UNIVERSITY COLLEGE LONDON

(UNIVERSITY OF LONDON)

SCHOOLING, TRAINING AND LABOUR MARKET OUTCOMES

by

Ioannis G. Doumenis

Thesis submitted for the Degree of Master of Philosophy

of the University of London.

January 1994

To my mother and the sweet

memory of my father.

Abstract

The subject matter of this thesis is the discussion of some economic and econometric issues arising from the effects of schooling and training decisions of individuals – with particular reference to married women – on the probabilities of the various outcomes they face in the labour market. Such outcomes include the states of employment (or unemployment), movement between these states (transition probabilities), the levels of earnings involved in the states.

Issues arising out of some non-structural experimental or "quasi"-experimental analyses of the effect of schooling and training programs on the agents' earnings are dealt with in Chapter 2 first. However, once this is done, we should be aware that not only are we unable to fully characterise what likely solutions would we get in terms of individual behaviour – since utility considerations are not analytically embodied or modelled in such a framework – but also that we are not confronting the evolutionary nature of the problem. That is, the choice of whether to remain in schooling or take a period of training is fundamentally dynamic in the sense that it will affect the whole stream of future earnings and employment outcomes. Clearly a dynamic programming approach would be natural. Structural models are dealt with next; we discuss how these place more "structure" on the way agents' utility is affected and how they are concerned with deriving parameter estimates of variables entering the utility functions. Structural models often seem complex. This may be due to an inability to include heterogeneity or state dependence in them. We turn to such issues next. Chapter 3 begins with an exposition of the literature relevant to a model of labour market transitions. This is followed by an empirical application of this model and the effect of schooling thereof using U.S. data. Chapter 4 develops that aspect of transitions in which emphasis is given to the possibility that job offer probabilities received by women are significantly determined by their level of education.

Acknowledgments

The present work was made possible by the award of a scholarship from the "Alexander S. Onassis" Public Benefit Foundation.

I wish to express my gratitude to my supervisor Professor Richard W. Blundell for guiding me throughout this research. I am also indebted to Dr. Jerzy Szroeter for endless hours on econometrics, going through ideas and providing moral support.

Dr. Pasquale Scaramozzino is also thanked for his very constructive criticisms while Dr. James Symons has devoted much of his time discussing further refinements of the model of transitions.

Ramses Abul Naga deserves special thanks for his suggestions concerning Chapter 4 and in so many other ways besides.

I also thank Victor Aguirregabiria, César Alonso, Dr. Manuel Arellano, Dolores Collado, Dr. Antonis Demos, Professor Costas Meghir, Dr. Ian Preston, Dr. Elias Tzavalis, Dr. George Vasilellis and others who helped me in so many different ways.

Warmest thanks also go to Venia Voyiatzi for her moral support and constant encouragement.

Last but not least, I would like to express my deepest gratitude to my mother who has stood by me in all circumstances and my brother, for being the little wise one !

Contents

Chapter 1

Introduction.

The purpose of this thesis is to discuss some issues pertinent to the schooling and training decisions of individuals and in particular married women and how these decisions relate to labour market outcomes. Some emphasis is placed on issues relating to the probabilities of moving between the states of employment and unemployment, frequently termed transition probabilities.

Firstly, we look at the way the training decisions of individuals relate to outcomes they would subsequently face in the job market. There is a long-standing debate (which is prominent in the U.S.) on this issue namely on the methodology we should adopt to investigate the effects manpower training may have on individuals' wages, duration of employment etc. There are two main schools in this debate; the experimental which focusses on the implementation of the problem having a group of participants into training and a group with which their earnings shall thus be compared to (comparison group), and the non-experimental for which a theoretical framework is developed in the absence of these groups. We analyse issues evolving around these two strands of training evaluation programs in the first section of Chapter 2.

However, we should be aware of the fact that the issues tackled there are "incomplete", one could say, in two ways. Since these models are concerned with typically only the reduced form (no utility considerations are analytically encompassed) they are rather inadequate in characterising the likely solutions that would ensue in terms of individual behaviour. Furthermore, since they relate to a static perspective, they simply ignore the evolutionary nature a decision problem of this kind entails. Individuals wishing to enrol for one more period into a training program, though, or participate in the labour market could, strictly speaking, be envisaged to act in a dynamic framework. Such a decision is intrinsically dynamic in the sense that it will affect the whole stream of earnings and employment outcomes. Issues of a dynamic nature that require even the arguments of the utility functionals of agents to be explicitly estimated and characterised is the theme of discussion in the second section of Chapter 2.

The models of this section in turn are often particularly complex since they are quite demanding from a computational viewpoint. Once we brief over some essentials of the main analytical tool of analysing dynamic discrete choices – dynamic programming – we give an illustration through an important paper which centres on the development and estimation of a structural model relating schooling with job decisions, Wolpin (1987). Wolpin introduces a finite horizon into the standard discrete choice search model and also allows for the probability of receiving job offers to be influenced by the length of stay (i.e. the duration) in an unemployment state.

This complexity qualification however may be hiding a lack of ability to incorporate individual heterogeneity, so important in the experimental or discrete choice estimation issues of the first section. Thus issues of heterogeneity as well as state dependence which can be seen as another form of "triggering" a dynamic framework are next discussed.

In Chapter 3 we look at a model of labour market transitions for a group of white married women, estimated using data from the American Consumer

Expenditure Survey (CES). This model is illustrated in order to demonstrate a way by which state dependence can enter. This is so since the modelling of transitions focusses on how a current labour supply status is influenced by the status in the previous period. Evidence of state dependence is immediately apparent from the descriptive statistics that are presented in this Chapter after an introduction to the CES data set.

The reduced form of this model is an attempt to consider the probabilities of moving between the two labour market states that we have considered in this study, as a product of two terms, one proxying the demand side and the other the supply. We shall see in Chapter 3 that conditional on individual assets – that are proxied by consumption – it is possible to "condition out" the supply side and consider only demand side effects. The model is designed to be a dynamic forward looking one, so that a complete characterisation of the solution would require that it is solved through backwards recursions, as is the case if one wishes to solve a dynamic programming problem. However, for the purposes of illustrating the effect education would have on transitions, we limit ourselves into looking at the reduced form only. Recovering the structural parameters of the utility function which would require some further elaborations is not attempted here. The aim is to look at the influence of education on labour market transitions in a form of the model that only the demand side has an effect. To this end we shall find in Chapter 3 significant effects from both College graduates as well as High School leavers.

This brings us to the next Chapter, 4, which is related to the model of labour market transitions that we had been discussing so far. The idea behind Chapter 4 is to provide a simple two period framework that could put a tiny bit more structure on the labour demand side of the model of transitions. The aim in this Chapter is to make the probability of job offers dependent upon the general level of economic activity, the types of agents that would wish to participate as well as their relative probabilities of being successful in the labour market.

Ideally, these job offer probabilities could be endogenous to the agent that makes the decision but since they can not be viewed under her control – at least at the point of writing! – we can only qualify them as being education related. In the static set-up that is used in this Chapter women decide to engage or not in a human capital augmenting activity in the first period, and face a certain probability of being offered a job in the labour market in the second. The model shares features that are similar in spirit to the Spence (1973) job market signalling model.

One of the insights that we would like to offer with the model of this Chapter is over some role through wealth or the distribution of assets on the participation in the labour market decision of married women. It is also clearly intended that such a formulation for the model would provide some basic link for the way the job offer distribution is modelled in both of the last two Chapters.

In Chapter 4 we shall see that under appropriate assumptions the level of assets would not be an argument into the participation decision. A word of caution, however, would point to the fact that the results on this irrelevance of assets proposition hinge on the assumption of risk neutrality that we adopt – simply for clarity – in the agents' utility formulation. Therefore, it is true that incentive effects may come into play if utility is non-linear and interactions between wealth and the likelihood of participation may present themselves if agents are risk averse.

The thesis sums up in a last Chapter which also contains some suggestions for further research.

Chapter 2

A Critical Look at the Literature.

This chapter will present an eclectic and brief overview on the literature relevant to the role of schooling and training decisions on various labour market outcomes. We shall discuss both non-structural and structural models of the decision to participate in the labour market, start a period of training or carry on schooling. Then we shall also turn our focus to note some issues that are usually not explicitly taken into account in either type of the mentioned models such as heterogeneity and state dependence. These latter issues are often pretty important in the models of the next two sections. Their exclusion from these models may typically lead to less reliable conclusions being drawn but it is often the case that it can be cumbersome to adequately incorporate them into any of these models.

We shall begin with analyses of experimental or quasi-experimental nature of the effect of schooling and training programs on the level of earnings and employment.

2.1 Panel data, cross section and experimental data issues.

In the present section we shall review some of the main economic implications of the impact of manpower training and schooling on individual earnings. More precisely we shall be interested in looking at studies that analyse the aforementioned impact both on the level of the wage rate as well as on the duration of employment. Over the last twenty years, there has been a major concern to report on any particular direction of policy recommendation with respect to this issue, but it seems so far that the subject remains victim of an ongoing controversy. Let us briefly note some elements of this debate.

The two main approaches in the recent years come under the headings of the *experimental* and *non-experimental* respectively. In the former, the subject matter of the analysis is to conduct an experiment whereby you recognise the need to have one comparison group and one for those who participate. At a rudimentary level, the approach entails comparing the effects on, possibly, earnings the programme had on its participants with the earnings of the comparison group. Ashenfelter (1978, 1983), Ashenfelter and Card (1985) and Bassi (1983a, 1983b, 1984) are but some of the main references. On the other hand, the second approach is within a purely econometric framework in which the model of participant outcomes in the absence of training needs to be specified and implemented so as to have in hand a satisfactory model of the process reproducing the data. A form of regression analysis is usually utilised to control for systematic differences in the characteristics (broadly defined) between "participants" and non-participants. A significant body of research has emerged in the past ten years on this strand of the literature which evolves a lot around the work of Heckman and his associates (e.g. Heckman and Robb (1983)).

Ashenfelter (1978) concentrates on an evaluation of the impact of alternative training programmes on the earnings of "classroom" trainees under the Manpower Development and Training Act (MDTA). The trainee cohort comprises of those people who started training in the first three months of 1964 so as to ensure their having completed training by the end of the year. The analysis in the paper offers some advantages and some disadvantages. As evidence in favour of the former comes the fact that unlike for a recent cohort, an early cohort's evolution can be monitored for many years after training in the labour market. In addition, if fruitful lessons in the early stages of the programme can be learnt, these can provide the yardstick upon which changes in public policy in this respect can be effected. However, how relevant it is to base present discussions concerning the conduct of public policy on the results of such early cohorts remains to be seen. Ashenfelter acknowledges this as a caveat impinging upon his framework.

The earnings function of the i^{th} individual in the t^{th} period used in Ashenfelter (1978) is:

$$y_{it} = \alpha + \sum_{j=1}^{k} \beta_j y_{i,t-j} + \sum_{j=1}^{k'} \beta_j'(A_i + t)^j + \epsilon_t + \epsilon_i + \epsilon_{it} \tag{1}$$

with $E(\epsilon_{it})=0$ while ϵ_i is individual specific component of the error term capturing perhaps ability, motivation and ϵ_t time specific pertaining to aggregate movements in earnings. A_i refers to the age of individual i in period t=0 and α and β denote parameters. Then Ashenfelter proceeds to examine the effect of training on earnings. Rewritting the k^{th} order difference equation for earnings as a first order difference equation in matrix notation and using superscripts c and p to denote comparison and participation groups respectively, viz:

$$z^c_{it} = Bz^c_{it-1} + d_{it} + b_i + u_{it} \tag{2}$$

$$z^p_{it} = Bz^p_{it-1} + d_{it} + b_i + u_{it} + R_t \tag{3}$$

with

$z_t = [\, y_t,\ y_{t-1},\ \dots,\ y_{t-k+1}]'$, $\quad d_{it} = \left[\alpha + \sum_{j=1}^{k'} \beta_j'(A_i + t)^j + \epsilon_t \right] \gamma$

$b_i = \epsilon_i \gamma$, $\qquad \gamma = [\,1,\ 0,\ 0,\ \dots,\ 0\,]'$, with k-1 zeroes.

R_t : incremental effect of training on trainee earnings in t^{th} period.

$= 0$, in periods before training

< 0 , likely to be during training period.

Note that the B matrix appearing in (2) and (3) above would look like:

$$B = \begin{bmatrix} \beta_1 & \beta_2 & \dots & \beta_k \\ 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 1 & \dots & 1 & 0 \end{bmatrix}$$

We need not consider the reparameterisations adopted in the paper in greater detail but it should suffice to stress the way he proceeds. Thus Ashenfelter using backward substitutions and (3) , (4) repeatedly and each time in lagged form derives expressions for comparison and participation groups. Using also

$p_i = 1$, for people becoming trainees in $(t-s+1)^{th}$ period

$= 0$, otherwise, rewrites :

$$z_{it} = p_i z^{p}_{it} + (1 - p_i) z^{c}_{it}$$

$$= B z_{it-1} + d_{it} + b_i + R_t p_i + u_{it}$$

$$= B^{s} z_{it-s} + d^{*}_{is} + b^{*}_{is} + R^{*}_{s} p_i + u^{*}_{is} \qquad (4)$$

where for example,

$$R^{*}_{s} = \sum_{\tau=0}^{s-1} B^{\tau} R_{t-\tau}$$

is the amount by which the earnings of participants are higher in period t than they would otherwise be, had they not trained.

There possibly are two ways to fit equation (4) to the data according to whether we exploit the form in line two or three of (4). Choosing to use the former is attractive if we are willing to eliminate the individual fixed effects b_i (by mere first differencing). Using the latter provides us with estimates of the training effects because R_s^* are in that formulation the parameters to be estimated.

We should note that there is a strong assumption being maintained throughout the analysis presented above. This is that the same earnings generating function is assumed to operate for both control and participation (trainee) groups. This serves a plausible "pedagogical" role in the treatment and is one of the main assumptions maintained in Bassi (1983) in her random effects estimation.

Furthermore, as noted above, one should be cautious on how one treats the individual specific effect b_i. Ashenfelter points to this clearly and also shows that if individual fixed effects are present and correlated with trainee participation after holding constant age and pre-training earnings levels, only then there will be bias in estimation. Subtracting the earnings function at the base period, $t-s$, from that prevailing at t we can remove the individual specific fixed effect b_i. This in turn implies that the effects on earnings of any such variables that are unchanging are in this manner being wiped out.

Using results developed after Ashenfelter's paper was published (see also Hsiao (1986)) it is possible to estimate the parameters of interest as well as the fixed effects in a consistent and efficient way by using OLS after having transformed the data by subtracting group means from each observation.

Rewritting the initial form of the earnings function, (1), without the autoregressive structure and without time-specific shocks, ϵ_t, for simplicity and concreteness[1] we can get a standard linear (error-components) model:

$$y_{it} = x'_{it}\beta + \epsilon_i + \epsilon_{it} \qquad i = 1,\ldots,N \text{ and } t = 1,\ldots,T$$

Thus x_{it} would be the $k \times 1$ vector of time varying regressors assumed strictly exogenous (that is uncorrelated with the full stream of past, current or future realisations of ϵ_{it}) and ϵ_i can also be a random effect but in this brief illustration we shall assume as in Ashenfelter (1978) that it is fixed across individuals. In an aggregate time series regression the individual specific training effect would be absorbed by the constant and would thus remain unidentified. On the other hand, in a cross-section if as Ashenfelter points out ϵ_i is correlated with trainee participation (i.e. at least with a subset of x_{it}) across individuals then a biased β estimator would ensue. If we can not find reasonable enough valid instruments then we have to restrict the error structure. If, say we are willing to believe that as an identifying restriction we can pose that there are no economy – wide effects on earnings then it is possible to identify the fixed effests. On the other hand if we eliminate the latter by differencing we can identify the difference in economy – wide errors that would then be present.

Finally, the paper reaches the following conclusions ; trainee groups faced an unforeseeable deterioration in their earnings the year before the advent of

[1] Ashenfelter in fact initially drops the autoregressive structure by assuming B=0 (so that B^*=0) and also that β'_j=0 fot j>1. He reports results for this and then results out of maintaining the autoregressive structure. In the latter case [op.sit. Table 5 Ashenfelter (1978)] estimated training effects are slightly increased probably due to the fact that the individual effects are negatively correlated with trainee status.

training and earnings foregone out of the training process itself. In addition, training is found to have increased the earnings of all trainee groups. However, apart from the fact that earnings are truncated whenever the Social Security taxable maximum is reached, the main limitation of the analysis is attributed on the selection procedure being adopted. This falls under the statistical treatment for controls against selection bias and merits further consideration in the present context [Bassi (1983a, 1983b, 1984)].

In Bassi (1983b) an effort is made to report on any possible effects the Comprehensive Employment and Training Act (CETA) of 1973 may have had on the posttraining earnings of participants to the programme. The approach as the author points out is concerned with producing estimates that are free of selection bias - whether this is as a result of nonrandom self-selection or selection by programme administrators. Many interesting results emerge out of her studies. Most importantly perhaps comes the finding that "CETA has had a positive and often significant effect on the earnings of participants and that women benefit more than do men" [Bassi (1983b) p.541] even though she can not rigorously explain this on the basis of her data. To reach that conclusion though she first has to turn to the nonrandom selection problem which arises if individuals are not assigned randomly between groups but their decision to participate into a programme may be correlated with unobservables in the earnings equation thus rendering OLS estimators inconsistent. She mentions the possibility of using a random effects estimator provided that all relevant observable characteristics are used as explanatory variables and the underlying earnings structure is the same between participants and nonparticipants in order to produce unbiased OLS estimates. The random effects model that is used by

Bassi is the following:

$$y_{it}=x_{it}\gamma + P_i\beta_t + \epsilon_i + \epsilon_t + \epsilon_{it} \tag{5}$$

In estimating the random effects model, Bassi includes among the regressors age cubed, education and the status of the head of the household. Her findings indicate that the estimates out of this estimator using earnings equations in levels are rejected because she finds statistically significant differences in the pre-programme earnings between the control and the participant groups.

If however a fixed effects estimator is used then any unobservable individual specific fixed effects can be removed by differencing the earnings equation over two periods one postprogramme, t and one prior to participation, s viz.

$$y_{it} - y_{is} = (\epsilon_t - \epsilon_s) + (X_{it} - X_{is})\gamma + P_i\beta_t + (\epsilon_{it} - \epsilon_{is}) \tag{6}$$

In this framework there might be evidence of what is referred to as "creaming" i.e only those individuals with the highest possible permanent income and the lowest possible transitory income will be chosen; since CETA administrators have been evaluated conditional on the postprogramme earnings and employment records of participants to it the net effect may be a correlation between ϵ_{is} and P_i (the dummy denoting participation). It is quite difficult detecting the presence of creaming and for that reason one two or perhaps more years before the advent of training can be used as the base period.

At this point we can mention Ashenfelter and Card's (1985) difference in differences estimator. This uses a linear components of variance structure to express the probability that an individual i works at time t, P_{it} in order to derive an estimate of the training effect. More precisely if one writes

$$P_{it} = \alpha_i + \beta_t + \theta d_{it} \tag{7}$$

with α_i an individual specific component, β_t a time (year) specific component, d_{it} an indicator for the post of year of training status and θ subsequently denoting the training effect then if conditioning on α_i there is independence between employment outcomes across years we can derive a consistent estimate of θ. This estimator hinges on the assumption that the participation decision is independent of pre-training outcomes. This is clearly a very strong assumption which may not necessarily hold in practice. In fact this relates to creaming noted above in Bassi's research and to the case where transitory shocks may influence pre-training earnings adversely and contribute towards a decision to enter training. Card and Sullivan (1988) present evidence in support of this, meaning that employment probabilities are not independent of training status given the individual effects α_i.

Thus to illustrate the difference in differences estimator let us denote by E_{it} an indicator variable equal to 1 if individual i is employed at t and by $\overline{E}_{ipre}$ and $\overline{E}_{ipost}$ the individual specific means of E_{it} in the pre-training and post-training periods respectively. Then the expected change in average employment rates for the trainee group between the two periods is

$$\mathrm{E}(\overline{E}_{ipost} - \overline{E}_{ipre} \mid d_{itr} = 1) = \beta^* + \theta \tag{8}$$

where β^* denotes the difference in the means of the year-specific β_t's between the pre- and post-training periods. The equivalent expression for the control group is

$$\mathrm{E}(\overline{E}_{ipost} - \overline{E}_{ipre} \mid d_{itr} = 0) = \beta^* \tag{9}$$

and thus the expected "difference in differences" of the average pre- and post-training employment rates for trainees and controls is

$$\mathrm{E}(\bar{E}_{ipost} - \bar{E}_{ipre} \mid d_{itr} = 1) - \mathrm{E}(\bar{E}_{ipost} - \bar{E}_{ipre} \mid d_{itr} = 0) = \theta \tag{10}$$

and thus we can use the relevant to the above expected values sample information to get a consistent estimate of θ. Card and Sullivan(1988) note however that the difference in differences estimator is not the most efficient linear estimator of θ in P_{it} but that a weighted least squares procedure is needed in order to efficiently estimate θ in this expression. This procedure is only mentioned in their paper without any further information as what likely weights are likely to be used so that θ is efficiently estimated. In the appendix to this chapter we briefly go through the weighted least squares procedure that would possibly achieve this.

