STEPPING STONE MOBILITY

Boyan Jovanovic
Yaw Nyarko

STEPPING STONE MOBILITY

ABSTRACT

People at the top of an occupational ladder earn more partly because they have spent time on lower rungs, where they have learned something. But what precisely do they learn? There are two contrasting views: First, the "Bandit" model assumes that people are different, that experience reveals their characteristics, and that consequently an occupational switch can result. Second, in our "Stepping Stone" model, experience raises a worker's productivity on a given task and the acquired skill can in part be transferred to other occupations, and this prompts movement. Safe activities (where mistakes destroy less output) are a natural training ground.

Boyan Jovanovic
Department of Economics
New York University
269 Mercer Street, 7th Floor
New York, NY 10003
and NBER

Yaw Nyarko
Department of Economics
New York University
269 Mercer Street, 7th Floor
New York, NY 10003

1. Introduction

"You must be able to walk before you can run."

The claim is that one must learn the simple before mastering the difficult. It clearly is true in the educational process, where learning a subject like mathematics starts the student off doing simple exercises, and then moves him on to more and more complex problems. It is equally true in the training process -- whether a person is learning how to play a musical instrument, how to drive a car, or how to operate a machine. But the claim also seems to hold as a description of the sequence of activities a person does over his or her entire working life: simpler tasks and jobs tend to precede more complex and difficult ones. Using several different measures of job complexity, Wilk and Sackett (1995) find that as people age, they tend to perform harder tasks. And so it would seem that the principle that one begins with the simple and then moves to the complex, should also apply to how people choose the sequence of jobs and occupations over their lifetimes. That is the claim of this paper: That occupations are like rungs on a ladder -- that they are, in other words, like "stepping stones".

The opening phrase also takes as given the obvious: walking is a very different activity from running. There are many examples of distinct activities that make up a career path, and Table 1 lists some (most of which we owe to Nachum Sicherman). In many, one must be in one in order to get to the next, while others describe typical careers. Some are likely to take place within one firm, while others can occur only across firms:

Professor → Dean	Actor → President	Police officer → Security expert
Player → Coach	Engineer → Manager	Lawyer → Politician (senator)
Lawyer → Judge	Police officer → Private Detective	Team member → Team leader
TA/RA → PhD	White Belt → Black belt	Priest → Cardinal
Soldier → General	Dancer → choreographer	IRS employee →
Associate → Partner	Kitchen worker → Chef	Tax consultant

Table 1: Examples of Occupational Ladders

1

Some would argue that occupations exist merely to reward better workers with "promotions", and that to focus on occupational ladders is to look at the surface, the veneer, and that this can only serve to hide the deeper forces underneath. We disagree on two grounds: First, an occupational change usually signifies some shift in the kind of work being done. For example, on a company's job ladder, the person at the top does not do the same thing as a person at the bottom. Occupational distinctions are thus more than labels. And second, the occupational ladder surely is a symptom of a deeper force -- human capital accumulation -- and studying how people move up the ladder cannot but improve our understanding of this more fundamental concept.

Climbing the occupational ladder often involves changing firms. Twenty years ago, Hall and Kasten noticed that most job-changers move to a higher-paying 3-digit occupational category. Since average pay in an occupation is widely known, why don't people enter the top occupations right away instead of starting with the lower paying ones? And does the answer have something to do with the tendency for people to undertake more complicated tasks as they gain experience?

The answer, of course, is that people at the top of the ladder earn more because they have spent time on lower rungs. Our model says that jobs that are "close" in information content will form rungs a ladder, but the safer ones should come first : On such jobs, mistakes destroy less output, and hence they are a natural training ground. A person must "work up" to being CEO -- or else we would expect them to make costly errors. Secondly, if mistakes are costlier on complex jobs, labor should flow from occupations with flatter learning curves and into occupations where learning curves are steep.

A different hypothesis for why a worker moves from job to job is that the accumulation of experience is accompanied by a sorting process in which employers and workers learn what skilled job each worker can do best. As they learn it, the worker is assigned to occupations where his kind of ability is needed. This is a model not of an accumulation of knowhow, but of the accumulation of information about a worker's innate traits -- his or her comparative and absolute advantage. We refer to this hypothesis at the "Bandit model".

We shall argue, however, that the Stepping Stone model is the natural job-ladders model, because it emphasizes growth of skill on one job and its subsequent transfer to another, better job. This notion is central to the models of Weiss (1971), Rosen (1972), and Sicherman and Galor (1990). Far from being an unpleasant surprise, they argue, occupational mobility can be a planned career move -- the worker earns more lifetime income by taking up more than one

occupation in sequence. We extend the work of these authors by providing a microfoundation for the notion of general skill, skill that is useful in a variety of endeavors, and we then derive some further implications of this view. Related information-theoretic models are Prescott and Visscher (1980) and O'Flaherty and Siow (1992). But these models are about the firm screening the worker, and not about the transfer of <u>acquired skill</u>.

Human capital. The term "human capital" has now come to mean so many things that it is too general for our purposes. Even the distinction between innate and acquired skill is too vague unless given specific content. Any acquired job-specific skill is human capital, but then so too is information about oneself or about features of a job or industry that may affect one's prospects there. <u>Any</u> piece of information that can be used to raise a person's lifetime earnings in a sense augments a person's stock of human capital. In a world in which people and jobs differ and in which these differences are not immediately and commonly known, one must be more specific about what human capital is, at least when asking questions about occupational choice and the behavior of earnings and other aspects of labor market histories.

We argue that different kinds of information acquisition have very different impacts on a person's labor market history, and we distinguish information into two types:
      (a) information about personal traits and about the content of jobs, and
      (b) information about how to perform a given task.

The first kind of information (a) allows a person to choose a job where he fits best. Sometimes it is the employer who will assign the worker to some position and who therefore needs to know a worker's ability which, together with the job requirements will determine whether the worker is suitable for the position. This is the kind of information that drives the Bandit model. The second kind of information (b) is the kind that augments the skill of a given person on a given task -- learning how to type, how to drive, how to cook, how to play tennis. This is the kind of information that drives the Stepping Stone model.

Our aim is twofold. First, we want to understand occupational mobility. But second, we want to use this exercise to deepen our understanding of the general concept of human capital by distinguishing (a) from (b), which we see as fundamentally different. And we shall show that not only do (a) and (b) differ conceptually, they can also have sharply conflicting implications.

Section 2 presents our version of the Stepping Stone model. Since it is important to clarify the differences between the different motives for labor movement, we briefly describe the sectoral shock model in section 3, and then the Bandit model in more detail in section 4. Section

5 lists some facts that help us discriminate between the models. Section 6 briefly indicates how the Bandit model and the Stepping Stone model can be combined into a hybrid. Section 7 provides some further discussions and observations. section 8 concludes.

2. The Stepping Stone Model

We begin with a partial equilibrium context. The appendix shows how it fits into a general equilibrium setting.

2.1. Production possibilities and the information structure.

A farmer lives for two periods. In each period, he can grow either apples or bananas, but not both. The production function for apples is

$$q_A = 1 - (x - z_A)^2, \tag{1}$$

where q_A is the number of apples grown, z_A is a decision the farmer must take, such as the amount of irrigation and fertilizer to apply, and x is a random variable, such as the weather, realized only after z_A has been chosen. The production function for bananas is

$$q_B = [1 - (y - z_B)^2], \tag{2}$$

where q_B is the number of bananas grown, z_B is again a decision, and y is a random variable. A multiplicative normalization constant can be added to this production function, but it would leave all of our results unchanged. The random variables (x, y) have a bivariate normal distribution. The unconditional variances of x and y are $\sigma_x^2 < 1$ and $\sigma_y^2 < 1$. Let $\rho \equiv \sigma_{x,y}^2 / \sigma_x^2 \sigma_y^2$ denote their squared correlation coefficient. Then the conditional variances are

$$\text{Var}\{y \mid x\} = (1 - \rho)\sigma_y^2, \quad \text{and} \quad \text{Var}\{x \mid y\} = (1 - \rho)\sigma_x^2, \tag{3}$$

and they do not depend on the values of x and y that are realized.

One period of work growing apples reveals x, and one period growing bananas reveals y. It also reduces the variance of the other random variable to its conditional variance. A person can not learn x and y by talking to or watching other farmers -- hands-on experience is both necessary and sufficient for learning to occur.

4

Let p be the price of bananas in terms of apples. We shall only consider equilibria in which p is a constant. Formally, this is valid if the x's and y's of agents are all independent of each other. In this case the strong law of large number implies that the per-capita output of apples and bananas will be constant over time, and that prices will actually be constant. The Appendix lays out an overlapping generations structure formally.

We shall measure the rate of interest, r, in units of the price of apples at different dates. Since p is constant, this rate will be the same as the banana-denominated interest rate, and will satisfy $(1 + r)^{-1} = \delta$ (the discount factor for second-period utility). We assume that the farmer will maximize <u>expected</u> discounted lifetime income. (In the appendix we assume preferences that imply an indirect utility function linear in lifetime wealth. These preferences assume that consumption occurs only in old age. Also, productivity rises as people age. So for prices to be constant over time, we need a stable demography, and so we assume that there are overlapping generations).

In each period, a farmer faces an occupation choice (A or B), and a production choice (the level of z), which must be set <u>before</u> x and y are observed. The risk-neutral apple producer will set $z_A = E(x)$, and the banana producer will set $z_B = E(y)$. These expectations are conditioned on the information that the producers have at the time. Since their first-period decisions are set on the basis of no signal at all, their expected outputs (measured in units of apples) from following this strategy are $1 - \sigma_x^2$, and $p(1 - \sigma_y^2)$, respectively.

There are four possible career paths for the farmer: AA, AB, BB and BA. The expected discounted lifetime incomes from these four paths are

$$
\begin{array}{ll}
V_{AA} = 1 - \sigma_x^2 + \delta, & V_{AB}(\rho) = 1 - \sigma_x^2 + \delta p[1 - (1 - \rho)\sigma_y^2], \\
V_{BB} = p[(1 - \sigma_y^2) + \delta], & \text{and} \quad V_{BA}(\rho) = p(1 - \sigma_y^2) + \delta[1 - (1 - \rho)\sigma_x^2],
\end{array} \tag{4}
$$

The $V_{i,j}$ are perfectly foreseen, because $\text{Var}\{y \mid x\}$ and $\text{Var}\{x \mid y\}$ are not random.

Everyone is the same *ex ante*. There is no innate absolute advantage, no innate comparative advantage. [The latter are the essence of the Bandit model, but they are entirely absent here.] The Stepping Stone model is one of learning how to <u>do</u> something, and not a model of learning about oneself or learning about the quality of the worker-job match.

Our argument revolves around ρ. This is an index of how <u>general</u> human capital is. While V_{AA} and V_{BB} do not depend on it, V_{AB} and V_{BA} are strictly increasing in ρ. This result reflects the role that ρ plays -- it is an index of how much human capital can be transferred between apple growing and banana growing. Indeed, when $p = 1$, $\rho = 1$ and $\sigma_x^2 = \sigma_y^2$, apples and bananas become the same good as far as production is concerned.

Proposition 1: Suppose that $\sigma_y^2 > \sigma_x^2$. Then V_{BA} is always dominated, <u>regardless of the value of ρ and p</u>.

Proof: when $p > 1$, $V_{BB} > V_{BA}$ and when $p \leq 1$, $V_{AA} > V_{BA}$. ∎

In our 2-occupation economy, ρ is an index of how general human capital is, because it is a measure how much one can learn about x from observing y and vice versa. In an economy with many occupations, there will in general be a different ρ for each pair of occupations, and after choosing his first job, the worker would in general face a different ρ for each possible occupational change. But as long as there are just two periods of life, Proposition 1 would continue to apply.

