

HOW HOUSEHOLD CONTEXT AFFECTS
SEARCH OUTCOMES OF THE
UNEMPLOYED IN KWAZULU-NATAL,
SOUTH AFRICA:
A PANEL DATA ANALYSIS

TARYN DINKELMAN*

HOUSEHOLD APPROACHES TO LABOUR market issues in South Africa are not at all new. This paper presents a review of some household themes that have been used to examine the relationship between individuals and the labour market, and employs a search model of the labour market, adapted for household effects, to understand why some South Africans are successful in their job search while others remain jobless.

We consider data from two waves of a household survey covering African individuals in the KwaZulu-Natal province of South Africa between 1993 and 1998. We ask what happens to people who are jobless – unemployed or not economically active in 1993 – five years later in 1998? Are there household features in particular which influence outcomes in the labour market by 1998? We are able to do this, controlling for unobserved household heterogeneity, using the KwaZulu-Natal Income Dynamics Study (KIDS).

* PhD. student in Economics (University of Michigan). This paper grew out of my Masters dissertation work, completed at the University of the Witwatersrand in 2002. The patient guidance of my advisor, Martin Wittenberg (University of Cape Town) and comments of an anonymous referee are gratefully acknowledged. All errors remain my own.

There are two reasons why this exercise is relevant. First, in a mass unemployment context where employment opportunities are not evenly distributed across the country, search is costly and does not guarantee that an individual will ever receive a job offer. The simple search model we use describes how household resources, information and composition affect search costs and benefits which determine labour supply choices. This paper motivates for the use of a search framework for South Africa, which recognizes the role that household information and resources play in enabling labour market matches.

Second, panel data is used to examine transitions between labour market states over time. Most existing South African literature uses available cross sectional household survey data. This is problematic, particularly if our interest is in the effect of households on individual labour market outcomes. It is difficult to identify the direction of causality between household formation and search success or failure from a cross section. However, with the KIDS data, we can condition on 1993 household characteristics that logically precede labour market outcomes in 1998, and so have more confidence in the household effects that are found to impact on search success. We specify fixed and random effects logit models to capture observed household effects and see how much unobserved household heterogeneity ‘explains’ search successes and failures after five years.

We find that search in 1993 does matter for escaping unemployment over time and, surprisingly, that urban searchers are less successful than those who begin in rural areas in 1993. Pensioners in the household tend to reduce individual success in finding a job, although these effects are different for men and women. Additionally, households have large unobserved effects on the search success of their members. The evidence presented shows that households do influence individual labour supply choices and outcomes.

The paper begins by outlining the definition of

unemployment used, and discusses the South African literature dealing with households and individual labour market outcomes. Section 3 sets out a stylised model of search and shows how households may be incorporated. Section 4 contains the econometric methodology while 5 presents potential data problems and discusses some descriptive statistics. Section 6 discusses results of the analysis of successful and unsuccessful searchers, and section 8 concludes.

1. UNEMPLOYMENT DEFINITIONS

Discussion around the appropriate definition of unemployment for South Africa is well known.¹ The issue derives from the theoretical distinction between voluntary and involuntary unemployment, and the observation that a large proportion of the South African jobless who want jobs are not searching for work because they are ‘discouraged’.²

Notwithstanding this debate, we use the narrow definition of unemployment in this paper, since the effects of search behaviour on successful labour market transitions are one of the points of interest. We do not ignore the discouraged unemployed, but choose to group them with the not economically active (NEA) group. We assume everyone in this group has the potential to transit into a job over time. The pool of jobless individuals then consists of searchers, non-searchers and all the other NEA of working age.

It is important to note at the outset that we are modelling factors that affect whether someone transits from no job to a job by 1998. Our search model provides some

¹ See Dinkelman and Pirouz (2002), ILO (1996) and Kingdon and Knight (2000) for outlines of this debate.

