Methodology in Labour Economics;

A Review of the Literature

Guyonne Kalb*

Melbourne Institute of Applied Economic and Social Research

University of Melbourne

March 2002

Abstract

The main objective of this paper is to review the different types of methodologies used in the literature on work incentive effects associated with social security policies. Three broad groups of approaches are studied. First, hazard rate modelling as is used in the analysis of exit rates (or durations), such as unemployment duration, employment duration or duration of benefit receipt; second, models of labour supply and welfare participation; and third, (quasi-)experimental evaluations.

1 Introduction

Judging by the size of the literature on work incentive effects of the social security system and the policies associated with this system, the issue is one of great importance. This paper looks at the methodologies used in this literature. It is important to draw on international research, because although some research exists in Australia, it is relatively limited, whereas an extensive literature exists on this topic in the U.S., Canada, the UK and other Northern European countries.

However, given the arrival of new longitudinal data on labour market issues and welfare participation in Australia, such as the Longitudinal Data Survey (containing a sample from the administrative data on the receipt of social security payments) and in the near future with the Household, Income and Labour Dynamics in Australia (HILDA) survey, there are opportunities for Australian research in this area to catch up.

The literature in this field makes use of a wide variety of methodologies. They can mainly be classified into three groups, since work incentive effects are mostly studied by:

- Hazard rate modelling as is used in the analysis of durations, such as unemployment duration, employment duration or duration of benefit receipt. This type of modelling is, for example, used to examine the effect of education levels on the duration of unemployment.
- Models of labour supply and welfare participation, which estimate the relationship between preferences for leisure and income on the one hand and personal and household characteristics on the other hand. In combination with information on the financial situation of a household at the different levels of labour supply, the expected level of labour supply can be derived in different (hypothetical) situations; and
- (Quasi-)experimental evaluations which compares outcomes in comparable treatment and control groups (in (quasi-)experiments) to study the effects of policy changes.

Each approach has its own methodology. Over the years, the methods used in the three approaches have been refined and extended, allowing more features of the real world to

be incorporated in the models used. In this paper, the development of the methodology and the resulting improvement of the research is discussed.

Section 2 deals with duration or hazard rate modelling and Section 3 deals with labour supply and/or welfare participation models. Finally, Section 4 discusses evaluations when (quasi-) experimental data are available. Section 5 provides concluding comments.

2 Hazard rate modelling

The hazard rate out of unemployment, or transition rate or exit rate as it is alternatively called, represents the probability of exit from unemployment in the next period conditional on being in unemployment in the current time period. The expected duration is a function of the hazard rate, for example, a high hazard rate implies a short duration of the unemployment spell is expected. Mortensen (1977) laid the foundation for the extensive use of hazard rate models in labour market research. He developed a structural job search model, which formed the economic basis for the use of the reduced form hazard rate models. The structural job search model involves the estimation of reservation wage and market wage equation, whereas the reduced form hazard rate model only implicitly compares market wage and reservation wage in its one equation. However, underlying the probability of exiting at each point in time is the probability that the offered wage rate is higher than the reservation wage rate. Therefore, the equation to be estimated will contain explanatory variables that are related to the reservation wage and the market wage.