So an agent will <u>never</u> move from a high variance occupation to a low variance one. The intuition for switching emerges from the following special case: Suppose $\rho = 1$. Then learning x means that we also learn y. But suppose σ_x^2 is much smaller than σ_y^2. Then the experimentation cost (in terms of foregone output) of learning y is much less if the young person produces apples. We shall now derive a general result formalizing this intuition. Define the following functions of ρ:

$$P_0(\rho) \equiv \frac{1 - \sigma_x^2}{1 - \sigma_y^2 + \delta(1 - \rho)\sigma_y^2} \quad ;$$

$$P_1(\rho) \equiv \overline{P}_1 \equiv \frac{1 - \sigma_x^2 + \delta}{1 - \sigma_y^2 + \delta} \quad ; \quad \text{and}$$

$$P_2(\rho) \equiv \frac{1}{1 - (1 - \rho)\sigma_y^2} \quad .$$

Then one can easily verify the following:

(i) $\quad V_{AB} \overset{>}{\underset{<}{=}} V_{AA} \quad$ as $\quad p \overset{>}{\underset{<}{=}} P_2(\rho)$;

(ii) $\quad V_{BB} \overset{>}{\underset{<}{=}} V_{AA} \quad$ as $\quad p \overset{>}{\underset{<}{=}} \bar{P}_1$; and

(iii) $\quad V_{BB} \overset{>}{\underset{<}{=}} V_{AB} \quad$ as $\quad p \overset{>}{\underset{<}{=}} P_0(\rho)$.

We provide a diagram of the functions P_i for $i = 1, 2, 3$, which uses some properties that we collect together in the lemma below:

<u>Lemma 1</u>: Suppose $\sigma_y^2 > \sigma_x^2$. Then (i) $P_0(\rho)$ (resp. $P_2(\rho)$) is strictly increasing (resp. strictly decreasing) and continuous in ρ; (ii) $P_0(0) < \bar{P}_1 < P_2(0)$; (iii) $P_2(1)=1< \bar{P}_1 < P_0(1)$; and (iv) there exists a ρ_c in $(0,1)$ such that

$$P_0(\rho_c) = \bar{P}_1 = P_2(\rho_c) \ ,$$

where

$$\rho_c = \frac{\delta\sigma_y^2 + \sigma_x^2(1-\sigma_y^2)}{\sigma_y^2(1-\sigma_x^2 + \delta)} \ .$$

<u>Proof</u>: Consists of easy algebra. ∎

Figure 1 uses the facts in the lemma to depict the functions P_i, indicating the optimal career paths in p - ρ space. Figure 1 shows the following: When $\rho < \rho_c$, there is no occupational switching, regardless of p. When the price is higher (resp. lower) than $\bar{P}_1$ agents spend all of their lives growing bananas (resp. apples). When p is equal to $\bar{P}_1$, agents are indifferent between spending their entire lives growing apples or growing bananas.

Now suppose that $\rho > \rho_c$. Then when p is very high (resp. low) it is optimal for agents to spend all their lives in banana (resp. apple) farms. But when p is at an intermediate level, in particular when p $\in$ [$P_2(\rho)$, $P_0(\rho)$] it is optimal to switch professions: agents will grow apples when young and bananas when old. So with the aid of Figure 1, we have proved
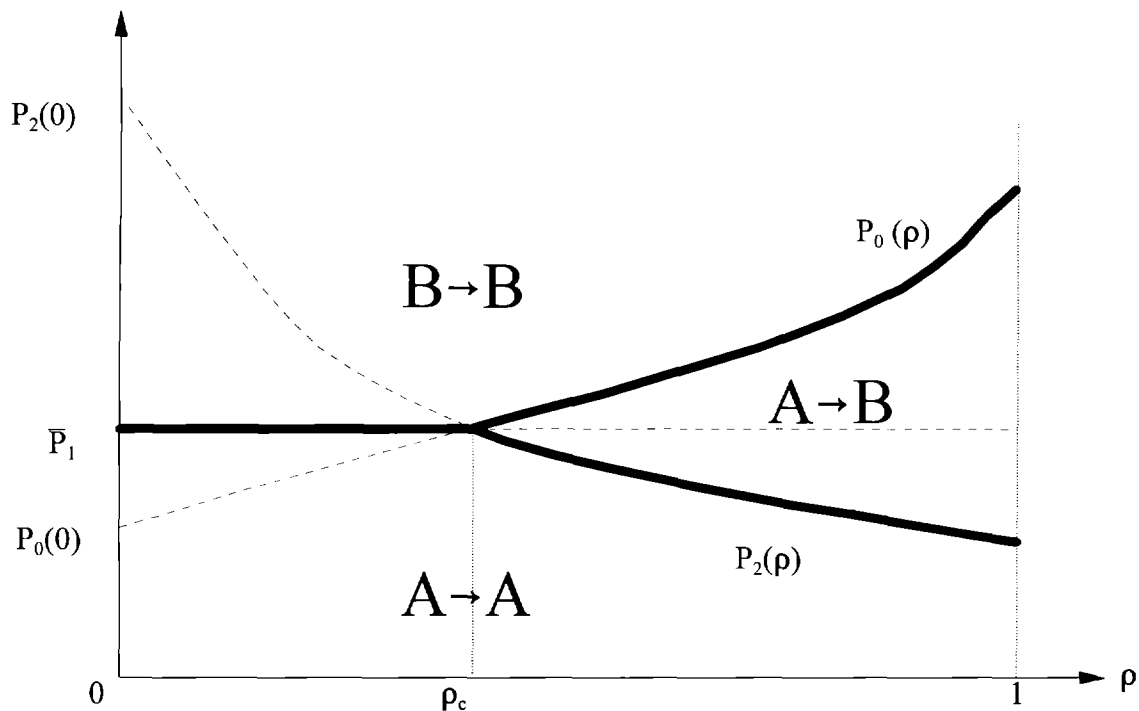
7

Figure 1:   Optimal Career Paths in p-ρ Space

Proposition 2: Occupational movement can occur only if $\rho \geq \rho_c$.

So, occupations will form adjacent rungs on an occupational ladder only if human capital is sufficiently transferable between them. Note too that while the condition $\rho > \rho_c$ is only necessary, it is also sufficient if in addition both goods are supplied in equilibrium. This is clear from figure 1 -- if $\rho > \rho_c$ and we are outside the $A \rightarrow B$ region, we must be in a region where only one good is produced.

Let us use the notation $i \rightarrow j$ to refer to an agent who chooses farm i when young and farm j when old where i, j = apple, banana. $B \rightarrow B$ agents produce respectively $1-\sigma_y^2$ and 1 units of bananas when young and old.   The $A \rightarrow B$ agents produce respectively 0 and $1-(1-\rho)\sigma_y^2$ units of bananas when old.   From Proposition 1 we there will be no $B \rightarrow A$. By assumption there is a unit mass of generation of agents. Fig. 2 depicts the supply curve of bananas for the regime $\rho < \rho_c$.

8

This supply curve is obtained by selecting a particular value of ρ that is <u>less</u> than ρ_c, and then varying p while holding ρ constant. In the figure, we use the definition $V^* \equiv$ Max $\{V_{AA}, V_{BB}, V_{AB}, V_{BA}\}$. This figure shows that supply responds much as it would in a static model; above some threshold price $\overline{P}_1$, everyone would wish to grow bananas. The equilibrium number of farmers of each type will then depend on the demand for apples and bananas. The general equilibrium aspects are detailed in the appendix.

Similarly, Figure 3 depicts the supply curve for the regime $\rho > \rho_c$. This supply curve is

obtained by selecting a particular value of ρ that is <u>greater</u> than ρ_c, and then varying p while holding ρ constant. Behind this supply curve is a shifting nature of switching behavior, as a function of p. Figure 4A depicts the fraction of the young workers that are switching. When Q_B = 0, all the young are in apples, and no one ever switches. Similarly, when $Q_B = 2 - \sigma_y^2$ (the maximal output of bananas), there is again no switching, since the young and the old are growing bananas. Movement is the largest at $Q_B = 1 - (1 - \rho)\sigma_y^2$. At that point <u>all the young</u> are switching from A to B! To the left of that point, all the young are in A, and only some switch to B, and to the right of this point, only some of the young are in A, and they all later switch to B.

Figure 4B shows there is a whole range of consumer utility functions leading to a whole range of <u>demand</u> functions and resulting equilibrium prices p for which all the young start off in A and then they <u>all</u> move to B. This is not an occurrence of zero measure in parameter space.
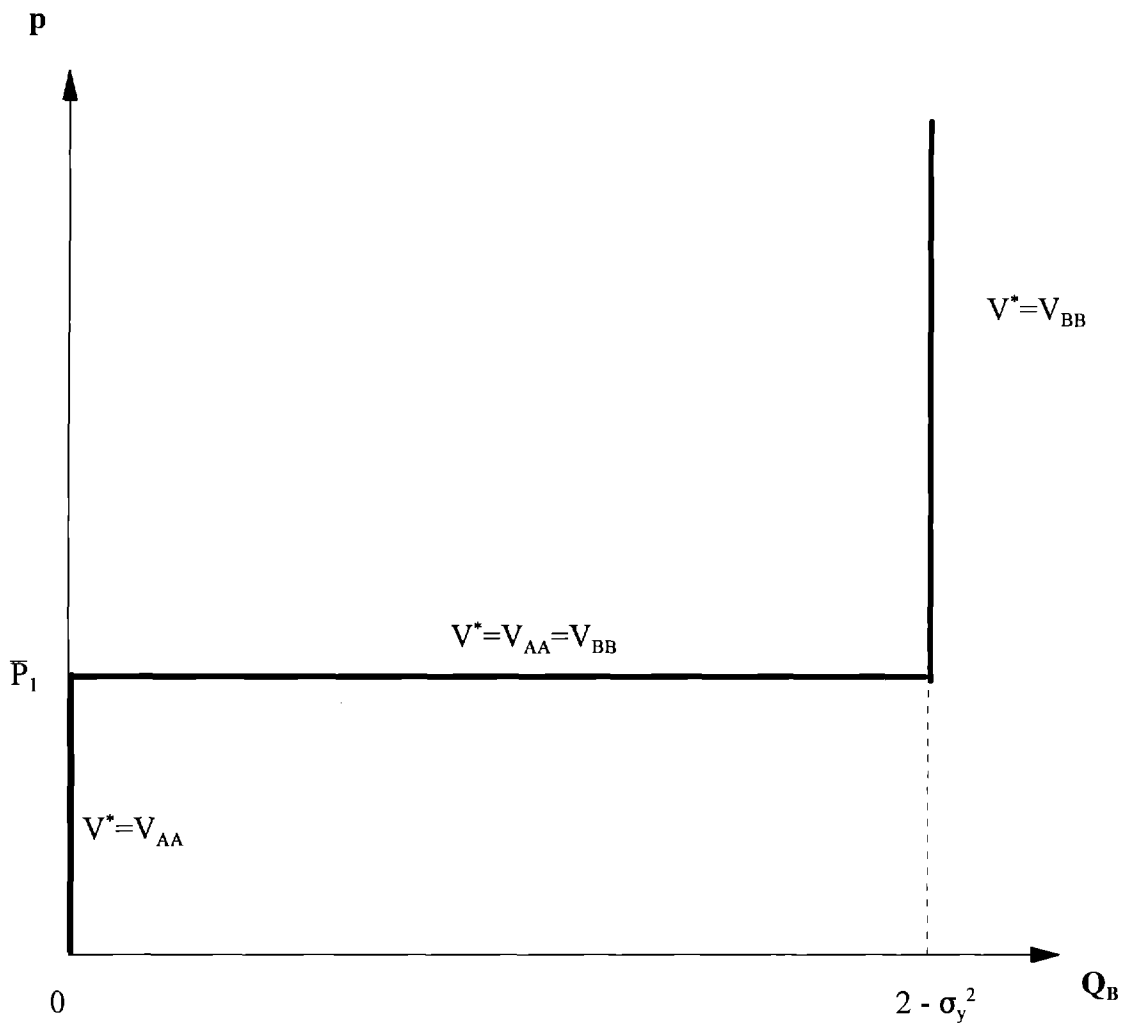
p

$V^*=V_{BB}$

$V^*=V_{AA}=V_{BB}$

$\overline{P}_1$

$V^*=V_{AA}$

0

$2 - \sigma_y^2$

Q_B

Fig. 2 : the Supply Curve of Bananas When $\rho \leq \rho_c$.