Finally, contamination bias may be another reason why one may come up with biased estimates of training effects since undetected CETA participants may progressively "contaminate" the comparison group over time.

At this point it is rather crucial to note that in the studies of Ashenfelter (1978), Bassi (1983, 1984) as well as in Kiefer (1979) most of the results are not only comparable but also similar to a certain extent.[2]

Hence apart from the fact that female participants to the programmes tend to benefit more than their male counterparts as we noted above, there is

[2] The studies use different estimation techniques and utilise different data sets: Ashenfelter compares a random sample from the Social Security Administration (SSA) records with participants in 1964 in the Manpower Development and Training Act (MDTA) : Bassi compares CETA paricipants in 1975 and 1976 to a random sample from the Continuous Longitudinal Manpower survey (CLMS)while Kiefer compared 1969 participants into MDTA Job Opportunities in the Business Sector (JOBS), Job Corps (JC), Neighborhood Youth Corps (NYC) with a sample out of the target population of the programmes.

unanimity in other findings also. Thus minority males experience some earnings gains while all of the studies find no programme benefits for white males. In addition minority women benefit comparatively less than white women from participation in the programs.

2.2 Structural Models and Schooling Related Labour Market Transitions

2.2.1 Structural Models

In this section we continue by expositting some issues relevant this time to structural models dealing with the effect of schooling and training programs on earnings and employment. The need to have such models presents itself if we think carefully of the nature of the problem we have to analyse.

In particular, the choice of whether to remain in schooling or take a period of training is fundamentally dynamic in the sense that it will affect the whole stream of future earnings and employment outcomes. The decision is also discrete so that it seems that a natural approach would be to use a dynamic programming framework in order that we tackle the problem. That approach is, however, in turn distinctly different from the approaches of the last section, relying much more heavily on the structure of a dynamic choice problem in order to identify responses.

Therefore, and as the title of this section emphasises, we shall here illustrate some of the issues encountered by representative structural models incorporating the decision to take up additional periods of schooling, broadly defined. What we usually mean when referring to structural models in this literature is that we are interested in estimating parameters inherent in the very basic relationships guiding behaviour. These are clearly the utility function and the relevant constraints.

Models of this kind fall into the subgroup of the Markovian Decision Processes that can occasionally be found in the recent micro-econometrics literature under the name dynamic stochastic discrete choice models.

The last almost ten years have experienced an unprecedented pace under which this literature on the specification and estimation of dynamic discrete choice models continuously expands but here for reasons of compactness we shall draw our attention on some key survey references and selectively discuss these studies in the light of issues on labour market transitions in connection with schooling and training decisions. Most of the material described here and an excellent introduction to the literature can be found in the works of Eckstein and Wolpin (1989a, 1990) and Rust (1992). Other important references in connection to specific aspects of the issues we sketch can be found in any of these articles.

As far as estimation methodologies of dynamic discrete choice processes are concerned, most of the recent important issues are discussed in the survey by Eckstein and Wolpin (1989a). We shall borrow some of their notation which happens to be fairly standard in the recent literature. Analysing the empirical lessons drawn from the application of the method of dynamic programming in economics, Rust (1992) is also a clear and thorough reference.

For a finite T (the finite horizon case), the appropriate solution method involves solving in a backward sequential sense Bellman's (1957) equation. Bellman's equation relates rewards to the problem in the current period to those that would accrue in future periods and is elegantly presented in Bertsekas (1976, 1987). In the finite horizon case when we assume that reservation wages are time varying it is economic theory which poses restrictions on the shape of the reservation wage path. In turn the shape of this path will affect the likelihood function since the latter is a function of observed labour force states. The difference with the infinite horizon case rests with the absence in the latter of a terminal solution for the reservation wage.

Using the indicator function $d_i(t)=1$ to signify that alternative (or more precisely, state) i is chosen in period t and $=0$ if it is not and notting that $\sum_{i=1,2} d_i(t)=1$ i.e. at any particular time period the available alternatives are mutually exclusive then at any t, $t=0,1,\ldots,T$ an individual has the objective of maximising

$$\mathrm{E}\left(\sum_{j=t}^{T}\beta^{j-t}\sum_{i=1,2}R_i(j)d_i(j)\mid\Omega(t)\right) \tag{11}$$

where E(.) is a mathematical expectations operator, $\beta\in(0,1)$ denotes the individual discount rate, $\Omega(t)$ is the individual's information set which includes all relevant past and current realisations of those variables that may affect the outcome of (11) and most importantly $R_i(j)$ is a random variable for the reward to the individual if i is chosen at t. In order to maximise (11) the optimal sequence of control variables $\{d_i(t)\}$, $i=1,2$ $t=0,1,\ldots,T$ must be chosen which is available when desicions are being taken. Since the $R_i(t)$'s are random variables, in order to proceed we need to define the form their maximum value would be expected to take at any period t. This is:

$$\mathrm{V}\big(\Omega(t)\big)=\sup_{\{d_i(t)\}_{i=1,2}}\mathrm{E}\left(\sum_{j=t}^{T}\beta^{j-t}R(j)\mid\Omega(t)\right) \tag{12}$$

where $R(j)=R_1(j)d_1(j)+R_2(j)d_2(j)$ is the actual reward or $-$in terms of utils$-$ utility derived at time j. The function V(.) is dependent upon information available at time t and is subject to the dynamic programming form

$$\begin{aligned}\mathrm{V}\big(\Omega(t)\big)=\max\Big\{&R_1(t)+\beta\mathrm{E}\big\{\mathrm{V}\big(\Omega(t+1)\big)\mid d_1(t)=1\big\}\,,\\ &R_2(t)+\beta\mathrm{E}\big\{\mathrm{V}\big(\Omega(t+1)\big)\mid d_2(t)=1\big\}\Big\}\end{aligned} \tag{13}$$

Had we have to choose among a larger number of alternatives then the current rewards of these alternatives, R_n , n=3,4,... together with their discounted future optimal valuation streams, $\beta\mathrm{E}\left\{\, . \;|d_n=1\right\}$ would have to be considered inside (13). This however can easily get out of hand in practical estimation matters and is what Bellman coined the "curse of dimensionality". Dynamic programming can in principle be used to compute often complicated sequential choice problems and optimal solutions to them in problems with non-stationary history dependent transition probabilities and time non-separable utility functions. However the curse of dimensionality is always a restriction for even a simple finite horizon problem.

If, say the state and control spaces each comprise of three elements – so this is really a pretty small problem – then after ten periods there are over 9^{10} possible histories. Then one should use backward induction to get to a solution of the problem[3], which means one should evaluate the utility function at each of these 9^{10} possible scenaria and perform 3^{19} maximizations. Then for each previous history, H_9 , 9^9 conditional expectations must be computed so that overall for a backward recursion from period 10 to 0 the overall burden would be proportional to the number of histories which rises at an exponential rate as the horizon lengthens. Therefore keeping both the number of alternatives as well as the elements in the state space as few as possible has been the main concern in many applications while problems involving the choice between two alternatives have been popular in the literature; examples include the decision to accept employment or to continue searching, fertility decisions, renew a patent or let it

[3]Here we proceed with the discussion on the assumption that a (closed form) solution is attainable. This, however, is clearly not always possible.

expire, replace a bus engine or not etc.

There has constantly been active concern towards identifying alternative ways of reducing the choice set size. Convenient distributional assumptions for the random elements having an impact on the model as well as for the contemporaneous reward functions for each alternative can significantly improve matters. For example an assumption which has been adopted in the literature [see for instance Berkovec and Stern (1991)] in order to get analytical solutions to the dynamic programme is to assume an additively separable into its deterministic and stochastic components reward function, together with i.i.d. extreme value error terms[4].

Thus, because of the dimensionality problem, most practical applications of the method are usually restricted to the subclass of Markovian Decision problems (MDPs). In this framework people have preferences which re represented by a utility functional comprised of the sum of state dependent utility functions at each period t, $u(i_t, d_t^o)$ together with their expectations described in terms of a Markov transition probability denoted by $\pi(i_{t+1} \mid i_t,\ d_t^o)$. The subscript o denotes that an optimal action has taken place. Now, it is known in the literature that with the very general conditions that are required in Blackwell's Theorem the solution to an MDP would be a decision rule which

[4] In Berkovec and Stern (1991) where the ϵ's assume to draw from independent extreme value distributions i.e.

$$F(\epsilon_i) = \exp\left\{ - e^{\frac{-\epsilon_i}{\tau}} \right\}$$

analytic solutions are derived to equal:

$$E\max_i \, [\overline{V}_i + \epsilon_i] = \tau\left[\gamma + \ln \sum_i e^{\left(\frac{\overline{V}_i}{\tau}\right)}\right]$$

where i denotes the choice of alternatives entertained by their model; full or part-time job or retirement, $\overline{V}_i = V_i - \epsilon_i$ and γ is Euler's constant ($\gamma = 0.577$).

would be a function at each t of the state at t i.e $d_t^o = g_t(i_t)$.

However Rust (1992) notes that even if one were to assume an additive error term to this that could possibly lead to a degenerate[5] statistical model since even though individuals are assumed to behave optimally in economic terms, in statistical terms they may be seen to randomly depart from such behaviour.

Allowing for unobserved to the econometrician states that the agents may happen to be in, Rust (1988) uses optimal decision rules of the form $d_t^o = g_t(i_t, \kappa_t)$ in order to develop the statistical framework for the esimation of MDPs with unobserved state variables. In function g, κ_t denotes the information signal about the state the agent has available to her which however is unobserved by the econometrician. From then on the decision problem can be solved by an appropriate selection of the sequence of decision rules[6] that would maximise the stream of expected discounted utility over the requisite time horizon.

Before going into the next Section which discusses an important application of employment transitions to schooling considerations we draw our attention once again to the computational difficulties involved out of estimating structural parameters from a dynamic programme.

It is usually the case that the computational problems involved out of having to numerically compute a programme under large dimensions motivate the following assumption on the joint transition density for (i_t, κ_t). It is known as the *Conditional Independence* assumption and is due to Rust who gets the following expression for the Markov density:

[5]This is not very useful as this is a model where a subset of the variables in it lead it to take a limiting distribution that takes the values of either 0 or 1. In this case since the discussion implies that the agent's decision can be perfectly predicted this means the resulting degenerate model would only take the value 1.

[6]or control variables as they alternatively also often termed in this way.

$$p(i_{t+1},\ \kappa_{t+1}|\ i_t,\ \kappa_t,\ d_i\ ) = q(\kappa_{t+1}|\ i_{t+1})\ \pi\ (i_{t+1}\ |\ i_t,\ d_i\ ) \tag{14}$$

where each alternative d_i belongs to D, the choice set and $q(.)$ is the cumulative distribution function of κ given the contemporaneously observed by the econometrician state variables. The usefulness of the CI assumption is evident. It implies that, the probability density for i_{t+1} only depends on i_t and d_i but not on the unobservables κ_t. In a sense the effect through the unobserved states κ has been "annihilated". However, a restriction that this assumption involves is that since i_{t+1} is a sufficient statistic for κ_{t+1} any pattern of statistical dependence that may exist between subsequent values of κ would channel itself exclusively through the vector of observables i_{t+1}. Rust has extensively utilised this assumption and tested its validity with an LM test. We shall not get to any more detail at this stage but it is of purpose to stress that he uses the assumption to establish some theorems that will enable him to write the likelihood function as products of Markovian densities and derive his *nested fixed point algorithm*

However, aside the heavy computational burden one can imagine this method would involve, it is nevertheless a useful one in analysing discrete choice problems. There are two commonly encountered types of discrete choice problems; those usually referred to as *optimal stopping* as well as those involving some *sequential choice.* The first would relate for example to decisions which in a sense bear no "undoing"; to mention an example an individual who indulges in active job search while unemployed and subsequently finds a job would in no circumstance viewed under the optimal stopping prism return to the non-employment state. However, this initial or in fact any previous activity of hers would be in her choice set when the agent's behaviour is modelled in a sequential choice framework. In these schooling decisions whereby agents can be allowed to leave a schooling state but can choose to return to it if at any time they perceive of this as a rational expected utility maximiser's decision, their behaviour could be modelled sequentially. In such cases state recall is a possibility.

2.2.2 The Transition from school to work: Wolpin (1987)

Let us now turn, after this brief introduction over the main tool of analysing discrete choice problems into an application involving the schooling decision. Wolpin (1987) makes a modification to the standard discrete choice search model by incorporating a finite horizon and a variable arrival of offers probability made explicitly dependent upon duration of unemployment, in his analysis of school-leavers' success into finding employment.

Individuals in his model are assumed to have decided upon an optimal consumption stream while at search. In a completely specified version of the model the search horizon's length would seem to depend upon other things on the initial asset position, the individual impatience and the form of the arrival probabilities. At each period in his model the individual receives an offer or does not. If he receives an offer he either has to accept or reject it in which latter case has to further continue to search at a fixed cost. The offers that one can receive arrive randomly at each period and they cannot be influenced by any conscious choice from his side[7]. In Wolpin's model these offers are i.i.d. with no recall. We shall further see below that he also allows the likelihood of offers to vary with unemployment duration. Once one offer happens to suit him and he accepts the job, the accepted job is assumed to last forever and financial market constraints are no longer binding. Individuals are constrained by the search horizon; thus if no acceptable offer has been received, and the individual reaches the terminal date of her horizon, she is obliged to accept the next offer available to her. His model bears the reservation wage property and if the arrival of offers probability is nonincreasing then it can be deduced that the reservation wage is monotonically declining with duration until the end of the search horizon.

We can see that the reservation wage for all $t < T$ in his model is given by the following expression which is explicitly dynamic since it includes the one-step ahead value function discounted back to today's terms:

$$w^r{}_t = -c + \delta V_{t+1} \tag{14}$$

where c is the cost of search for an offer to be received one period ahead, δ the subjective discount factor and T the search horizon. A job would be accepted at any t if $w_t \geq w^r{}_t$ and if w_t is i.i.d. then the value of unemployed search at t takes the form

$$\begin{aligned} V_t &= \alpha_t \; E \max[w_t, -c + \delta V_{t+1}] + (1-\alpha_t) \; E \, [-c + \delta V_{t+1}] \\ &= \alpha_t \; [E(w_t \mid w_t \geq w_t{}^r) \; Pr \; (w_t \geq w_t{}^r) \\ &\qquad\qquad + w_t{}^r Pr \; (w_t < w_t{}^r)] + (1-\alpha_t) \; w^r{}_t \end{aligned} \tag{15}$$

where α_t is the arrival probability at t. To solve for (15) Wolpin exploits the finite horizon and his assumption that as soon as the individual reaches the

---

[7]If they can, in fact, be influenced by factors that appear at the individual's environment and be controlled by her they might be seen as being endogenous to the model. In such a case the model presented in Chapter 3 would in fact endogenise the arrival of offers probability.

terminal date of search she is willing to accept the first offer that is open to her, to get

$$V_T = [\alpha_T E(w) - (1-\alpha_T)c] \ + \delta(1-\alpha_T)[\alpha_{T+1}\ E(w) - (1-\alpha_{T+1})\ c] \ + \ldots$$
$$+ \delta^r\ (1-\alpha_T)(1-\alpha_{T+1}) \cdot \ldots \ldots$$
$$\ldots (1-\alpha_{T+R-1})\ [\alpha_{T+R}\ E(w) - (1-\alpha_{T+R})\ c] \qquad (16)$$

Wolpin wants to numerically calculate reservation wages and the sequence of offer probabilities by utilising the data on unemployment durations and accepted wages. He is also interested in getting income net of search costs as well as the distribution of offers. He adopts normal ($w_t = \bar{w} + u_t$) and lognormal ($w_t = \tilde{w}\, e^{u_t}$) distributions for w_t needed in (15) and uses the properties of the truncated normal distribution to rewrite the value of search in terms of normal offers and lognormal ones. For example in the normal case one can get

$$V_t = \alpha_t\Big[\bar{w} + \eta_t\, \Phi\left(\tfrac{\eta_t}{\sigma_u}\right) + \sigma_u\, \phi\left(\tfrac{\eta_t}{\sigma_u}\right)\Big] + (1-\alpha_t) {w^r}_t \qquad (17)$$

where $w_t = \bar{w} + u_t$ and $\eta_t = {w^r}_t - \bar{w}$. He incorporates duration dependence into his model by allowing the offer probability sequence to be explicitly driven by it,

$$\alpha_t = \Phi(m_0 + m_1 t) \qquad (18)$$

Thus a negative (positive) m_1 implies that offer probabilities decline (grow) with the unemployment duration. Finally, since given a homogeneous sample of individuals the reservation wage cannot be greater than the minimum observed wage in the given time period, so that the entire reservation wage profile can be influenced by the global minimum observed wage in the sample, Wolpin would like to control for it. For this reason in order to reduce the influence of observed wage outliers on the estimated parameters he assumes that accepted wages are measured with error. In particular for the latter he assumes

$$w_t = \overline{w} + \theta_t \qquad \text{and} \qquad \ln w_t = \ln \widetilde{w} + \theta_t \tag{19}$$

for the normal and lognormal cases respectively, the composite error term θ_t is defined as $\theta_t = u_t + \epsilon_t$ where $\epsilon_t \sim N(0, \sigma_\epsilon{}^2)$ and u_t is assumed independent of ϵ_t. He thus writes the sample likelihood of I individuals the first K of which experience unemployment enduring for $t_i - 1$ periods , (i.e. $\forall\ i = 1...K$) the remaining $I - K$ having an incomplete spell lasting for l_i (now $i = K + 1^{th}$ to I^{th} individuals) in the normally distributed offer case as:

$$L = \prod_{i=1}^{K} \prod_{j=1}^{t_i - 1} \left[\alpha_j\ \Phi\left[\frac{\eta_j}{\sigma_u}\right] + (1 - \alpha_j) \right]$$

$$\times \left\{ \alpha_t^i \left[1 - \Phi\left(\frac{\eta_{ti} - \rho\left(\frac{\sigma_u}{\sigma_\theta}\right)\theta_t^i}{\sigma_u \sqrt{1 - \rho^2}} \right) \right] \frac{1}{\sigma_\theta}\, \phi\left[\frac{\theta_t^i}{\sigma_\theta} \right] \right\}$$

$$\prod_{i=K+1}^{I} \prod_{j=1}^{l_i} \left[\alpha_j\ \Phi\left[\frac{\eta_j}{\sigma_u}\right] + (1 - \alpha_j) \right]. \tag{21}$$

In the likelihood $\rho = \frac{\sigma_u}{\sigma_\theta}$ where $\sigma_\theta = \sqrt{\sigma_u^2 + \sigma_\epsilon^2}$, thus $1 - \rho^2$ gives the fraction of the wage variance accounted for by measurement error. Wolpin does not present his results out of the normality of offers assumption since he says that results are as unreasonable as actually giving negative mean offer and costs of search estimates while the latter are unrealistically large also in magnitude. An estimated discount factor of almost zero meaning that individuals are not so impatient with a spell of unemployment does not help the normal case either.

His results from the lognormal case are as follows: The search horizon is estimated to extend to 54 weeks after graduation taken from the fact that reservation wages had been positive over that many weeks. Mean weekly wage offers are estimated at \$188 about 17 less than the mean of observed offers. The arrival rate is found to be as low as 0.012 the first week after graduating, reaching 0.009 one year after and measurement error for this specification is found to account for only 1.2% of total offer variation.

In a version of his model that could admit a limited amount of heterogeneity, Wolpin incorporates a functional form for the wage equation that takes into account the individual's NLS ability score and father's education, thence,

$$\ln w = \ln\widetilde{w} + \beta_1 AFQT + \beta_2 FS$$

where $AFQT = 1$ if the individual's NLS ability score is above average and zero otherwise and $FS = 1$ if the individual's father is a high school graduate or more and zero otherwise. As expected, mean offers vary with these two dummies ranging from 173 for low ability school leavers with low education father to 250 for those with the reverse characteristics. In addition this model explains a 9.2% of the total variation in offers; that is greater than was the case in the simpler model. The cost of search is estimated higher than in the previous case at \$223 per week. The arrival rates are proven to be the dominant factor – their decline being sharper than the decline of the reservation wages – in producing the decline over the horizon in the transition rates.

2.3 Heterogeneity and State Dependence in Employment.

The dynamic structural models of the previous section seem quite complex and at first sight sometimes prohibitive at least as far as estimation methods are concerned. This qualification, however true sometimes may be, can hide a lack of ability to incorporate individual heterogeneity, so important in the experimental or discrete choice estimators of the first section.