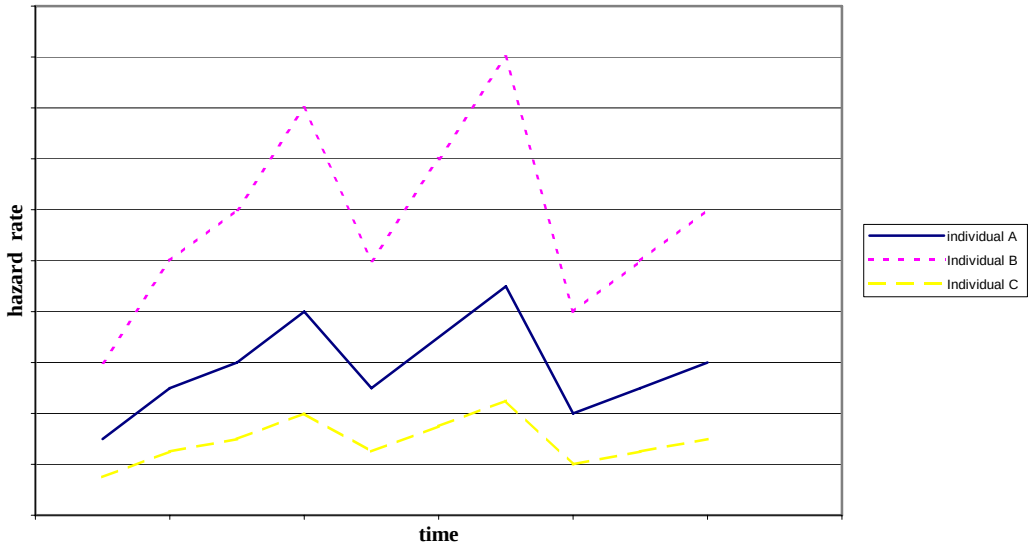
2.1 Proportional hazard function

The hazard rate models most commonly used involve proportional hazard functions, which means that the parts to represent the time dependence (the baseline hazard), and the dependence on individual characteristics are independent of each other. Such a specification imposes a pattern of exit rates through time, which is the same for everyone. Different individuals have hazard rates at different levels, but at each point in elapsed time the ratio of the exit rates of the individuals is the same. Figure 1 illustrates this. Individual A in figure 1 has a hazard rate which at all points in time is two times as high as the hazard rate for individual C and only half of the hazard rate for individual B.

2.2 Base line hazard

Several different specifications for the baseline hazard have been used. Currently the piecewise constant baseline hazard is popular because of its flexibility. Other specifications may impose constant baseline hazards, monotonically increasing or decreasing baseline hazards, or other fixed patterns. The piecewise constant approach, however, does not a priori impose a particular shape on the baseline hazard. The shape of the baseline hazard determines whether there is positive, negative or no duration dependence. Negative (positive) duration dependence means that the baseline hazard is downwards (upwards) sloping over time.

Figure 1 Proportional hazard rates



To estimate a piecewise constant baseline hazard, time intervals covering the range from 0 to the maximum duration observed are defined. The number and position of the time intervals can be chosen freely as long as at least one person is observed to exit in each time interval. For each interval a parameter is estimated, which indicates how likely exit is in that particular interval compared to exit in the other intervals. The advantage is that a researcher can start with a large number of intervals and reduce intervals by merging intervals, when parameters (indicating the level of the exit rate) for adjoining time intervals are not significantly different from each other.

Duration dependence can be made even more flexible by using an alternative specification, which allows the duration dependence to vary over the business cycle (Cockx and Dejemeppe, 2000). This means that under different economic

circumstances, hazard rates may depend in different ways on the elapsed time in a spell. They find that in a recession, men over 28 years old experience negative duration dependence, whereas under improved labour market conditions, positive duration dependence is found. This supports the idea that in a downturn employers rank potential employees by their unemployment duration, employing those who have been shortest unemployed. In an upturn this ranking procedure is less important. For younger men no duration dependence is found in good or bad labour market circumstances. This is possibly explained by the fact that education attainment is a more informative measure of a young person's productivity than unemployment duration.

2.3 Unobserved heterogeneity

Another important issue in the development of hazard rate models has been the way in which researchers deal with unobserved heterogeneity. It is unlikely that all differences between individuals are captured by the characteristics of a person that we observe in a data set. In such circumstances it is said that there is unobserved heterogeneity. If the researcher cannot control for the differences in exit rates that occur as a result of these unobserved differences between individuals, duration dependence will be biased downwards. In the early stages of spells, a group of individuals with mixed characteristics will be at risk of exiting. However, those individuals in the group with the most favourable characteristics will exit in the early stages of the duration, leaving behind those who are less likely to exit. As a result, it appears as if hazard rates later in the spell are lower than hazard rates earlier in the spell. However, this is not a true duration dependence effect. Instead it represents changes in the distribution of unobserved characteristics in the population yet to exit from unemployment.

In the early days of research in this field, many researchers did not control for unobserved heterogeneity. In the more recent literature, it is widely acknowledged that unobserved heterogeneity is an important issue and several methods for dealing with unobserved heterogeneity are available.

