration is its persistence, possibly reflecting a continuing dynamic adjustment process, even these papers revert to the simple comparative-static model in order to produce their main results. In the Harris-Todaro paper, models are formulated in which the expected urban wage is made equal to the rural wage (that is, in which migration would be zero), and the resulting equilibrium urban unemployment rate is investigated in the light of various tax and minimum wage policies. Even the original Todaro paper concentrated upon a state in which migration is zero except for an "equilibrium" amount which is calculated by dividing the rate of urban job creation by the employment rate and then subtracting the rate of natural urban population growth.

This reluctance to work with the migration phenomenon itself may be attributable to the fact that dynamic models such as equation (1), plausible as they certainly are, are not derived from any formal theory, but are simply stated as initial hypotheses. The observation that a rational household must take account of the presence of unemployment in the city, and thus must compare rural income to an expected value for urban income can only be used to derive an equilibrium model for which migration is zero when the returns from the two alternatives are equal. It cannot be used to explain a stable rate of migration when the returns are unequal. Ordinary rates of migration can only occur if many individuals who could migrate do not do so even if the expected urban wage substantially exceeds the rural wage. Of course, a variety of independent variables come immediately to mind which could be used to account for this sluggishness (and many of these play an important role in econometric studies of migration), but none of them are introduced explicitly into the theory. As a consequence, equation (1) states, in effect, that migration decisions are based upon a rational comparison of income alternatives, but that a large fraction of the population irrationally fails to act on this comparison. That the size of this fraction should vary inversely with the difference between V_u and V_r is plausible, but it is nowhere explained by the theory itself. In short, optimization theory, not being defined for disequilibrium, is simply inadequate to the task.

A Learning Model

The theory which we plan to use here is incorporated in the Bush-Mosteller stochastic learning model as presented by Cross [1]. The approach is entirely behaviorist in that, as in most psychological learning models, "decisions" of individuals are treated as random variables whose likelihoods are dependent upon each individual's own previous experience and not upon any kind of explicit expected income or utility calculations. The variables are therefore similar in spirit to those found in sample survey studies of migration,[2] except that we will not use migration probabilities as the dependent variables but focus instead upon a workers' choice of location: the individual worker, i, will be found in the urban sector during period t with a probability P_t^i and in the rural sector with a probability $1-P_t^i$. The value of P_t^i is determined from experience, and there is no presumption that the worker "knows" anything about the market beforehand.

In this regard, learning models address the problem of uncertainty (in this case the uncertainty of employment in the city) in quite an unconventional way. Traditional maximization approaches require that uncertainty be handled with statistical estimation and search procedures, many of which employ quite sophisticated techniques.[3] In some cases, these even demand some prior information as to the probability distributions which characterize a market. Whereas optimization models therefore require greater and greater sophistication on the part of individual decision-makers as the importance of uncertainty grows, this learning model will make essentially the same assumptions under both uncertainty and certainty. The mathematics found in the two alternative approaches are often similar, but the interpretations are wholly different.

Since the purpose of this paper is to focus upon the economic factors which may influence migration, we will concentrate here upon the economic experience of the worker. A more general model could employ variables reflecting such factors as age, education, family size, or population density, in just the same way in order to derive a more sophisticated view of the migration process. For the purposes of our simple model, we will characterize the objective situation as follows. There is one major urban center to which migration is possible. If worker i locates himself in this city, he will find "modern sector" employment which pays a high wage, W_u, with a probability q_i. If he fails to find such a job there still may exist various forms of marginal employment which will pay a very low wage W_o. If he does not live in the city, rural employment guarantees an income of W_r.[4] In general, we have $W_u > W_r > W_o$. The likelihood q_i should depend upon quite a number of variables such as the age of the worker, the length of his stay in the city, and, most important, whether or not he had a modern sector job in the previous period t-1. In the face of the potential complexity in determining q_i, however, it is common to assume (following Harris-Todaro) that jobs are distributed randomly, and since we are only concerned in this note with demonstrating the usefulness of a learning model, we will preserve this assumption. This makes q_i equal to the employment rate itself, $\frac{E}{S}$, where E is the number of modern sector jobs.