10

Fig. 3 : The Supply Curve of Bananas When $\rho > \rho_c$

Fraction
of young
that move
from A→B

0            $1-(1-\rho)\sigma_y^2$ $2-\sigma_y^2$ Q_B
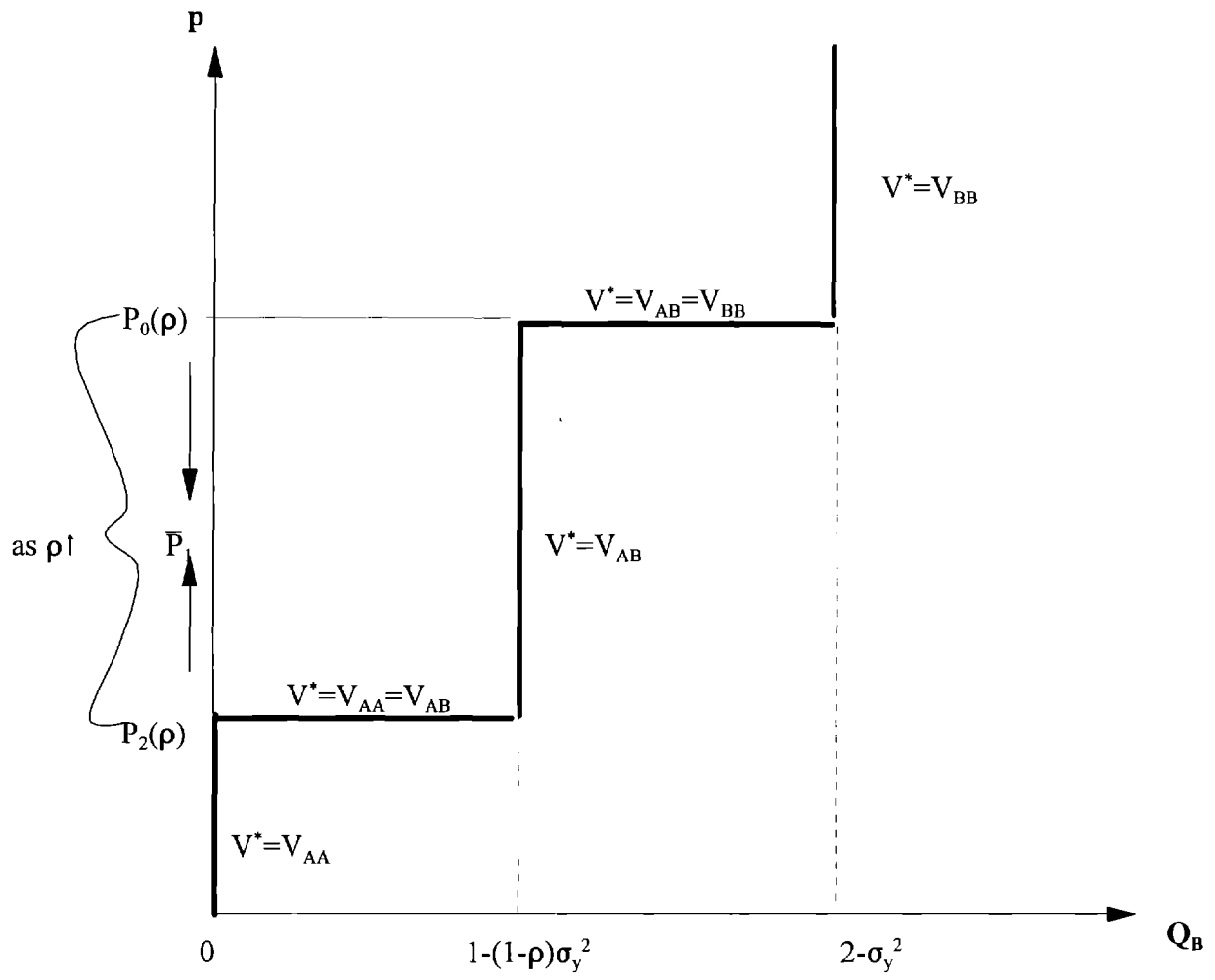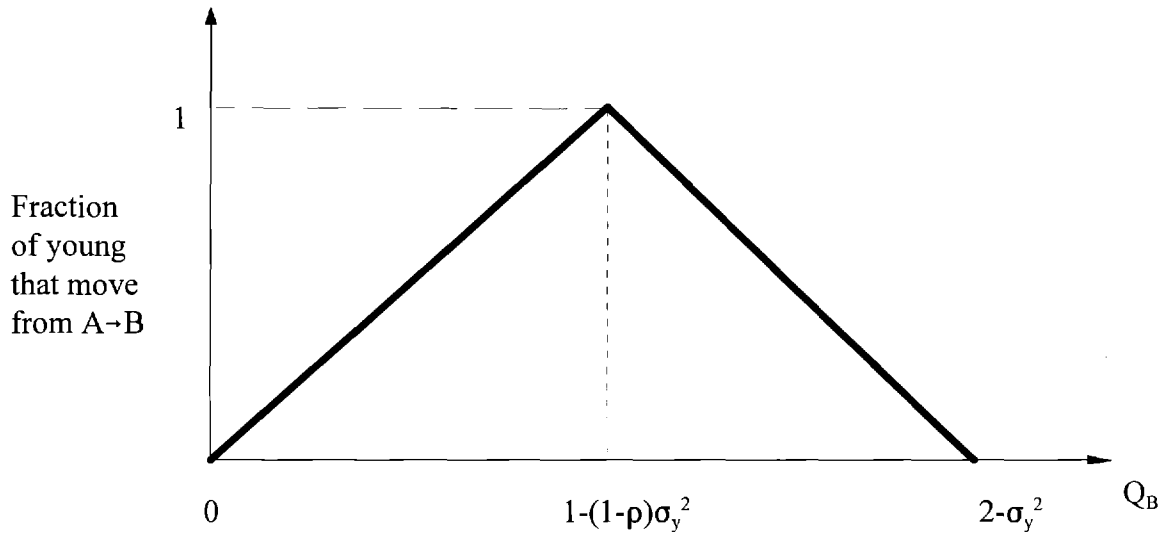
Fig. 4A: Stepping Stone Movement: Fraction of Young That Move as Function of Output

Fraction
of young
that move
from A→B

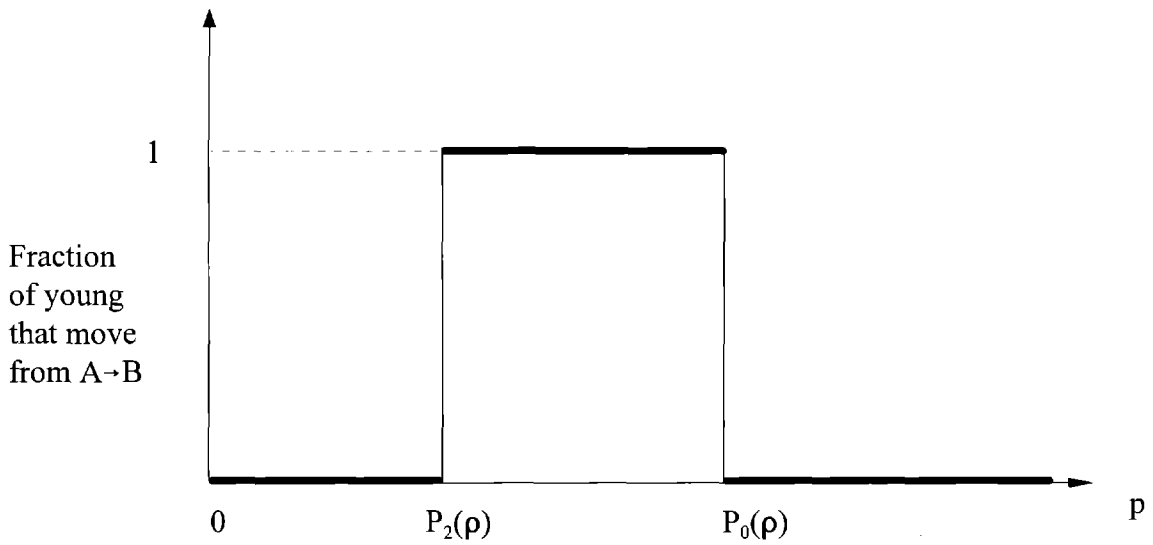0 $P_2(\rho)$ $P_0(\rho)$ p

Fig. 4B: Stepping Stone Movement: Fraction of Young That Move as Function of Price

12

2.2. Age-earnings profiles and promotions in the Stepping Stone model.

The Stepping Stone model (and the Bandit model of section 4) both assume that career choices are made by the worker. The models describe workers' supply -- which jobs they are willing to undertake, in what sequence. Every job is open to every worker, and it is the wage structure that induces the worker to make the equilibrium choice of job. In reality it often is the firm that decides what task a worker will perform. The firm often actively trains a worker, evaluates the worker's abilities, and then assigns him to whatever task it deems is appropriate for that worker. How consistent is the model with this observation?

There is an interpretation in which firms enter at zero expected profit, organize tasks A and B under one roof, allocate workers to tasks A and B, and pay wages. Under this interpretation, q_A is the number of efficiency units of service A that the worker supplies, and q_B is the number of efficiency units of service B. Then p is interpreted as the price per efficiency unit of service B, and the relative price per efficiency unit of skill A is normalized at one. Wages could either be piece rates (so that the worker gets q_A or q_B) or time rates, in which case the worker would get his expected output conditional on his experience. Profit-maximizing behavior would dictate that the firms' allocation choices coincide with the equilibrium choices described here.

The deterministic approach to human capital has emphasized the growth of wages with job-tenure and with labor market experience. Typically, one specifies expected output as a function of experience, and then adds a disturbance with no specific rationale. Our theory gives rise to implications about all the moments of wages, and hence it complements the traditional approach. To bring this point out more clearly, Figures 5A, 5B, and 5C describe lifetime expected earnings in the three types of equilibria in the Stepping Stone model that can occur when $\rho > \rho_c$. Which one prevails depends on where exactly p is.

$ \uparrow

$p(1-(1-\rho)\sigma_y^2)$

Average earnings in B

$1-\sigma_x^2$

Average earnings in A

1

2

AGE
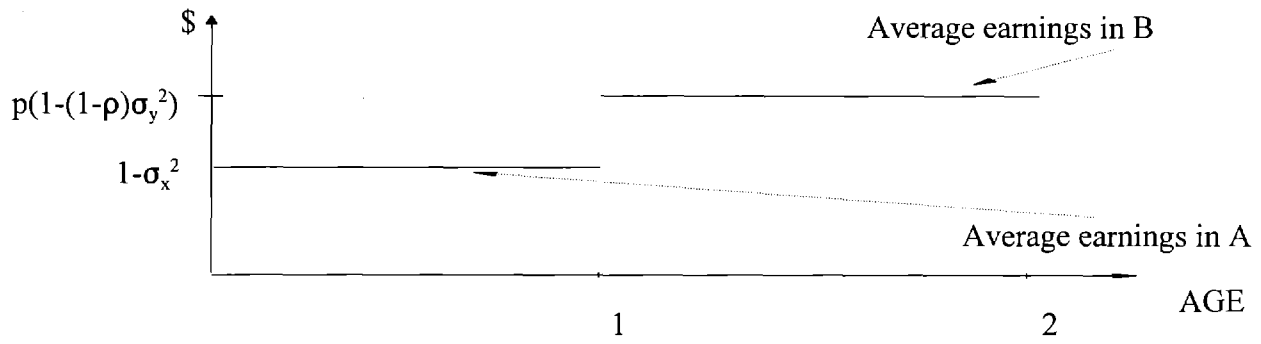
Fig: 5A. $p \in (P_2(\rho), P_0(\rho))$ -- All Start in A. All Move to b

$ \uparrow

Average earnings in B

1

$1-\sigma_x^2$

Average earnings in A

1

2

AGE
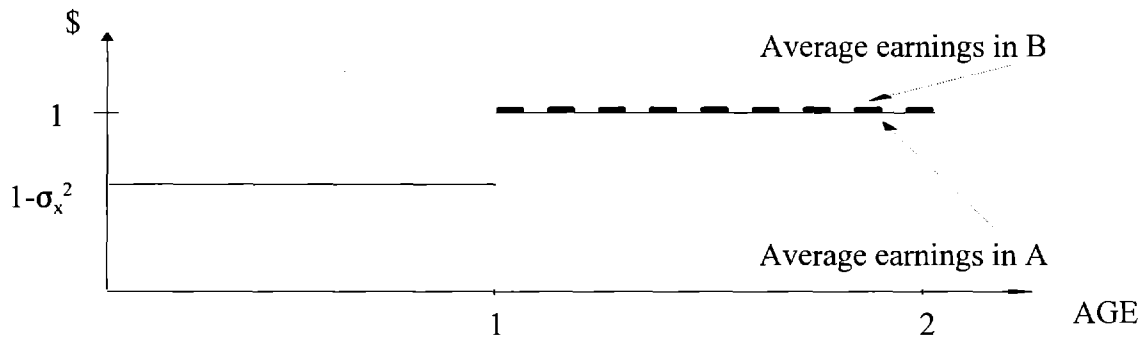
Fig: 5B. $p = P_2(\rho)$ -- All Start in A. Some Move to b

$ \uparrow

p

Earnings in B of B-young

$p(1-(1-\rho)\sigma_y^2)$

Earnings in B of A-young

$1-\sigma_x^2$

$p(1-\sigma_y^2)$

Earnings of A-young

Earnings of B-young

1

2

AGE

Fig: 5C. $p = P_0(\rho)$ -- Some start in A, some start in B, all end up in B.