However, separating duration dependence and unobserved heterogeneity is difficult to achieve. Identifying the effect that each has on the hazard rate often requires assumptions on the functional form of either duration dependence or unobserved heterogeneity. Barrett (2000) estimates hazard rate models with two alternative specifications for the unobserved heterogeneity. The effects of the personal

characteristics are robust to this change in specification, but the duration dependence parameters alter markedly after the change: from indicating negative duration dependence to no or hardly perceivable duration dependence. A popular approach in the literature is to specify both duration dependence and unobserved heterogeneity nonparametrically, which means a flexible pattern can be estimated that is not restricted by the choice of a distributional form. The distribution for the unobserved heterogeneity is approximated by a discrete multinomial distribution, which allows for a discrete number of different values for each heterogeneity term (see Heckman and Singer, 1984). The number of different discrete values (or points of support as they are usually called) needed is determined in estimation. Baker and Melino (2000) find that the problem with this approach is that it tends to overestimate the dispersion of the unobserved heterogeneity distribution by adding too many points of support in the discrete multinomial distribution. As a result, there is a positive bias to the duration dependence and a bias away from zero on the coefficients of observed heterogeneity. Baker and Melino propose an alternative method, which penalizes specifications with many points of support when choosing between adding another point of support or stopping.

These problems in separating duration dependence and unobserved heterogeneity is unfortunate for policy makers, because there is a difference in the type of policies one would use to address the two different problems. If duration dependence is the cause for long term unemployed not exiting from unemployment then the key to preventing this is to prevent people becoming long term unemployed. On the other hand, if unobserved heterogeneity is the cause for low exit rates, then it is important to identify these unemployed people early in their unemployment and target them with appropriate policies. In that case it would be important to find out what is underlying this unobserved heterogeneity which distinguishes these long term unemployed people from other unemployed people.

A recent paper addresses this issue by looking at time series data on aggregate outflows from different duration classes (Van den Berg and Van Ours, 1996). This allows them to separate the two effects. In U.S. data, they find that except for white men, duration dependence is dominated by unobserved heterogeneity. Earlier work has discussed the value of panel data where repeated spells can be observed for one person. Assuming that unobserved heterogeneity does not change over time (but is person specific), the

longitudinal information for an individual provides a tool for separating unobserved heterogeneity from duration dependence. This is discussed in Chamberlain (1982) and Heckman and Singer (1982). Applications using this multiple spell approach can be found in Olsen and Wolpin (1983) and in Newman and McCullough (1984) who look at fertility and in Jain and Vilcassim (1991) who study purchase timing. If multiple spells of unemployment were observed for each individual, this approach could be applied to separate duration dependence and unobserved heterogeneity in the duration analysis of unemployment spells.

Research has also been extended to analysis of multiple types of spells[1] (such as, following sequences of unemployment and employment durations), which may be correlated with each other. This requires a multivariate distribution function of the unobserved heterogeneity terms. Van den Berg (1997) discusses several possible options and the range of values for the correlation coefficient that each of the options can attain.

To estimate a structural job search model, data with information on reservation wages and offered wages are needed. Often, not enough information is available to identify such a model. Progress in these types of models has involved identifying a reservation wage path instead of assuming a stationary reservation wage over the whole length of the duration (Van den Berg, 1995), which allows the reservation wage to change as the unemployment spell progresses.

3 Labour supply and/or welfare participation models

Blundell et al. (2000) show that behavioural modelling is important when examining the effects of policy changes. Behavioural modelling makes use of the estimated parameters of a model, which describes the relationship between labour supply, wage rate, other income and individual characteristics. The main underlying assumption is that individuals choose a level of labour supply and income, conditional on the attainable options that optimise their utility. Gross income at the different levels of labour supply is calculated using the relevant labour supply, wage rate and other income. A behavioural model depends on an accurate representation of benefit and taxation rules to calculate what the net income is at all levels of gross income. The model can then be used to predict what the change in a person's labour supply behaviour will be as a result of a policy change. The effect of any policy change affecting the gross to net income

transformation can be calculated. In Blundell et al. (2000) it is shown that inclusion of behavioural effects (changes in labour supply) in a microsimulation model results in a predicted cost of extending Family Credit to the more generous Working Families Tax Credit in the UK that is 14 per cent lower than would have been predicted without allowing for behavioural changes. Creedy and Duncan (2001) provide an extensive discussion of the use of labour supply models in microsimulation.

3.1 Types of labour supply models

Empirical labour supply models have seen major changes over time. With the increased computing power available to researchers, previously impossible or extremely cumbersome approaches have now become achievable.