In this section we shall make a note on two related issues that frequently crop up when we want to describe the evolution of behaviour of agents moving between alternative states. These are heterogeneity and state dependence. But let us start with describing what would be implied in their absence. In our framework of participation probabilities if we were to assume independence that would clearly imply:

$$P(\ \text{wwork}^i_t = 1 \mid \text{wwork}^i_{t-1} = 1) = P(\ \text{wwork}^i_t = 1\ ) \tag{22}$$

i.e whether or not a spouse is observed to have worked in the previous period gives us not much information about her work status in the current period. However, the two main reasons why we would expect (22) not to hold in practice are:

i) *Heterogeneity:* This postulates that unobservable attributes among people that may affect them differently are responsible for influencing their attitudes towards work.

ii) *(True) State dependence:* As its name implies it is often the case that present status influences future status.[8] True is in parenthesis to emphasise the fact that apparent state dependence can be viewed as a result of unobserved heterogeneity and not necessarily pure attachment to a particular labour force state. Hence beyond the idea that what we do is often simply the result of habit formation[9] more importantly this property tries to examine a potential link in observable modes of behaviour between states.

Heckman and Willis (1977) considered heterogeneity in their sequential labour force participation study for married women and provide evidence to suggest that

$$P(\text{wwork}^i{}_t=1 \mid \text{wwork}^i{}_{t-1}=1) \geq P(\text{wwork}^i{}_t=1) \qquad (23)$$

That is in the Heckman and Willis study state dependence as related to an individual's labour force participation sequence implies that the chance of a woman who was employed in period $t-1$ to be employed at t also, is higher than her unconditional probability of working at t. However as we shall see in Chapter 3 state dependence also implies that a woman who was employed at $t-1$ is more likely to still be having employment at t than one who was not working at $t-1$. The formulation that they employ assumes that unmeasured variables follow a components of variance model i.e. that the individual has a permanent component as well as a transitory component which is assumed serially uncorrelated. They furthermore considered the formulation of what they call a "beta-logistic" model. In that, they try to examine the presence or

[8]For example a woman being employed in the current period may signal a higher ability to her employer as compared to an unemployed one (stigma effect) such that it is likely that she will be working in the next period too.

[9]One of the recent prominent studies that develops this form of heterogeneity bringing in the importance of habit formation in a rigorous framework has been that of Constantinides (1990). He uses the idea with a view to shed some light into the resolution of the "Equity Premium Puzzle".

otherwise of heterogeneity and because the beta distribution can be suitably parameterised so as to take different shapes, they recognise and exploit this property to look at labour force participation probabilities. Thus their panel data discrete choice model leads them into a U shaped distribution for participation probabilities meaning that the majority of women either have participation probabilities of zero or of one. The limitations of their model as they themselves recognise are the assumptions that the individual response probabilities are assumed constant over time and state independent. The beta logistic model cannot account for the fact that some significant determinants of labour force participation e.g infants have a varying influence over time and is therefore misspecified. They believe that sequential labour force participation is to a great extent due to apparent state dependence caused by heterogeneity but may also be affected by true state dependence. A more general perspective for the heterogeneity process since it seems that it is rather less trivial to tackle it accurately, requires that more structure is given to the process describing the unmeasured components. An extensive account for that latter need is given in Heckman (1981, Rosen ed.) who utilises urn model theory to illustrate how heterogeneity and state dependence might arise and implements suggestions on how they could be modelled.

Among numerous other papers that have tried to account for heterogeneity Willis and Rosen (1979) deserves mentioning. Their starting point is to recognise that people are different in terms of their capacities to finance a given amount of schooling and further education, in their views about the future, with respect to their preferences and a host of other such things. Thus we might say that the benefits and costs out of completing a given curriculum are random variables across the population and are determined by things either directly observable to the econometrician or unobservable.

Let us however now bring in some more notation and try to describe some of the issues glossed over above. We assume first that women make relevant decisions conditional on the state they are in i.e employment or unemployment and that they take these decisions in equally spaced time intervals. The expected lifetime utility a married woman i enjoys if she is employed in period t is denoted as ν_{1it} and this is a function of all those demographic variables that can and/or are expected to affect her decisions. These variables may include the presence or birth of children, geographic attachments, levels of education and stock of human capital considerations for the wife and others etc. Similarly, let ν_{0it} refer to the case where she is unemployed at t. Accordingly, she may accept a job at period t if there is some utility to be derived in doing so rather than staying unemployed, i.e. if the difference in utilities $V_{it} = \nu_{1it} - \nu_{0it}$ is a positive quantity or not in which case she should rather not take up the job. A reasonable way to proceed is by a simple assumption concerning the decomposition of the difference in utilities V_{it} into two components. The first, ψ_{it} is a function of variables that are observable to the econometrician while those remaining unobservable are ϵ_{it}. Thus the difference in utilities can be written as:

$$V_{it} = \psi_{it} + \epsilon_{it} \tag{24}$$

Heterogeneity is then simply due to ϵ_{it}, the unmeasured disturbance which is influenced by individual uncertainty and factors that can be observable to the individual but not to the econometrician.

If we further let d_{it} capturing the fact that she works at time t (value 1) or not (value 0) then we can specialise the function V_{it} into:

$$V_{it}=Z_i\beta+ \delta \sum_{\tau<t} d_{i\tau} + \epsilon_{it} \quad , \tag{25}$$

$$i = 1,...I \quad t = 1,....T \quad \mathrm{E}(\epsilon_{it}) = 0, \ \mathrm{E}(\epsilon_{it}\epsilon_{i\tau}) = \sigma_{t\tau}, \ \mathrm{E}(\epsilon_{it}\epsilon_{k\tau}) = 0$$

δ in (25) denotes a constant depreciation parameter of the effect of past work experience on the difference in utilities, V_{it}. Affordable degrees of generalisation could be imposed on (25) but for our purposes we keep the model simple. Seeing the above model as an urn model where each woman makes independent draws from her own urn let ϵ_{it} have a components of variance structure i.e.

$$\epsilon_{it} = \zeta_i + \eta_{it} \tag{26}$$

where ζ_i is a realisation of a mean zero random variable that is individual specific and η_{it} an i.i.d zero mean random variable. Then if we also let δ=0 this model essentially corresponds to that of Heckman and Willis(1977). Women make independent draws from their urns which differ in their composition. A model with true state dependence would arise if they were rewarded each time on their picking up a red ball with one more red. This translates into (25) in having Z_i=1 and no ζ_i component.

We should however mention at this stage that the most common type of state dependence encountered in the recent literature on the field is through *endogenous experience (human capital).* This is the approach which has been followed in a lot of work on the issues of labour force participation (e.g. Eckstein and Wolpin 1989a, 1989b) and on other applications (Eckstein and Wolpin 1989a is a comprehensive source). The direction that is followed in Eckstein & Wolpin (1989b) for example is to present a structural dynamic model of married women's labour force participation in which work experience or cumulative participation in the market shall be endogenous. Their model is complicated by the fact that they need to allow for non-separabilities in the utility function. Thus they allow the utility function to include the number and age distribution of children on their own and interacted with participation, they interact consumption with participation, schooling and participation and finally experience to be interacted with participation to allow for intertemporal non-separability.

However, state dependence could be modelled by employing a slightly different framework using the information on *transitions* and this incidentally will occupy us here. Using transitions, state dependence can easily be seen as labour supply status being influenced by previous period's status. Eckstein and Wolpin (1989b) for example concentrate on how labour market participation or past experience has an effect on future wages (even though as yet we have not explicitly used a wage equation) which in turn affect future participation in the labour market and work decisions. This induces dynamic aspects on labour force participation. Results by Berkovec and Stern (1991) at least as far the retirement decisions of older men in the National Longitudinal Survey are concerned indicate that these dynamics play an important role on their labour force participation decisions. Their dynamic model even performs better than that in a static context when compared using crude statistics based on sums of squared residuals. In either a dynamic or static model they find that it is important to allow for unobserved heterogeneity forms as well (job specific effects and unobserved person effects in a labour force participation model).

Appendix A: Efficient Estimation of Employment Probabilities in Card and Sullivan (1988)

In the present appendix we show the form the weights would bear in a GLS estimation that would efficiently estimate the difference in differences of employment probabilities in Card and Sullivan (1988).

The estimator of θ, the training effect, adopted in their paper can be expressed in the following manner:

$$\hat{\theta} = \left(\frac{\sum_{tr} E_{i2}^{tr}}{N_{tr}} - \frac{\sum_{tr} E_{i1}^{tr}}{N_{tr}} \right) - \left(\frac{\sum_{con} E_{i2}^{con}}{N_{con}} - \frac{\sum_{con} E_{i1}^{con}}{N_{con}} \right)$$

$$= \left(\frac{\sum_{tr} (E_{i2}^{tr} - E_{i1}^{tr})}{N_{tr}} \right) - \left(\frac{\sum_{con} (E_{i2}^{con} - E_{i1}^{con})}{N_{con}} \right)$$

$$= \left(\sum_{tr} w_{itr} \, (E_{i2}^{tr} - E_{i1}^{tr}) \right) - \left(\sum_{con} w_{icon} \, (E_{i2}^{con} - E_{i1}^{con}) \right)$$

where *tr* and *con* denote trainees and controls respectively, subscript 2 post-training earnings, 1 pre-training earnings and the weights are equal in both cases to the inverse of their respective observations $\left(w_{itr} = \frac{1}{N_{tr}} \text{ and } w_{icon} = \frac{1}{N_{con}} \right)$.

In GLS the weights take the form

$$w_{itr} = \frac{1}{\mathrm{Var}\,(E_{i2}^{tr} - E_{i1}^{tr})} \text{ and } w_{icon} = \frac{1}{\mathrm{Var}\,(E_{i2}^{con} - E_{i1}^{con})} \tag{A.1}$$

Now, $\quad E_{i2}^{tr} - E_{i1}^{tr} = E\left(\bar{E}_{i2} - \bar{E}_{i1} \mid D_{it} = 1 \right) + u_i^{tr}$

and $\quad E_{i2}^{con} - E_{i1}^{con} = E\left(\bar{E}_{i2} - \bar{E}_{i1} \mid D_{it} = 0 \right) + u_i^{con}$

for, say, normally distributed with zero mean and constant variance disturbances (u_i's), so that subtracting the difference in earnings of controls from that of the trainees we obtain

$$\underbrace{(E_{i2}^{tr} - E_{i1}^{tr}) - (E_{i2}^{con} - E_{i1}^{con})}_{y_i} = \theta + \underbrace{(u_i^{tr} - u_i^{con})}_{\epsilon_i} \text{ or simply}$$

$y_i = 1\theta + \epsilon_i$ with 1 an $N \times 1$ vector of ones ($N = N_{tr} + N_{con}$) so that the estimator would take the form

$$\hat{\theta}_{WLS} = \underbrace{\left[1'V(\epsilon)^{-1}1\right]^{-1}}_{C} \left[1'V(\epsilon)^{-1}y\right] \tag{A.2}$$

It is then straightforward to observe from (A.2) that the weights would be of the form (A.1) multiplied by the sum of the squares of the variance of the errors i.e.

$$w_{itr} = C\frac{1}{\text{Var}\,(E_{i2}^{tr} - E_{i1}^{tr})} \text{ and } w_{icon} = C\,\frac{1}{\text{Var}\,(E_{i2}^{con} - E_{i1}^{con})} \tag{A.3}$$

Then it is clear that to apply GLS to (A.2) one needs to calculate Var (ϵ_i). This assuming independence between u_i^{tr} and u_i^{con} and between pre – and post earnings for trainees and controls can be written as

$$\text{Var}\ (\epsilon_i) = V(E_{i2}^{tr}) + V(E_{i1}^{tr}) + V(E_{i2}^{con}) + V(E_{i1}^{con})$$

where for example $V(E_{i2}^{tr}) = Pr^{tr}(\text{job})\left[1 - Pr^{tr}(\text{job})\right]$

$$= (\alpha_i + \beta_t + \theta)(1 - \alpha_i - \beta_t - \theta) \quad \text{where } \mathrm{D}_{it} = 1$$

while, $\quad V(E_{i2}^{con}) = (\alpha_i + \beta_t)(1 - \alpha_i - \beta_t) \qquad \text{since } \mathrm{D}_{it} = 0$.

We can similarly derive expressions for the pre – training outcomes and thus apply GLS.

Chapter 3

Labour Market Transitions and Schooling of Married Women in the U.S.

3.1 Introduction

In this section we shall provide a brief introduction on the motivation for the model of labour market transitions. The main aim of the analysis here is to estimate what Card and Sullivan (1988) denote as the accession probabilities for an individual (a married woman in this respect) in the labour market. This is simply the estimated probability of an individual's movement into employment from unemployment between two consecutive periods. The period of measurement in the present Chapter shall be taken to be a quarter so that women's movements in the labour market statuses are described between two consecutive quarters. Layoff probabilities – that is, losing a job between two consecutive periods – and retention probabilities (maintaining a job) – can also be incorporated at ease within this simple framework. More precisely we shall look at those probabilities in the sample of women that were not engaged in market work in the previous quarter (sample denoted further below as wwl=0) including both the sample of those women who were employed in the previous quarterly interview (sample denoted wwl=1) and those who were not. The model that we shall be developing draws also from Blundell (1990) and has been used to incorporate life cycle consumption considerations in Blundell *et.al.* (1992).

Table Ia

Transitions by Year of Interview[10](%)

	1|0	0|0	1|1	0|1	Mean.Unem(%)
1980	5.3	94.7	100	0	7.6
1981	7.4	92.6	97.4	2.6	7.4
1982	6.2	93.8	97.8	2.2	9.1
1983	8.6	91.4	96.6	3.4	10.1
1984	8.1	91.9	97.8	2.2	7.8
1985	6.1	93.9	97.7	2.3	7.3
1986	9.4	90.6	98.5	1.5	7.0
1987	8.4	91.6	98.0	2.0	6.4
1988	7.3	92.7	98.0	2.0	5.7

Before we embark into this let us look at some transitions from the CES dataset that we use in this work. In the first three tables some transition information is given from the raw data. Firstly, in Table Ia we give (percentage) number of cases observed within every of the four possible two period work histories we can observe in the sample and by reference year. The aggregate mean unemployment rate in the U.S. for the same years can also be

[10]In every i|j pair i refers to the variable wwork and j to wwl(i.e wwork's value at the previous interview). For example under the column 1|0 the number of unemployment to employment transitions are listed for every year.

found in this Table. This is informative since in this way we could have a picture on whether the micro-data tell us anything that at the descriptive level can be rationalised in terms of this aggregated information. Clearly, work histories for more than two periods could similarly be devised but this does not directly provide a sample equivalent information to the estimation content of the model of this chapter. We can observe the following in Table Ia; only 5.3% of the total sample experienced an accession transition while no person was observed losing her job in the sample in 1980. For this reason we decided to drop the observations pertaining to the year 1980 from the estimation that follows later on. In addition we may notice that as unemployment increases up until 1983, so does the volume of observed transitions; we can see that transitions to both directions increase in percentage terms. This is certainly more plausible in the case of one's losing her job $(0|1)$ but not so for the sample accession rate $(1|0)$. However,we can also observe that the sample accession rate peaks in 1986 at a point where unemployment is at a downturn which would seem more economically sound.

In Table Ib similar information is plotted for the four possible transitions. This time the sample is drawn conditional on those spouses having completed High School education and then also for those who are classified in the data set's description of variables as College Students (this sample also includes graduates).

Table Ib

Sample Transitions by Wife's Education Groups

High School Education

YEAR	1|0	0|0	0|1	1|1
1980	4	86	0	148
1981	22	310	23	715
1982	23	349	24	793
1983	26	337	27	742
1984	23	276	16	735
1985	18	257	22	671
1986	10	80	5	284
1987	23	273	20	909
1988	18	167	13	615

College Education

YEAR	1|0	0|0	0|1	1|1
1980	4	57	0	183
1981	21	236	17	862
1982	13	182	15	909
1983	27	201	29	922
1984	21	222	22	1029
1985	12	201	17	993
1986	8	83	6	425
1987	22	192	21	1177
1988	10	176	16	803

Finally in Table Ic sample transition information is given conditional on the wife being white or black.

Since it is clear that our subsample consists almost entirely of white women we decided to consider only those in the estimation. Finally Table Id gives similar information by interview number. Both white as well as black women are included in the Table so that we can draw comparisons in terms of race too. This table is not only informative of the sample sizes that we are dealing with and which are small relative to the original sample numbers (around 2.2%) but also indicates between which interviews the bulk of transitions took place. That clearly for both transition directions takes place between the fourth and the fifth interview. Women in the sample are also likely to lose their jobs with an overwhelming majority (92%) between these two interviews.

<u>Table Ic</u>

Sample Transitions by Wife's Race Groups

White

YEAR	1|0	0|0	0|1	1|1
1980	7	138	0	308
1981	43	539	38	1459
1982	36	525	40	1595
1983	50	538	55	1531
1984	45	473	38	1643
1985	31	459	38	1556
1986	18	156	11	674
1987	42	450	40	1941

1988	22	328	26	1313

Table Ic *(continue)*

Black

YEAR	1|0	0|0	0|1	1|1
1980	1	4	0	17
1981	2	14	3	86
1982	0	8	0	100
1983	2	12	3	97
1984	0	18	2	74
1985	0	12	1	86
1986	0	4	0	33
1987	2	12	2	107
1988	5	16	0	76

Table Id

Transitions by Interview number

Interview	1|0	0|0	0|1	1|1
3	94	1361	16	4383
4	81	1296	9	4468
5	134	1171	278	4271
Total	309	3828	303	13122

3.2 The CES data set.

In this section we make reference to the Consumer Expenditure Survey Database (CES) as this is described in Attanasio, Koujianou and Weber (1989) and use it to perform some preliminary exercises on the labour market behaviour of married women.

In the original tapes for the CES database there is information about two independent samples of the U.S population. The first sample which is in fact more to our concern here consists of approximately 4,500 households which are being interviewed on a quarterly basis about family characteristics as well as income, employment and purchases. Out of these households about eighty percent are interviewed again the following quarter, the remaining twenty percent being replaced by a random group.

The nature of the database is such as to allow us to first order it by the identifier for each household (NEWID) and within each household by the number of times it was interviewed (INTNO) hence creating a short time series (four interviews) wherein we have information about all households whose interview number was not missing. In the data set encountered here, information for every household is provided within the course of one year. The month assigned for each interview is not necessarily the same for different households. Thus the picture one gets if the data were sorted by an identifier for each household, the quarter each interview takes place and month, year is that the first observation for the "first" household is collected in March 1980 and the last observation from the "last" household in November 1988. Attanasio *et al.*

discard the other households since this seemed a plausible way for the treatment of missing observations. They justify this on the grounds that one may not be sure how to tell apart a household whose variable is missing from those households for which this variable takes a value of zero. However, they note the seriousness of the missing observations issue by bringing attention to the fact that a significant proportion in the sample reported zero consumption on food both at home and out. These households do not report consumption on other commodities available to them either, and should be treated as missing observations.

Next and most importantly we needed to create the dependent variable in this study. This is if the wife in the household works or not (wwork) and was constructed from the composition of earners variable (EARNCOMP) by merging together all possible categories for this variable where a spouse is present[11]. This variable is limited to taking values of zero or one according to whether she is observed to be working or not. Within the household identifiers we then generate the one period lag of wwork, denoting it wwl. By construction therefore we generate some missing values in wwl since in the first interview there is no information about the spouse's previous work status. A similar procedure is adopted to defining the variable for the work status of the husband and its first period lag, hwork and hwl respectively. Total number of observations on wwork are 23416 and the same for hwork while wwl and hwl are each observed 17562 times the rest 5854 being the missing values.

We proceeded by considering different logit regressions which are

[11]Incidentally, the EARNCOMP variable refers to the composition of earners; 1. Ref. person only, 2. Ref. person and spouse, 3. Ref. person spouse and others, 4. Ref. person and others, 5. Spouse only, 6. Spouse and others, 7. Others only and 8. No earners. On the basis of this definition, therefore, we created wwork by allowing it to draw from categories 2,3,5 and 6.

appropriate here since we have a discrete dependent variable to estimate. Since it will be of interest to focus on labour market transitions and provide some framework to analysing them we report the different regressions by the value of the wife's work status variable in the previous period. Hence to illustrate with the wife's case several logit estimations were run for the dependent variable wwork against various explanatory variables and the results in each case where reported according to whether the spouse was or was not working in the period before. Before we comment on some results, let us give a brief description of the variables that were utilised in the estimation as well as to a few that for reasons we shall briefly discuss in passing, we originally experimented with but did not include them in the final estimation of the model.

First, some of the variables that we use from Attanasio *et al.* to create most of the variables and dummies used here are: AGEREF (age of reference adult), AGE2 (age of second adult), ASCOMP1 (number of males 16+), ASCOMP2 (number of females 16+), ASCOMP3 (number of males aged 2 to 15), ASCOMP4 (number of females aged 2 to 15), inf (number of infants), EDUCREF (education of reference person), EDUCA2 (education of second adult), ORIGIN1 (origin of reference person), ORIGIN2 (origin of second adult), REGION (standard region), SEXREF (sex of reference person; this indicates that the reference person in our data set is not exclusively always male) .

Now totexp1 refers to total (household) expenditure lagged once and is used as the consumption variable. all2_15 groups together all children in the household unit aged 2 to 15 irrespective of sex. Regional dummies MidW, South and West are self- explanatory. Race dummies for both reference and second adults are also used and these are racerfwh, racerfbl, rac2wh, rac2bl where wh and bl denote white and black respectively. For the education dummies used for the reference adult we chose to group into one category all those who either never attended school or whose education is elementary and this appears as edrfelnv. edrf2 refers to high school education but less than high school graduate, edrf4 to College education but less than College graduate, edrf5 to College graduate education while edrf6 to more than four years of College. Finally two education dummies are used for the spouse according to the education level attained being High School and College (denoted ed2HS and ed2C respectively). In the first instance the husband's dummies were included in addition to the wife's but since they were almost perfectly collinear with hers only the latter ones were considered in estimation. The estimation results are to be found in Sections 3.3 and 3.5 below.