Migration

Beginning with the likelihood P_t^i, we use the actual experience of the individual during t to modify this likelihood to a new value P_{t+1}^i. Naturally, both P_t^i and P_{t+1}^i must be bounded between 0 and 1, and we expect P_{t+1}^i to vary positively with the degree of success which is encountered at the location which is chosen. A modification of the well-known Bush-Mosteller learning model provides the simplest function which meets these conditions. For example, if i lives in the city and finds a job, then we write

$$P_{t+1}^i = P_t^i + \alpha_i(W_u)(1-P_t^i) \tag{2a}$$

where the function $\alpha_i(W)$ describes the rate of learning as a function of the reward (wage) magnitude. We emphasize that "learning" here is not to be interpreted in the sense of "finding out," and that the worker is not being described as someone attempting to estimate q_i. Instead, $\alpha_i(W)$ simply reflects the empirical observation that actions which are met with success tend to be repeated. $\alpha_i(W)$ has the general properties

$$0 < \alpha_i(W) < 1, \ \alpha_i'(W) > 0, \ \alpha_i''(W) < 0,$$

but it is most convenient here to approximate this function with the linear form $\alpha_i'W + \alpha_0$ where α_i' is the slope of the function $\alpha_i(W)$ in the vicinity of W.

Even though $P_{t+1}^i > P_t^i$ in this example, the worker may nevertheless go back to the rural sector at the end of the period--that is, he quits and goes home with probability $1-P^i_{t+1}$. In this regard we are already departing from the properties of optimization theory. In both the Todaro and Harris-Todaro models, it is assumed that anyone who migrates to the city stays there so long as the expected urban economic opportunities are greater. In fact, however, some out-migration, even among the urban employed, is a common phenomenon in less-developed countries, and a dynamic model ought to reflect that fact.

The value which is taken by P^i_{t+1} depends upon whether the worker lives in the city and whether he gets a job. If the worker fails to find an urban job, then:[5]

$$P^i_{t+1} = P^i_t + (\alpha_1^! W_o + \alpha_o)(1-P^i_t) \tag{2b}$$

If he lives in the rural sector, earning W_r with certainty; a similar formula is applied to $(1-P^i_t)$, the probability of staying in the rural sector, and this reduces to:

$$P^i_{t+1} = P^i_t(1-\alpha_1^! W_r - \alpha_o) \tag{2c}$$

Combining 2a - 2c with their associated likelihoods[6] and simplifying, we can obtain an expected value for P^i_{t+1};

$$E[P^i_{t+1}] = P^i_t + \alpha_i^! P^i_t(1-P^i_t)(\overline{W}_i - W_r) \tag{3}$$

where $\overline{W}_i \equiv q_i W_u + (1-q_i)W_o$.

If N represents the total population available for urban/rural employment (assumed to be fixed for the purposes of this note), then the expected urban population at time t is given by

$$S(t) = \sum_{i=1}^{N} P^i_t,$$

and the expected urban population at time t+1 is given by

$$S(t+1) = \sum_{i=1}^{N} E[P^{i}_{t+1}]$$

Expected migration, M(t), is the difference S(t+1) - S(t) and, using equation (3), this becomes:

$$M(t) = \sum_{i=1}^{N} \alpha_i' P^i_t(1-P^i_t)(\overline{W}_i - W_r) \tag{4}$$

Finally, accepting the Harris-Todaro assumption that every city dweller has the same chance of employment, $(q_i = \frac{E}{S})$, then equation (4) becomes:

$$M(t) = (\overline{W} - W_r) \sum_{i=1}^{N} \alpha_i' P^i_t(1-P^i_t). \tag{5}$$

Since the summation term is positive whatever the values of the individual probabilities, this implies that $M(t) > 0$ whenever the expected urban wage exceeds the rural wage, and this is the main proposition which we wished to obtain.[7]

Properties of the Migration Function

Since the migration model given by equation (1) was not explicitly derived from any underlying dynamic adjustment theory, there has naturally been considerable debate over its most appropriate form. Todaro, for example, uses the size of the urban labor force, S, as the base from which to measure the rate of migration. Zarembka [9] has objected to this specification on the grounds that it is the rural population that provides the migrants, and that therefore (N-S) should be used as the base. In fact, equation (4) does not support either of these positions. In order to get a simple picture of the operation of equation (4), let us assume temporarily that all individuals are identical. This would make $P_t^i = \frac{S(t)}{N}$ for all i. Now equation (5) becomes:

$$M(t) = (\overline{W} - W_r)\ \alpha'(1 - \frac{S(t)}{N})\ S(t) \qquad (5')$$

or

$$\frac{M(t)}{N-S(t)} = \frac{S(t)}{N}\ \alpha'\ (\overline{W} - W_r)$$

This formulation differs from that of both Todaro and Zarembka in that the learning process applies to the entire population. In this model, M(t) is the net summation of rural-urban migration and urban-rural migration, whereas Zarembka and Todaro only considered the effects of wage differentials on the rural population, presuming that those who move to the city never go home.

The migration process described by equation (4) will eventually lead to a stable population distribution. Net migration will reach zero if $\overline{W}_i = W_r$ for all i and the comparative static properties of an economy in such an equilibrium can be evaluated as usual.[8] Even in this equilibrium, of course, many individual workers are changing location: it is only the net flow of migrants which is zero.