Three types that can be distinguished are:

a) The first generation of labour supply models, which approximates the non-linear budget constraint by a linear constraint in order to derive labour supply equations that could be easily estimated. Average net wage rates or marginal wage rates in the observed labour supply points are used in the approximations.

b) A next generation of models looked for the optimal labour supply at each of the separate linear segments and at each of the kinks of a budget constraint that was non-linear but piecewise linear. Burtless and Hausman (1978) are the first to use this approach, which allows researchers to incorporate tax and benefit systems more accurately into labour supply modelling. Moffitt (1986) gives an extensive description of the method and provides an overview of papers that have applied this method. A disadvantage of this approach is that one has to impose restrictions on the parameters so that the Slutsky conditions are satisfied[2] in a broad region. Failing to do so may result in a model that is statistically not meaningful (such as negative probabilities of being in particular labour supply points). MaCurdy et al. (1990) explain that this restriction of the parameter space may bias substitution effects upwards and income effects downwards. In addition, for a large number of separate linear segments, the method can become econometrically complicated and computationally cumbersome to apply. This means that, for example, to use it in labour supply models for couples, where the budget constraint is determined by the wage rates, benefit withdrawals and tax rates of two persons, can be extremely difficult. As a result, few papers have

appeared using this method to estimate behavioural models of labour supply for couples.

c) The third generation of models has simplified estimation by only considering a limited number of discrete labour supply points instead of the full range of possible hours. One of the first to use this approach is Zabalza et al. (1980). From a theoretical point of view, the choice from a limited number of labour supply points can be defended by the fact that in real life people normally do not get to choose their labour supply freely between zero and a maximum number of hours. Some studies have compared the capability of different models to capture features of observed labour supply (for example, Van Soest et al. 1990; Tummers and Woittiez, 1991). They found that discrete labour supply models generally do a good job.

3.2 Comparing specifications

Within each of the three strands of labour supply models used in the literature, several variations are possible, increasing the number of possible specifications further. Unfortunately, relatively few papers make comparisons between alternative model specifications to study the results' sensitivity to the chosen model. Exceptions are Mroz (1987), Gerfin (1993), Van Soest (1995), Blomquist (1996), and Kawaguchi (1994).

Mroz (1987) uses U.S. data to find that a wide range of responses to changes in wage rates and other income can be found for married women's labour supply using different specifications for the labour supply model. He uses a restrictive sample only including married women with working husbands, where both the wife and the husband are between 30 and 60 years old. Female labour supply is estimated using a reduced form equation where male labour supply is considered exogenous. Mroz (1987) examines the importance of controlling for self-selection into the labour force, the assumption on exogeneity of the women's labour market experience and wages, and the endogeneity of children and non-wife income. He specifies a large number of different models, tests the different specifications, and compares the outcomes with regard to the wage and income elasticities. He finds that the three most important assumptions are the way self-selection into the labour force is treated, endogeneity of wages, and use of labour force experience as an instrumental variable to control for the endogeneity of wages. Assumptions that were found to be unimportant are the exogeneity of non-wife income

and the number of children, controls for non-proportional income taxes and self-selection into the labour force once labour force experience is treated as endogenous. Mroz finds that amongst the models that cannot be rejected are models that generate elasticities of female labour supply that are no larger than those for men. This is an interesting finding, because it is generally believed (and found in empirical analysis) that women's labour supply is more responsive to wage or income changes than men's labour supply.

Gerfin (1993) compares three models where the unobserved market wages for non-workers are treated differently, using Swiss data for married women. He estimates a simultaneous wage and labour supply model, which is how such a model should be estimated ideally. The results of this model are then compared with two commonly used alternative specifications, where the unobserved wages for non-workers are replaced by their expected values. The expected values are obtained from a separately estimated wage equation. One version of the model only replaces the unobserved wages and the other version also replaces the observed wages (of the workers) by their expected values. From the three models, the elasticity of labour supply with regard to wages and income can be simulated. A comparison shows that the elasticities are quite different in the three specifications. The model where only the unobserved wages are replaced by expected values is most different with regard to the wage elasticities, whereas the model where all wage values are imputed is most different with regard to the income elasticities.

Van Soest (1995) compares four specifications of a labour supply model for couples in the Netherlands. One of the comparisons is between a model where wages are imputed by replacing unobserved wages by the expected value and a model where unobserved wages are integrated out of the likelihood function. The latter means that the uncertainty in predicting wages is taken into account. The differences in the resulting elasticities are not as large as in Gerfin's (1993) paper, but the elasticities seem to be slightly biased upwards as a result of not accounting for prediction errors. A more important inclusion in the model is the hours restriction in the form of a penalty function for part-time hours. Leaving the hours restriction out of the model more severely biases the own wage elasticities upwards, while the cross wage and income elasticities are biased slightly downwards. Accounting for random preferences to allow for unobserved differences in the preference for leisure hardly has an effect on any of the elasticities.