As we have noted only in passing above, the main body of the analysis here will be concentrating on the transitions into and out of employment of the spouse and will look at the effects of different variables on her labour market behaviour. It has to be remembered that in what follows the states of employment and unemployment are considered while that of being out of the labour force is not available in this study.

3.3 Descriptive statistics and first step estimation.

This section will provide background information on some basic descriptive statistics that shall help us understand the model of the next section.

Next it concerns itself with estimating using logit specifications the accession and retention probabilities for the sample of married women that we specified above.

Since the hypothesis of state dependence in between labour market states is instrumental in this work we felt it was essential to first provide with some basic non-parametric information on the relevance of this hypothesis in the raw data that we use. In the next table the hypothesis of independence between the labour market status of women in consecutive interview periods is tested by means of a Pearson test.

Table 3

prev.int.	work now 0	1	Total
0	3828	309	4137
	92.67	2.30	23.56
1	303	13122	13425
	7.33	97.70	76.44
Total	4131	13431	17562
	100.00	100.00	100.00

Pearson chi2(1) = 14326.3000 Pr = 0.000

Table 3 *(continue)*

2 int.ago	work now 0	1	Total
0	2461	371	2832
	89.36	4.14	24.19
1	293	8583	8876
	10.64	95.86	75.81
Total	2754	8954	11708
	100.00	100.00	100.00

Pearson chi2(1) = 8340.8648 Pr = 0.000

It is then seen that the null of independence is heavily rejected. Then women's labour market status is tested between the current and two interviews before dates and the null of independence between these statuses is firmly rejected as well. Notice that even though rejection is at the 0% level that in fact the value of the χ^2 is almost twice when testing in between consecutive periods meaning that the degree of state dependence as reported by the raw data is stronger in this case.

As a first step we proceeded by estimating two logit regressions one on each of the two relevant subsamples. That is, first a logit on the subsample of women that were not engaged in employment during the previous interview was estimated with the characteristics vector including the following: the number of infants in the household, inf, 2 education dummies for the wife, ed2HS, ed2C as well as the age of the spouse.

In addition to these consumption lagged once was also included as a regressor with a view – as we shall see as we go along – to have a continuous decision (consumption) interacting with a discrete one (choice to participate in the labour market or not). Most importantly, however, we should mention that the aggregate Unemployment Rate for the U.S. for the period January 1980 to November 1987 is included in the regression[12]. This as one might expect should be a demand side variable and we want to consider it in the estimation. Initially, other demographic dummies were also used i.e. origin as well as time dummies but because of near or perfect collinearity with other regressors we decided not to include them. At the second step a logit regression was run including the same regressors but this time imposed onto the subsample of women who were employed during the previous interview.

From these two regressions we take estimates which we shall denote by β and γ respectively as it may be customary to refer to them that way throughout what follows. These estimates are depicted in Table II. Note that the estimations are relevant for the sample of white married women aged 20 to 64 inclusive, for the years 1981 to 1988.

[12]One should be aware that the lagged unemployment rate (hence 1980 to 1987 is used) should be included in the analysis on data from 1981 to 1988 otherwise we might confuse causes with consequences of the effect of the aggregate unemployment rate on transitions.

Table II

Sample ww1=0

Iteration 4: Log Likelihood =-1003.2597

Logit Estimates

Number of obs =	3748
chi2(7) =	19.81
Prob > chi2 =	0.0060

Iteration 4: Log Likelihood =-1003.2597

wwork	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
totexp1	-.000124	.0000563	-2.211	0.027	-.0002349	-.0000141
un1 |	.023663	.043	0.545	0.586	-.0615291	.1088548
AGE2 |	.07151	.07112	1.005	0.315	-.0679316	.2109559
AGE2sq	-.0011	.0009	-1.241	0.215	-.0028764	.0006466
inf |	-.2412601	.178	-1.355	0.175	-.5902231	.107703
ed2HS	.9630778	.5154	1.869	0.062	-.0473563	1.973512
ed2C	1.154789	.5191681	2.224	0.026	.1369089	2.172669
_cons	-4.3952	1.485347	-2.959	0.003	-7.307362	-1.483022

Sample ww1=1

Iteration 4: Log Likelihood =-1324.1334

Logit Estimates

Number of obs = 11986

chi2(7) = 60.98

Prob > chi2 = 0.0000

wwork	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
totexp1	-.0001	.0000476	-2.139	0.032	-.000195	-8.50e-06
un1	-.11043	.0413951	-2.668	0.008	-.1915704	-.0292882
AGE2	.0943	.0582123	1.620	0.105	-.0198249	.2083861
AGE2sq	-.0011	.0007378	-1.541	0.123	-.0025833	.0003091
inf	-.838	.1452261	-5.769	0.000	-1.122546	-.553213
ed2HS	.707	.3400803	2.080	0.038	.040615	1.37384
ed2C	1.09	.3434459	3.177	0.001	.4179781	1.764397
_cons	2.232	1.180652	1.891	0.059	-.0822158	4.546321

The vector of estimated coefficients was utilised in the next part which entailed reading blocks of the data from the two subsamples of the data set.

Where we use the sample of women that were not engaged in market work in the previous interview (ww1=0) we observe that the probability of theirs working during the current interview period has a sample mean as low as 7.7 per cent. This can plausibly be justified on the grounds that the sample of those women suffers from poor labour market characteristics as well as from any stigma effects and is correspondingly that low. On the other hand the transition (retention) probability estimate for the women who were employed at the previous interview has a corresponding mean of 98 per cent because these women enjoy labour market characteristics valuable enough to guarantee them a less discontinued working history. We can also note the fact that the sample (or actual) and predicted estimates coincide apart from the fact that the standard deviation is as it should be expected to be higher in the raw data.

Table III

Predicted Probabilities under alternative estimated coefficients[13]

Variable	Obs	Mean	Std. Dev.	Min	Max
β	3748	.0765742	.0184983	.0056796	.1233955
β_1	11986	.0809266	.0180342	.0080866	.1261123
β_{01}	15734	.0798898	.0182397	.0056796	.1261123
γ	11986	.9760554	.0123796	.7740755	.9896138
γ_0	3748	.972301	.0161694	.7384747	.989172
γ_{01}	15734	.9751611	.0134748	.7384747	.9896138

Now, by predicting the values for the dependent variable across the different households, we obtain β in the smaller subsample and γ in the larger. To check, however, that everything was working in good order the two different vectors of estimated values were imposed each also on their complementary samples. Thus the data were also read using β on the sample of previously

[13]To avoid even further unnecessary notational complications we give summary statistics of the vectors under which the predicted probabilities are derived. Where there is no subscript this denotes that this is the probability derived in the logit estimations in the beginning. Where there is, this denotes the relevant subsample that each probability vector was read from. For example the row corresponding to γ_0 refers to summary statistics for the probability vector read by employing γ derived from the logit estimate in Table II but imposed on the "smaller"(3748 obs.) subsample.

employed women and γ on that of previously unemployed ones. In addition for further check they were also imposed on the whole sample the results of this being in Table III. From the Table it is easy to check that for example β_1 has a higher mean than β since in the larger sample that the former is calculated in, workers with better characteristics will dominate over those with poorer thus driving the mean up. When the pooled sample is used β_{01} is correspondingly lying in between the values for β and β_1 as it would be expected. Similar reasoning can be applied in the case of the probabilities estimated using alternative γ's.

Following this we need to use β and γ obtained above both on the sample ww1=0. Making use of the former estimate on the sample of women who were not working before we obtain estimates of the accession probability P10 some statistics of which can be found in Table IV below:

Table IV

Statistics on Series p10

p10

	Percentiles	Smallest		
1%	.0253884	.0056796		
5%	.0395484	.0105823		
10%	.0526796	.0136364	Obs	3748

25%	.0665503	.0149514	Sum of Wgt.	3748
50%	.0787524		Mean	.0765742
		Largest	Std. Dev.	.0184983
75%	.0893243	.1193725		
90%	.0974565	.1208472	Variance	.0003422
95%	.1038831	.1227112	Skewness	-.6204427
99%	.1133235	.1233955	Kurtosis	3.491321

These statistics on the vector P10 are nothing more than utilising β of the logit regression performed initially since the sample used is the same.

In addition, some descriptive statistics for P11, the retention probabilities, can be found in Table V that follows

Table V

Statistics on Series p11

p11

	Percentiles	Smallest		
1%	.9286844	.7740755		
5%	.9512656	.798278		
10%	.9622495	.8096078	Obs	11986
25%	.973335	.8618325	Sum of Wgt.	11986
50%	.9794371		Mean	.9760554

		Largest	Std. Dev.	.0123796
75%	.9836836	.9893457		
90%	.9858375	.9893944	Variance	.0001533
95%	.9867346	.9893977	Skewness	-3.259758
99%	.9882023	.9896138	Kurtosis	24.46843

Observing simply the sample means of the two series in the tables above we could reason as follows; in the case of P10 the sample mean is the one obtained using β in the ww1=0 sample as before (0.076574).

3.4 The Model

We shall be interested in developing a dynamic forward looking model of discrete labour market transitions. The core of the model is also to be found in Blundell *et. al.* (1992). The set up has agents living a finite life of T years and having the discrete choice over employment in the current period (S_t=1) or not – which we specifically associate with unemployment here – (S_t=0). We do not employ a three state model since we do not observe the "out of the labour force " state in our data.

To proceed we need to specify an indicator function describing the existence or absence of job offers. Hence, let the discrete random variable $J_t \in [0, 1]$ denote this and therefore we can write the arrival and staying-on (or layoff) rates as

$$\Pr(J_{t+1}=1 \mid S_t=i) = \alpha_{1i} \qquad i=0,1 \tag{1}$$

i.e. if i=0 then α_{10} denotes the "accession" arrival rate that is the arrival of offers from unemployment while if i=1 it denotes the staying on rate α_{11}. Hence, the origin state is the right subscript and the destination the left.

For the accumulation of assets first consider the usual framework of a per period budget constraint in the form

$$A_{t+1} = (1 + r_t)(A_t - c_t + w_t h_t + Y_t)$$

where A_t denote net real assets (or wealth) at the beginning of period t, r_t the real interest rate, h_t the number of hours in employment and w_t the real wage. Non-earned income is denoted by Y_t. Assuming a finite horizon and no bequests then also A_{T+1}=0. However, here we shall slightly modify the assets equation of motion since we model participation and transitions and shall replace h_t above by S_t. Thus if a woman is employed at t this would inflate her assets at t+1 by the total real returns on her wage, $(1+r_t)w_t$ and would not have an effect otherwise, viz.,

$$A_{t+1} = (1 + r_t)[A_t - c_t + w_t(S_t=1) + Y_t] \tag{2}$$

According to this equation, an agent at the beginning of period t receives (or not) an offer and observes its remuneration. She has to decide whether to accept it or not and how much to consume at a snapshot in t. The justification for the presence of $(1+r_t)A_t$ in (2) is also fairly standard.

We can thus write a more specialised form of the dynamic programming problem maximising an appropriate value function over the control variables which are c_t and S_t. We can however simplify the structure by maximising only over c_t and writing down two Bellman equations according to whether the spouse is employed (S_t=1) or not (S_t=0). Hence for each state and time period we have:

$$V_T(S_T = 1) = R_T(c_T) \;;$$

the reward or felicity function at T as a function of real consumption at T.

$$V_T(S_T = 0) = l \quad ;$$

the time invariant positive value of leisure and for t<T assuming optimal per period consumption allocations the corresponding value functions are:

$$\begin{aligned} V_t\Big(S_t{=}1;\ \mathrm{A}^e_{t+1}\Big) = R_t(c_t) + \delta\Big\{ [1\text{-}\alpha_{01}]\ \mathrm{E}_t\max[V_{t+1}(S_{t+1}{=}1;\ \mathrm{A}^e_{t+1}), \\ V_{t+1}(S_{t+1}{=}0;\ \mathrm{A}^e_{t+1})] \\ +\alpha_{01}\ \mathrm{E}_t[V_{t+1}(S_{t+1}{=}0;\ \mathrm{A}^e_{t+1})\Big\} \end{aligned} \qquad (3)$$

$$\begin{aligned} V_t\Big(S_t{=}0;\ \mathrm{A}^u_{t+1}\Big) = l + \delta\ \Big\{\alpha_{10}\ \mathrm{E}_t\max[V_{t+1}(S_{t+1}{=}1;\ \mathrm{A}^u_{t+1}), \\ V_{t+1}(S_{t+1}{=}0;\ \mathrm{A}^u_{t+1})] \\ +\ [1-\alpha_{10}]\ \mathrm{E}_t[V_{t+1}(S_{t+1}{=}0;\ \mathrm{A}^u_{t+1})\Big\} \end{aligned} \qquad (4)$$

δ is a personal time invariant discount factor and the superscripts on A_{t+1} refer to employment, e and unemployment, u. Writing the problem as we did we assume an additive time separable von-Neumann Morgenstern utility function for the within period utility functions or felicities (Deaton (1992)).

To develop the framework in order to study transitions we need to write an equation for $\Pr(S_{t+1}{=}1 \mid S_t{=}i)$, i=0 for the accession probability and =1 for the retention. To achieve this a woman needs to receive an offer now, J_{t+1}=1 <u>and</u> jointly with this evaluate if it is to her interest to be employed or not given her recent employment history (S_t=i). To do the latter she needs to evaluate the gain in utility by looking at the difference between the optimal reward at t of the lifetime utility from working and that from not working i.e. (3) – (4). Thus, a woman's transition probability from a state i=0,1 to employment can be written given the above discussion as:

$$\Pr\,(S_{t+1}{=}1 \mid S_t{=}i\,) =$$

$$\Pr\Big(J_{t+1}{=}1,\; V_{t+1}(S_{t+1}=1) \;>\; V_{t+1}(S_{t+1}=0) \mid\; S_t{=}i\Big) \tag{5}$$

and if the unobseved structure in preferences as this is entered through the value functions is assumed independent from this determining the arrival of offers then we can write (5) as :

$$\Pr\,(S_{t+1}{=}1 \mid S_t{=}i\,) =$$

$$\Pr\Big(J_{t+1}{=}1 \mid S_t{=}i\Big)\Pr\Big(V_{t+1}(S_{t+1}=1) > V_{t+1}(S_{t+1}=0) \mid S_t{=}i\,\Big)$$

$$=\alpha_{1i}\,\Pr\Big\{\,V_{t+1}(S_{t+1}=1) \;>\; V_{t+1}(S_{t+1}=0) \mid\; S_t{=}i\,\Big\} \tag{6}$$

The inequality in {.} being conditioned on S_t=i means that in evaluating her "next step" she must make use of all relevant information that is presently in her choice set and this is described by her current labour force status.

However for reasons that shall become apparent as we go along we shall see that we need to make two simplifying assumptions to relax this conditionality in order to achieve identification. Hence individual specific preference errors are not allowed and also time specific preference errors are assumed to be serially uncorrelated. In the framework we were discussing in the introductory section on Heterogeneity and State Dependence the first condition translates to having no ζ_i component and the second to having $E(\epsilon_{is}\epsilon_{jt})$=0 $\forall$ i,s,j,t (see 2.3.25 and 2.3.26) which precludes the possibility of having state dependent preference errors to give us:

$$\Pr(S_{t+1}=1 \mid S_t=i) = \alpha_{1i} \Pr\Big\{ V_{t+1}(S_{t+1}=1) > V_{t+1}(S_{t+1}=0)\Big\} \tag{7}$$

Since under the last two assumptions $\{.\}$ is now the same for all individuals we can thus by getting the ratio for i=0 in (11) to i=1 identify $\frac{\alpha_{10}}{\alpha_{11}}$ and subsequently obtain estimates of it.

3.4.1 Motivating the model

To give some motivation for the model consider writting it in the form:

$$\Pr(\text{wwork}_t^i=1|\text{ww1}_{t-1}^i=0,\ z_{it},\ c_{it-1}) = \alpha_{1|0}(z_{it})\Psi(z_{it},c_{it-1}) \tag{8}$$

$$\Pr(\text{wwork}_t^i=1|\text{ww1}_{t-1}^i=1,\ z_{it},\ c_{it-1}) = \alpha_{1|1}(z_{it})\Psi(z_{it},c_{it-1}) \tag{9}$$

Equation (8) depicts the accession probability for an individual i between periods $t-1$ and t given her characteristics vector z_{it} and consumption lagged once, c_{t-1}. Then clearly as shown in this equation we are interested in parameterising this function multiplicatively into two components; the first,$\alpha_{1|0}$ what is usually referred to in the literature as the arrival rate of offers is taken to be a function of only the relevant individual demographic characteristics and not of the consumption variable. The latter is allowed to enter only through function $\Psi(.,.)$ which is assumed to be common among the two subsamples and thus enters in an identical fashion in (8) and (9). We shall try to explain why this Ψ function enters in the same way in both (8) and (9) by looking into the α

and Ψ components more carefully. It is thus intended that the αs distinguish the demand from the supply side of the model as the latter is proxied by the Ψ components.

More precisely the α's are meant to represent the demand side of offers "arriving" to a worker and are a function of her characteristics z while the Ψs denote the supply side of the model and are useful for utility comparisons. That is through Ψ a woman will take into account her demographic characteristics as well as her consumption expenditure undertaken in the last period in order to calculate the extent to which she would be willing to work in this current period. Her preference, therefore, for employment at t given her vector of demographic characteristics z_{it} as well as her consumption expenditure at $t-1$, c_{t-1}, which might be seen as a "proxy" for her assets is entered through $\Psi(.,.)$.

To give some intuition for this model as an example consider the case where $\alpha_{1|0}(z_{it}) = \alpha_{1|1}(z_{it}) = 1$ so that every individual irrespective of labour market state, faces the same absolutely certain chance of receiving one offer at period t. That is the demand for labour for all individuals irrespective of their most recent work status dictates that they shall be exposed to one offer at any t. The arrival and the layoff rates are then not a function of one's previous period employment status. Her work supply considerations then will coincide in the two cases meaning that since there is no "filtering" of individual labour supply considerations through the demand for labour, the probability of transitting to employment would be identical to one's preferences to take up a paid job. A moment's thought would reveal that this would be the case when there is no state dependence in the labour market with the case for state dependence arising clearly when $\alpha_{1|0}(z_{it}) \neq \alpha_{1|1}(z_{it})$. In consequence of the no state-dependence assumption both the accession and retention probabilities would be equal to the same common term $\Psi(z_{it},c_{it-1})$ i.e. to individual i's likely preference over employment. This, however, is a strong if not a false assumption given that what we often come across in the labour market is that a current employment status may in fact influence a future one; i.e. the opposite. The hypothesis of "true" state dependence [Heckman & Willis (1977)] is the subject matter of ongoing research and concerns also the present one. It is in quotes to distinguish it also from what they note as "apparent" state dependence caused by unobserved heterogeneity.

However, before saying anything further let us also look at equation (9) which correspondingly refers to the retention probability for the same individual during the same time span. It is equivalently split into an α and a Ψ component the α now taking the role of ensuring that offers would be forthcoming so that individual i in (9) maintains employment between t and $t-1$.

From the independent[14] logit regressions described earlier we therefore get an estimate of the LHS of (8) which is β and one for the LHS of (9) which is γ. The question that could be raised now is the following; how is it possible to get an identical function Ψ for (8) and (9) since we initially have to obtain estimates of the coefficients of the vector (z_{it}, c_{it-1}) and these clearly would be expected to vary significantly among the two different subsamples ? In other words given that we obtain β in (8) and γ in (9) how is it possible to ignore that this shall alter our preference for employment side in the two equations ? We clearly have to be able to say why this is in fact identical in (8) and (9) since otherwise we will not be able to simplify the structure of the model for estimation. To answer

[14]The terminology "independent" here is not ideal but it is simply meant to distinguish the logit regression of the first step from the "conditional" (two-step) estimation of the second.

this question consider we focus on one representative individual (even though we shall not hinge on any particular notions of representativeness) drawn from the sample ww1=0. For this particular individual and in fact for any individual from the sample that was drawn from we can write the above two equations as follows:

$$P_{1|0}(\beta, z_{it}, c_{it-1}) = \alpha_{1|0}(z_{it}) g(\beta, z_{it}, c_{it-1}) = h(\beta, z_{it}, c_{it-1}) \tag{10}$$

$$P_{1|1}(\gamma, z_{it}, c_{it-1}) = \alpha_{1|1}(z_{it}) g(\beta, z_{it}, c_{it-1}) = h(\gamma, z_{it}, c_{it-1}) \tag{11}$$

The new notation should be clear enough; The LHS of (10) and (11) are the LHS of (8) and (9) respectively and $g(.,.,.) \equiv \Psi(.,.)$ for people in a given subsample. We shall note that this in fact is identical for any particular individual drawn from either subsample. As for $h(.,.,.)$ appearing in these equations we shall explain below. In other words once she is chosen from her own sample ww1=0 she maintains the information on her demographic characteristics in (10) through β. However if she could freely move between samples and say become part of the sample of those who retain their jobs[15] we would notice the following ; her preference for employment as this is in fact expressed by the g(.,.,.) function above would be the same (being the same individual) but the action now is in the demand side and through the filtering process her probability of maintaining her employment[16] is given by P_{11}.

