STOCHASTIC ASPECTS OF THE LABOR FORCE PROCESS

Haskel Benishay, Northwestern University*

I. Introduction

This paper develops a model of the labor force-employment process within the framework of which the variability which is expected to occur normally in the number of labor force participants, in the number of employed, in accessions and in retirements is highlighted and emphasized. The basic approach consists of grafting a stochastic labor force-employment process onto a stochastic population process. The major improvements over previous work are the following three:[1] (1) the development of results under more realistic assumptions in regard to accessions and retirements (2) the explicit incorporation of employment into the model albeit under simplified assumptions (3) the provision of an outline of a more realistic model which includes multiple entries and exits from the labor force and employment.

II. Notation

The notation and assumptions of the model as well as the description of the population-labor force-employment process are presented below.

Let the time interval from point in time t-1 to point in time t be t, so that t represents either a time interval or its end point. Generally, the symbol t will represent a time interval when flows are treated and the end point of this interval when stocks are considered.

Let E, V, before parentheses represent the expectation and variance of the variable in parentheses.

N_t is an independent exogenous random variable representing the number of live births during interval t.

S is an endogenous random variable representing the number of people alive at a point in time t.

F is an endogenous random variable representing the number of individuals in the labor force at point in time t.

D_d is the possible span of life at birth, for any individual. D_d is a discrete variable representing intervals (units) of time which can consist only of integers which range between a minimum of 1 (one) and a maximum of m. The probability that $D_d = i$ where i takes on consecutively, values from 1 to m, is P_{di}, i.e., $P(D_d=i)=P_{di}$, i=1...m.

d deaths is an endogenous random variable representing the number of individuals who die during a time interval.

a_{lv} represents live accessions to the labor force during a time interval. Lower case v represents life here and henceforth. Within the context of this model a_{lv} is an endogenous random variable.

P_{ali} is the probability that an individual will join the labor force (become an accession) i periods from birth. e is the first possible number of time intervals after birth at which and after which accessions occur (say 14 years after birth). The i in this case can assume values from e to m-1. D_{al} is the random variable defined by $e \le i \le m-1$, and by P_{ali}.

r_{lv} represents live retirements from the labor force during time intervals. It is an endogenous variable determined within the model.

P_{crli} is the conditional probability that an individual will retire from the labor force i periods after birth given that he has already acceded to the labor force.

F_v is an endogenous random variable representing the number of live individuals who are in the labor force at the end of time intervals.

w_{lvF} is an endogenous random variable representing the number of live individuals in the labor force who are becoming employed for the first time during a time interval.

---

*Portions of this presentation and related materials were presented in a faculty seminar in Northwestern University Economics Department in spring 1967, and at the U.S. Department of Labor, Bureau of Labor Statistics in the summer of 1967. Comments made in response to these presentations were useful in providing a perspective in regard to the relative importance of various phases of this work. Support for some of the research embodied in this presentation was provided by the U.S. Department of Labor, Manpower Administration.

[1]This presentation is made against the background of a previous paper with similar but more limited aims [1].

W_{lvF} is an endogenous random variable representing the number of live people within the labor force who are employed at the end of a time interval.

U is an endogenous random variable representing the number of live people within the labor force who are unemployed at the end of a time interval.

P_{wli} is the probability that an individual will become employed, i periods from time of birth. It is assumed that the first possible number of periods (after birth) at which and after which an individual may become employed is e, the same number of time intervals after birth at which first accession to the labor force may occur (14 years after birth). Thus i in this case can also assume only discrete integer values between e and m-1. Additionally, the probability P'_{wli}, that an individual is employed at t, is the cumulative probability of P_{wli}. D_{wli} is the random variable defined by possible values $e \le i \le m-1$ and probabilities P_{wli}.

The exogenous variables under consideration are assumed serially and mutually independent, with one exception. The conditional probabilities of retirement at i are P_{crli} for members of of the labor force at i-1 and zero for non-members at i-1. It is logically impossible to retire from the labor force without having entered first. Retirement from the labor force is thus dependent on previous entry into it.

The exogenous variables in this model are:

1. Births, N.
2. The life span distribution, D_d.
3. The distribution of accession times, birth to accession D_{al}.
4. The distribution of (first) employment times, birth to first employment, D_{wl}.
5. The conditional probabilities of retirements, P_{crli}.