An interesting approach is taken by Blomquist (1996). He carries out Monte Carlo simulations to investigate the properties of different estimators. In one specification, the endogenous wage equation and labour supply are estimated simultaneously, accounting for all the details of the nonlinear budget constraint (caused by taxes and benefit payments). In the other specifications, the budget constraint is linearized in the optimal point, which simplifies estimation. This gives a wage rate and other income to use as inputs in a labour supply equation. The use of Ordinary Least Squares to estimate the labour supply equation would result in a bias, therefore, an instrumental variables method is used. Blomquist estimates two versions of this instrumental variables specification using different instruments. He does a Monte Carlo simulation to explore the behaviour of the different estimators in different circumstances and finds that no estimator is the best under all circumstances. Instead the optimal estimator mainly depends on the sample size and the type of measurement errors in the data. He proposes to complement actual estimates in empirical studies with a simulation study that generates data similar to the real data and to examine the effects of imposing measurement errors similar to those in the real data on the performance of the different estimators. This might help with choosing a preferred approach and give an indication of how the estimates should be interpreted. Unfortunately, identifying the measurement errors present in the real data might not always be easy to achieve.

Another specification issue regards the type of utility function that is chosen as the utility being maximized by the household. Usually, a unitary model is selected, where it is assumed that the household maximizes one common utility function, which depends on the amount of leisure time of the different members and on total consumption. More recently, models are being proposed where the different members in a household are maximizing individual utility functions, which may be dependent on the other member's overall well-being. Kawaguchi (1994) estimates four models of household labour supply, each based on a different type of utility function. The specifications are the neoclassical utility function (where there is a common household utility), wife as secondary earner model (where the husband's labour supply is pre-determined and independent of the wife's labour supply), the intra household trade model (where there is no joint consumption of goods and where non-market time is traded within the household at the market price), and the Nash bargaining model (where the household bargains over the allocation of non-market time based on a threat point, for example, the

level of utility of each person when separated). He then tests whether the restrictions necessary for the first three models are fulfilled. The last model cannot be tested. He finds that the conditions necessary for the first three conditions are all rejected. However, the sample used is rather restrictive in order to simplify estimation (e.g. only couples where husband and wife are working, the wife is between 20 and 40 years old and the husband is between 20 and 45 years old are included).

The above five studies show that different specifications may result in different outcomes. A wide variety of specifications is examined. One thing that three of the above studies have in common is that they all identify the way of dealing with the unobserved wages as an important issue. The impact that the choice of a particular specification in several of the above comparative studies has, indicates that a systematic and thorough comparison of alternative models can be important when, for example, specifying a model for policy simulations. Such a comparison should allow several aspects of the specification to be changed and assessed on the effect these changes have on the results generated by the model.

3.3 Neglected issues

Welfare participation and labour supply are mostly analysed separately. However, the decisions made regarding hours of work and welfare participation are very likely to be taken simultaneously, given that eligibility for welfare depends on earned and unearned income. In many labour supply models it is assumed that everyone who is eligible for benefits will participate in welfare. The question is whether this is a reasonable assumption. Duclos (1997) finds that even after accounting for other possible sources of differences between eligibility rates and take up rates of Supplementary Benefits in the UK, the take-up rate still is 20 per cent lower than the eligibility rate. In addition, inclusion of a possible negative effect on the household utility level if the household participates in welfare provides a possible explanation for households working for an income that is only slightly more than welfare participation would have provided.

Another important issue in labour supply, which is often neglected, is involuntary unemployment. Not everyone who is willing to work can find employment. In a recent paper by Holzer and Stoll (2000), the employer's demand for the labour supply of welfare recipients is explored in four large metropolitan areas in the U.S. They find that even if the labour market for welfare recipients is quite strong, it is uneven in terms of

the individual recipient's access to employment, depending on the recipient's location and race (minorities are less likely to be hired). This labour market can also be substantially weakened during an economic recession.

In addition to involuntary unemployment, many people cannot work the exact number of hours they would prefer. Most data sets only contain information on actual labour supply and do not ask for preferred hours of work. As a result, observed labour supply is often influenced by labour demand.