In other words her preference structure is presumed to carry over intact when she is in ww1=1 and the blend with demand in this camp in fact dictates that all individuals with existing previous period employment should face a

[15]Assuming, for convenience, this takes place with random replacement of one individual from the ww1=0 sample so as not to alter their size.

[16]Given that she is now in the sample of women with "better" characteristics but in fact this is the accession probability for her.

higher chance of finding a job; the demand for such kind of women is signalled through γ. Finally it is also clear that a diametrically similar argument could be run had we wanted to compare an individual's preferences from the wwl=1 sample in the wwl=0 one. In that case we would read the g functions with γ's and the h ones the same as before. That is the reason why we can assume that the function Ψ is the same between the samples for the same individual and in fact because specifically of the fact that conditioning on the assets we can 'retrieve' the same individual. This is crucial from both an economic as well as an econometric point of view. It is crucial as far as the former is concerned since it imposes a restrictive preference structure for employment by requiring that both types of women have identical preferences. It is crucial from an econometric point of view in terms of identification.

3.4.2 Identification

For the identification of all but one of the parameters of the model we need to focus in each period on the fraction of people that are working and those that are not. More specifically variations in the data on transitions, the nonlinearities in the model as well as the finiteness of the horizon interplay to achieve identification of the coefficients of interest $\left(\text{here the coefficients of } \frac{\alpha_{10}}{\alpha_{11}}\right)$. People differ with respect to their most recent observable participation episodes as this is discretely proxied by S_t. They are then exposed to a demand for labour exogenous arrival (or layoff) rate which affects their assets position through the effect it is expected to have on consumption levels. Hence, the sufficient factors required for identification of the arrival to staying in rates is provided by the equation of motion for the accumulation of assets A_{t+1} . Also that the unobserved structure in preferences (as entered through the value functions) is assumed independent from this determining the arrival of offers. Finally, individual specific errors are not allowed and time specific preference errors are serially uncorrelated.

3.5 Estimation

As we have already said above in the first step we get estimates β for (8) and γ for (9). Since the estimates obtained are from different samples we clearly need to use any relevant information that is apparently pertinent to each estimate on its own sample on the wwl=0 set; i.e. we need to use β and γ obtained above, but presently on the sample of women that were not engaged in work in the previous interview(wwl=0). Thus doing this we get

$$P_{1|0}(\beta, z_{it}, c_{it-1}) = \frac{\alpha_{1|0}(z_{it})}{\alpha_{1|1}(z_{it})}\ P_{1|1}(\gamma, z_{it}, c_{it-1})$$

$$= \rho(z_{it})\ P_{1|1}(\gamma, z_{it}, c_{it-1}) \tag{12}$$

which is the equation that is estimated. From equation (12) we can note that what we can identify is a vector of estimated parameters for the ratio $\rho(z_{it})$. For the present two specifications for the ratio of the two arrival rates have been adopted; the first was to consider $\rho(z_{it}) = z_{it}\delta$, a linear restriction term, while also the ratio taking an exponential form and thus $\rho(z_{it}) = e^{z_{it}\delta}$ was considered in the estimation as well. It would be interesting also to have been able to separately identify parameter estimates for each arrival rate but here we focus on obtaining estimates that will identify their respective ratio. This is informative enough as we noted in passing earlier since information on this ratio helps to draw conclusions concerning any possible state dependence that shall be present. We do not explicitly accomodate any apparent state dependence caused by heterogeneity and use a vector of demographic variables in our parameterisation of the model.

Table IX

Maximum Likelihood Estimation

Linear Restriction($Z_i\delta$) Non-Linear Restriction($e^{Z_i\delta}$)

Number of obs = 3748

Linear Restriction		Non-Linear Restriction	
F(6, 3741)	= 1575.43	F(6, 3741)	= 1662.32
R-square	= 0.7165	R-square	= 0.7272
Adj R-square	= 0.7160	Adj R-square	= 0.7268
Root MSE	= .00986	Root MSE	= .15242

p10	Coef.	Std. Err.	Coef.	Std. Err.
un1	.0027	.0001	.0335	.0018
AGE2	.0026	.0002	.0390	.0026
AGE2sq	-5e-5	2.1e-6	-7e-4	3.3e-5
inf	-.0133	.0005	-.172	.0071
ed2HS	.0435	.0009	.854	.0137
ed2C	.0520	.0009	.948	.0139
_cons	-.0072	.0034	-4.023	.0551

An immediate reaction would point to the rather surprisingly high determination of the coefficient estimates. Before we look into this let us however continue. Most variables have signs as predicted by theory. In a labour supply context, infants, for example, would be expected to exert a negative effect on wife's participation indicating the likely fact that they have an increase in her shadow wage of time at home so that home time would increase at the expense of market experience.

Note, however, the demand side interpretation of the coefficient estimates since these are with the supply-side "conditioned out". Hence, a woman with small children may face a higher turnover rate so that the demand for such women is relatively smaller than those without infants. In terms of the demand for labour then, she should be expected to face a lower probability of entering employment due to precisely that reason. In addition her level of education seems to be positively related to her probability of entering employment from unemployment and this is shown by the higher coefficient of ed2C as compared to that of ed2HS.

Now, probably due to the large sample size but more importantly due to the inherent two-step estimation procedure adopted we come up with this evident over-precision in the coefficient estimates. Deriving the standard errors that would incorporate this estimation procedure is looked at in the Appendix to this Chapter. In the next Table we report them for the above model in the non-linear restriction case. The significance of the above estimates drops tremendously.

Table X

Dependent Variable: $\ln\left(\dfrac{P_{1|0}(\widehat{\beta}, z_{it}, c_{it-1})}{P_{1|1}(\widehat{\gamma}, z_{it}, c_{it-1})}\right)$

Variable	Corrected S.E.	t
un1	.038799474	.8634
AGE2	.06597841	.5911
AGE2sq	.00080623	-.8682
infants	.18063574	-.9522
ed2HS	.47452108	1.7997
ed2C	.47534459	1.9943

We thus see that only the influence of education is significant after we correct the standard errors. This could in turn mean that the educational dummies are doing a reasonable job at picking up the effect on the probability of an accession transition, persons possessing actually characteristics of job keepers might have.

As a final step, we graph the probabilities discussed above, and not their logarithmic ratio, over time; this is done in Figures 1 to 3. The continuous lines in these graphs connect the median points for the respective years. We can observe that the conditional accession probability is following an almost identical pattern to the Logit one except for the Years 1983 to 1986 where the conditional is everywhere and above the Logit accession probability.

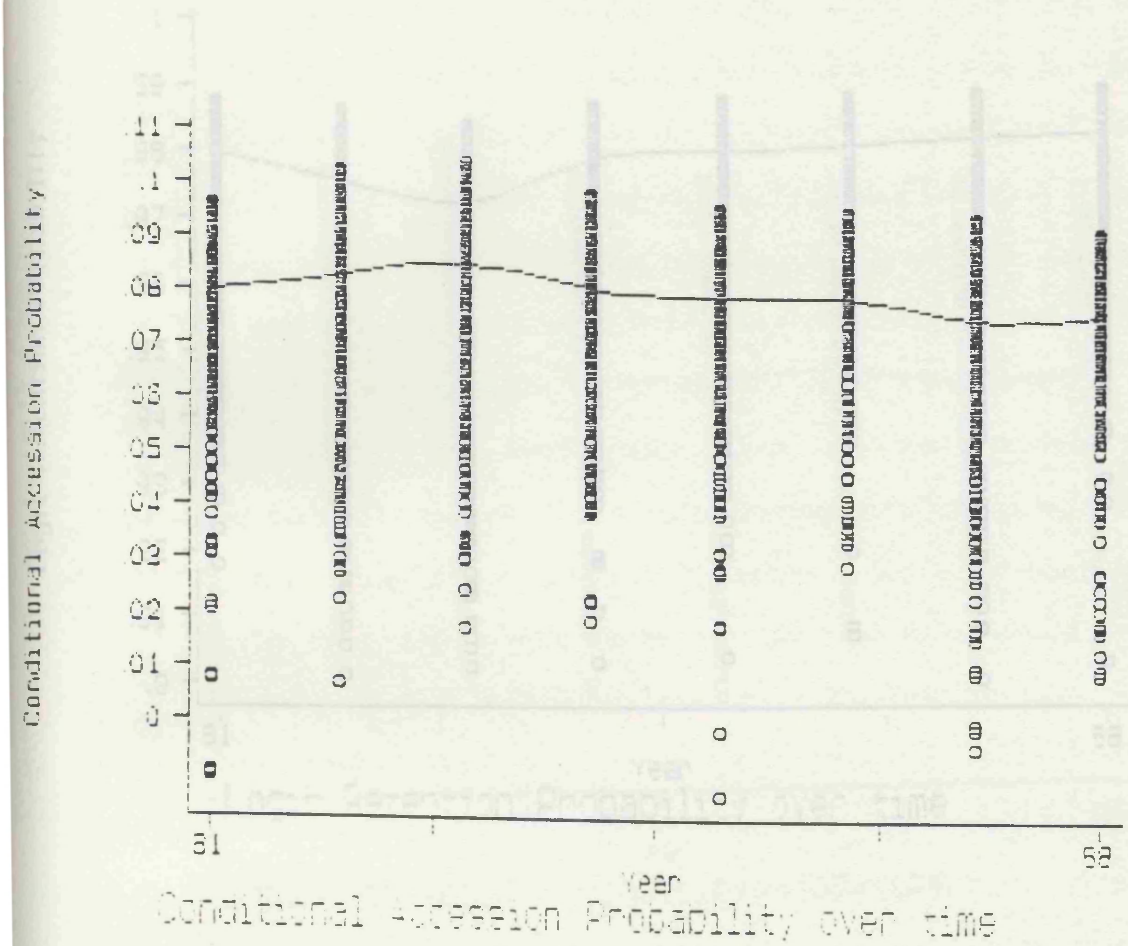
Conditional Accession Probability
.11
.1
.09
.08
.07
.06
.05
.04
.03
.02
.01
0
81
88
Year
Conditional Accession Probability over time

Figure 1.

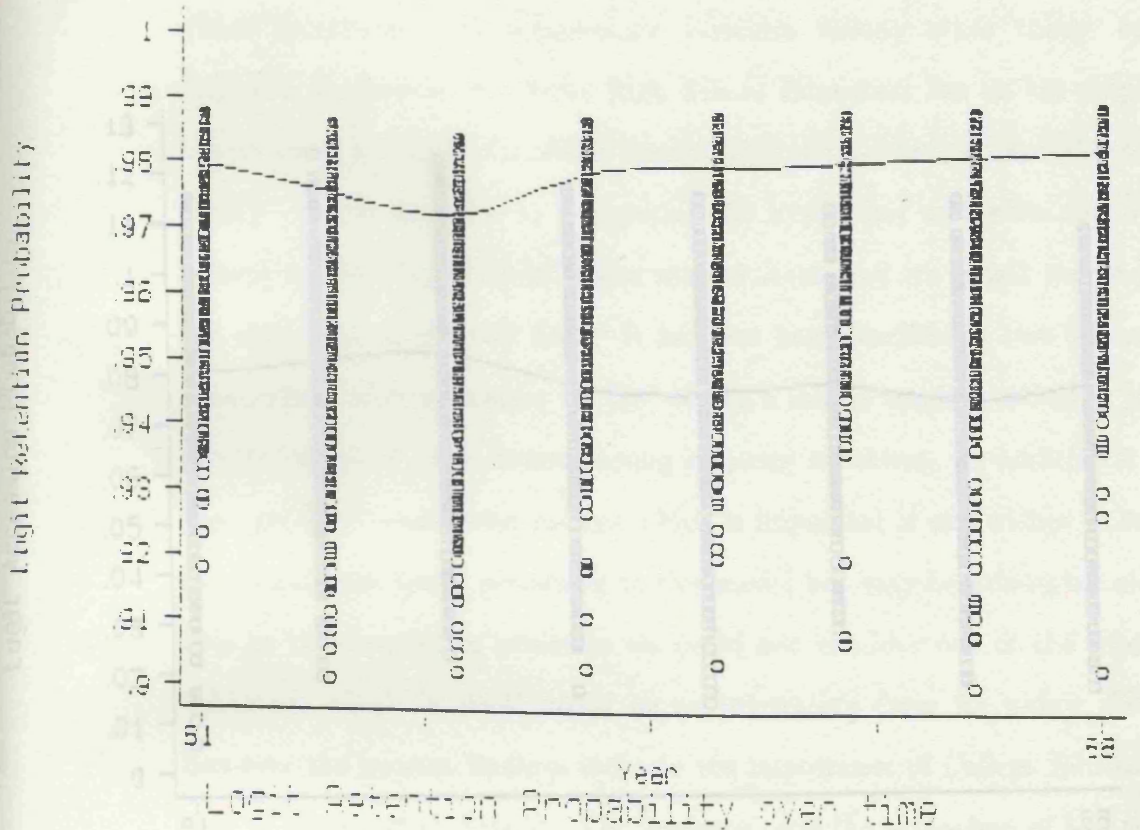
Logit Retention Probability
1
.99
.98
.97
.96
.95
.94
.93
.92
.91
.9
81
88
Year
Logit Retention Probability over time

Figure 2.

3.6 Conclusions

The present Chapter illustrated a model of labour market transitions which incorporated as explanatory variables among other things education dummies to describe the effect High School Education has on the employment behaviour of a sample of married women from the Consumer Expenditure Survey (CES). It was desirable to incorporate the hypothesis that state dependence is present among (consecutive) labour market states and we in fact saw that this is the case even in the raw data. It has not been possible to test for any racial interactions with education on the women's labour market activities since the sample consisted of an overwhelming majority of whites. In addition it has not been possible to use asset income which is important if one wishes to test some functional form issues pertaining to this model but may be attempted elsewhere. Due to the two states available we could not consider out of the labour force behaviour which in itself would be an interesting focus for policy discussions. However the present findings indicate the importance of College Education over that of simply High School. The credence over the utilisation of business cycle effects through the economy-wide unemployment rate on labour market transitions, is still an open question. For example the economy-wide mean unemployment rate estimates seemed to suggest that an employed person may lose his job but can always get another one. Furthermore, we saw (coefficient of totexp1 in the first step estimation) that there exist wealth effects through consumption on transitions. In future work it is intended that such wealth effects could be more adequately treated. Finally, in the Appendix to this Chapter we derive the standard errors that are valid for the second step estimation procedure

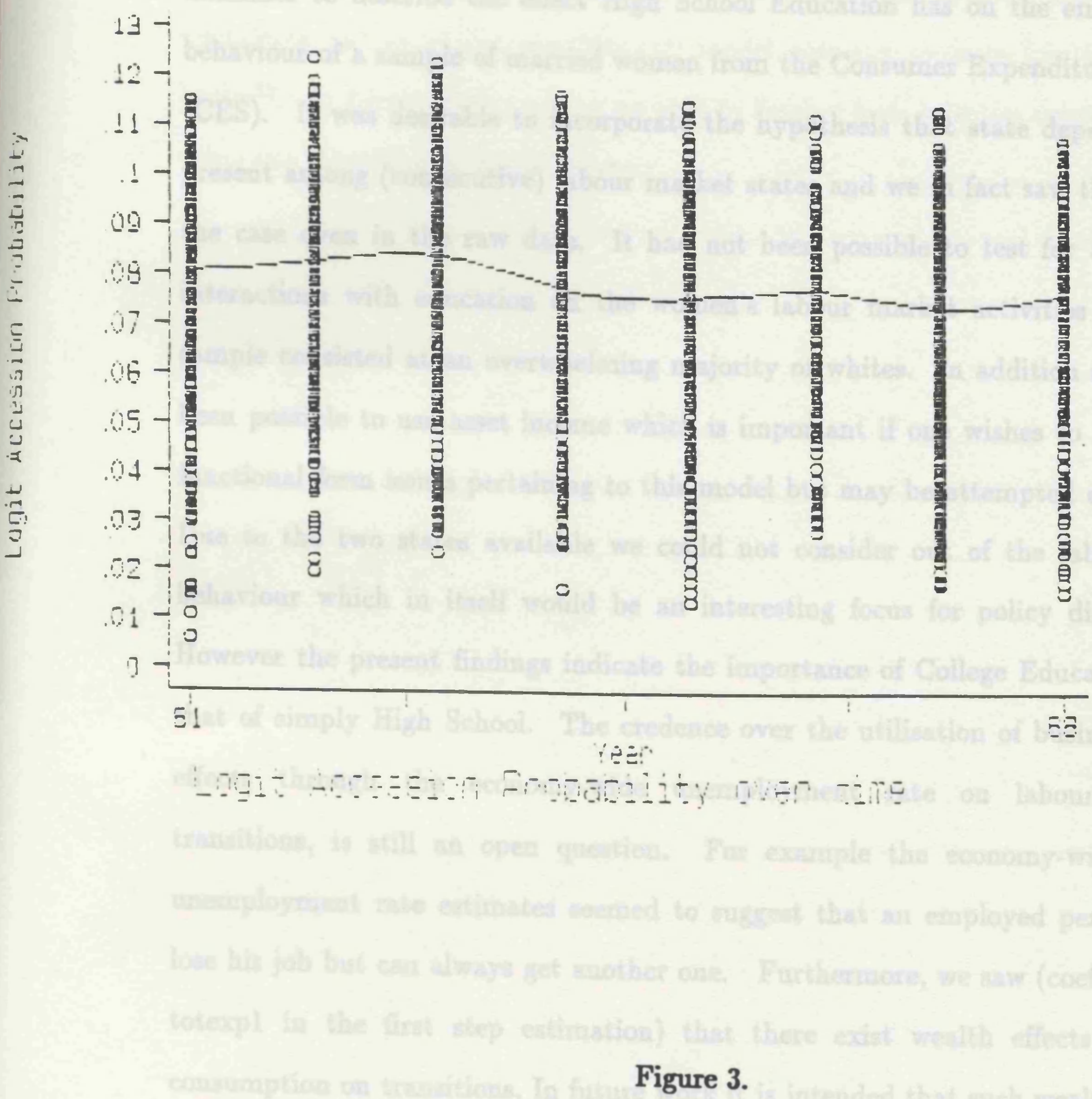

Figure 3.

3.6 Conclusions

The present Chapter illustrated a model of labour market transitions which incorporated as explanatory variables among other things education dummies to describe the effect High School Education has on the employment behaviour of a sample of married women from the Consumer Expenditure Survey (CES). It was desirable to incorporate the hypothesis that state dependence is present among (consecutive) labour market states and we in fact saw that this is the case even in the raw data. It had not been possible to test for any racial interactions with education on the women's labour market activities since the sample consisted at an overwhelming majority of whites. In addition it has not been possible to use asset income which is important if one wishes to test some functional form issues pertaining to this model but may be attempted elsewhere. Due to the two states available we could not consider out of the labour force behaviour which in itself would be an interesting focus for policy discussions. However the present findings indicate the importance of College Education over that of simply High School. The credence over the utilisation of business cycle effects through the economy-wide unemployment rate on labour market transitions, is still an open question. For example the economy-wide mean unemployment rate estimates seemed to suggest that an employed person may lose his job but can always get another one. Furthermore, we saw (coefficient of totexp1 in the first step estimation) that there exist wealth effects through consumption on transitions. In future work it is intended that such wealth effects could be more adequately treated. Finally, in the Appendix to this Chapter we derive the standard errors that are valid for the second step estimation procedure that we pursue. After controlling for this we saw in the previous section how dramatically the accuracy of the coefficient estimates drops. This may be due to the fact that the model can not capture at all different points in time the effect through the variables we specified but perhaps some other unobservable (at least at the time of writing this) is driving it. However, even utilising cross-sections containing history of the variables this model does not seem to fare that much better[17]. In future work we may be able to further look into the existing model using the General Household Survey.

[17]In some preliminary work in progress utilising cross-sections from the General Household Survey of the U.K. The results seemed more promising but due to time limitations we only hope to develop them in the future.

Appendix B: Consistency of the estimator and computation of the standard errors

In this section we show consistency of the estimator and also derive the standard errors for this form of model we believe is intuitively more appealing from an economic view-point. The approach utilises elements from Pagan(1984) and Arellano & Meghir(1992) and can be seen as a variant of the Delta Method (individual subscripted though). However, unlike in the latter paper, here the dependent variable is also observed with error and we wish to control for its error structure. This structure then is the one pertaining to the estimate of the ratio of arrival rate probabilities.