² For different age cohorts, between 45% and 60% of the unemployed were ‘discouraged’ – not seeking work because there were no jobs – in 1999. Figures from author’s own calculations, October Household Survey 1999.

structure for thinking about these processes. However, since we have only 2 observations per individual and a less-than-optimal 5 year gap in between, we cannot say that individuals move from their 1993 labour market state into their 1998 positions, without any intervening movements between states. We separate the searching unemployed from all others without jobs in 1993 to identify whether that particular labour supply choice affects subsequent labour market outcomes, but maintain the hypothesis, to be discussed in section 3, that if an individual wants a job, search is necessary at some point.

2. THE HOUSEHOLD, THE INDIVIDUAL AND THE LABOUR MARKET: ISSUES ARISING FROM THE SOUTH AFRICAN LITERATURE

There are four emerging themes linking the household, the individual and the labour market in South Africa: households as *private safety nets*, households as *productive units*, households as *information networks* and households as *work/search culture-generators*.

Firstly, growing out of the literature which links poverty with labour market outcomes, the household has been regarded as a *private safety net* for individuals without jobs.³ These people are able to share the wage income and other transfer payments flowing in to the household, but in the process, may experience reduced incentives to look for work.

Klasen and Woolard (1998, 2000) examine how household structure adapts to deal with these unemployed. They assume the individual's employment position is exogenous and consider the optimal residence decision - to set up one's own household versus joining the household of a parent or relative - as a consequence of this labour market status. They find no evidence that access to pension incomes and remittances in the household reduces labour supply, and conclude that people who attach themselves to rural

³ See Borat *et al.* (2001) for work in this area.

households have very few desirable labour market characteristics (Klasen and Woolard, 1998:20).

A serious problem with the Klasen and Woolard analysis is that with their data, they are unable to temporally separate the point of household formation and the point of employment/unemployment. In our analysis, we are better placed to assume that household formation precedes individual outcomes in the labour market, since we have data on the former from 1993 and the latter from 1998.

Under the second household effects theme, the composition of the household precedes and conditions labour supply choices. In the spirit of Becker (1965), the household may be viewed as a *productive unit* that must be sustained. The production of household goods may differentially influence male and female labour supply decisions. In this vein, Bhorat and Leibbrandt (2001) model the effect of household composition variables on the labour market participation decision (Bhorat and Leibbrandt, 2001:121). Using cross-sectional variation from a 1995 household survey, they find that with more household helpers, fewer child-rearing responsibilities and less income for sharing, women choose to enter the labour force and search for work. A lack of pension and other income sources is also important for increasing male participation (Bhorat and Leibbrandt, 2001:119-120).

In another paper, also employing cross-sectional variation in households and individuals, Bertrand, Mullainathan and Miller (2003) focus on the labour supply choices of working age males, in response to an increase in household income. They find quite convincing evidence that pension income in the household allows individuals – particularly eldest sons – to reduce their labour supply, as the flow of resources from younger to older adults reverses when older adults – particularly females – reach pensionable age.

Both of these papers reverse the causal link between households and labour supply choices, taking the form of the

household as the exogenous variable. This is the approach we take. In our analysis, we are additionally able to control for some household level heterogeneity by having two observations for each household at two different points in time. The one gap in the Borat and Leibbrandt paper which we also seek to address is the possibility that household-level variables directly affect the probability of search success. We include variables in our analysis which might explain how households matter for individuals being successful searchers, rather than simply labour market participants.

The third strand of household level effects on individual labour market outcomes extends this idea. In this approach, the household is a *privileged source of information* for the unemployed, if it contains existing labour market links – e.g. an employed person in the household or a migrant worker. Since information is costly to acquire, and the chances of being a successful searcher in a mass unemployment environment are small, individuals rely on their friends and relatives in the household to bring information back to them.

Wittenberg (1999) uses the 1995 OHS data to show that when unemployment is high, school qualifications which everyone has access to matter little, location in relation to job markets matters a lot, and labour market activities of others in the household matter to some extent for determining outcomes in the labour market.⁴ In another paper, Wittenberg (2001) shows that access to employment is mediated through non-market interactions within household and neighbourhood networks. Tracking these effects down in a static framework off the OHS 1995, the results suggest that having an older employed household member increases the probability of working for both genders, with larger effects in rural areas.