Most of the above discussion relates to cross-sectional data, neglecting any longer-term effects and assumes people are not considering the impact of current decisions on the future. Ziliak and Kniesner (1999) conclude in a recent paper that studies on work incentive effects of income taxes should use longitudinal data, which would allow for time-nonseparable labour supply. They argue that a progressive tax on wage and interest income makes this important. In addition, cross-sectional data neglects influences on labour supply from the stage of the life cycle someone is in. This may be in particular important for women who have children or are planning to have children. Furthermore, a model based on cross-sectional data cannot incorporate the importance of working in a period on lifetime earnings, where temporarily dropping out of the labour market can set back one's career. This consideration is probably most important for people with high education levels and/or a large amount of work experience. Unfortunately, the amount of longitudinal data is limited especially in Australia.

4 Evaluations and Experiments

The U.S. has a long history of experiments with taxation and benefit systems. There is an extensive literature on the methodology associated with the analysis of experiments and evaluation of policy changes.

Most evaluations of experiments or changes centre around the average impact that a particular change has on, for example, unemployment duration. Heckman et al. (1997) argue that the distribution of effects over programme participants is equally relevant in assessing a programme or changes in policy. Examples of other criteria in the evaluation that may be of interest are the proportion of programme participants benefiting from it, the proportion of the total population benefiting from the programme, quantiles of the impact distribution and the distribution of gains over subgroups.

Blundell and Costa Dias (2000), in their overview of evaluation methods for non-experimental data, distinguish five categories of evaluation methods. Four experimental approaches, including the pure randomised social experiment, the natural experiment (or differences-in-differences approach), the matching method and the selection method; and the structural simulation approach (which is non-experimental). In the first experimental approach, individuals are randomly assigned to a treatment and control group. This makes evaluation very simple, since in principle one could compare the mean values of the variable of interest in the two groups to obtain the effect of a policy. However even in this case, Hotz et al. (1999) recommend the collection of detailed pre-training or pre-policy change data, which would facilitate the extrapolation of results to other populations than the one studied. In addition, information on the exact nature of programmes would make the evaluation of programmes more useful to policy makers. It is also important to distinguish between the effect of the 'intention to treat' (that is, being allocated to the treatment group) and the effect of the treatment on those who are actually treated when evaluating experimental data. Heckman et al. (1999) warn that not all evaluations are best done using an experimental approach. For example, the experimental approach will not provide the estimated parameters of interest in ongoing programmes or when close substitutes to the evaluated programme are available.

The three other experimental methods are based on finding comparable individuals to those in the treatment group for comparison. First, the natural experiment in which a control group is constructed with similar characteristics and observed in a comparable period. Second, the matching method which is based on selecting a comparable non-treatment individual for every treatment individual based on a group of selected factors, the matching variables. It is assumed that the matching variables can be selected in such a way that people with the same values of these factors react in the same way to policy changes. Finally, the selection method (or Heckman approach) which aims to take the bias of being selected for treatment out of the estimated treatment effect. This approach needs at least one factor that contributes to being selected for treatment but not to the effect of the treatment. This allows researchers to estimate the part of the error term in the outcome equation, which is correlated with selection into treatment.

The advantage of the (quasi-)experimental approaches is that fewer assumptions are needed and that the choice for a particular economic theory (or behavioural model) is less important than in the structural simulation approach. By comparing the behaviour

of appropriate groups the effects of particular policies can be deduced. However, the disadvantage of this approach is that the results are specific to the policy evaluated in the experiment and it is impossible to extrapolate them to alternative policy changes in the same way as can be done using more structural simulation analyses.

The most important issue in the (quasi-)experimental approach is to compare appropriate groups, which have common time effects across groups and no composition changes within each group. This can be difficult to achieve[3]. The availability of high quality data to construct these groups is of major importance and no method can make up for bad data (Heckman et al., 1999). Especially in quasi-experiments where policy changes have occurred over time to particular groups, one has to take care to control for other possible changes that occurred around the same time. However, these experiments provide an opportunity to compare behaviour before and after the change or (the change in) behaviour of one group with another group. Blundell and Costa Dias (2000) give an example of a control group, which at first sight may seem a reasonable choice, but after some further exploration reveals a major shortcoming. Eissa and Liebman (1996) evaluated the 1986 expansion in the U.S. Earned Income Tax Credit (EITC) for sole parents by constructing an artificial control group. They compared the change in labour supply of women with children with the change in labour supply of women without children before and after the policy change. Blundell and Costa Dias point out that the participation of women without children is at a much higher level (about 95 per cent) than the participation of sole parents. As a result, the control group cannot be expected to increase its labour force participation by a similar amount as the sole parent group in response to an improved state of the economy. The above results are therefore likely to be an overestimation of the true effects.