To proceed thus assume that the model can be written in terms of the simplistic notation that has been adopted throughout the paper, i.e. the "true" model has the form

$$P_{1|0}(\beta, z_{it}, c_{it-1}) = \rho(z_{it}) P_{1|1}(\gamma, z_{it}, c_{it-1})$$

and for a log-normally distributed disturbance term u we can write

$$P_{1|0}(\widehat{\beta}, z_{it}, c_{it-1}) = \rho(z_{it}) P_{1|1}(\widehat{\gamma}, z_{it}, c_{it-1})\, e^{u} \tag{A.1}$$

$$\text{where } \rho(z_{it}) = e^{\delta' z_{it}}.$$

We shall derive the standard errors for the above model. When we take logs it takes the form

$$\log[P_{1|0}(\widehat{\beta}, z_{it}, c_{it-1})] = \delta' z_{it} + \log P_{1|1}(\widehat{\gamma}, z_{it}, c_{it-1}) + u$$

Thus we can perform OLS on:

$$\log\left(\frac{P_{1|0}(\widehat{\beta}, z_{it}, c_{it-1})}{P_{1|1}(\widehat{\gamma}, z_{it}, c_{it-1})}\right) = \delta' z_{it} + u \tag{A.2}$$

Therefore suppressing the t subscripts and using $P_{1|0}(\hat{\beta})$ and $P_{1|1}(\hat{\gamma})$ with an aim to economise on unnecessary notational repetitions we have[18]:

$$\hat{\delta} = [\sum_i z_i z_i']^{-1}[\sum_i z_i \log \frac{P_{1|0}(\hat{\beta})}{P_{1|1}(\hat{\gamma})}]$$

which using a Taylor series approximation around $\hat{\beta}$ and $\hat{\gamma}$ can be written as:

$$\hat{\delta} = [\sum_i z_i z_i']^{-1}[\sum_i z_i \log \frac{P_{1|0}(\beta)}{P_{1|1}(\gamma)}] +$$

$$[\sum_i z_i z_i']^{-1}\Big\{\sum_i z_i\big\{[\log P_{1|0}(\hat{\beta}) - \log P_{1|0}(\beta)] - [\log P_{1|1}(\hat{\gamma}) - \log P_{1|1}(\gamma)]\big\}\Big\}$$

$$= \delta + [\sum_i z_i z_i']^{-1}\Big\{\sum_i z_i\big\{[\partial \log \frac{P_{1|0}(\beta)}{\partial \beta'}]_* (\hat{\beta} - \beta) - [\partial \log \frac{P_{1|1}(\gamma)}{\partial \gamma'}]_+ (\hat{\gamma} - \gamma)\big\}\Big\} \qquad \text{(A.3)}$$

where $*$ means evaluated at a point between $\hat{\beta}$ and the true value,β and $+$ similarly for a point lying between $\hat{\gamma}$ and γ. Before we proceed with the computation of the standard errors we can easily deduce even from (A.3) that the estimator of the state dependence ratio, $\hat{\delta}$, is a consistent estimator of the true value δ. This is so since $\hat{\beta}$ and $\hat{\gamma}$ are consistent ML logit estimators and the terms preceding them inside the curly brackets are clearly finite as $N_1 \to \infty$ since they contain log vectors of probabilities whose sum over the i persons is finite.

Let us now establish some notation for the cross-section dimension. If N_1 denotes the observations on the women that experienced accession transitions,

[18]It has to be reminded that $\hat{\delta}$ is evaluated at this subset of the data which excludes consumption expenditure in CES, i.e at z_{it}, whereas the accession and retention probabilities have been estimated on the comlete data matrix which includes consumption.

(1|0) and N_2 the observations on those who retained their employment, (1|1), $N_1 + N_2 = N$ we want to look at the asymptotic standard errors of the 2^{nd} step estimator $\hat{\delta}$ which is derived if we condition the estimation on the N_1 subsample and is further documented in the main text.

So $\sqrt{N_1}(\hat{\delta} - \delta) \overset{as.eq}{\sim}$

$$[\frac{1}{N_1}\sum_i z_i z_i']^{-1}\left\{\frac{1}{N_1}\sum_i z_i \left[\partial \log \frac{P_{1|0}(\beta)}{\partial \beta'}\right]_* \sqrt{N_1}(\hat{\beta} - \beta)\right\} -$$

$$[\frac{1}{N_1}\sum_i z_i z_i']^{-1}\left\{\frac{1}{N_1}\sum_i z_i \left[\partial \log \frac{P_{1|1}(\gamma)}{\partial \gamma'}\right]_+ \sqrt{N_1}(\hat{\gamma} - \gamma)\right\}$$

$$\equiv [\frac{1}{N_1}Z'Z]^{-1}[\frac{1}{N_1}Z'P_\beta]\sqrt{N_1}(\hat{\beta} - \beta) - [\frac{1}{N_1}Z'Z]^{-1}[\frac{1}{N_1}Z'P_\gamma]\sqrt{N_1}(\hat{\gamma} - \gamma) \tag{A.4}$$

with some more desired notational simplifications!

However,since $\hat{\beta}$ and $\hat{\gamma}$ draw from two independent logistic distributions it is well known that:

$$V_{\sqrt{N_1}(\hat{\beta} - \beta)} \equiv As.Cov(\sqrt{N_1}(\hat{\beta} - \beta)) = \left\{\frac{1}{N_1}\sum_{i \in N_1} \frac{f_i^2 z_i z_i'}{F_i(1 - F_i)}\right\}^{-1} \tag{A.5}$$

and $$V_{\sqrt{N_2}(\hat{\gamma} - \gamma)} \equiv As.Cov(\sqrt{N_2}(\hat{\gamma} - \gamma)) = \left\{\frac{1}{N_2}\sum_{i \in N_2} \frac{\overset{\times}{f}_i^{\,2} z_i z_i'}{\overset{\times}{F}_i(1 - \overset{\times}{F}_i)}\right\}^{-1} \tag{A.6}$$

where F_i and f_i denote the cdf and pdf respectively of the logistic distribution for the estimation of $\hat{\beta}$ while the similar symbols bearing $\times$ denote those for $\hat{\gamma}$. Note once more that the two estimators are drawn from subsamples independent from one another; since β refers to those women who were not observed in employment in the previous interview and γ to those who were.

Furthermore since

$$\frac{N_1}{N_2} V_{\sqrt{N_2}(\hat{\gamma}-\gamma)} = V_{\sqrt{N_1}(\hat{\gamma}-\gamma)} \tag{A.7}$$

and on the assumption that $\frac{N_1}{N_2} \to k$, finite and non-zero we can readily substitute (A.5) and (A.7) in (A.4) to further derive the Asymptotic Covariance matrix for $\sqrt{N_1}(\hat{\delta}-\delta)$. This is:

$$V_{\sqrt{N_1}(\hat{\delta}-\delta)} = [Z'Z]^{-1}[Z'P_\beta]V_{\sqrt{N_1}(\hat{\beta}-\beta)}[Z'P_\beta]'[Z'Z]^{-1} +$$

$$[Z'Z]^{-1}[Z'P_\gamma]\frac{N_1}{N_2} V_{\sqrt{N_2}(\hat{\gamma}-\gamma)}[Z'P_\gamma]'[Z'Z]^{-1} \tag{A.8}$$

Appendix C: Derivation for a more general set-up.

If the estimator $\hat{\delta}$ however can be expressed in the form

$$\hat{\delta} = f(Z,\hat{\beta},\hat{\gamma}) \tag{B.1}$$

then in a similar way to that used to derive $\hat{\delta}$ given a specific functional form as above,

$$\hat{\delta} = \delta + [\frac{\partial f}{\partial \beta'}]_*(\hat{\beta}-\beta) + [\frac{\partial f}{\partial \gamma'}]_+(\hat{\gamma}-\gamma)$$

or

$$\hat{\delta}-\delta = [\frac{\partial f}{\partial \beta'}]_*(\hat{\beta}-\beta) + [\frac{\partial f}{\partial \gamma'}]_+(\hat{\gamma}-\gamma) \tag{B.2}$$

$\Leftrightarrow$

$$\sqrt{N_1}(\hat{\delta}-\delta) = [\frac{\partial f}{\partial \beta'}]_*\sqrt{N_1}(\hat{\beta}-\beta) + [\frac{\partial f}{\partial \gamma'}]_+\sqrt{N_1}(\hat{\gamma}-\gamma) \tag{B.3}$$

Denote the Asymptotic Covariance matrix of $\sqrt{N_1}(\hat{\beta}-\beta)$ by $V_{\hat{\beta}}$ and that of $\sqrt{N_2}(\hat{\gamma}-\gamma)$ by $V_{\hat{\gamma}}$ then provided $V_{\hat{\beta}}$ is independent of $V_{\hat{\gamma}}$ (as in the problem here) and the relevant boundedness condition for the ratio between the two sample sizes is again appropriately taken into account then

$$V_{\sqrt{N_1}(\hat{\delta}-\delta)} = [\frac{\partial f}{\partial \beta'}]_* V_{\hat{\beta}} [\frac{\partial f}{\partial \beta'}]_*' + [\frac{\partial f}{\partial \gamma'}]_+ \frac{N_1}{N_2} V_{\hat{\gamma}} [\frac{\partial f}{\partial \gamma'}]_+' \tag{B.4}$$

while with dependence between $\hat{\beta}$ an $\hat{\gamma}$ (B.4) becomes:

$$\begin{aligned} V_{\sqrt{N_1}(\hat{\delta}-\delta)} = {} & [\frac{\partial f}{\partial \beta'}]_* V_{\hat{\beta}} [\frac{\partial f}{\partial \beta'}]_*' + [\frac{\partial f}{\partial \gamma'}]_+ \frac{N_1}{N_2} V_{\hat{\gamma}} [\frac{\partial f}{\partial \gamma'}]_+' \\ & - [\frac{\partial f}{\partial \gamma'}]_+ \sqrt{\frac{N_1}{N_2}} \, (ACov\hat{\beta},\hat{\gamma})' [\frac{\partial f}{\partial \beta'}]_*' \\ & - [\frac{\partial f}{\partial \beta'}]_+ \sqrt{\frac{N_1}{N_2}} \, (ACov\hat{\beta},\hat{\gamma}) [\frac{\partial f}{\partial \gamma'}]_*' \end{aligned} \tag{B.5}$$

where $(ACov\hat{\beta},\hat{\gamma}) \equiv As.Cov\Big(\sqrt{N_1}(\hat{\beta}-\beta) \, , \, \sqrt{N_2}(\hat{\gamma}-\gamma)\Big)$.

Chapter 4

Women's Labour Force Participation: A Human Capital Approach.

4.1 Introduction.

The present Chapter presents a simple theoretical framework for looking at the employment decisions of married women. It conceptually relates to – or rather has been triggered by – the developments of a model of labour market transitions as these relate to savings behaviour using the CES data set from the U.S. [Blundell *et.al* (1992)] or to the examination of the effect of education utilising the same data set as we saw in the previous Chapter. By utilising a static set-up where women engage in a human capital augmenting activity in the first period while can choose between participation or abstention from the labour market in the second, we are able to develop a model that is capable of putting a little bit more structure on the labour demand side. The resulting formulation for the offer probabilities in the body of the Chapter is thus dependent upon the general level of economic activity, the types of agents that wish to participate and their relative probabilities of being successful in the labour market.

One of the insights we wish to offer in the present Chapter and through the model developed in the next section is over some role through wealth or the distribution of assets on the participation in the labour market decision of married women. In developing the model we can also provide a link with the way the job offer distribution is modelled in the transitions model. To do this though we need to establish some basic working notation first.

We let individuals be of the usual two types: high and low ability. Other things being equal, high ability (H) individuals have a higher success probability than low ability (L) individuals in competing for jobs in the market.

Ability	H	L
Fraction	λ	$1-\lambda$
Success		
Probability	p_H	p_L

Table 1

If a low ability person has a success probability p_L of finding a job, then H has a corresponding probability $p_H = \delta\ p_L$, where $\delta > 1$. These success probabilities as we shall see will be dependent amongst other things, on the number of jobs on offer that we shall denote by s. We shall also let individuals differ according to their wealth endowments, y, where y is a random variable assumed to derive for simplicity from a uniform[1] distribution over the interval $[a,b]$. It thus follows that:

$$f(y) = \frac{1}{(b-a)} \tag{1}$$

[1]One could change this distributional assumption to a more realistic one – in terms of economic arguments for the distribution of wealth –, such as Pareto or Exponential. However, for the purposes of the chapter in its current form this should not alter the main conclusions. In further elaborations of the present work we will take this point and possible consequences of it into consideration.

and that $F(z)=\frac{[z-a]}{[b-a]}$, is the proportion of individuals with wealth holdings below or equal to z. Wealth and ability are assumed independently distributed so that an uncomfortable assumption that the rich tend to be the higher ability individuals can be ruled out.As we shall see in the next section even though in Blundell *et.al.* (1992) and Chapter 3 here, the success probabilities above are identified as one-to-one mappings of the macroeconomic activity, in this Chapter they will be derived to be functions of this and other important quantities such as the proportion of high ability women, the magnitude by which they are more able than their low ability counterparts etc. But let us see how this evolves.

This Chapter is thus organised as follows: In section 2 we show that in an economy where there are equal opportunities for everyone (and at zero cost) to acquire education the demand for jobs is independent of the distribution of wealth. In section 3 we allow assets and the level of their skills to enter the participation decision. Section 4 provides some brief conclusions.

4.2 An economy with equal opportunities for education.

This is a model for married women. There are two periods and no time discounting. Individuals face a choice between two alternatives. In order to become skilled labourers they have to undergo a process which will augment their human capital holdings in period 1, at the end of which they face a probability p_i, i = L,H of finding a job. The availability of one job offer takes up one further period and is rewarded with a fixed (at least in the short run) real wage w^2. No job offer forces workers to stay unemployed in period 2 and earn a payoff of k. The alternative choice individuals can opt for consists of staying at home (unemployed) and earn a payoff of 2k. We assume $w > 2k$ right from the outset so that women have an incentive to get a job. These choices are sketched in Figure 1.

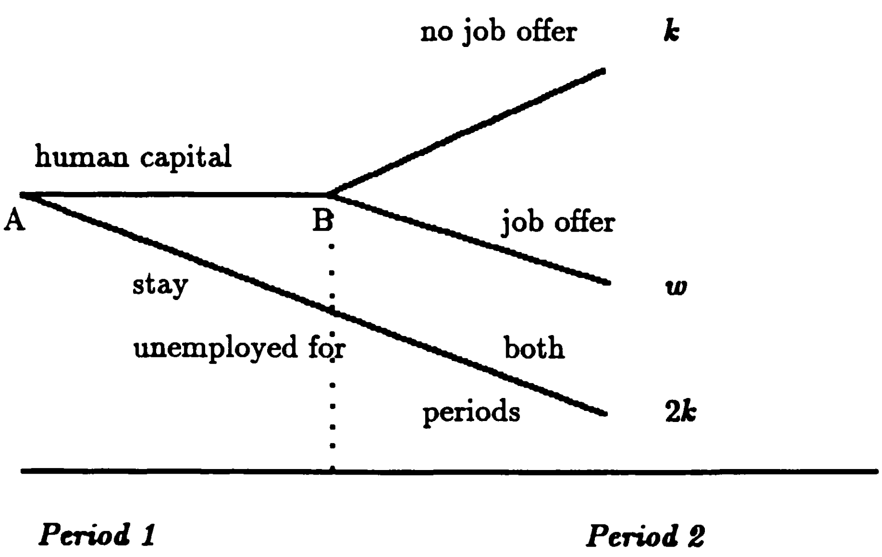

Figure 1

Women enter in node A possessing different abilities (reflected in p_H and p_L) which can be influenced (and ideally increased) by investing in human capital in the first period. We can view these probabilities as exogenous parameters of the utility function depending, for the moment, on genetic factors. Below we shall

[2]This assumption is crucial as far as $s \leq 1$ which is central to the set-up of this model. i.e. not all of the people who look for jobs will find one.

see how the number of job offers can also become an argument in them. If women decide to take up education, then after its completion they face the possibility of a job offer with a certain non-zero probability. In this sense, investing in human capital would be an element to the sequential decision of women to participate in the labour market.

For an L-type person with wealth y, her expected utility of finding a job – assuming risk neutrality –, is:

$$u_L(p_L, w, k, y) = p_L(w+y) + (1-p_L)(k+y) \tag{2}$$

and her reservation utility is

$$u_L(k, y) = 2k + y = u \tag{3}$$

while for an H-type person with identical wealth y

$$u_H(p_H, w, k, y) = p_H(w+y) + (1-p_H)(k+y) \tag{4}$$

We should stress that due to the risk neutrality assumption the *argmax* of either (2) or (4) each subject to her respective reservation utility is independent of y. If utility is non-linear, incentive effects come into play; e.g. poor people can no longer forego their home production in first period – less likely to want to put their chance of getting a job in jeopardy. We decided here to keep the framework simple but we should point to the fact that wealth would enter the participation decision should we assume agents are risk-averse.

The expected payoff to H is ex-ante higher since we have assumed that $p_H = \delta p_L$, with $\delta > 1$. However her reservation utility is identical to that of an

L-type person with equal wealth:

$$u_H(k,y) = 2k + y = u \tag{5}$$

There are s jobs in the market place that women wishing to enter employment have to compete for. We normalise s to lie between 0 and 1 so that s can also be viewed as a demand for labour rate[3]. Then the success probabilities of H and L must satisfy the following equality:

$$\lambda p_H + (1-\lambda)p_L = s \tag{6}$$

assuming both H and L participate, while $\lambda p_H = s$ if only H do.

Now, since $p_H = \delta p_L$,

$$[1-\lambda+\lambda\delta]\ p_L = s$$

so that

$$p_L = \frac{s}{[1-\lambda+\lambda\delta]} \tag{7}$$

which has a very simple explanation. That is, the success probability of an L-type person increases with the number of job offers, and falls as the fraction λ of high ability individuals rises. Likewise,

$$p_H = \frac{\delta\ s}{[1-\lambda+\lambda\delta]} \tag{8}$$

To derive the labour supply schedule in this simple framework we find the values of s, s^*_H and s^*_L that induce H and L to look for jobs. For an L person with wealth y, s^*_L is the value of the admission rate that leaves her indifferent between participating or not. This is easily deduced by taking

[3]or simply proxy some macroeconomic activity indicator, e.g tending to 1 in booms and collapsing to 0 in recessions.

$$w\frac{s^*_L}{[1-\lambda+\lambda\delta]} + k\left\{1 - \frac{s^*_L}{[1-\lambda+\lambda\delta]}\right\} = 2k$$

so that

$$s^*_L = \frac{k}{w-k}\,[1-\lambda+\lambda\delta] \tag{9}$$

while for a high ability individual with identical wealth we have to reason as follows. To find the reservation employment level for the high ability population means that we cannot be in the case $s \geq s^*_L$ anymore as we were up to the present moment since no H are willing to participate at such a level. Thus when $s < s^*_L$, only H participate and we need to specify the lowest value for s that would induce their participation.

For H to participate, equations (4) and (5) indicate that p_H should be at least equal to

$$(p_H \mid s < s^*_L) = \frac{k}{w-k}$$

and if $s < s^*_L$

$$s < \frac{k}{w-k}\,(1\text{-}\lambda+\lambda\delta) \text{ from (9).}$$

Thus,

$$(p_H \mid s < s^*_L) = \min\left(1, \frac{s}{\lambda}\right) \quad \text{so that}$$

$$\frac{k}{w-k} = \min\left(1, \frac{s}{\lambda}\right).$$

(i.e. if $s \geq \lambda$, $p_H = 1$ while if $s < \lambda$, $p_H = \frac{s}{\lambda}$.)

Therefore the lowest value of s which induces participation by H is such that $\frac{s^*_H}{\lambda} = \frac{k}{w-k}$

i.e. $\quad s^*_H = \frac{\lambda k}{w-k}$

We can therefore summarize the discussion of this section in the following table. From this table it is clear that there is one more equilibrium condition as is usually the case in Spence type models.

s \ p	p_H	p_L
$s \geq s^*_L$	$\frac{\delta s}{1-\lambda+\lambda\delta}$	$\frac{s}{1-\lambda+\lambda\delta}$
$s^*_H \leq s < s^*_L$	min $(1, \frac{s}{\lambda})^+$	L does not participate.
$s < s_H^*$	H and L do	not participate.

+: $\lambda p_H \leq s$. If $\lambda p_H = s \Rightarrow p_H = \frac{s}{\lambda}$.

Table 2

Since H has a higher success probability than L, it follows that she is willing to accept a job at a lower reservation rate than L, thus $s^*_H < s^*_L$. If the relative gain $(w-k)$ from having a job rises, then at any value of s, the supply of labour rises, i.e. s^*_H and s^*_L fall. Finally a rise in the fraction of high ability individuals, λ, increases the values s^*_H and s^*_L since it makes competition "tighter". We draw the labour supply schedule in Figure 2. When everyone enjoys equal opportunities in education, the fraction of individuals with education will be non-decreasing in the level of labour demand (or level of macroeconomic activity) s.