3. The Sectoral Shock Model

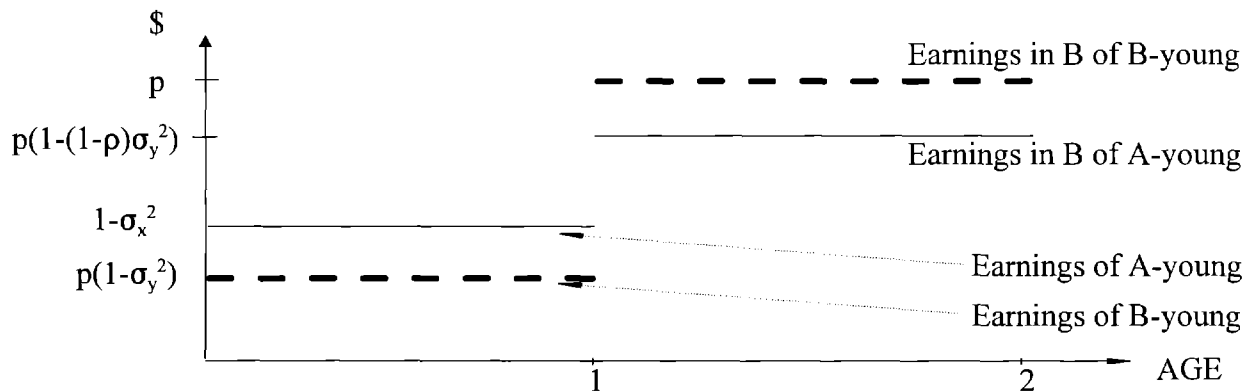Suppose the automobile sector makes cars using engineers and manual labor as inputs. If an increase in the demand for cars causes an inflow of labor of both types into the automobile sector, we will see a net inflow of labor, and the gross flow of workers will equal the net flow. In fact, however, gross flows among sectors, both of jobs (Davis and Haltiwanger 1992, Table 1) and of workers (Jovanovic and Moffitt 1990, Table 1) are much bigger than the net flows. To explain this with a sectoral shock story requires that the shock -- say the invention of a new mass-production technique -- increase the demand for engineers, and displace the unskilled workers. (Krusell *et al* (1995) point out that the falling relative price of equipment may have caused the skill premium in wages to rise in recent years). We should then see an inflow of engineers, an outflow of unskilled workers, and therefore the gross flow will exceed the net flow of labor between the automobile sector and the rest of the economy.

This refinement of the sectoral shock still does not explain occupational ladders. In an occupational choice equation, sectoral shocks explain "time dummy" effects, not "age" effects. The latter are the focus of the Stepping Stone model.

4. The Bandit Model of Occupational Choice

What follows is a two-period version of the model analyzed first by Johnson (1978), and its central implication, as described by Miller (1984, p. 1088) is "that the inexperienced gravitate towards risky ventures. Because younger workers are less experienced than their older counterparts, they are more willing to try jobs where success is rare; older people who have quit them already know they themselves are unsuitable."

Recall that in the Stepping Stone model, such behavior is <u>impossible</u>. In the Bandit model, it happens for some, if not most parameter values. However, as suggested by an example in Waldman (1984), we shall see that it does not happen for <u>all</u> parameter values.

4.1.   The 2-period Bandit model.

Let there be two occupations, A, and B. Let α and β denote a person's productivity in occupations A and B respectively. As in Roy's (1951) model, there is a distribution of the vector (α, β) in the population. Both Johnson and Miller assume that this distribution is normal:

$$(\alpha,\beta) \quad \sim \quad N\left((E\alpha, E\beta), \begin{bmatrix} \sigma_\alpha^2 & Cov\,(\alpha,\beta) \\ Cov\,(\alpha,\beta) & \sigma_\beta^2 \end{bmatrix}\right).$$

Let

$$r \;=\; \frac{Cov^2(\alpha,\beta)}{\sigma_\alpha^2\sigma_\beta^2}$$

be the squared correlation coefficient. We shall repeatedly use the following result:

Lemma 2: Under the above assumptions,

$$E(\alpha|\beta) \;=\; E\alpha + c\,(\beta - E\beta) \qquad \text{where} \qquad c \;=\; [Cov\,(\alpha,\beta)/\sigma_\beta^2\,]$$
$$E(\beta|\alpha) \;=\; E\beta + d\,(\alpha - E\alpha) \qquad \text{where} \qquad d \;=\; [Cov\,(\alpha,\beta)/\sigma_\alpha^2\,]$$

$$Var\,(\alpha|\beta) \;=\; (1\text{-}r)\,\sigma_\alpha^2 \qquad\qquad \text{and} \qquad Var\,(\beta|\alpha) \;=\; (1\text{-}r)\,\sigma_\beta^2\,.$$

Proof: [Chow (1983, p. 12)] ∎

In the appendix we also show that when $\sigma_\alpha^2 < \sigma_\beta^2$, the parameters can generate any c in (-1,1) (and nothing else) and any d ∈ $\mathbb{R}$.

At the outset, agents know the population distribution of (α, β), but not their individual (α, β) type. A lifetime is just two periods here, so we assume that after a period an individual learns completely his productivity in the occupation he experienced. If α and β are correlated, this will allow him to infer something about his productivity in the other occupation. For example, a young person in occupation A would, after one period, learn his α. He then will revise his prior distribution over β to be the normal distribution with mean $E(\beta|\alpha)$ and variance $Var\,(\beta|\alpha)$, as given in Lemma 2. Note that just like the Roy model, this is a model where some workers have a comparative advantage in one occupation, and some in the other. But some workers also have an absolute advantage over others. Nevertheless, this is a model in which assignments are based entirely on comparative advantage, and we shall have more to say about this in section 7.1.

Assume that the individual is risk neutral, and that he discounts second period income by the factor δ. Suppose that his earnings equal his productivity. Further, let the price of the skill β relative to that of α be unity. (If it were not, we could define a new β whose mean and standard deviation equal that of the old β, multiplied by the price of that skill, and the analysis applies). Then a person with $\alpha > \beta$ has a comparative advantage as an apple grower, and a person with $\alpha < \beta$ has a comparative advantage as a banana grower. The costs of switching occupations are assumed to be zero.

4.2. The lifetime decision problem.

If, when young, the individual worked in occupation A, then he will have learned α by the end of the period, and he will switch occupations if and only if $\alpha < E[\beta|\alpha]$. Otherwise he will remain in occupation A, where he will earn α in the second period. Let V_α be the expected present value of starting out in occupation A. Then,

$$V_\alpha = E\alpha + \delta E[\text{Max} \{\alpha, E[\beta|\alpha]\}] . \tag{5}$$

By a symmetric argument, the expected present value of starting out in occupation B is

$$V_\beta = E\beta + \delta E[\text{Max} \{\beta, E[\alpha|\beta]\}] . \tag{6}$$

The following result characterizes the first period optimal decision:

Proposition 3: For each value of r there exists a critical number $\Delta^c(r)$ such that

$$V_\beta \overset{>}{\underset{<}{=}} V_\alpha \quad \text{as} \quad E\beta - E\alpha \overset{>}{\underset{<}{=}} \Delta^c(r).$$

Further, $\Delta^c(r)$ is strictly negative for r in $[0,1)$; $\Delta^c(1) = 0$; and $\Delta^c(r)$ is continuous in r.
Proof: See appendix. ∎

This is illustrated in fig. 6 below.

$\Delta \equiv E\beta - E\alpha$

Region where B is chosen first

Typical Johnson (1978) parameter vector

J

0

1

r

W

$\Delta^c(r)$

Region where A is chosen first
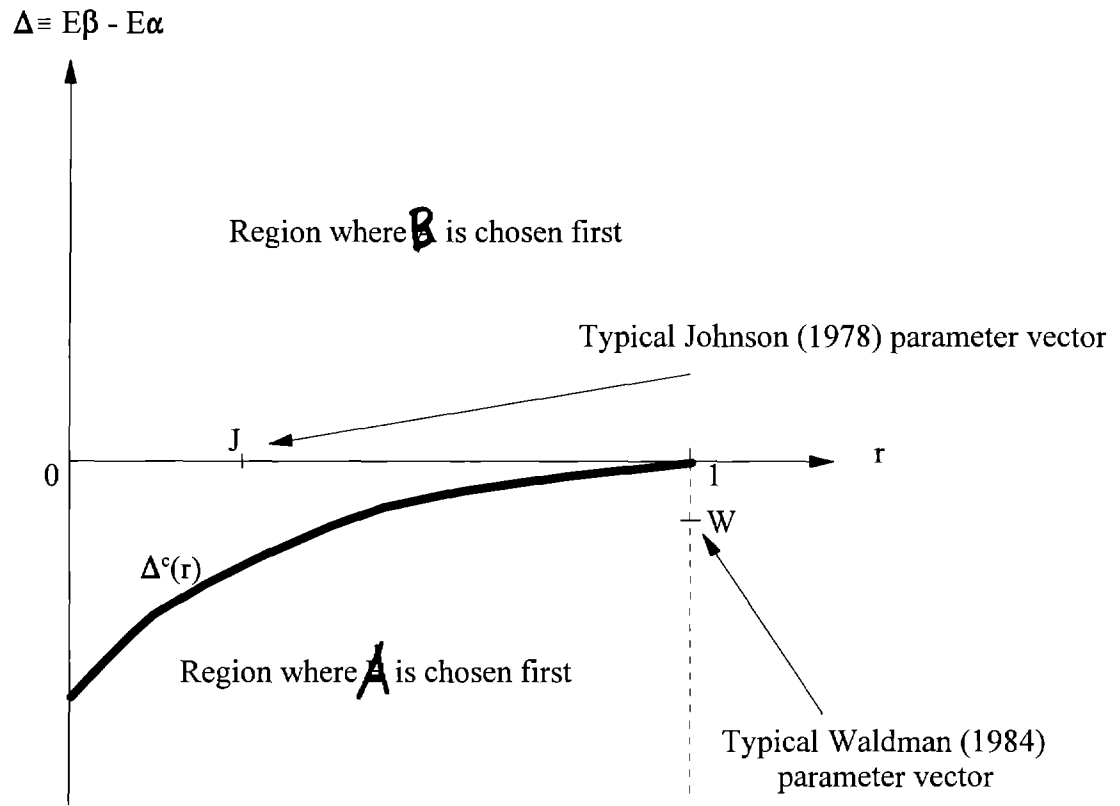
Typical Waldman (1984)
parameter vector

Fig: 6.   The Optimal Date 1 decisions in $((E\beta\text{-}E\alpha)$, r) space

We can now prove Miller's assertion. Assume that the mean payoff in the two occupations is the same, but that occupation B is riskier. Then the young worker will indeed prefer to start out in occupation B. [This of course is a direct implication of Proposition 3 (by setting $\Delta \equiv E\beta - E\alpha = 0$).]

Corollary 1: (Johnson 1978, Proposition 1 -- The young prefer the riskier occupation). Suppose that $E\alpha = E\beta$, $\sigma_\alpha^2 < \sigma_\beta^2$ and $r<1$. Then $V_\alpha < V_\beta$.

We now have, for the second period decision,

<u>Lemma 3</u>: Unless $Cov(\alpha, \beta) = \sigma_\alpha^2$ or $Cov(\alpha, \beta) = \sigma_\beta^2$, the inequalities

$$\alpha < E[\beta|\alpha] \qquad \text{and} \qquad \beta < E[\alpha|\beta]$$

both occur with positive probability. That is, the inequalities are both met generically in the parameter space.