⁴ These effects may also be due to wider community level effects rather than individual household effects (Wittenberg, 1999:38). This is something we note, but do not deal with here.

In addition to characterising the household as an information network, two points emerge from these papers. Firstly, it is useful to look at different degrees of labour force attachment, rather than concentrate on only those in the labour market, because the decision to search can become dependent on the prevailing level of unemployment (Wittenberg 1999:47). For this reason and the reasons provided in section 1, our analysis looks at all of those who have no jobs in 1993 – the searching unemployed, the non-searching unemployed and the other NEA. Secondly, the importance of analysing labour market dynamics is made clear. In piecing together three consecutive household surveys and analyzing age cohorts over these periods, Wittenberg (1999) moves some way towards describing dynamic effects. However, if two data sets tell the same story, it is not possible to establish whether this is because the labour market was operating in the same way over time, or whether there was a considerable amount of change in jobs at the individual level, which averaged out at the aggregate level (Cichello *et al.*, 2000:2). Again, we try to deal with the question of how household effects matter for individual labour market outcomes over time by using the KIDS panel survey.

The fourth household effects theme grows out of the third. The *culture of the household* has been argued to matter for choices about whether to participate in the labour market and for the probability of locating a job. This household culture might include work or search ethic, special methods of search (*e.g.* some households might not have any history of members queuing for jobs), motivation, literacy, language skills and family access to transport, among other things (Wittenberg, 2001:6). Identifying this kind of household culture externality is, however, very difficult, as it is not observed in any direct form. It might only be captured as part

of a residual.⁵ In section 6, we account for as much of the observed household effects as possible using observed household variables, and identify how much of the residual may be due to leftover household culture effects that include but are not restricted to the above-mentioned.

This discussion of the South African literature demonstrates that researchers have attempted to link household contexts with individual choices and outcomes within the labour market using other data sets. Although this is an active area of research, the approach has not yet been taken extensively to the KIDS data. The KIDS has been used in a variety of analyses focused almost entirely on either the household or the individual.⁶ In this paper, we focus on bringing the two parts - households and individuals - together within a search framework.

3. THE THEORETICAL MODEL

All search models are populated by agents who have imperfect information about the location and type of available jobs and wage offers, and who make decisions about how much information it is worthwhile to invest in.⁷ In the labour market, price signals alone are unable to match jobs and workers. Firms and potential employees must search for each other to overcome frictions of time, space and

⁵ Indeed, using 1995 data, Wittenberg (2001:13) finds that about 25 per cent of the remaining variation in work probability between individuals is not accounted for by identifiable household features in the context of his model specification.

⁶ For example: Maluccio *et al.* (1999), Carter and May (2000) and Leibbrandt and Woolard (2000) address poverty, inequality and welfare at a household level, while employment and earnings dynamics at the individual level have been considered by Cichello *et al.* (2000) and Keswell (2000).

⁷ Such information is also imperfect and comes at a price for firms, but we are not primarily concerned with search on the demand side of the labour market in this paper.

information, before such matches are possible. Search is a necessary activity of any individual wanting a job in any labour market, although the type of search activity may take very different forms. How much to invest in information, or how much search is optimal before accepting a job offer or leaving the labour force, is the important question that job search models try to answer.⁸

The model we use is a stylised interpretation of how search choices are made. Individuals looking for jobs are assumed to face a known wage offer distribution, represented by cumulative density function $F(W)$. They sample offers from this distribution with replacement at cost (c), which is assumed to be a constant cost per observation. This cost includes direct out-of-pocket costs - travel to an employment agency, purchase of newspapers - as well as the opportunity cost of not spending time in alternative activities: housework or leisure. Only one offer may be accepted and there is no recall of offers. The worker must choose a search strategy which helps her decide when to accept an offer (Mortensen and Pissarides, 1999:2571).

The discounted value of being in search unemployment (U) satisfies the following Bellman equation⁹:

$$rU = b - c + P[W - U] \tag{1}$$

where r is a discount factor, b is any actual income flow received in unemployment (unemployment insurance and other income transfers) plus the imputed value of alternative

⁸ The seminal article on search in the labour market is Stigler (1962). Detailed discussions of the development and application of search models, as well as variations in these models, can be found in Devine and Kiefer (1991), Lippman and McCall (1976a and 1976b) Mortensen (1986), and Mortensen and Pissarides (1999).