In evaluations it is important to recognize the possibility of differences in effects of programmes on different types of people or under different economic circumstances. In addition, the analysis should allow these different effects to play a role in the decision of a individual to participate in a programme (the self-selection process). Accounting for this endogeneity of participation is important to determine the effect of a programme. As a result, care should be taken when extrapolating the estimated effects of a programme to a different situation, if the experiment has not been set up to allow such generalisations of results. Furthermore, implementation of an experimental

programme on a large scale may change the prices and opportunities of everyone in the population, changing general equilibria and/or displacing other workers.

The most appropriate approach to follow when evaluating non-experimental data depends on the available data and the policy parameter of interest. Sometimes a combination of methods can be most suitable. However, the approach taken has to be decided on a case-by-case basis, there is no 'method of choice' (Heckman et al., 1999; Blundell and Costa Dias, 2000).

5 Conclusion

The field of labour economics benefits to a large extent from developments in econometrics. Research in this field is based on advanced econometric techniques to enable disentanglement of opposite effects and allow the isolation of the issue of interest. For example in policy evaluations, a researcher will discover that the ideal experimental design of two perfectly comparable groups with just one difference between them (the policy variable of interest) does not often occur. However, the application of econometric techniques has enabled researchers in this situation to draw sensible conclusions notwithstanding the lack of ideal data.

In other areas of labour economics where the complexity of taxation and social security systems (which are recognized as major components in the labour supply decision of individuals) complicates modelling of the real world, research has been helped to a large extent by applying new econometric and computational techniques and by more powerful computers allowing the use of these techniques.

This article has given an overview of and references to current methodology used in this field. It aims to be a starting point, given that the techniques used in this field are continuously evolving and improving and often need to be adapted to suit a particular situation at hand.

Endnotes

* I should like to thank Jeff Borland and John Creedy for helpful comments.

1 This is different from the above paragraph, where multiple spells of one type are discussed.

2 This is similar to imposing quasi-concavity on the preferences.

3 See Blundell and Costa Dias (2000, p. 429).

References

Baker, M. and A. Melino (2000), 'Duration dependence and nonparametric heterogeneity: A Monte Carlo study', *Journal of Econometrics*, 96, 357-393.

Barrett, G.F. (2000), 'The effect of educational attainment on welfare dependence: Evidence from Canada', *Journal of Public Economics*, 77, 209-232.

Blomquist, S. (1996), 'Estimation methods for male labor supply functions; How to take account of nonlinear taxes', *Journal of Econometrics*, 70, 383-405.

Blundell, R. and M. Costa Dias (2000), 'Evaluation Methods for Non-Experimental Data', *Fiscal Studies*, 21(4), 427-468.

Blundell, R., A. Duncan, J. McCrae and C. Meghir (2000), 'The Labour Market Impact of the Working Families' Tax Credit', *Fiscal Studies* 21(1), 75-104.

Burtless, G. and J. Hausman (1978), 'The Effect of Taxation on Labor Supply: Evaluating the Gary Income Maintenance Experiment', *Journal of Political Economy*, 86(6), 1103-1130.

Chamberlain, G. (1982), 'Heterogeneity, omitted variable bias and duration dependence', in J.J. Heckman and B. Singer (eds.), *Longitudinal analysis of labor market data*, 3-38.

Cockx, B. and M. Dejemeppe (2000), 'Duration and Calendar Time Dependence of the Exit Rate out of Unemployment in Belgium. Is it True or Spurious?', mimeo, Université Catholique de Louvain.

Creedy, J. and A. Duncan (2001), 'Behavioural Microsimulation With Labour Supply Responses', in J. Creedy *Taxation and Economic Behaviour, Introductory surveys in Economics*, volume I, 165-216.

Duclos, J.-Y. (1997), 'Estimating and testing a model of welfare participation: the case of supplementary benefits in Britain', *Economica*, 64, 81-100.