A random birth process and a random death process take place together to form the population process. Each time interval a random number drawn from N, determines births. Next, N random draws from the life span distribution, D_d, determine the length of life i for each of the N births. This process is viewed as continuing repeatedly until process equilibrium (steady state) is reached and beyond into a phase in which various stages of the process occur continually through time. Every period births, and deaths from births in previous periods, occur. Every period individuals born and not yet dead comprise the living population. The labor force-employment process is "grafted" onto the population process. Every period accessions come into being from birth cohorts which came into the world 14 to 30 years or so before this period. Every period individuals who have joined the labor force in the past, retire. Period after period individuals are becoming employed for the first time. Period after period individuals originating from preceding birth cohorts come to the end of their life.

III. The Endogenous Variables: Concrete Definitions

The life span random variable D_d may be viewed as a succession of dichotomous variables X_{di}, applicable to a succession of points in time at i periods from birth, for each of which the two possibilities for an individual at birth are: 1. dying i periods after birth, defined identically as unity (one). 2. dying not after i periods, defined identically as zero. Clearly the probability of unity is P_{di} the probability of zero $(1-P_{di})$. Also, $E(X_{di})$, $V(X_{di})$, are simply P_{di} and $P_{di}(1-P_{di})$ respectively, as is true for the binomial distribution. This view becomes useful for the technical definition of, d, deaths.

Deaths, d, is conceived as a sum of the m terms below, each one of which being the number of individuals born i periods before t and dying during t.

(1.1)

$$d = \sum_{j=1}^{N_{t-1}} X_{d1j} + \sum_{j=1}^{N_{t-2}} X_{d2j} + \cdot \cdot \cdot + \sum_{j=1}^{N_{t-m}} X_{dmj}.$$

Each term above is the sum of a random number N of random draws from X_{di}.

The technical definition of S, the number of individuals alive at t, requires a definition of a second binomial variable, Y_{di} whose parameter is Q_{di}, where Q_{di} is the sum of the previously defined P_{di} from P_{di} to P_{dm}, i.e.:

(1.2)

$$Q_{d1} = P_{d1}+P_{d2}+P_{d3}+ \cdots +P_{dm}$$

$$Q_{d2} = P_{d2}+P_{d3}+ \cdots +P_{dm}$$

$$Q_{d3} = P_{d3} + \dots + P_{dm}$$

$$\vdots$$

$$Q_{dm} = P_{dm}$$

Y_{di} is a binomial variable whose value of unity represents the state of being alive after i-0' periods from time of birth and whose value of 0 represents the state of being dead after i-0' or less where 0' is a number as close to zero as is conceivable.

Q_{di} is the probability that an individual will remain a live member of the population, S, after i-0' periods from the time it is born, while $(1-Q_{di})$ is the probability that an individual will no longer be in S, i-0' periods or less from birth. Births are thereby dichotomously divided into those who die and leave the population after i-0' periods or less and those who still remain part of the population after i-0'.

The number of people in the population can also be expressed as a sum of terms each of which being a random number of random draws in the explicit way below.

(1.3)

$$S = \sum_{j=1}^{N_{t-1}} Y_{d1j} + \sum_{j=1}^{N_{t-2}} Y_{d2j} + \dots + \sum_{j=1}^{N_{t-m}} Y_{dmj}.$$

Each term is a sum of a randum number of births at t, t-1...t-m+1 which are randomly "still alives" at point in time t.

Accessions, a_{lv}, is similarly conceived to be a sum of terms originating from birth cohorts before period t which produce accessions in period t. For the purpose of expressing a_{lv} under this view additional notation and discussion are necessary.

X_{alvi} is a binomial variable whose value of unity represents the state of being alive and entering the labor force i periods from time of birth and whose value of zero represents not being in this state. The probability of X_{alvi} given previous assumptions, is P_{ali} $Q_{d(i+1)}$ and the range of i is between e and m-1. Accessions, a_{lv}, can thus be viewed as the sequence of terms below:

(1.4)

$$a_{lv} = \sum_{j=1}^{N_{t-e}} X_{alvej} + \sum_{j=1}^{N_{t-(e+1)}} X_{alv(e+1)j} + \dots + \sum_{j=1}^{N_{t-(m-1)}} X_{alv(m-1)j}$$

Each term is a sum of randum number N_{t-i} of random draws from X_{alvi}.