⁹ Bellman's principle asserts that the present decision in a sequence of decisions maximises the discounted current net return plus expected future stream of returns, assuming that decisions in the future are made optimally, where the expectation taken is conditional on current information (Mortensen, 1986: 856).

activities in this state of joblessness, c is cost of search per period and P is the probability of the individual receiving a job offer. W and U are, respectively, the expected values of being employed or unemployed (Mortensen, 1986:856).

In deciding whether to accept a job offer at w drawn from $F(W)$, the individual compares the discounted value of the stream of future W 's in that job to the expected value of income to be received in continued search, U . The expected value of the stream of future income, if the individual accepts a job offer, satisfies:

$$rW = w - P_{JD}[W - U] \quad (2)$$

where w is the offered wage rate, P_{JD} is the probability of the job terminating at a future point in time, and $[W-U]$ is the net value of the job over search unemployment to the individual.¹⁰

If the wage offer realised is higher than the expected value of future search, then the job is accepted; however, if the expected value of search in the next period is higher, then the value of search in this period will be higher if the job offer is refused (Mortensen and Pissarides, 1999:2571). The optimal stopping strategy therefore says: continue to search in this period if you expect to get a better offer for the present discounted value of future income in the next period.

To accept a job offer, the value of the offer must at least match the value of continuing in unemployment – this is the reservation wage property shown in (3):

$$rW(w^*) = rU = w^* \quad (3)$$

Substituting (3) in (1), we have

$$w^* + c = b + P[W-U] \quad (4)$$

which has the intuitive explanation that the worker stops

¹⁰ Including a term to indicate the possibility of the job terminating at some future point is a variation on the basic search model (Pissarides, 2000).

searching and accepts a job when the marginal costs of additional search on the left hand side are equal to the marginal benefits from that extra search on the right hand side (Mortensen, 1986: 58).

Additionally, $rU = w^* \geq b$: the reservation wage or the imputed value of search must be at least as large as the value of alternative activities so that the worker does not prefer to remain out of the labour force. Thus, the final equation we have is that which describes the discounted value of being out of the labour force (O):

$$rO=b \tag{5}$$

We make the simplifying assumption that the value of non-wage activities in unemployment and out of the labour force is the same (b).

From the value functions in (1), (2) and (5), the labour market status of any individual at one point in time may be described: a worker will have $W>U$, a searching unemployed person will have $U>O$ and $U>W$ for any received job offers and someone NEA will have $O>U$. We can also think through how the reservation wage (w^*), and by implication the optimal stopping point of search, changes when the environment of search changes:

If b increases, w^* rises - the higher the opportunity cost of taking a job offer, the fewer job offers will be accepted and the longer the period of unemployment will be.

If P increases, U and w^* rise. Now some lower job offers will no longer be acceptable. When the chances of getting a job are higher, the unemployed can afford to be choosy.

If c or r increase, w^* will fall - the more costly it is to search, the less search will be undertaken and lower job offers will become more acceptable. If, as c or r increase, w^* falls below b , then the worker stops searching and drops out of the labour force completely.

An implication of the model is that if search costs are high enough or the probability of finding a job low enough,

then U may be lower than O even though W might be higher than the value of the alternative activity. This scenario probably represents the case of the discouraged unemployed.

Search models portray the individual as a forward-looking agent, aware of the uncertainties surrounding the location of a job offer and considering the value and costs of investing in information before continuing with search. Clearly, the values of U , W and O depend on individual characteristics like age, gender and education. We include these variables in our empirical analysis. But these value functions are also affected by household context.

Search models in general have not explicitly addressed the role of the household in the search choice.¹¹ It is possible to see how they would matter in a search model by assuming that b depends to some extent on the household. The most obvious way in which this could occur is in the form of income transfers within the household, between family members. Parents (either working or pensioners) might provide unemployed children with an income to survive and finance search activities. This income effect might also reduce search, as O increases.