Eissa, N. and J.B. Liebman (1996), "Labor supply response to the earned income Tax Credit", *Quarterly Journal of Economics* 111(May), 605-637.

Gerfin, M. (1993), 'Simultaneous discrete choice model of labour supply and wages for married women in Switzerland', *Empirical Economics* 18, 337-356.

Heckman, J.J., R.J. LaLonde and J.A. Smith (1999), 'The Economics and Econometrics of Active Labor Market Programs', in the *Handbook of Labor Economics*, volume 3A, eds. O. Ashenfelter and D. Card, North-Holland, Amsterdam.

Heckman, J.J. and B. Singer (1982), 'Social science duration analysis', in J.J. Heckman and B. Singer (eds.), *Longitudinal analysis of labor market data*, 39-110.

Heckman, J. and B. Singer (1984), 'A method for minimizing the impact of distributional assumptions in econometric models for duration data', *Econometrica*, vol. 52, no. 2, pp. 271-320.

Heckman, J.J., J. Smith, and N. Clements (1997), 'Making The Most Out Of Programme Evaluations and Social Experiments: Accounting For Heterogeneity in Programme Impacts', *Review of Economic Studies* 64, 487-535.

Holzer, H.J. and M.A. Stoll (2000), 'Employer demand for welfare recipients by race', Discussion Paper no. 1213-00, Institute for Research on Poverty, University of Wisconsin-Madison.

Hotz, V.J., G.W. Imbens and J.H. Mortimer (1999), 'Predicting the Efficacy of Future Training Programs Using Past Experiences', technical working paper 238, National Bureau of Economic Research, Cambridge, U.S.

Jain, D.J. and N.J. Vilcassim (1991), 'Investigating household purchase timing decisions: a conditional hazard function approach', *Marketing Science* 10, 1-23.

Kawaguchi, A. (1994), 'Testing neoclassical and non-neoclassical models of household labour supply', *Applied Economics*, 26, 9-19.

MaCurdy, T., D. Green, and H. Paarsch (1990), 'Assessing Empirical Approaches for Analyzing Taxes and Labor Supply', *The Journal of Human Resources*, 25(3), 415-490.

Moffitt, R. (1986), 'The econometrics of piecewise-linear budget constraints; A survey and exposition of the maximum likelihood method', *The Journal of Business and Economic Statistics*, 4(3), 317-328.

Mortensen, D.T. (1977), 'Unemployment insurance and job search decisions', *Industrial and Labour Relations Review* 30(4), 505-517.

Mroz, T.A. (1987), 'The sensitivity of an empirical model of married women's hours of work to economic and statistical assumptions', *Econometrica*, 55(4), 765-799.

Newman, J.L. and C.E. McCulloch (1984), 'A hazard rate approach to the timing of births', *Econometrica* 52(?), 939-961.

Olsen, R.J. and K.I. Wolpin (1983), 'The impact of exogenous child mortality on fertility: a waiting time regression with dynamic regressors', *Econometrica*, 51(?), 731-751.

Tummers, M.P. and I. Woittiez (1991), 'A Simultaneous Wage and Labor Supply Model with Hours Restrictions', *Journal of Human Resources*, 26, 393-423.

Van den Berg, G.J. (1995), 'Explicit expressions for the reservation wage path and the unemployment duration density in nonstationary job search models', *Labour Economics* 2, 187-198.

Van den Berg, G.J. (1997), 'Association measures for durations in bivariate hazard rate models', *Journal of Econometrics* 79, 221-245.

Van den Berg, G.J. and J.C. van Ours (1996), 'Unemployment Dynamics and Duration Dependence', *Journal of Labor Economics* 14(1), 100-125.

Van Soest, A. (1995), 'Structural Models of Family Labour Supply; A Discrete Choice Approach', *The Journal of Human Resources*, 30(1), 63-88.

Van Soest, A., I. Woittiez, and A. Kapteyn (1990), 'Labor Supply, Income Taxes, and Hours Restrictions in The Netherlands', *The Journal of Human Resources*, 25(3), 517-558.

Zabalza, A., C. Pissarides, and M. Barton (1980), 'Social Security and the Choice between , full-time work, part-time work and retirement', *Journal of Public Economics*, 14, 245-276.

Ziliak, J.P. and T.J. Kniesner (1999), 'Estimating Life Cycle Labor Supply Tax Effects', *Journal of Political Economy*, 107(2), 326-359.