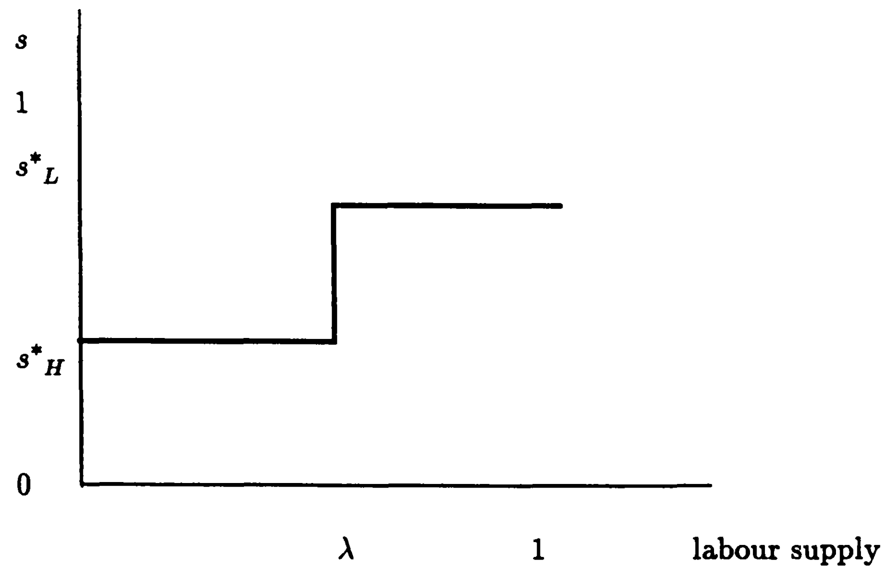

Figure 2

The equilibrium employment schedule for this economy will satisfy the 'min' condition: Employment = min {supply, demand} = e, viz.

$$e = \begin{cases} 0 \text{ if } s < s^*_H \\ \lambda \text{ if } \lambda \leq s \leq s^*_L \\ s \text{ otherwise} \end{cases} \tag{11}$$

The fraction of individuals who find jobs as a function of the level of macroeconomic activity s is drawn in Figure 3.

In an economy such as this with two levels of ability, no one will choose to enter employment until the level of macroeconomic activity reaches a minimum value, here s^*_H. When the demand for labour reaches level λ, the entire high ability population will get jobs. However until the activity rate rises to s^*_L, there will not be any new labour force participation by women – assuming this is consistent with random selection of workers –.

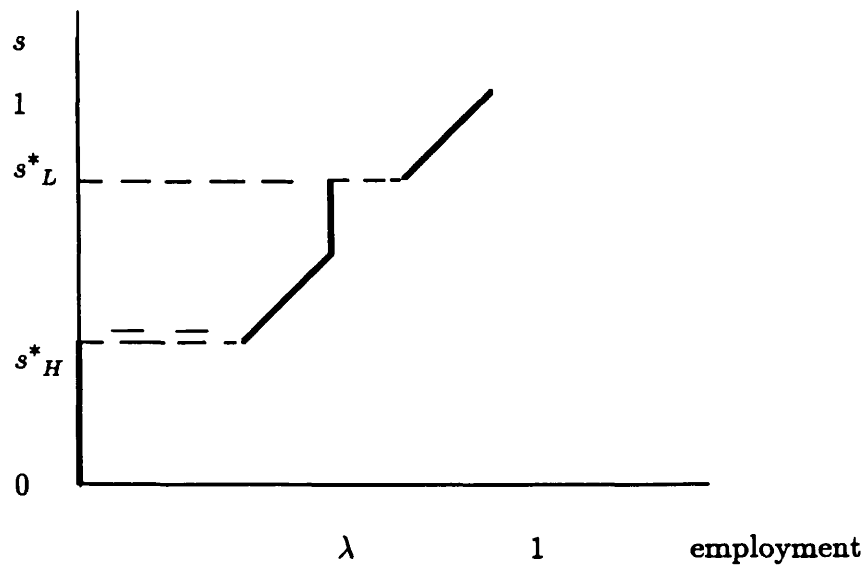

Figure 3

As can be seen from equations (9) and (10), in an economy where there are equal education opportunities for everyone, the distribution of wealth will have no role to play in the decision to take up a job. This should not be seen as a surprising result: if individuals can not increase their success probabilities by acquiring the skills or even generate those signals to the employers that are important in order to have any luck in the job market, then their decisions to compete for jobs will most probably only mirror their innate abilities. This will always be the case unless their alternative options to the workplace differ.

4.3 A role for assets in the participation decision.

We have established in the model above that in the case where acquiring education was irrespective of being rich or poor, wealth holdings did not directly enter the consideration to take up a job provided ability and wealth were not correlated. However if we postulate that women can be allowed to purchase private education which would increase their probability of being offered a job, then high and low ability agents can increase their respective success probabilities to:

$$p_{H\theta} = p_H + \theta \quad \text{and} \quad p_{L\theta} = p_L + \theta \tag{12}$$

The assumption that θ enters additively (rather than multiplicatively) in the success probability is used to simplify the model below, so as to let the benefit of being skilled constant across the population.

If we were to adapt the Grossman and Stiglitz[4] (1976) framework to the present one, then the demand for labour would not be publicly observed and θ would contain information about the random variable s. In our case it is usually the case that many firms make no secret about their decisions to expand their places such that they advertise their vacancies and universities their lectureship posts etc. However, it is often also the case that we observe internal appointments which find their way through arrangements among people working for the same firm. In the latter case the Grossman – Stiglitz framework may be applicable. However we can offer suggestions for alternative interpretations for the variable θ.

[4]Their model addresses the efficiency issue that has occupied a vast amount of research in the workings of the financial markets.

The simplest explanation for the existence of θ is that it summarizes an acquired level of skills which increases the probability of being offered a job. In an empirical context this could be made a function of relevant on the job characteristics or to be allowed to vary among high and low ability individuals according to availability of data on different types of jobs. In this context θ may also appear to capture skills and training of superior quality acquired by having already worked for some well reputed establishment in the past.

The second interpretation for θ, (rather well suited in the case for the distribution of public goods that are not subject to competition) would be to view it as a bribe, or the power to lobby. If rich people tend to be the first to have telephone lines and prosperous businessmen are last to pay their telephone bills, it could be that they can do so because they can bribe the appropriate supplier or government official. The size of the bribe would act as a proxy for the price in a Walrasian market: as the supply of the public good becomes less abundant, or as demand increases, the size of the bribe would have to rise[5].
Let the cost of acquiring a level of skills θ be $c(\theta)$. We assume that there are decreasing returns to scale in the production of skills, so that the first and second derivatives of the cost function with respect to θ are both positive, $c(0)=0$, $c(1)\to\infty$ and θ is defined over the range $[0,\ 1-p_H]$[6]. Consider the net benefit of purchasing a level of skills θ, $B_i(\theta)$ to an individual of type i; $i = L,H$:

$$\begin{aligned} B_i(\theta) = (p_i+\theta)(w+y)+(1-p_i-\theta)(k+y)-c(\theta)-p_i(w+y) \\ -(1-p_i)(k+y)\ ,\ i=L,H \end{aligned} \tag{13}$$

or

---

[5] θ could also proxy contacts, e.g. the son of a prime minister is likely to be offered an attractive job in order to please his father!

[6] Clearly for $p_{H\theta}$ to be a well defined probability $-p_H \leq \theta \leq 1-p_H$. However, if $\theta \in [-p_H,\ 0)$ the woman would rather do without the purchase of skill.

$$B_i(\theta) = \theta(w-k) - c(\theta), \quad i = L,H \tag{14}$$

Thus the quantity of skills θ^* produced in this economy is defined via the first order condition (dropping the subscript):

$$B'(\theta^*) = (w-k) - c'(\theta^*) = 0 \ , \quad i = L,H \tag{15}$$

This is illustrated in Figure 4.

Differentiating (15) one finds

$$\frac{d\theta}{dw} = \frac{1}{c''(\theta)} > 0 \tag{16}$$

so that in an economy where skills are produced subject to decreasing returns (c'' > 0), when the relative benefit $w-k$ of getting a job increases, the quantity θ^* of skills produced will also increase. Therefore, a low wage economy (i.e. one with a low value of $w-k$) might not have a sector which produces skills.

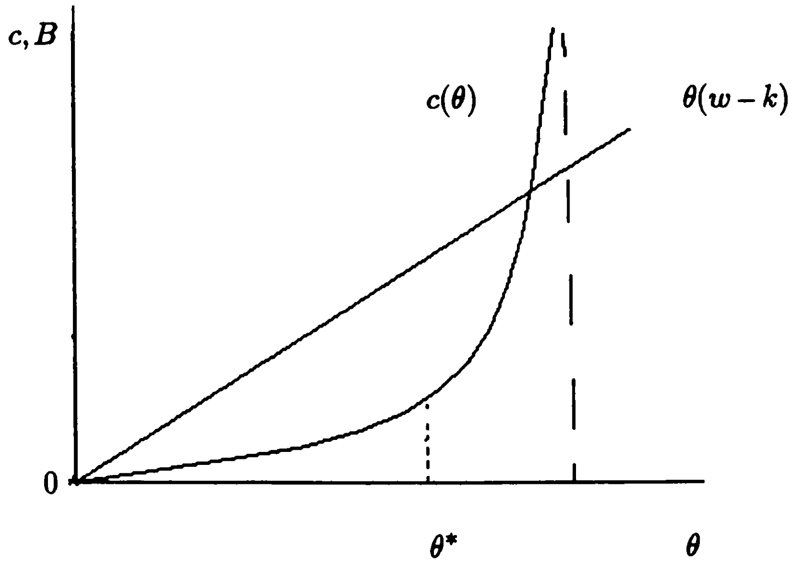

Figure 4

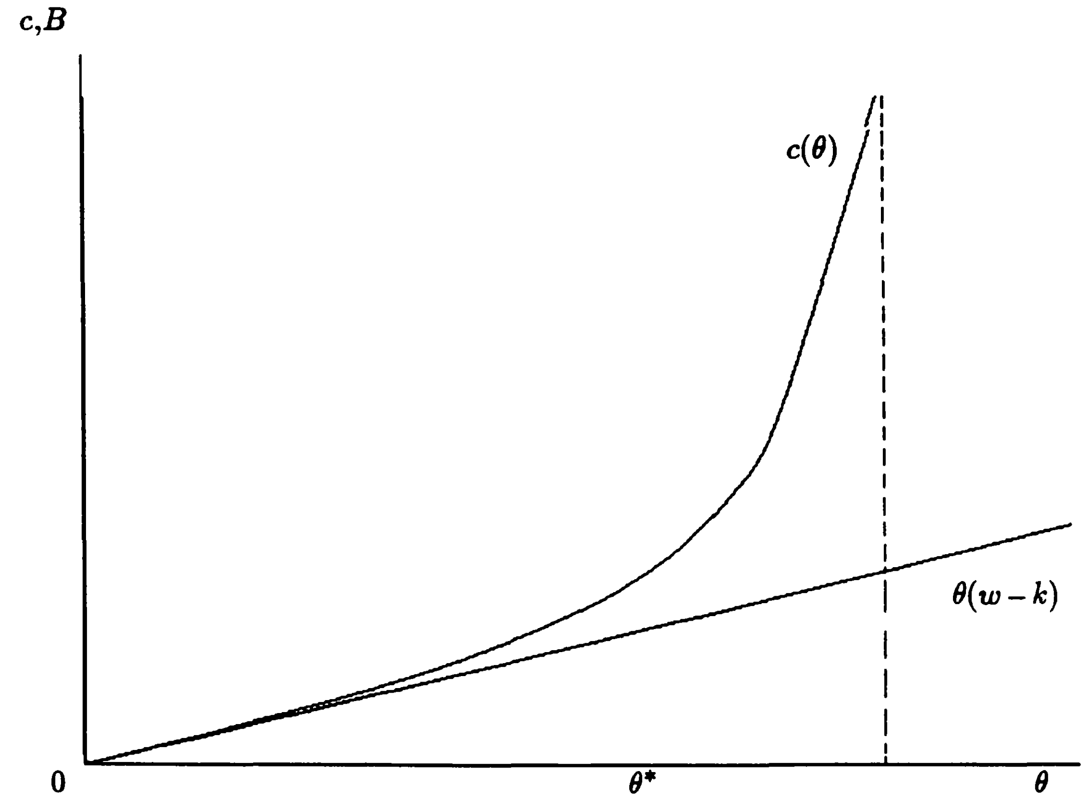

Figure 5

The low wage economy has no skill sector.

For such an economy equation (14) has a corner solution at θ=0. It is equivalent to the economy of the previous section, where all individuals had access to the same level of public information. To give an example consider the following cost function:

$$c(\theta) = \frac{1}{(1 - p_H - \theta)^{\beta}} \quad , \ \beta > 0, \qquad 0 \leq \theta \leq 1 - p_H$$

then from (13) we have that

$$\theta^* = 1 - p_H - \left\{\frac{\beta}{(w-k)}\right\}^{\frac{1}{\beta+1}}, \ \beta > 0$$

θ^* is increasing in $w-k$, $\theta^* \rightarrow (1-p_H)$ as $w-k$ tends to infinity and $\theta^* \rightarrow -p_H$ as $w-k$ approaches β[7].

As the gains from participation increase, the demand for skills, and thus the cost of purchasing skills (e.g. the costs of joining a training programme if they are not subsidised to do so) both rise. The ability of certain individuals to acquire skills will then reflect itself on the success probabilities, and thus on their participation decisions. The distribution of wealth will then have a role to play in deriving women's employment decisions. We can summarise the above discussion in the following:

Proposition 1:

In an economy experiencing decreasing returns to scale in the production of skills, $\lim_{\theta \to 0} c'(\theta) > w-k$, that is when the wage differential between the two sectors, $w-k$ is very low, $\theta = 0$ and therefore there is no skill producing sector.

From (1), the proportion of individuals who can purchase the required skills is

$$1 - F(c(\theta^*)) = \frac{b - c(\theta^*)}{b-a} \tag{17}$$

When wealth and ability are independently distributed, the success probabilities are defined via the following equality:

[7]In the latter case p_H would also have to be zero for it to be valid. However, as we saw in the previous footnote this is a possibility that would not be entertained.

$$(1-\lambda)\left\{ F(c(\theta))\, p_L + [1-F(c(\theta))][p_L+\theta]\right\} +$$
$$\lambda\left\{ F(c(\theta))\, p_H + [1-F(c(\theta))][p_H+\theta]\right\} = s \tag{18}$$

The LHS of (18) consists of the following terms: the low ability population $(1-\lambda)$ from which a fraction $(1-F(.))$ purchases skills and has a success probability $p_L+\theta$, the remaining $F(.)$ percent having a probability to succeed p_L. The λ high ability individuals are divided into two similar subgroups (in similar proportions): the skilled individuals have a success probability $p_H+\theta$, while the fraction $F(.)$ of unskilled individuals have a success probability p_H.

Solving (18), we get that

$$p_L = \frac{s-\theta[1-F(c(\theta))]}{[1-\lambda+\lambda\delta]} \tag{19}$$

Thus an L-type woman with wealth below $c(\theta^*)$ sees her chance of being employed reduced by an amount $\frac{\theta[1-F(c(\theta))]}{[1-\lambda+\lambda\delta]}$. As would be intuitively plausible her chance of finding employment would diminish by the fraction of those women who can purchase skills, $(1-F(.))$ as well as by those possessing a higher ability, λ. However, note, that a change in θ (induced by, say, a change in w) has an ambiguous effect on p_L. An increase in θ widens the gap between the opportunities open for skilled and those open for unskilled women to get a job. However, since an increase in θ also implies that fewer people will be able to purchase skills, it does not always follow that the success probability of an unskilled individual will fall.

From (19), the success probability of an H-type woman with funds below $c(\theta^*)$ is

$$p_H = \frac{\delta\left\{s - \theta\,[1 - F(c(\theta))]\right\}}{[1 - \lambda + \lambda\delta]} \tag{20}$$

Likewise from (10),(17) and (18), for low and high ability women who purchase skills, we have the following two success probabilities:

$$p_{L\theta} = p_L + \theta = \frac{s - \theta[1 - F(c(\theta))]}{[1 - \lambda + \lambda\delta]} + \theta \tag{21}$$

$$p_{H\theta} = p_H + \theta = \frac{\delta\left\{s - \theta[1 - F(c(\theta))]\right\}}{[1 - \lambda + \lambda\delta]} + \theta \tag{22}$$

Differentiating (21) and (22) with respect to θ, we conclude that an increase in the level of skills leads to an increase in the success probabilities of skilled individuals.

Now for use in what shall follow we denote by LS the low and by HS the high ability individuals who can purchase skills, while by LU and HU their unskilled counterparts. To construct the participation schedule, we need to derive the reservation values of the level of economic activity s, for LS, HS, LU and HU. Therefore the value s^*_{LU} solves:

$$\frac{s^*_{LU} - \theta[1 - F(c(\theta))]}{[1 - \lambda + \lambda\delta]}\,w + \left\{1 - \frac{s^*_{LU} - \theta[1 - F(c(\theta))]}{[1 - \lambda + \lambda\delta]}\right\}k = 2k \tag{23}$$

thus

$$s^*_{LU} = \frac{k}{(w - k)}\,[1 - \lambda + \lambda\delta] + \theta[1 - F(c(\theta))] \tag{24}$$

Compared to the model of the previous section, we can now see that the reservation value for a low ability person is increased by $\theta[1 - F(c(\theta))]$. In addition through

$$\frac{\partial s^*_{LU}}{\partial \theta} = [1 - F(c(\theta))] - \frac{\theta}{b-a}\frac{dc}{d\theta} \gtrless 0$$

we see that an increase in the skill level has an ambiguous effect on the reservation value a low ability unskilled woman has for entering the market. Below, we shall see that a rise in θ can have an *un*ambiguous negative effect on s^*_{HU} given that a certain condition is met. The reservation employment rates for LS & HU vary according to whether $w - k \gtrless w_{RD}$. We therefore first derive HU's reservation employment level assuming that after LU, LS exit first and then LS's given that HU are those who exit the job market first.

$\boxed{\text{Case 1: LS exit first}}$ (i.e. groups LU and LS have exited)

In this case (18) becomes

$$\lambda \left\{ F(c(\theta))\, p_H + [1 - F(c(\theta))][p_H + \theta] \right\} = s$$

so that

$$p_H = \frac{s}{\lambda} - \theta\left(1 - F\big(c(\theta)\big)\right)$$

To find s^*_{HU} we solve

$$\left[\frac{s^*_{HU}}{\lambda} - \theta\,(1 - F(c(\theta)))\right](w-k) = k \;\Rightarrow$$

$$s^*_{HU} = \frac{\lambda k}{w-k} + \lambda\theta\,(1 - F(c(\theta))) \tag{25}$$

It is also evident from (25) that

$$\frac{\partial s^*_{HU}}{\partial \theta} = \lambda\left[1 - F(c(\theta)) - \frac{\theta}{(b-a)}\frac{dc}{d\theta}\right].$$

Rearranging this expression we can observe that acquiring skill level θ can cause a reduction to s^*_{HU} iff $\frac{b-c(\theta)}{\theta} < \frac{\partial c}{\partial \theta}$. The above are summarised in the following proposition:

Proposition 2:

(i) A low ability unskilled woman's reservation employment, s^*_{LU}, is higher to that of simply a low ability person's , s^*_L, by the total level of skills acquired by those who can obtain them.

(ii) A rise in the skill level will have an ambiguous effect on the employment reservation values of low ability unskilled workers with the overall effect depending on whether the proportion of people who can acquire skills dominates over the "density of skills" required.

(iii) A rise in the skill level can have a negative effect on the employment reservation value of high ability unskilled workers if and only if the marginal cost of acquiring skills strictly exceeds the maximal remaining wealth per unit of skill.

Case 2: HU exit first (i.e. groups LU and HU have exited)

Again (18) becomes:

$$(1-\lambda)\,[1-F(c(\theta))][p_L+\theta] + \lambda\,[1-F(c(\theta))][p_H+\theta] = s$$

so that

$$p_L = \frac{s-\theta(1-F)}{(1-F)[1-\lambda(1-\delta)]}$$

and using that $p_{L\theta} = p_L + \theta$ and the fact that this group needs to consider also their cost of purchasing skills, c(θ),

$$\left(\frac{s^*_{LS}}{\lambda} + \theta F\big(c(\theta)\big)\right)(w-k) = k + c(\theta) \;\Rightarrow$$

$$s^*_{LS} = \left[1 - F\big(c(\theta)\big)\right]\left[\frac{[k + c(\theta)]\,[1 - \lambda(1-\delta)]}{w - k} + \lambda\theta(1-\delta)\right] \tag{26}$$

The last group to exit is HS. In this case (18) becomes

$$\lambda\ [1 - F(c(\theta))][p_H + \theta] = s \ \Rightarrow$$

$$p_H = \frac{s}{\lambda(1 - F(c(\theta)))} - \theta \ \text{, so that}$$

$$s^*_{HS} = \frac{\lambda\left[1 - F\big(c(\theta)\big)\right][k + c(\theta)]}{w - k} \tag{27}$$

and to show that $s^*_{LS} < s^*_{LU}$, we can express s^*_{LS} as

$$s^*_{LS} = s^*_{LU} - \ [1 - \lambda(1-\delta)]\,\frac{\left[\theta\big(1 - F(.)\big)\,(w - k) - c(\theta) + F(.)\ (\ k\ + c(\theta))\right]}{w - k} \tag{28}$$

where $F(.) \equiv F\big(c(\theta)\big)$. Because the benefit received by acquiring skills $B(\theta)$ = $\theta(w - k) - c(\theta)$ is never negative, it follows that the second term on the R.H.S. of (28) is positive so that indeed $s^*_{LS} < s^*_{LU}$.