<u>Proof</u>: The marginal distributions of α and β are normal, and so all points on the line have positive density. Then by Lemma 1, it is clear that there will exist some intervals on the line where each of the inequalities mentioned in the statement of this lemma will hold. Since a positive probability will be assigned to these intervals, this lemma is true. ∎

Lemma 3 implies that conditional on there being some young workers in both occupations, after a period of experience some of the labor will flow in both directions. But if young people are all the same *ex ante*, they must be indifferent between the two occupations. That is, we must have $V_\alpha = V_\beta$. When this is true, the model gives rise to flows of labor in both directions:

Proposition 4: (Occupational switching in both directions). Suppose that $V_\alpha = V_\beta$. Then generically in the space of parameters $\{\sigma_\alpha^2, \sigma_\beta^2, Cov(\alpha, \beta)\}$, labor will flow in both directions between the two occupations.

Proof: The additional restriction $V_\alpha = V_\beta$ does not imply either equality in the first line of Lemma 3 ∎

19

From Lemma 2, we know that if we consider $E[\alpha|\beta]$ as a function of β then a graph of it against β will have slope c with $|c|<1$. Suppose that $E\alpha=E\beta=1$. Then the point of intersection of such a graph with the 45 degree line is at $\beta=E[\alpha|\beta]=1$. We therefore obtain figures 7A and 7B:

4.3.   Age-earnings profiles and promotions in the Bandit model.

A potential difficulty with the Bandit model as a model of occupational ladders stems from Corollary 1, and from the fact that top-paying occupations also have a higher variance of earnings. Suppose B is the high-variance occupation, so that $\sigma_\beta^2 > \sigma_\alpha^2$. If "top-paying" is interpreted as a statement about unconditional means so that $E(\beta) \geq E(\alpha)$, then Corollary 1 implies that all young workers would want to start at the top, and this is counterfactual. Indeed, suppose that $E\alpha = E\beta = 1$ so that all the workers begin in B. It should be clear from fig. 7A that those who move from B to A do worse than those that remain in B in the second period. In other words, those that move would appear to have been "demoted," when we look only at second period wages. This is illustrated in fig. 7B. A further problem with this explanation of occupational mobility stems from Corollary 1: B is the riskier activity, and so the variance of earnings over people shrinks dramatically over the life cycle. But in fact this variance grows with labor market experience (e.g. Mincer 1974, O'Neill 1995). It is easy to show that in the Stepping Stone model, if ρ is sufficiently less than 1, variance of earnings in the cohort of workers can rise over the life cycle, directly as a consequence of Corollary 1 -- safe activities are chosen first.

However, "top-paying" is a statement about the earnings of those workers that end up choosing B. In this case it is possible that $E(\beta) < E(\alpha)$, and Corollary 1 no longer applies. Indeed, as suggested by the Bandit model analyzed by Waldman (1984), it is possible to find parameter values in the Bandit model under which

(i)   the safer occupation A is sampled first, that
(ii) people that switch from A to B raise their earnings in the process, and
(iii) those who switch from A to B earn more than those who remain in A.

Such predicted behavior is what one would associate with a career that begins in A, and ends in B for those who are "promoted". Inspecting fig. 6 shows that the following easily verified result (which implies (i) - (iii) above):
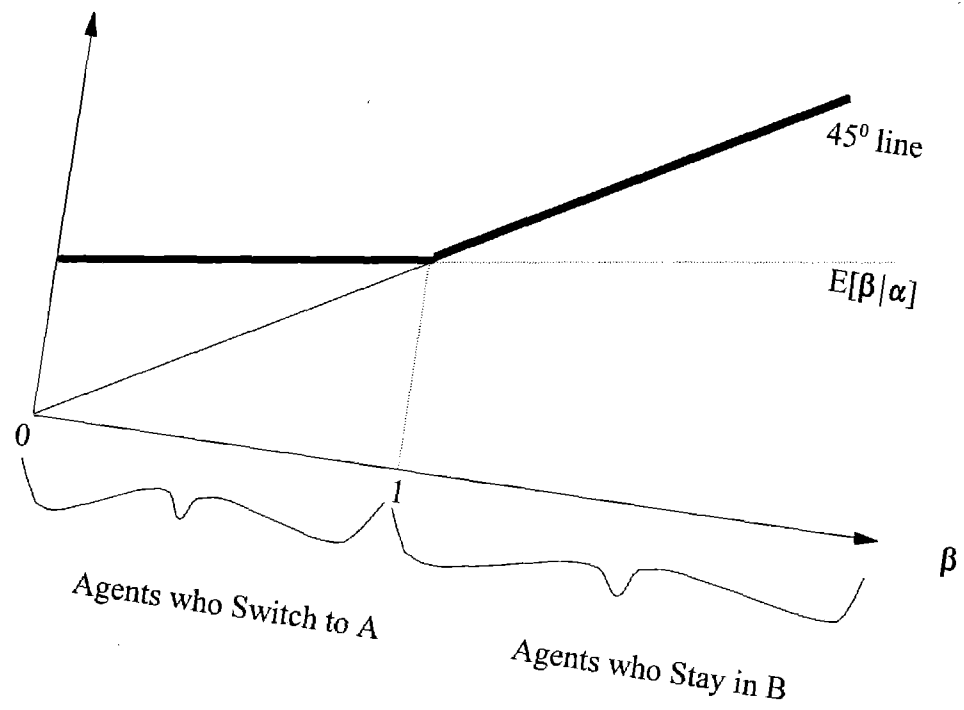
20

Fig. 7A: Selection in the Bandit Model.

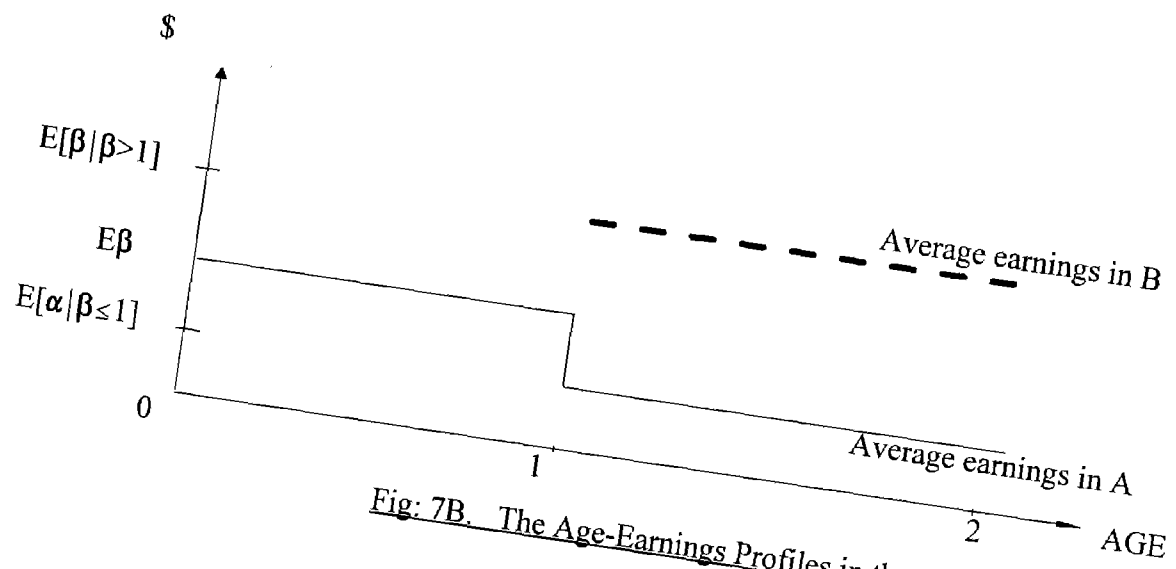Fig: 7B. The Age-Earnings Profiles in the Bandit Model

21

Corollary 2 (Waldman) : Suppose that $E\alpha > E\beta$ and $\sigma_\alpha^2 < \sigma_\beta^2$. Then for r close enough to 1, $V_\alpha > V_\beta$.

Figure 6 shows why Johnson and Waldman reached opposite conclusions about the predicted occupational choice time-path in the Bandit model.   Both of were correct of course. They just used different parameter values.

5. Evidence: Stepping Stone or Bandit?

The most commonly examined evidence is that on the relation between seniority and wages, and many authors have asked whether it reflects the growth of skills (as occurs in the Stepping Stone model) or the effect of selection by characteristic of individuals and jobs (as occurs in the Bandit model)-- e.g., Altonji and Shakotko (1987), Topel (1981). We will not add to this, but rather discuss other implications that emerge from our analysis.

5.1   Evidence favoring the Stepping Stone model:

Safe tasks first, risky ones last: (The Bandit model implies the opposite) A natural measure of complexity is the steepness of the learning curve. In our model, a novice's expected productivity in occupations A and B is $1 - \sigma_x^2$ and $1 - \sigma_y^2$ respectively, whereas his terminal productivity is equal to unity in both. So the learning curve is steeper in B. At first blush this may seem counterfactual, since we know that people tend to focus their learning early on in life. But this is exactly what the Stepping Stone model says -- early jobs prepare us for later jobs. By the time he directs his first operation, a Brain Surgeon has held other jobs -- Intern, Surgical Assistant -- that have prepared him for the task at hand, so that we would not expect much further learning. [In keeping with this, we found only a modest learning curve in angioplasty operations in Jovanovic and Nyarko (1995a)]. The same is true of a Chief Executive Officer; by the time he assumes the position, he has typically held a whole host of jobs. The issue is one of sample selection -- if we forced a novice out of medical schools to perform actual operations on the brain, we would see a very steep learning curve!

A different explanation for why people do safe jobs first is based on incomplete insurance markets, and risk-aversion: People get richer as they age, and they can tolerate the greater risk. Moreover, in evaluating any hypothesis on the effects of age, one must include time effects that control for technological change and other forces that are unrelated to the personal growth of

human capital that we are discussing here. One useful data set is the cross-sectional survey of Indian farmers studied by Foster and Rosensweig (1995).

Share of Rabi-Season Acreage Planted to Different Crops by Age of Head

Age	Wheat	Pulses	Other*
25	0.25	0.42	0.21
35	0.42	0.18	0.27
45	0.45	0.18	0.25
55	0.42	0.20	0.28
65	0.57	0.14	0.20
75	0.60	0.16	0.16

* Includes sorghum, corn, other cereals, oilseeds, sugarcane, and cotton. A residual category consisting primarily of vegetables is not shown.

$$\text{share of wheat} = \underset{(4.21)}{0.25} + \underset{(4.21)}{0.004}\,\text{Age} + \underset{(3.34)}{0.0004}\,\text{Land} \qquad \text{Adj. } R^2 = 0.02, \quad n = 4{,}118$$

Table 2

We can distinguish the risk-aversion/imperfect capital markets hypothesis from ours with the Foster and Rosensweig data. The gave us the information in Table 2 and the accompanying regression. The data are from the 1970-71 round of the National Council of Applied Economic Research Additional Rural Incomes Survey. They describe rural households in India from a national probability survey begun in 1968-69. From the viewpoint of Table 2 in which there is a 50-year variation in age, these are essentially cross-section data. For 4826 households, the data provide information by crop on acreage planted during the two main agricultural seasons in India (Kharif and Rabi) as well as detailed information on the household including age of the household head and landholdings which are correlated with wealth. These can be used to proxy for any wealth effect on choice of crop. The table and regression focuses on the Rabi season crops. Table 2 shows that young farmers plant more pulses (a type of lentil) and later on in life shift toward planting wheat.

23

The regression shows that this is not caused by a wealth effect. A form of expertise that is transferred from pulses to wheat is knowing how to deal with external markets: Wheat is more likely to be a cash crop and hence riskier than pulses, and pulses may offer a more gradual introduction to external markets.

Is it reasonable to associate costliness of mistakes with the concept of complexity of an activity? The learning curve is steeper in B in the sense that the ratio of ideal to initial expected performance, $1/\sigma_y^2$, is higher. We find it natural to think of B as the more complex activity. Complexity goes hand in hand with steepness of learning if one follows Kremer (1993) and defines a complex activity as embodying a greater number of tasks. [See equation (12) of Jovanovic and Nyarko (1995a) for a learning curve formula when there are N "stages to an activity. This formula exhibits greater steepness when N is larger]. If we accept this, then our model implies that simple activities should precede complex ones. In support of this implication, Wilk and Sackett (1995) find that people do hold more complex jobs as they age. Their measure is Gottfredson's (1986) Occupational Aptitude Patterns Map, an index of cognitive ability needed to perform the activities on a job: Vocabulary, reading, math, perceptual speed, and so on. The index assigns each job a "complexity integer" on a scale of 1 to 10. Chemists, physicians, and engineers are 10's, while yarn sorters, general laborers, and baker helpers are 1's. Wilk and Sackett's first table shows that in the NLSY sample, over a span of 12 years, this index of the cohort rose by more than a full category, from 3.35 to 4.84.