The value of b could also depend in non-measurable ways on the composition of the household, and might vary by gender. For example, the benefit gained from remaining out of the labour force and raising children might be different for men and women. With other females in the household, the value of this activity - and thus the opportunity cost of entering search unemployment - for a particular women in the household is reduced.

Another way in which households could impact on rU is through c . Search costs are incurred because information has to be paid for, so if the household can provide information to the unemployed individual at lower cost - perhaps by

¹¹ Although Pissarides (2000:ch7) does discuss the added and discouraged worker effects in his general equilibrium search framework.

virtue of being situated in a location where search costs are lower - then c falls and consequently the asset value of being in searching unemployment increases. For example, if a household has one employed member, then the access to information this member provides can reduce c for other unemployed members. However, reducing c also raises the reservation wage and so might prolong the length of time an individual searches before accepting a job offer.

In the South African labour market, anecdotal evidence has also suggested that labour market connections, possibly deriving from the household, can act to increase P for certain workers. This is sensible within a search framework where firms also face costs of finding suitable workers; and the implication is that in a labour market with massive frictions and very little information about worker quality, households with some labour market connections may be better placed to enable successful search activities of their members (Wittenberg, 2001). If these jobs offer higher wages than other jobs found without such contacts, then rW may also be affected by being in particular households.

Finally, if we accept that household culture affects individual choices, then being in a household in which members are motivated to search could increase the perceived value of P for other unemployed people in the household – thus inducing them to search rather than remain NEA.

With this search model in mind, we discuss the econometric methods used to answer the question: which individual and household factors influence the transition from joblessness to employment?

4. METHODOLOGY

We model the probability of an individual being a successful searcher in 1998 – that is, the probability that they will find a job – based on individual and household characteristics

which affect value functions (1), (2) and (5) of our search model. We do this using the matched panel data to control for unobserved group heterogeneity at the household level.

To be a successful searcher, individuals must find jobs that meet their reservation wage. They must have $W > U$ or, if a job is handed to them without search, $W > O$. Since we do not observe individual value functions but only the outcomes of these choices, we must specify the following latent variable model (Greene, 1997: 615-635, 880-882):

$$y_{ij}^* = x_{ij}'\beta + h_i'\gamma + (h_i m_{ij})'\lambda + \varepsilon_{ij} \quad (6)$$

with $\varepsilon_{ij} = u_i + \omega_{ij}$

and $y_{ij} = 1$ if and only if $y_{ij}^* > 0$

where $y_{ij} = 1$ indicates that individual j in household i is employed in 1998, $y_{ij} = 0$ indicates joblessness (unemployed or NEA) in 1998, and y_{ij}^* is the latent variable representing the outcome of the comparison of the value functions W , U and O . The right hand side of the reduced form equation (6) includes variables expected to matter for w , b , c , P , P_{JD} , W and U : x_{ij} are the individual characteristics of individual j in household i , h_i are the measurable household features of household i which will be the same for all individuals in this group, m_{ij} is a gender dummy (0=female, 1=male) producing interaction effects for men and household level variables and ε_{ij} is the disturbance term capturing everything that cannot be measured.¹² We assume ε_{ij} has a logistic distribution and so run logit models.¹³

We specify that the residual term can be decomposed

¹² Thus we model $\Pr(y_{ij}=1 | x_{ij}, h_i, m_{ij}) = \Pr(y_{ij}^* > 0 | x_{ij}, h_i, m_{ij}) = \Pr(x_{ij}'\beta + h_i'\gamma + (h_i m_{ij})'\lambda > -\varepsilon_{ij} | x_{ij}, h_i, m_{ij})$.

¹³ The logit model is the preferred alternative for dummy dependent variables – it forces probabilities to be between 0 and 1 and allows variables to have non-linear effects on employment probabilities (Greene, 1997:873-876). This model, estimated by maximum likelihood techniques, is more frequently used to model labour market outcomes than simple linear probability models.

into a household level effect, u_i , and all remaining idiosyncratic error, ω_{ij} . If $u_i=0$, there is no household level effect and the results of a simple logit model will produce unbiased coefficients. However, if belonging to a particular household matters for labour market outcomes, then we need to use a fixed effects (FE) or random effects (RE) logit to deal with u_i .