Live retirements r_{lv}, is also conceived as a sum of terms which result from past birth cohorts. X_{rlvi} is defined as a binomial variable whose value of unity represents being alive, having already acceded to the labor force, and retiring. The probability P_{rlvi} of X_{rlvi} can be derived from some previous and a few additional definitions. Define P_{crli} as the conditional probability that an individual retires at i periods from birth, given that he has already acceded and not yet retired. Thus, the probability that an individual acceded and is retiring alive at i is for e+1:

(1.51)

$$P_{rlv(e+1)} = P_{crl(e+1)} P_{ale} Q_{d(e+2)}$$

for e+2:

(1.52)

$$P_{rlv(e+2)} = P_{crl(e+2)} [P_{ale}(1-P_{crl(e+1)}) + P_{al(e+1)}] Q_{d(e+3)}$$

for e+3:

(1.53)

$$P_{rlv(e+3)} = P_{crl(e+3)} [P_{ale}(1-P_{crl(e+1)}) (1-P_{crl(e+2)})$$

$$+ P_{a1(e+1)}(1-P_{cr1(e+2)})$$

$$+ P_{a1(e+2)}]Q_{d(e+4)}$$

$$\cdots \qquad \cdots \qquad \cdots$$

for (m-1):

(1.5(m-1))

$$\begin{aligned} P_{r1v(m-1)} = P_{cr1(m-1)} & [P_{a1e}(1-P_{cr1(e+1)}) \\ & (1-P_{cr1(e+2)}) \\ & \cdots (1-P_{cr1(m-2)}) \cdot \\ & + P_{a1(e+1)}(1-P_{cr1(e+2)}) \\ & (1-P_{cr1(e+3)}) \\ & \cdots (1-P_{cr1(m-2)}) \\ & + \cdots + P_{a1(m-2)}]Q_{dm}. \end{aligned}$$

The logic of the above expression is straight-forward. The conditional probability of retiring (first term) is multiplied by the probability of having acceded but not yet retired (terms in square parenthesis) which in turn is multiplied by the probability of being alive (last term). The probability of having acceded but not yet retired by i after birth is essentially the cumulative probability of having acceded by i less the probability of having already retired by i.

With the help of the preceding results, r_{1v}, live retirements, can be written as:

(1.6)

$$r_{1v} = \sum_{j=1}^{N_{t-e-1}} X_{r1v(e+1)j} + \sum_{j=1}^{N_{t-e-2}} X_{r1v(e+2)j} + \cdots + \sum_{j=1}^{N_{t-m+1}} X_{r1v(m-1)j}.$$

And each term, again, is a random number N_{t-i} of random draws from X_{r1vi}.

Given the previous results, the labor force, F_v, can be conceived also as the sum of random sums of random variables where the number of terms in each random sum is the random number of births in t-1. X_{Fvi}, is defined as a binomial variable representing the state of being in the labor force and alive i periods after birth. The probability P_{Fvi} of X_{Fvi} can be gleaned from previous results. Being in the labor force means having entered but not left the labor force and being alive. The probability of such an event has already been provided partially in expressions (1.5i) above. Delete from each of these expressions the first term, P_{cr1i}, and what remains is the probability P_{Fvi} of X_{Fvi}, i.e., of having entered the labor force, not having left it, and being alive.

The labor force, F_v can be viewed as a sum of random draws from a random variable as were other endogenous variables.

(1.7)

$$F_v = \sum_{j=1}^{N_{t-e}} X_{Fvej} + \sum_{j=1}^{N_{t-(e+1)}} X_{Fv(e+1)j} + \cdots + \sum_{j=1}^{N_{t-(m-1)}} X_{Fv(m-1)j}.$$

And now we address ourselves to employment. We treat two aspects of employment, employment inflow and employment stock, i.e., the number of people becoming employed each period and the number of employed people outstanding at a point in time. The former is represented by w_1, the latter by W_1. Two assumptions are made in respect to employment: (1) employment is a state into which an individual enters once (hence the 1 in w_1). (2) First entries into the state of employment occur e periods after birth at which entry into the labor force is assumed to commence.

As has been indicated earlier, we represent the probability of becoming employed i periods from time of birth by P_{w1i}. Consequently, P'_{w1i} the probability of being employed at i, as a result of entering the labor force at i or before i is the cumulative probability of joining the employed anew (becoming employed for the first time) at i and at periods smaller than i.