Expressing also s^*_{HS} as a function of s^*_{HU}:

$$s^*_{HS} = s^*_{HU} - \lambda\left[\frac{\theta(\ 1 - F(.))(w - k) - c(\theta) - F(.)(k + c(\theta))}{(w - k)}\right] \tag{29}$$

and comparing equations (28) and (29) it is almost immediate to see that $(s^*_{HU} - s^*_{HS}) < (s^*_{LU} - s^*_{LS})$. From equations (25) and (26) we have that:

$$s^*_{HU} - s^*_{LS} = \frac{\lambda k - (1 - F(.))\left[\big(k + c(\theta)\big)\,[1 - \lambda(1-\delta)] - \lambda\theta\delta(w - k)\right]}{w - k} \gtrless 0 \tag{30}$$

Equation (30) tells us that the difference in the reservation admission rates between HU and LS can be either positive or negative. In the model of the previous section as well as in the case of the low wage economy, s^*_{HU} is always smaller than s^*_{LS} (since $\theta = 0$). Note that $s^*_{HU} - s^*_{LS}$ is likely to be widened by an increase in δ. In addition,

$$\frac{\partial(s^*_{HU} - s^*_{LS})}{\partial(w-k)} = \frac{\left(1-F(.)\right)\left(k+c(\theta)\right)[1-\lambda+\lambda\delta]-\lambda k}{(w-k)^2} \tag{31}$$

meaning that the gap between s^*_{HU} and s^*_{LS} can inrease if there is a fairly large proportion of individuals who cannot purchase the required skills[8]. Let k comprise of all those benefits that accrue to them by staying unemployed (shadow price of leisure together with unemployment compensations, say). Further denote by w_{RD}, that wage reservation differential which equates s^*_{HU} to s^*_{LS} (see equation (30) above). The labour force participation decision will thus be a function of the wage rate but we shall sketch in Figure 6 three cases of interest. To understand the construction of the graphs the essential conditions one should have in mind as discussed in the preceeding paragraphs are summarised in the following proposition:

Proposition 3:

(i) For women possessing identical skills (i.e both U or both S)

$s^*_{Hi} < s^*_{Li}$, $i = U,S$.

(ii) $s^*_{LS} \gtrless {}^*_{HU}$ *as* $w-k \lessgtr w_{RD}$.

(iii) $(s^*_{HU} - s^*_{HS}) < (s^*_{LU} - s^*_{LS})$.

(iv) $s^*_{LS} < s^*_{LU}$.

[8] Rearranging (31) we can see that the condition for it to be positive is that $F < \frac{c+\epsilon}{c+k}$, – where ϵ is slightly smaller than $k - \left(\epsilon = k - \frac{k\lambda}{[1-\lambda(1-\delta)]}\right)$.

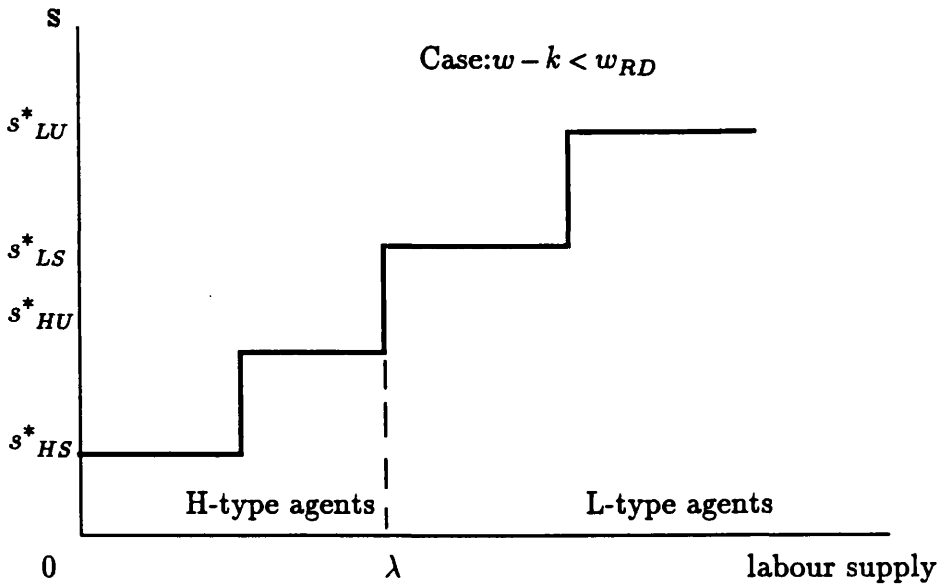
s
Case:$w - k < w_{RD}$
s^*_{LU}
s^*_{LS}
s^*_{HU}
s^*_{HS}
H-type agents
L-type agents
0
λ
labour supply

Figure 6a

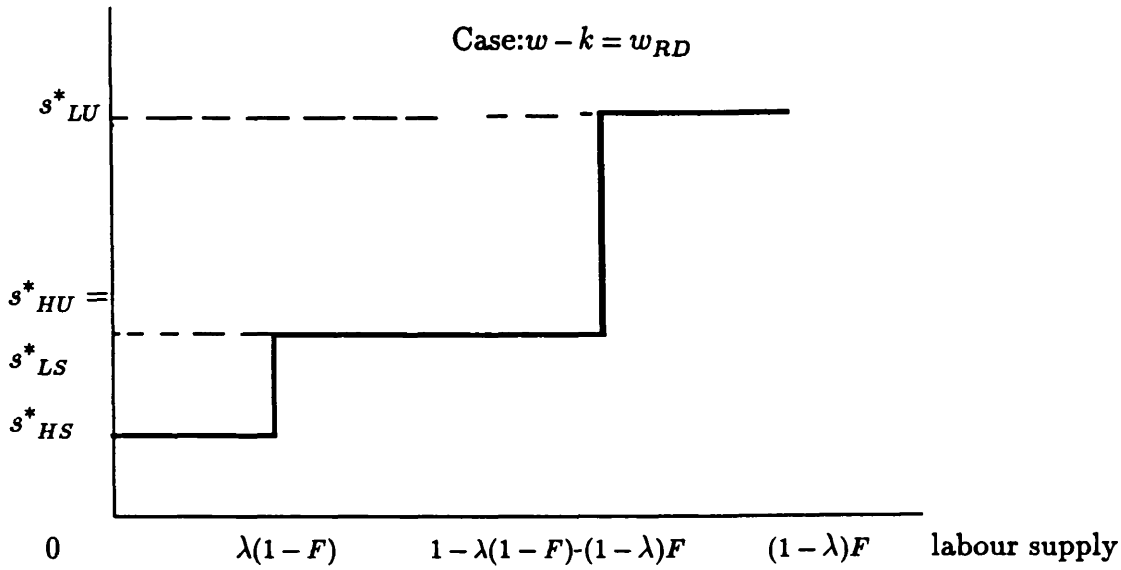
Case:$w - k = w_{RD}$
s^*_{LU}
s^*_{HU} =
s^*_{LS}
s^*_{HS}
0
$\lambda(1-F)$
$1-\lambda(1-F)$-$(1-\lambda)F$
$(1-\lambda)F$
labour supply

***Figure* 6b**

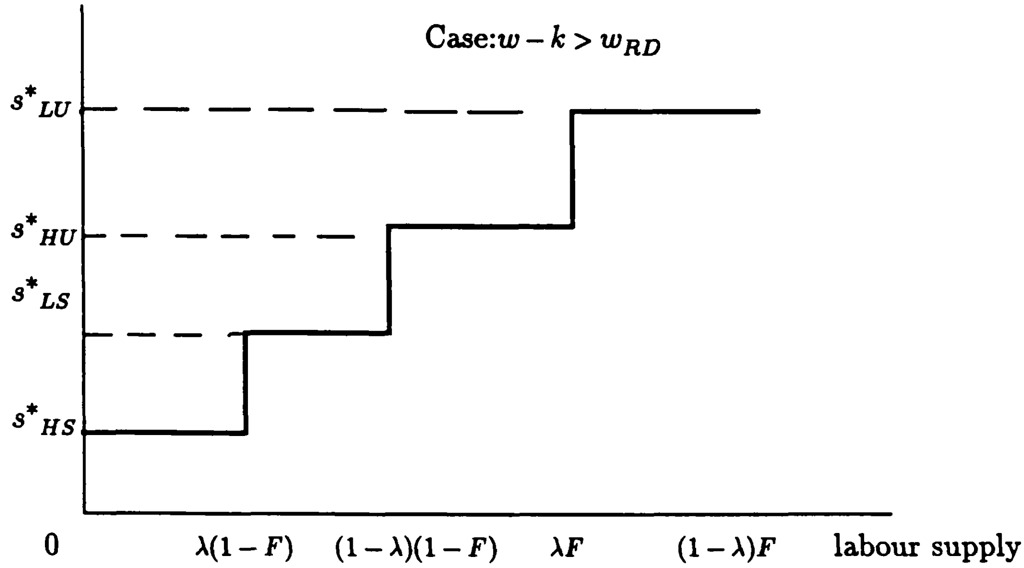

Figure 6c

Let us look first at part (a); Case : $w - k < w_{RD}$. Thus up to a fraction of the λ, high ability women, only their proportion of skilled workers would wish to participate. The proportion of unskilled women would then wish to enter the market at a higher level s^*_{HU} once the first group has been exhausted (till λ). The step function reaches its maximum when the level s^*_{LU} is reached where everyone would be willing to participate. Part (b) is the case where the wage is just sufficient to persuade the HU and LS populations to have identical reservation values and thus we have one less step in the function while (c) could be seen as a more realistic case in a society where ability and skills are each rewarded on their merits and ability takes precedence over skills for women possessing the <u>same</u> level of skills. However for different skill levels this is not necessarily the case (e.g Proposition 3(ii) for Case: $w - k > w_{RD}$; also see Fig 6(c)).

4.4 Conclusions.

The aim of the analysis in this chapter has been to present a simple framework whereby women choose to participate in the labour market taking into account the possibility that their abilities to succeed in this task can be related to their level of education. We sketched a way that their probabilities to receive one job offer can be in fact dependent upon the economic environment, the proportion of women who possess higher abilities and the extent to which they have higher chance of being employed in comparison to their less endowed – in terms of ability – fellow seekers.

We saw in Section 2 that when our framework allows for equal opportunities in education then the distribution of wealth does not enter the decision to accept a job. This decision, ceteris paribus, thus may only be reflecting women's different abilities. In Section 3 however assets are allowed to enter this choice.

Chapter 5

Conclusions and Ideas for Further Research.

5.1 Conclusions.

As suggested by the empirical evidence from a variety of studies it is not an easy undertaking to draw definite conclusions about the effects of schooling and training on labour market outcomes. The studies that we have considered are quite heterogeneous in sources and methods and the differences may significantly affect the findings. However, it is possible to draw some main conclusions from their findings and ours.

Firstly, we saw in the section on the evaluation of training and experimental data issues that among other findings it seems that the main benefit recipients of those programmes are their female participants more than their male counterparts. Also, that minority males experience some earnings gains while there seem no benefits that can be found for white males in these studies. White women also seem to benefit comparatively more than minority women from programme participation.

Next we considered some issues pertaining to discrete choice problems and their connection to our theme of schooling and labour market outcomes. In the models we encountered, individual behaviour was adequately captured in the sense that a finite set of decision rules is more easily comprehensible and simpler to analyse from a policy perspective. Wolpin's work on the transition from school to work was drawn as an illustration for the way the standard search model was extended to encompass such decision.

Next we illustrated further by a model of labour market transitions which was focusing on movements between the states of unemployment and employment. The purpose had been to incorporate among other things, education dummies to describe the effect College and High School education have on the employment behaviour of a sample of married women from CES. State dependence was present and quite strong. We saw that College education is important over and above that from High School in terms of the probability of entering employment from unemployment, both considering as well as conditioning out the supply side effects.

The last Chapter focussed on that aspect of the model of transitions on which emphasis is centred on the possibility that the arrival rates that women receive are significantly determined by their level of education. We saw that it had been possible to design a simple framework for women's labour force participation in which their innate abilities to succed in the market place are a function of the economic environment. Further, they also depended on the proportion of women who possess higher abilities and the extent to which they have higher chance of being employed in comparison to their less endowed – ability wise – fellow seekers.

When people enjoy equal opportunities in education we saw in Chapter 4 that the distribution of wealth does not enter the decision to take up a job and this decision, ceteris paribus, may simply only reflect women's different abilities. Even though as we stressed we feel this result would depend upon the assumption of risk-neutrality a further section allows assets to enter the decision to participate.

5.2 Ideas.

In terms of carrying the research beyond the lines that this thesis has suggested, we can offer – at least – three possible routes that ourselves may further investigate.

The first one concerns the nesting of the model of transitions within another model. More specifically, since the two unconditional accession and retention probabilities depend on a "common" supply side we can instead of eliminating it when we perform a division, introduce another function of preferably alternative functional form that may contain extra parameters with the aim to achieve at least functional form identification. Then as long as for some values of the parameters (preferably zeros) the new function becomes constant (one, say) we can test the validity of these values. Let the parameters be $\pi_1, \ldots, \pi_k$. If $\pi_1 = \pi_2 = \ldots = \pi_k = 0$ then the "new" model would nest the model of Chapter 3. Then one could apply GMM to the new model to get $\hat{\pi}_1, \ldots, \hat{\pi}_k$, say, and would test the hypothesis through a Wald test or likelihood ratio test.

The second suggestion concerns the endogenisation of the arrival of offers probabilities. In other words, it may be possible to impose credit market imperfection assumptions and adverse selection issues in terms of costs of effort experienced between the high and low ability populations which would achieve this aim.

Finally, the incorporation of a continuous explanatory variable – such as earnings which, however, was not available to us in CES – into the model of transitions would give us the opportunity to link it with some investigations of the efficiency of sub-optimally weighted least squares estimators in models of heteroscedasticity of unknown but ordered according to size form. The latter would allow us to bring the model of transitions to bear with some interesting new developments on an intriguing aspect of heteroscedasticity.

References and Bibliography.

Abowd, John M. and Card, David (1989): "On the Covariance Structure of Earnings and Hours Changes.", *Econometrica*, 57, 411-445.

Abul Naga R. (1992): *"Poverty and the Economics of Information"*, mimeo, LSE.

Amemiya, Takeshi (1985): *"Advanced Econometrics."*, Basil Blackwell, Oxford.

Arellano, Manuel and Meghir Costas (1992): "Female Labour Supply and On-the-Job Search: An Empirical Model Estimated Using Complementary Data Sets, *Review of Economic Studies*, 59, 537-557.

Arrow, Kenneth J. and Kurz, Mordecai (1970): *"Public Investment, the Rate of Return and Optimal Fiscal Policy."*, Ch.2, Resources for the Future, Inc., the Johns Hopkins Press, Baltimore, Maryland.

Ashenfelter, Orley (1978): "Estimating the Effect of Training Programs on Earnings.", *The Review of Economics and Statistics*, 60, 47-57.

_ (1983): "Determining Participation in Income-Tested Social Programs.", *Journal of the American Statistical Association*, 78, 517-525.

Ashenfelter, Orley and Card, David (1985): "Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs.", *The Review of Economics and Statistics*, 67, 648-660.

Attanasio, Orazio P., Koujianou Penny and Weber Guglielmo (1989): "Consumer Expenditure Survey Database.", mimeo, The Center for Economic Policy Research, Stanford University.

Barnow, Burt (1987): "The Impact of CETA Programs on Earnings.", *Journal of Human Resources*, 22, 157-193.

Bassi, Laurie J. (1983a): "Estimating the Effect of Training Programs with NonRandom Selection.", unpublished Ph.D. dissertation, January, Department of Economics, Princeton University.

_ (1983b): "The Effect of CETA on the Postprogram Earnings of Participants.", *Journal of Human Resources*, Fall, 539-556.

_ (1984): "Estimating the Effect of Training Programs with Non-Random Selection.",*The Review of Economics and Statistics,* 66, 36-43.

Bellman, Richard E. (1957): *Dynamic Programming.* Princeton, New Jersey: Princeton University Press.

Berkovec, James and Stern, Steven (1991): "Job Exit Behavior of Older Men.", *Econometrica*, 59, 189-210.

Bertsekas, Dimitri P. (1976): *Dynamic Programming and Stochastic Control.* Academic Press Inc., New York.

Bjorklund, Anders (1989): "Evaluations of Training Programs: - experiences and proposals for future research.", mimeo, Swedish Institute for Social Research.

Blundell Richard W. (1990): "*Evaluating Structural microeconometric models of labour supply* ", invited lecture, Sixth World Congress of the Econometric Society, Barcelona, August, mimeo, UCL.

Blundell Richard W., Magnac Thierry and Meghir Costas (1992): *"Savings and Labour Market Transitions"*, mimeo, UCL.

Blundell Richard W. and Walker Ian (eds.) (1986): *"Unemployment, search and labour supply"*, Cambridge University Press.

Burdett, Kenneth N., Kiefer, Nicholas M., Mortensen, Dale T. and Neumann, George R.(1984): "Earnings, Unemployment and the Allocation of Time Over Time.",*Review of Economic Studies*, 51, 559-578.

Card, David and Sullivan, Daniel (1988): "Measuring the Effect of Subsidized Training Programs on Movements in and out of Employment.", *Econometrica*, 56, 497-530.

Constantinides George (1990): "Habit Formation: A Resolution of the Equity Premium Puzzle, *Journal of Political Economy,* 98-3, 519-543.

Devine, Theresa J. and Kiefer, Nicholas M. (1991): "*Empirical Labor Economics: The Search Approach.*", Oxford University Press.

Eckstein, Zvi and Wolpin, Kenneth I. (1989a): "The Specification and Estimation of Dynamic Stochastic Discrete Choice Models.", *Journal of Human Resources*, 24, 562-598.

Eckstein, Zvi and Wolpin, Kenneth I. (1989b): "Dynamic Labour Force Participation of Married Women and Endogenous Work Experience.", *Review of Economic Studies*, 56, 375-390.

Gönül, Füsun (1989): "Dynamic Labor Force Participation Decisions of Males in the Presence of Layoffs and Uncertain Job Offers", *Journal of Human Resources*, 24, 195-220.

Grossman, Sanford J. and Stiglitz, Joseph (1976): "Information and competitive price systems.", *American Economic Review*, 66, 246-253.

Ham, John C. and LaLonde, Robert J. (1990): "Using Social Experiments to Estimate the Effect of Training on Transition Rates.", In J. Hartog, G. Ridder and J. Theeuwes (eds.) *Panel Data and Labor Market Studies.* Elsevier Science Publishers B.V. (North-Holland).

Hausman, Jerry A. and Taylor, W. E. (1981): "Panel Data and Unobservable Individual Effects.", *Econometrica*, 49, 1377-1398.

Heckman, James J. (1981): "Heterogeneity and State Dependence." In S. Rosen (ed.) *Studies in Labor Markets.* University of Chicago Press, Chicago, 91-140.

_ and Robb, Richard (1983): "Alternative Methods for Evaluating the Impact of Training on Earnings.", mimeo, University of Chicago.

_ and Willis, Robert J. (1977): "A Beta-Logistic Model for the Analysis of Sequential Labor Force Participation by Married Women.", *Journal of Political Economy*, 85, 27-58.

Kiefer, Nicholas M. (1979): "Population Heterogeneity and Inference from Panel Data on the Effects of Vocational Education.", *Journal of Political Economy*, 87, S213-S226.

_ (1988): "Economic Duration Data and Hazard Functions.", *Journal of Economic Literature*, 26, 646-679.

LaLonde, Robert J. (1986): "Evaluating the Econometric Evaluations of Training Programs with Experimental Data." *American Economic Review*, 76, 604-620.

Mortensen, Dale T. and Neumann George R. (1984): "Choice or Chance.", In

G.R.Neumann and N.C. Westergard-Nielsen (eds.) *Studies in Labor Market Dynamics*, Springer-Verlag, Heidelberg.

Pagan Adrian (1984): "Two Stage and Related Estimators and Their Applications.", *Review ofEconomic Studies,* LIII, 517-538.

Ridder, Geert (1986): "An Event History Approach to the Evaluation of Training,Recruitment and Employment Programmes.", *Journal of Applied Econometrics*, 1, 109-126.

Rust, John (1992): "Do people behave according to Bellman's Principle of Optimality?", Working Paper, The Hoover Institution.

Spence, M. (1973): "Job market signalling.", *Quarterly Journal of Economics*, 87, 355-373.

Stiglitz, Joseph E. (1975): "The theory of screening, education and the distribution of income.", *American Economic Review*, 65, 283-300.

Willis, Robert J. and Rosen, Sherwin (1979): "Education and Self-Selection.", *Journal of Political Economy*, 87, S7-S36.

Wolpin, Kenneth I. (1987): "Estimating a Structural Search Model: The Transition from School to Work.", *Econometrica*, 55, 801-19.