Finally, the functional form of equation (4) has an important dynamic implication. For any constant value of $\overline{W} - W_r$, the character of the migration function is that of a logistic curve (indeed, the logistic is often characterized as "the learning curve"). For small values of P_t^i (equivalent to a predominantly rural population), M(t) is correspondingly small. M(t) is larger for larger P_t^i's, reaching a maximum value when the population is approximately equally distributed between rural and urban components (that is, $P^i(1-P^i)$ is maximized when $P^i=0.5$). M(t) falls again if the population shifts still further. Since most less-developed countries are predominantly rural, we would conclude that they are still in the rising phase of this process, and that if the values of $\overline{W}_i - W_r$ are maintained at present levels, migration will not only continue, but will accelerate.

Although the quotation from Harris-Todaro given earlier conveys a recognition that an acceleration in migration may in fact be taking place, their use of an equilibrium model diverts attention away from this condition and suggests only that migration is a continuing response to disequilibrium. The view obtained from equation (4) is much more pessimistic. Present investment and wage policies in less-developed countries are often designed to maintain a $\overline{W} - W_r$ differential in the face of migration. Emphasis is put upon the expansion of urban job opportunities as a means of holding down unemployment despite constant or even rising urban wage rates. According to the dynamic implications of our learning model, such attempts to accommodate development policies to current levels of migration are entirely hopeless; maintenance of the $\overline{W} - W_r$ differential in the face of an accelerating tide of rural-urban migrants will become a practical impossibility, and inevitably, $\overline{W}$ will, one way or another, be driven down to W_r.

Information

Although it is not entirely in keeping with the strict behaviorism embodied in the learning model we have used, many readers of this paper have inquired about the possibility that some workers are influenced by the successes and failures of others. If a rural worker learns (in the sense of "finding out") that an acquaintance has found high-paying urban employment, then he, too, may be more inclined to relocate in the city. In our notation, P_t^i may be increased to some larger value $(P_t^i)'$ where the extent of the increase is a function of the wage differential $W_u - W_r$. Using a learning function, β, defined similarly to α, we could write

$$(P_t^i)' = P_t^i + \beta(W_u - W_r)(1-P_t^i) \tag{6a}$$

On the other hand, information that the acquaintance is unemployed would discourage relocation in proportion to the wage differential $W_r - W_o$:

$$(P_t^i)' = P_t^i[1-\beta(W_r-W_o)] \tag{6b}$$

The effectiveness of any information flow is, in part, a function of quantity; the more workers there are in the city, the more occasions there will be for the choices of rural inhabitants to be influenced. This fact is most easily introduced by making the learning rate a function of the fraction of the population which is currently residing in the city. Using our linear approximation, this makes the learning rate equal to $\frac{S(t)}{N}$ times a constant parameter β'.

Since a proportion $\frac{E}{S(t)}$ of urban workers are employed, and $1-\frac{E}{S(t)}$ are not, the expected value of $(P_t^i)'$ may be obtained from equations (6a) and (6b) using these weights:

$$E[(P_t^i)'] = P_t^i + \beta'\frac{S(t)}{N}(1-P_t^i)(\overline{W}-W_r) + D(t) \qquad (7)$$

where $D(t) \equiv \beta'\frac{S(t)}{N}(1-\frac{E}{S(t)})(1-2P_t^i)(W_r-W_o)$.

If all workers were identical, so that $P_t^i = \frac{S(t)}{N}$, then equation (7) would be identical to equation (3) except for the third term on the right. This last term has the same sign as $(1-2P_t^i)$, and reflects a central tendency which is usually found in learning models in which "regret" plays an essential part. Since $(P_t^i)'$ is bounded by 0 and 1, and since it varies monotonically with "payoff" (realized urban wage minus W_r), the function relating $(P_t^i)'$ to the realized urban wage must have a point of inflection, whatever specific functional form may be used to define the model. As a consequence, contradictory pieces of information do not have symmetric influences over extreme values of P_t^i. According to such a model, a rural worker who is not inclined to move to the city anyway (P_t^i is low) is unlikely to be influenced by the news that those who have are unemployed. (From equation (6b), $(P_t^i)'-P_t^i$ is small if P_t^i is small.) On the other hand, news of others' success may go far toward "changing his mind." (From equation (6a), $(P_t^i)'-P_t^i$ is large if P_t^i is small.) Therefore these two opposing pieces of information do not cancel each other out, but are inclined to raise P_t^i on balance. For similar reasons, contradictory pieces of information will operate to reduce large values of P_t^i. The term D(t) incorporates this effect.

Equation (7) was derived from the perspective of rural workers, but we could as well have proceeded from the cases of urban workers who "remember" or have "learned" of rural employment opportunities, and who then compare W_r to the urban conditions around them.