Climbing the occupational ladder: Hall and Kasten (1976), Miller (1984), Sicherman and Galor (1990), Baker, Gibbs, and Holmstrom (1994), and Biddle and Roberts (1994) all provide vertical rankings of occupation, based mostly on pay, and they all find that workers gradually move up the occupational ladder. This also means that their earnings tend to rise, as they also do in the Stepping Stone model. Earnings rise with experience in the Bandit model too, but there workers would, according to Corollary 1, want to start at the top of the ladder right away, because the highest paid occupations also tend to be more risky.

The transfer of skills from blue-collar work to white collar work: In the NLSY data Neal (1995) finds one intra-firm occupation change for every six spells of full-time employment. For white collar workers, a large fraction of these moves are from a given occupation into management. For blue collar workers, most changes are movements between craftsmen, operative, and laborer, but still with significant movement into management. The Stepping Stone model says that this is because skills are transferred from blue collar work to white collar work. This claim gets support from Keane and Wolpin (1994) who study the impact on wages of prior experience of people that

have moved in both directions. They find (Table 5) that blue collar experience "raises" white collar wages by more than white collar experience raises blue collar wages.

5.B   Evidence favoring the Bandit model:

Complexity and innate ability: Wilk and Sackett (1995) and Wilk, Desmaris, and Sackett (1995) also found a definite tendency for individuals with higher cognitive ability (an innate trait) to gravitate towards the more complex activities over time, suggesting that a filtering or screening process goes on with the accumulation of experience. In keeping with this, Murnane, Levy, and Willett (1995) find that cognitive skills become increasingly related to wages as people age. These findings strongly favor the bandit model, since in the Stepping Stone model people are all *ex ante* the same.

Sociologists' studies of occupational ladders: Sociologists have studied occupational ladders from the point of view of a career path, but their evidence does not seem to show a rise in the complexity of work as people age. For example, Wilensky (1961) defines a career as a succession of related jobs, arranged in a hierarchy of prestige through which persons move in an ordered, more or less predictable sequence. He divides job transitions as orderly if the skills and experience gained on one job are directly functional to performance in the next job and jobs are arranged in a hierarchy of prestige. He then defines a disorderly career as one in which less than a fifth of the work life is in functionally related, hierarchically ordered jobs. Wilensky's Table 1 (p. 526) shows that 63 % of the of the sample has orderly careers. So clearly there are occupational ladders, and this perhaps favors the Stepping Stone model. But the study of complexity of work is not supportive of the that model. The complexity scale most widely used by Sociologists is in the Dictionary of Occupational Titles (DOT), put out by the Department of Labor scores occupations based on measures of complexity. There are three dimensions, dealing with the complexity of relationships to data, people, or things. Cohn and Schooler (1983) have linked the concept of a career to the concept of complexity; the latter being based on the DOT and on their own surveys of the people working in an occupation. Among the 3100 men in their sample, [p. 111, Figure 5.2]. shows that the correlation between complexity in 1964 and complexity in 1974 is 0.77. However, they do not show that complexity of work increases over one's lifetime. Finally, using data from a survey of the Institute for Social Research done in 1967, Baron, and Bielby (1982) also get mixed results on the relation between age and complexity.

25

6. The Hybrid Model

A person's productivity depends on innate traits and on the amount of acquired expertise, and a future study that tries to quantify how important each force is will need a hybrid of the Stepping Stone and Bandit models. We now sketch one such hybrid. The production functions now are:

$$q_A = \alpha[1 - (x - z_A)^2] \quad \text{and} \quad q_B = \beta[1 - (y - z_B)^2] . \tag{7}$$

The variables are defined as before. Assume <u>no</u> signal confusion, so that a period's experience in growing apples reveals both α <u>and</u> x fully, and similarly, a period's experience with banana growing reveals β and y. Further, the vector (x,y) is independent of the vector (α,β).

6.1. <u>Symmetric occupations with specific learning</u>.

The simplest case is one of complete symmetry between the two occupations. Assume that α and β are independent and normally and identically distributed, with mean k and with variance s^2. Moreover, assume that x and y are also independent (so that $\rho = 0$ and learning is occupation-specific) and identically and normally distributed, with variance equal to σ^2. In this case, <u>all</u> the movement is due to mismatch, and learning acts as a deterrent, since it creates occupation-specific knowledge. Suppose, moreover, that $p = 1$. Both occupations look identical ex-ante. Someone who has entered A when young will switch to B if and only if

$$\alpha < k(1 - \sigma^2) . \tag{8}$$

The left-hand side of this inequality is the person's earnings in A should he decide to stay there, and the right-hand side gives his expected earnings should he switch to B. Let Φ denote the standard cumulative normal integral. Since a similar inequality applies to those who enter B when young, the economy wide rate of movement is equal to the probability that the above inequality is satisfied, and this equals $\Phi(-k\sigma^2/s)$. Here, σ is an index of occupational-specific human capital and is decreasing in σ. Further,

$$\lim_{\sigma \to 0} \Phi(-k\sigma^2/s) = \tfrac{1}{2}, \tag{9}$$

a result that we would have expected given symmetry and the absence of switching costs.

6.2. <u>Symmetric α and β (as in 6.1.); $\rho = 1$, and $\sigma_x^2 < \sigma_y^2$</u>.

In this case, we expect some Stepping Stone movement, or at least we would get it if α and β were degenerate. Let p be such that the mean productivity of novices is the same in the two

26

occupations. That is, p satisfies

$$k(1 - \sigma_x^2) = pk(1 - \sigma_y^2). \tag{10}$$

Since $p = 1$, one period's experience teaches the worker how to perform both jobs perfectly. But α and β are still independent. For the sake of illustration (this is not an equilibrium argument), assume that young workers are, as in 6.1., equally divided between occupations A and B. Then occupational switchers consist of those who when young worked in A, but who found out that their $\alpha < pk$, and those who when young worked in B, but found that $\beta < p^{-1}k$. The rate of movement now is

$$\frac{1}{2} \Phi\left[\frac{(\sigma_y^2 - \sigma_x^2)k}{(1 - \sigma_y^2)s}\right] + \frac{1}{2} \Phi\left[\frac{-(\sigma_y^2 - \sigma_x^2)k}{(1 - \sigma_x^2)s}\right]. \tag{11}$$

In the absense of Stepping Stone forces (i.e., when $\sigma_x^2 = \sigma_y^2 = 0$), since α and β are symmetric movement due to Bandit forces cancel out. In particular, the two terms in (11) are equal to each other and are both equal to 1/4. Now introduce Stepping Stone forces (i.e., $\sigma_x^2 < \sigma_y^2$). Then the first term exceeds the second. The difference between the two terms can loosely be called Stepping Stone movement.

7. Discussion and Extensions

7.1. Nonlinear pricing of efficiency, and absolute advantage.

In the Stepping Stone model we assumed that in A and in B there is a constant equilibrium price (1 and p respectively) per efficiency unit. If I produce twice as much q_A, my equilibrium reward will be twice as large. Since everyone in the Stepping Stone model is *ex ante* the same, no one has an absolute advantage over anyone else. In the Bandit model, we normalized both prices to (1, 1). In the Bandit model some workers also have an absolute advantage. Nevertheless, because prices are linear in skill in both models, to determine the equilibrium assignments, only the ratio of expected productivities in the second period, and only the ratio of expected lifetime incomes of career paths matters in the first period. In this sense assignments are based entirely on comparative advantage. But this assumption is inappropriate in situations in which one can not substitute the quantity of service for the quality of service. For example, suppose that for some

27

reason a restaurant must order its apples from a single source so that every restaurant can deal with only one apple producer. There is no reason why an apple producer that is twice as efficient should get twice the reward. Sattinger (1975), Rosen (1981), and Becker (1981) discuss settings in which the linearity fails. In their settings skill is exogenous. In ours skill is acquired though experience.(Kremer (1993, section 2) discusses skill-augmenting education in such a model). Still, our analysis should apply to this setting as well. Instead of being q_A and q_B, the rewards to the two activities would equal $p_A(q_A)q_A$ and $p_B(q_B)q_B$ respectively,

where $p_A(.)$ and $p_B(.)$ would be two non-linear price functions for the level of each type of skill. (In the models of Sattinger, Becker and Kremer, the $p_i(.)$ functions are increasing because a more productive worker is in equilibrium combined with more productive complementary resources). The spirit of the analysis remains the same.

7.2. Observable worker characteristics.

Workers in the model are *a priori* identical, but it isn't hard to indicate how one would allow for ex-ante differences in workers' characteristics. Even if producers of apples and bananas should differ in some way, it is natural to maintain the assumption that they all produce the same product, and that as a result, they all face the same prices. In a model in which information enhances productivity (just like in the Stepping Stone model) Rosensweig (1995, p. 154) asks how schooling may affect productivity; he argues that schooling can tighten priors around optimal decisions, and it can raise the informational content of each use of the technology. And in a Bandit model, Johnson (1978) and MacDonald (1980) argues that schooling raises earnings by informing workers about where his optimal assignment is -- that is, schooling tightens priors about comparative advantage.

7.3. Terms of trade and wage differentials.

In the last 20 years U.S. manufacturing employment has shrunk, and the skilled-unskilled wage differential has increased. Some have examined the possibility that this was caused by growth of manufacturing productivity abroad (Berman, Bound, and Griliches 1994). A well-documented transition on a job ladder is the flow of labor from blue collar to white collar work (see section 5.1). Since $A \rightarrow B$ is the direction in which labor flows in our model, we can think of A as blue collar (e.g., manufacturing), and B as white collar (e.g., some services) If the rest of the world learns how to manufacture more efficiently, the demand for A should fall so p should rise. We should see a rightward move along the second step on the supply curve in figure 3. We can therefore analyze the effect of such changes in the world's productivity through their effect on p.

28

7.4. Internal labor markets.

We have not discussed firms at all, and yet we know that a lot of occupational movement happens inside firms, and that firms tend to promote from within. From the results of Neal (1995) we know that there is at least one intra-firm occupation change for every six spells of full-time employment, and not every such spell represents an occupational change either. So internal occupational movement is important, and the present model could generate a distinction between internal and external occupational movement by introducing firms, and some reason why there would be firm-specific components to the (x, y) vector.

7.5. Valuing stepping stone mobility.

In the model of section 2, to what extent does the option to change jobs raise GNP? In the context of a Bandit model Jovanovic and Moffitt (1990) found the figure to be 8% of GNP. If we value the two kinds of output at their equilibrium prices -- 1 for apples and p for bananas -- then when $\delta = 1$ (so that the gain in lifetime value equals the gain in lifetime output), the number is easy to calculate. The increased output is just the number of young that move (see Figure 4), multiplied by the lifetime utility gain associated with moving: $V_{AB} - \max\{V_{AA}, V_{BB}\}$.

7.6. The advantages of countrywide occupational diversity.

The Stepping Stone model and the Bandit model both imply that with a richer menu of occupations to choose from, the worker's lifetime productivity will be higher. A country that specializes in either apples or bananas will not be able to afford its nationals the gains from either kind of mobility -- a disadvantage of specialization.

8. Conclusion

The Stepping Stone model explains why the complexity of work increases over one's lifetime. We have used this model to discuss the concept of occupational mobility. In contrast to the Bandit model, the quality of the technologies is known in the Stepping Stone model; what is uncertain best to operate the technology. The Stepping Stone model says that activities that are informationally "close" will form a ladder, but the safer ones should come first : They are a natural training ground because in such activities, mistakes imply smaller foregone output. The model also says that labor will flow from occupations with flatter learning curves and into occupations where learning curves are steep, as long as learning is sufficiently transferable between occupations.

The general message of this paper is however, much broader than the question of occupational mobility that we have focused on, and it is this: To understand human capital, it helps to distinguish information about personal traits and about the content of jobs on the one hand, from information about how to perform a given task on the other. The informational approach to human capital has so far stressed the first kind of information, information about parameters pertaining to traits of people and jobs. The Stepping Stone model illustrates that information theory can also illuminate the process by which people acquire general as well as job-specific skills.