Thus:

(1.81)

$$P'_{wle} = P_{wle}$$

(1.82)

$$P'_{wl(e+1)} = P_{wle} + P_{wl(e+1)}$$

$$\ldots \qquad \ldots$$

(1.8(m-1))

$$P'_{wl(m-1)} = P_{wle} + P_{wl(e+1)} + \ldots + P_{w(m-1)}.$$

The probability of the binomial variable whose value of unity stands for newly joining the employed, being in the labor force, and being alive, i periods from birth, to which we shall refer as, X_{wlvFi}, is simply the product $(P_{wli})(P_{Fvi})$. Thus the total number of live people 'inflowing' into state of being employed within the labor force, w_1, can be expressed as follows:

(1.9)

$$w_{1vF} = \sum_{j=1}^{N_{t-e}} X_{wlvFej} + \sum_{j=1}^{N_{t-e-1}} X_{wlvF(e+1)j} + \ldots + \sum_{j=1}^{N_{t-m+1}} X_{wlvF(m-1)j}.$$

Similarly, the probability of the binomial variable whose value of unity stands for having joined the employed, being in the labor force, and being alive i periods from birth, to which we shall refer as X_{WlvFi}, is the product $(P'_{wli})(P_{Fvi})$. Consequently, the stock of live people who are in the labor force and employed is expressed as:

(2.0)

$$W_{1vF} = \sum_{j=1}^{N_{t-e}} X_{WlvFej} + \sum_{j=1}^{N_{t-e-1}} X_{WlvF(e+1)j} + \ldots + \sum_{j=1}^{N_{t-m+1}} X_{WlvF(m-1)j}.$$

Unemployment, U, is a by product of preceding expressions. The variable, X_{Ui} represents the state of being, at i periods after birth, unemployed, in the labor force and alive. Its probability is the product of $(1-P'_{wli})(P_{Fvi})$ the term in the first parenthesis of the product being the probability of being unemployed by i, the term in the second representing the probability of being alive and in the labor force.

U can then be expressed as:

(2.1)

$$U = \sum_{j=1}^{N_{t-e}} X_{Uej} + \sum_{j=1}^{N_{t-e-1}} X_{U(e+1)j} + \ldots + \sum_{j=1}^{N_{t-m+1}} X_{U(m-1)j} .$$

IV. Some Results

The expectation and variance of the endogenous variables treated in the preceding section have the same general form. Hence, we shall provide the expectation and variance of one variable, death, d, and indicate that the other results may be expressed similarly.

The expectation and variance of deaths was derived in a previous paper as follows [1,2]:

(2.2)

$$E(d) = E(N) \sum_{i=1}^{m} P_{di}$$

(2.3)

$$V(d) = V(N) \sum_{i=1}^{m} (P_{di})^2 + E(N) - E(N) \sum_{i=1}^{m} (P_{di})^2.$$

The expectation and variance of the other endogenous variables are similarly derived. The difference being the substitution of the probabilities of the respective variables for P_{di}.

V. Some Further Developments

The treatment of r_{1v} and a_{1v}, presented earlier, points the way to a possible solution of some of the problems connected with the construction of a logical and probabilistic framework for the phenomena of multiple entries and retirements into and out of the labor force as well as multiple entries into and exits from the state of employment. It is useful to note that accessions are assumed here to occur concurrently with retirements, (at least from e+1 to (m-1) for each cohort), an assumption which is not usually made in traditional work in this area. It may be also interesting to note that retirements are expressed, for the case of concurrent entry and retirement, in such a way as to take account of the fact that retirements in t-1 cannot retire again in t. In the same spirit we assume in further extensions of this work that first entry, first retirement, second entry, second retirement, etc., both into and out of the labor force and into and out of employment, can occur simultaneously. We then treat second entry in relation to first retirement, second retirmenet in relation to second entry, etc., essentially, in the same fashion as r_{1v}, first retirement, was treated in this work in relation to, a_{1v}, first accession. This provides a sensible framework for the evaluation of the phenomena of multiple entries and exits into and out of the labor force and employment. Work in this direction has been conducted and will be continued.

Additional work has been and is being carried out to investigate the serial correlations and the cross and cross-serial correlations of the endogenous random variables, with the hope that it may become useful in forecasting labor force, employment, and their components.

References

[1] Benishay, H., "On the Construction of Stochastic Working Life Tables," Proceedings of the American Statistical Association, Business and Economics Statistics, 1965. Pp. 332-337.

[2] Benishay, H., "Parameters and Relations of Stochastically Lagged and Disaggregative Time Series," Econometrica, 1967.