We can now introduce information flows into the original model by replacing P_t^i in equation (3) by $(P_t^i)'$ from equation (7), and evaluating the combined model. This combined model will have qualitative properties different from the original model only to the extent that the term D(t) is significant.

If experience is the great teacher, and information as we have described it is largely discounted as rumor, then the learning rate β' is small compared to α', D(t) is insignificant, and we can use the original model as described by equation (4). Several readers of this paper have suggested to me that this is especially likely to be the case in less-developed economies characterized by poor information channels and widely dispersed rural populations.

If P_t^i is small for most workers, as would be the case in predominantly rural economies, $D(t) > 0$, implying a net flow of immigrants to urban areas which is greater than that derived from the simple model. In the same economies, of course, $\frac{S(t)}{N}$ is also small, and to the extent that this samll sample size slows the learning process, the entire effect of information is diminished.

Finally, the introduction of D(t) can modify the conclusion that equilibrium is achieved when $\overline{W} = W_r$. For example, if $W_u - W_r = W_r - W_o$, equation (7) indicates that $E[(P_t^i)'] - P_t^i = 0$ when $\overline{W} = 2W_r \sqrt{\frac{E}{N}}$, and the combined model will provide $\overline{W} = W_r$ in equilibrium only if $E = .25N$. For lower values of E, the influence of information will be to increase the equilibrium urban population over that implied by the original model.

<u>Footnotes</u>

1. See, for example, the papers found in Day and Groves [2], and the references therein.
2. For a summary of these studies, see [8].
3. See, McCall [4] or Phelps [5].
4. We could, with no less, distinguish a wage for the rural employed from a still lower rural unemployment wage. In keeping with the models already cited, however, we presume that all rural workers are equally employed.
5. Since W_0 corresponds to all sorts of urban activities, legal or not, other than "modern sector" employment, we would make $W_0 > 0$. Thus $P^i_{t+1} > P^i_t$ even for one who fails to find a regular job in the city (although naturally P_{t+1} is much larger for one who is employed). Even when $W_r > W_0$, we make $P^i_{t+1} > P^i_t$ for the jobless worker, reflecting a basic principle in the model that past behavior is what actually determines future behavior and not consideration of missed opportunities. In other words, the longer an unemployed worker stays in the city, the greater the likelihood that he will stay one more period. This is entirely consistent with the proposition that the longer the worker is unemployed, the less likely he is to stay in the city indefinitely. For further discussion of this point, see Cross [1] pp. 247 and 248.
6. Equation 2a applies with likelihood $P^i_t \cdot q_i$

   Equation 2b applies with likelihood $P^i_t(1-q_i)$

   Equation 2c applies with likelihood $1-P^i_t$

7. Incidentally, the term V_r in the denominator of the argument of (1) which is meant to restrict Todaro's function to proportionate wage differences has its counterpart in equations (4) and (5) in the value of α_1' which, as an approximation to the slope of $\alpha_1(W)$, declines as W_r increases.

8. It may be worth adding that in the determination of $\overline{W}_i$, one should take account of the fact that many urban workers are already more or less established in their jobs, so that, for them, q_i is large, or even equal to 1. The unemployment rate among the rest of the urban population may be much higher (and $\overline{W}_i$ correspondingly lower) than gross employment/population figures would suggest. W_0 is an equally important influence. The availability of marginal employment opportunities can raise $\overline{W}_i$ considerably by making the state of the urban unemployed bearable.

References

1. J. G. Cross, "A Stochastic Learning Model of Economic Behavior" Quarterly Journal of Economics, May 1973, pp. 239-266.

2. R. H. Day and T. Groves, eds. Adaptive Economic Models. Academic Press, Inc., New York, 1975.

3. J. R. Harris and M. P. Todaro, "Migration, Unemployment and Development: A Two-Sector Analysis" American Economic Review, March 1970, 60, pp. 126-142.

4. J. J. McCall, "Economics of Information and Job Search" QJE, February, 1970, pp. 113-126.

5. E. S. Phelps, "Money Wage Dynamics and Labor Market Equilibrium" in Phelps, et al., Microeconomic Foundations of Employment and Inflation Theory, pp. 124-166.

6. J. E. Stiglitz, "Wage Determination and Unemployment in L.D.C.'s" Quarterly Journal of Economics, May 1974, 88, pp. 194-227.

7. M. P. Todaro, "A Model of Labor Migration and Urban Employment in Less Developed Countries" American Economic Review, March 1969, 59, pp. 139-148.

8. M. P. Todaro, "Internal Migration in Developing Countries: A Survey" NBER Conference on Population and Economic Change in Less Developed Countries, September 30 to October 2, 1976.

9. P. Zarembka, "Labor Migration and Urban Unemployment" Comment" M. P. Todaro, "Reply" American Economic Review, March 1970, 60, pp. 184-188.