References:

Altonji, J., and Shakotko, R.
    (1987) "Do Wages Rise with Job Seniority?" Review of Economic Studies, 54: 437 - 459.
Baker, G., Gibbs, M., and Holmstrom, B.
    (1994) "The Internal Economics of the Firm -- Evidence from Personnel Data," Quarterly Journal of Economics 109, no. 4: 881 - 920.
Baron, J., and Bielby, W.
    (1982) "Workers and Machines: Dimensions and Determinants of Technical Relations in the Workplace," American Sociological Review 47 (April): 175-188.
Becker, G.
    (1981) Treatise on the Family. Cambridge: Harvard University Press.
Berman, E., Bound, J. and Griliches, Z.
    (1994) "Changes in the Demand for Skilled Labor within U.S. Manufacturing: Evidence from the Annual Survey of Manufactures,"Quarterly Journal of Economics 109, no. 2: 367 - 398.
Biddle, J. and Roberts, K.
    (1994) "Private Sector Scientists and Engineers and the Transition to Management," Journal of Human Resources 29, no. 1: 82 - 107.
Chow, G.
    (1983) Econometrics New York: McGraw-Hill.
Cohn, M. and Schooler, C.
    (1983) Work and Personality: An Inquiry in to the Impact of Social Stratification New Jersey Ablex Publishing Corporation.
Davis, S. and Haltiwanger, J.
    (1992) "Gross Job Creation, Gross Job Destruction, and Employment Reallocation," Quarterly Journal of Economics 107, no. 3: 819 - 864.
Foster, A. and Rosensweig, M.

(1995) Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture," Journal of Political Economy 6, no. 103: 1176 - 1209.

Gottfredson, L.

(1986) "Occupational Aptitude Patterns Map: Development and Implications for a Theory of Job Aptitude Requirements," Journal of Vocational Behavior 29: 254-291.

Hall, R. and Kasten, R.

(1976)"Occupational Mobility and the Distribution of Occupational Success Among Young Men," American Economic Review, Papers and Proceedings 66, no. 2: 309 - 315.

Johnson, W.

(1978) "A Theory of Job Shopping," Quarterly Journal of Economics: 261 - 277

Jovanovic, B. and Moffitt, R.

(1990) "An Estimate of a Sectoral Model of Labor Mobility," Journal of Political Economy 98, no. 4: 827 - 52.

Jovanovic, B. and Nyarko, Y.

(1995) "A Bayesian Learning Model Fitted to a Variety of Empirical Learning Curves," Brookings Papers on Economic Activity Microeconomics Issue, no.1.

Keane, M. and Wolpin, K.

(1994) "The Career Decisions of Young Men," unpublished, New York University.

Kremer, M.

(1993) "The O-Ring Theory of Economic Development," Quarterly Journal of Economics 108, no. 3: 551 - 576.

Krusell, P., Ohanian, L., Rios-Rull, J. and Violante, G.

(1995) "Capital-Skill Complementarity and Inequality," unpublished paper, University of Rochester.

MacDonald, G.

(1980) "Person-Specific Information in the Labor Market," Journal of Political Economy 88, no. 3: 578 - 97.

Miller, R.

(1984) "Job Matching and Occupational Choice" Journal of Political Economy 92, no. 6: 1086 - 1120.

Mincer, J.

(1974) Schooling, Experience and Earnings Brookfield, VT: Ashgate Publishing.

Murnane, R., Levy, F. and Willett, J.

(1995) "The Growing Importance of Cognitive Skills in Wage Determination," NBER working paper #5076.

Neal, Derek,

(1995) "The Complexity of Job Mobility Among Young Men," unpublished, University of
Chicago.

O'Flaherty, B. and Siow, A.

(1992) "On the Job Screening, Up or Out Rules, and Firm Growth," <u>Canadian Journal of
Economics</u> 25, no. 2: 346 - 67.

O'Neill, D.

(1995) "Wage Dispersion over the Life Cycle: Implications for Models of Individual Wage
Determination," unpublished, University of Newcastle upon Tyne.

Prescott, E.C., and Visscher, M.

(1980) "Organization Capital" <u>Journal of Political Economy</u> 88 (June): 446 - 61.

Rosen, S.

(1972) "Learning and Experience in the Labor Market," <u>Journal of Human Resources</u> 7: 325
- 42.

Rosen, S.

(1981) "Substitution on Production," <u>Economica</u>.

Rosensweig, M.

(1995) "Why Are There Returns to Schooling?" <u>American Economic Review</u>, Papers and
Proceedings 85, no. 2: 153 - 58.

Roy, A.D.

(1951) "Some Thoughts on the Distribution of Earnings," <u>Oxford Economic Papers (New
Series)</u> 3: 135 - 146.

Sattinger, M.

(1975) "Comparative Advantage and the Distribution of Earnings and Abilities,"
<u>Econometrica</u> 18: 455 - 68.

Sicherman, N. and Galor, O.

(1990) "A Theory of Career Mobility," <u>Journal of Political Economy</u> 98, no. 1: 145 - 76.

Topel, R.

(1991) "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority," <u>Journal of
Political Economy</u> 99, no. 1: 145 - 176.

Waldman, M.

(1984) "Job Assignments, Signalling, and Efficiency," <u>Rand Journal of Economics</u>, 15, no.2
(Summer): 255 - 270.

Weiss, Y.

(1971) "Learning by Doing and Occupational Specialization," <u>Journal of Economic Theory</u>
3: 189 - 98.

Wilensky, H.

  (1961) "Orderly Careers and Social Participation: the Impact of Work History
  on Social Integration in the Middle Class," <u>American Sociological Review</u> 26: 521-39.

Wilk, S., Desmaris, L. and Sackett P.

  (1995), "Gravitation to Jobs Commensurate with Ability: Longitudinal and Cross-Sectional
  Tests", <u>Journal of Applied Psychology</u> 80, no. 1: 79 - 85.

Wilk, S. and Sackett, P.

  (1995) "A Longitudinal Analysis of Ability-Job Complexity Fit and Job Change,"
  unpublished, Department of Management, The Wharton School.

<u>Appendix</u>

Proof of Proposition 4: Define $\Delta \equiv E\beta - E\alpha$ and let $\alpha_0 \equiv \alpha - E\alpha$, and $\beta_0 \equiv \beta - E\beta$. From Lemma 2,

$$V_\alpha = E\alpha + \delta E\text{Max}\{\alpha, E[\beta|\alpha]\} = (1+\delta)E\alpha + \delta E[\text{Max }\{\alpha_0, \Delta + d\alpha_0\}].$$

Similarly,

$$V_\beta = (1+\delta)E\beta + \delta E[\text{Max }\{\beta_0, -\Delta + c\beta_0\}].$$

Consider as fixed σ_β^2 and σ_α^2 with $\sigma_\beta^2 > \sigma_\alpha^2$. Consider the expressions above functions of Δ and r. Define $D(\Delta,r) \equiv V_\alpha - V_\beta$. Then

$$D(\Delta,r) = (1+\delta)\Delta + \delta\Big[E[\text{Max }\{\beta_0, -\Delta + c\beta_0\}] - E[\text{Max }\{\alpha_0, \Delta + d\alpha_0\}]\Big]. \tag{12}$$

Using the symmetry of the normal distribution and the fact that $E\alpha_0 = 0$ it is easy to check that $E[\text{Max }\{\alpha_0, d\alpha_0\}] = |d-1|E[|\alpha_0|]/2 = |d-1|\sigma_\alpha/(2\pi)^{1/2} = |r^{1/2}\sigma_\beta - \sigma_\alpha|/(2\pi)^{1/2}$. Similarly, $E[\text{Max }\{\beta_0, c\beta_0\}] = |r^{1/2}\sigma_\alpha - \sigma_\beta|/(2\pi)^{1/2} = \sigma_\beta - r^{1/2}\sigma_\alpha$, (where the last equality follows from r<1 and $\sigma_\beta > \sigma_\alpha$). Hence

$$D(0,r) = \{1/(2\pi)^{1/2}\}\left\{\sigma_\beta - r^{1/2}\sigma_\alpha - |r^{1/2}\sigma_\beta - \sigma_\alpha|\right\}$$

$$= \left\{ \begin{array}{ll} \{1/(2\pi)^{1/2}\}\left\{(\sigma_\beta + \sigma_\alpha)(1 - r^{1/2})\right\} & \text{when } r \geq \sigma_\alpha^2/\sigma_\beta^2 \\ \{1/(2\pi)^{1/2}\}\left\{(\sigma_\beta - \sigma_\alpha)(1 + r^{1/2})\right\} & \text{when } r \leq \sigma_\alpha^2/\sigma_\beta^2 . \end{array} \right. \tag{13}$$

It is easy to see that $\sigma_\beta^2 > \sigma_\alpha^2$ implies that

$$D(0,r) > 0 \text{ for all r in } [0,1) \text{ and } D(0,1)=0. \tag{14}$$

Assume for now that all the terms in the definition of $D(\Delta,r)$ above are differentiable with respect to Δ. Then it should be clear that $|\partial E[\text{Max }\{\beta_0, -\Delta + c\beta_0\}/\partial\Delta| \leq 1$ and $|\partial E[\text{Max }\{\alpha_0, \Delta + d\alpha_0\}]/\partial\Delta| \leq 1$ so $\partial D(\Delta,r)/\partial\Delta \geq (1+\delta) - 2\delta = 1-\delta > 0$. In particular, $D(\Delta,r)$ is strictly increasing in Δ. By replacing derivatives with derivatives from the left and right, it is easy to check that differentiability is not required for this latter result.

Hence, using the obvious continuity of D in Δ, there exists for each r in [0,1], a unique critical value $\Delta^c(r)$ of Δ such that $D(\Delta,r)>0$ for $\Delta> \Delta^c(r)$; $D(\Delta^c(r),r)=0$; and $D(\Delta,r)<0$ for $\Delta< \Delta^c(r)$. From (14), $\Delta^c(r) < 0$ for all r in [0,1). Finally, from (13), $D(0,1) = 0$. Since $\Delta^c(1)$ is by definition the unique value of Δ such that $D(\Delta,1) = 0$, we conclude that $\Delta^c(1)=0$. ∎

34

Extending the argument to a general equilibrium context.

There is a representative consumer with income I' to spend on c_A units of apples, c_B units of bananas and M units of a composite good M with prices p_A, p_B and 1 respectively. The consumer therefore solves the problem

$$\underset{\{c_A, c_B, M\}}{\text{Max}} \quad u(c_A, c_B, M) \qquad \text{s.t.} \qquad p_1 c_1 + p_2 c_2 + M \le I. \qquad (15)$$

We suppose the agent has Cobb-Douglas utility

$$u(c_A, c_B, M) = c_A{}^\mu c_B{}^\nu M^{1-\mu-\nu} \quad . \qquad (16)$$

This results in the demands for apples and bananas given by

$$c_A = \mu I / p_A \quad \text{and} \quad c_B = \nu I / p_B \quad . \qquad (17)$$

The relative price of bananas with respect to apples is therefore

$$p \equiv p_B / p_A = (\nu/\mu) \cdot (c_B/c_A)^{-1} \quad . \qquad (18)$$

which is the relative demand curve of bananas (relative to apples) as a function of the relative price of bananas. By adjusting the values of p_B and p_A we can arbitrarily adjust the levels of apple and banana demands keeping the ratio fixed.

In figures 8 and 9 below we draw this demand curve. We draw on the same diagram the supply curve of bananas relative to apples (and where we ignore the sections where apple production is zero so relative banana production is infinite). In figure 8 we see that when $\rho=0$ there is a unique equilibrium with price $p = \bar{p}_1$ and relative output of bananas equal to $Q_B{}^*/Q_A{}^* = \nu/(\mu \bar{p}_1)$, [where from section 2, $\bar{p}_1 \equiv (1-\sigma_x{}^2+\delta)/(1-\sigma_y{}^2+\delta)$].

When $\rho=1$, the equilibrium price depends upon the position of the demand curve, and in particular on the value of ν/μ. Figure 9 depicts three such demand curves, D, D' and D", corresponding to relatively low, medium and high values of ν/μ. These yield equilibrium relative prices of $p=1$, p in $(1,\bar{p})$ and $p=\bar{p}$ respectively [where from section 2, $\bar{p} = (1-\sigma_x{}^2)/(1-\sigma_y{}^2)$].
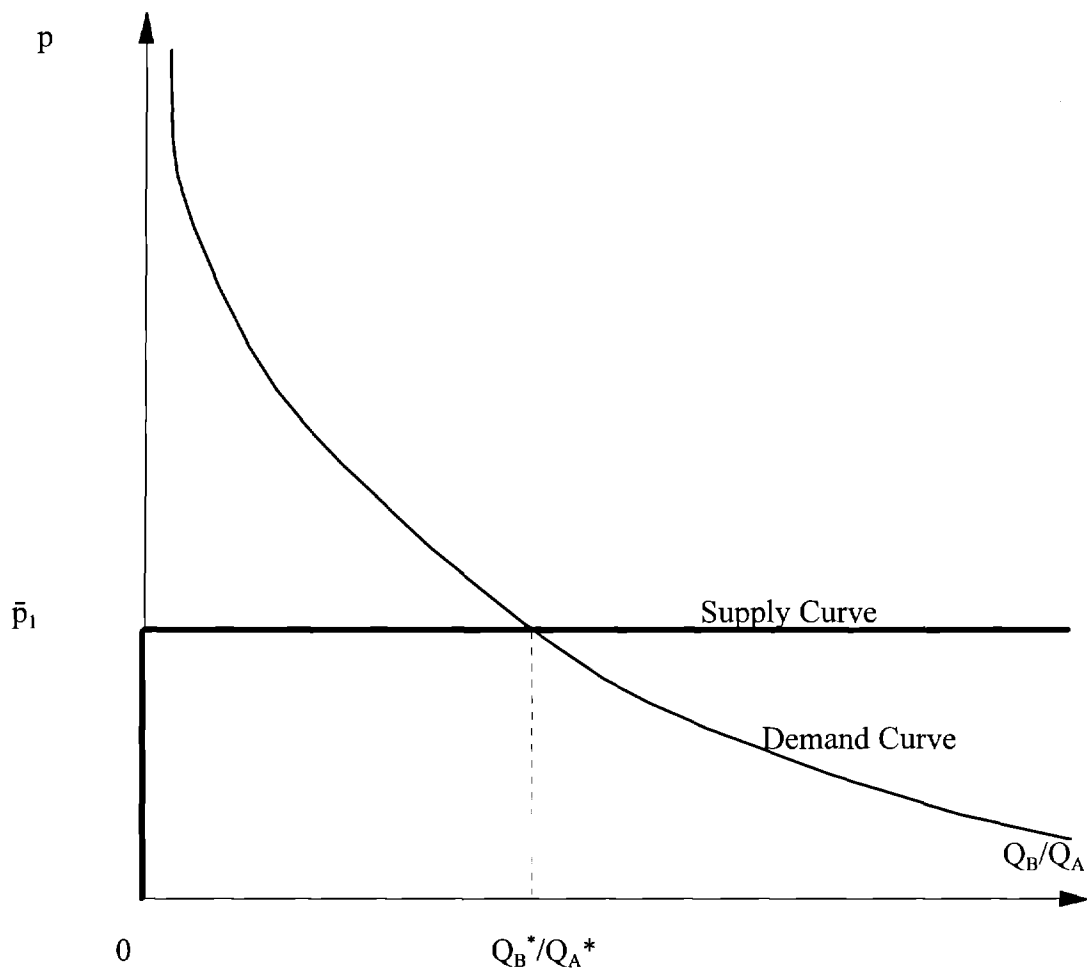
35

Fig. 8 : The Relative Supply and Demand Curves of Bananas when $\rho=0$
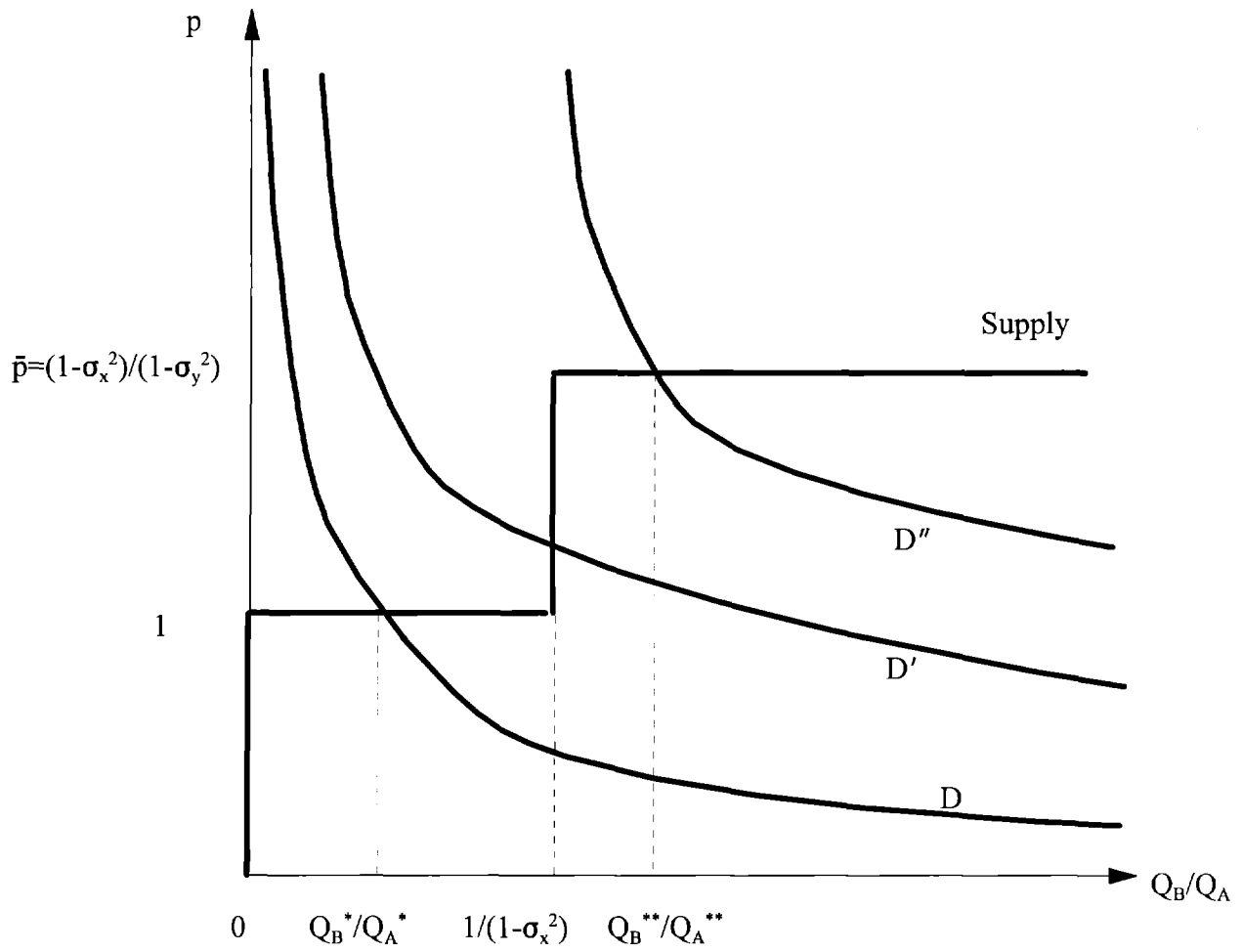
Fig. 9 : The Supply Curve and Three Different Demand Curves for Bananas when $\rho=1$.

The Farmer's Intertemporal income/savings decision and Risk Neutrality

In the text we assume that the farmer will maximice lifetime <u>expected</u> income. We now show that this assumption is not inconsistent with the preferences described above in this appendix. Suppose that all consumption of an individual takes place <u>after</u> the output when old has been received. In particular there is no consumption when young. Suppose that before consumption the old have income of I (made up of income received when old plus income when young, the latter perhaps with interest). Recall that in the appendix we have the representative consumer with utility function $c_A{}^\mu c_B{}^\nu M^{1-\mu-\nu}$ over the three goods c_A, c_B and M with prices p_A, p_B and 1. It is easy to check that the indirect utility function is linear in I. Hence the farmer will be risk neutral in income I. This means that in making his labor market decisions, the farmer will want to maximize expected income.

Random versus Non-Random Prices

In deriving the optimal career path of individuals we assume that prices are constant and the same in each period. So long as we assume a steady state in career choices of young and old, we shall obtain equal distributions of prices in each period. Next suppose that there is a continuum of agents so that no individual can influence the price by his own career choice. Then we may replace the price p with its expected value and arrive at the same conclusions. If in addition the x's and y's of agents are all independent of each other, we can invoke the strong law of large numbers to conclude that the prices will actually be constant.

Some More Remarks On Lemma 2:

Define

$$M \equiv \{(\sigma_\alpha{}^2, \sigma_\beta{}^2, \text{Cov}(\alpha,\beta)) \in \mathbb{R}_{++} x \mathbb{R}_{++} x \mathbb{R} \mid \sigma_\alpha{}^2 < \sigma_\beta{}^2, \text{ and } \text{Cov}^2(\alpha,\beta) < \sigma_\alpha{}^2 \sigma_\beta{}^2\}. \qquad (19)$$

Then M is the set of parameters that which could define the variance-covariance matrix of the normal distribution and which obeys our hypothesis that $\sigma_\alpha{}^2 < \sigma_\beta{}^2$. The expressions c and d in Lemma 2 are functions of the parameter vector $m = (\sigma_\alpha{}^2, \sigma_\beta{}^2, \text{Cov}(\alpha,\beta)) \in M$. For emphasis we shall index these by m so that $c(m) \equiv [\text{Cov}(\alpha,\beta)/\sigma_\beta{}^2]$ and $d(m) \equiv [\text{Cov}(\alpha,\beta)/\sigma_\alpha{}^2]$, respectively.

Lemma 2':

a. (i) For all m ∈ M, $|c(m)| < 1$; and (ii) $\forall$ c ∈ $\mathbb{R}$ such that $|c| < 1$ there exists an m ∈ M such that c(m)=c.

b. (ii) For all d ∈ $\mathbb{R}$, there exists an m ∈ M such that d(m)=d.

Proof: (a) (i) For m∈M, $c(m)^2 = [\text{Cov}^2(\alpha,\beta)/\sigma_\beta{}^4] \leq \sigma_\alpha{}^2/\sigma_\beta{}^2 < 1$, so part a.i follows. (ii) Fix any c

such that $|c|<1$ and define $m=(\sigma_\alpha^2, \sigma_\beta^2, \text{Cov}(\alpha,\beta))$ as follows: Set $\sigma_\alpha^2 = c^2$; define σ_β^2 to be any number in $(\sigma_\alpha^2, 1)$ and define $\text{Cov}(\alpha,\beta)$ by $\text{Cov}^2(\alpha,\beta) = c^2 \sigma_\beta^4$ and sgn $\text{Cov}(\alpha,\beta) = $ sgn c. Then clearly $\sigma_\alpha^2 < \sigma_\beta^2$; and $\text{Cov}^2(\alpha,\beta)/\sigma_\alpha^2\sigma_\beta^2 = \sigma_\beta^2 < 1$. Hence $m \in M$. Also, $c(m)^2 \equiv \text{Cov}^2(\alpha,\beta)/\sigma_\beta^4 = c^2$ so $c(m) = c$.

(ii)   Fix any $d \in \mathbb{R}$. Define $m=(\sigma_\alpha^2, \sigma_\beta^2, \text{Cov}(\alpha,\beta))$ as follows: Define $\sigma_\beta^2 = d^2$, set σ_α^2 to be any number in $(0, \min\{d^2, 1\})$, and define $\text{Cov}(\alpha,\beta)$ by $\text{Cov}^2(\alpha,\beta) = d^2\sigma_\alpha^4$ and sgn $\text{Cov}(\alpha,\beta) = $ sgn d. Then clearly $\sigma_\alpha^2 < \sigma_\beta^2$; and $\text{Cov}^2(\alpha,\beta)/\sigma_\alpha^2\sigma_\beta^2 = \sigma_\alpha^2 < 1$. Hence $m \in M$. Also, $d(m)^2 \equiv \text{Cov}^2(\alpha,\beta)/\sigma_\alpha^4 = d^2$ so $d(m) = d$. ∎