In the FE logit, the group effect is assumed to matter in an additive way. We essentially add a constant term into the model which must be estimated – this constant is u_i . We make use of Chamberlain's conditional logit to estimate the probability of a jobless individual in 1993 being employed by 1998 with fixed household level effects (Greene, 1997:899-901):¹⁴

$$\Pr(y_{ij} = 1) = \frac{e^{x_{ij}'\beta+h_i'\gamma+h_i'\lambda_m+u_i}}{1 + e^{x_{ij}'\beta+h_i'\gamma+h_i'\lambda_m+u_i}} \quad (7)$$

Greene (1997:900) shows that calculating this probability for one individual in the household requires maximising the probability *conditional* on everyone else in the household. For example, if we have one male ($j=1$) and one female ($j=2$) observation in the i^{th} household, then the conditional probability of observing the male in a job in 1998 is given by:

$$\Pr(y_{i1} = 1 | \sum y_{ij} = 1) = \frac{\Pr(y_{i1} = 1 | y_{i2} = 0)}{\Pr(y_{i1} = 1 | y_{i2} = 0) + \Pr(y_{i1} = 0 | y_{i2} = 1)} \quad (8)$$

The unconditional probabilities for this man and women are given in (9) and (10):

¹⁴ Note that the FE logit only uses groups in which there is some variation. If some households had all members employed in both years, these would not be used to compute the likelihood function of employability. These groups (481 of them) contribute nothing to the equation and so are omitted at the time of estimation. Additionally, we are not able to correct for cluster design in the FE logit.

$$\Pr(y_{i1} = 1) = \frac{e^{x_{i1}'\beta+h_i'\gamma+h_i'\lambda_m+u_i}}{1 + e^{x_{i1}'\beta+h_i'\gamma+h_i'\lambda_m+u_i}} \quad (9)$$

$$\Pr(y_{i2} = 1) = \frac{e^{x_{i2}'\beta+h_i'\gamma+u_i}}{1 + e^{x_{i2}'\beta+h_i'\gamma+u_i}} \quad (10)$$

and substituting these into (8) and simplifying, we get:

$$\Pr(y_{i1} = 1 \mid \sum y_{ij} = 1) = \frac{e^{x_{i1}'\beta+h_i'\lambda_m}}{e^{x_{i1}'\beta+h_i'\lambda_m} + e^{x_{i2}'\beta}} \quad (11)$$

Similarly, we could find $\Pr(y_{i2}=1 \mid \sum y_{ij} = 1)$ if the male in the household is unemployed.

By conditioning on the sum of the two observations, the FE model drops out the household level effects constant across individuals in the same group and identifies the individual coefficients β for men and women and the male interaction terms for the household variables λ_m . However, by sweeping household level heterogeneity out of the model, we lose the baseline household coefficients γ and the fixed effect u_i , and cannot estimate how much of ε_{ij} is accounted for by u_i , which is one of the things we are interested in.

To identify how much households do matter in an unobservable way – the contribution of u_i to ε_{ij} – we run a random effects (RE) logit. The RE logit assumes that the u_i 's are randomly distributed across households but constant over time. Identification of the coefficients on x_{ij} , and h_i , as well as a measure of $u_i/(u_i+\omega_{ij})$, is achieved by using all of the within and between group variation (Greene, 1997:896-899; StataCorp, 2001, vol. 4:434 - 455)¹⁵.

One potential problem arises from the necessary assumption of the model that the u_i 's must be uncorrelated with the x_{ij} 's, h_i 's and m_{ij} 's (Greene, 1997:632-633). Since some of our independent variables are observed household features, it is difficult to argue that there could be no

¹⁵ The ratio $u_i/(u_i+\omega_{ij})$ is shown as ρ in the results Tables.

correlation with household level unobservables. For example, it is hard to think about how a household with more educated adults and more adults working would not also have an unmeasured ‘motivational’ effect on jobless individuals in the group.

Nevertheless, our RE results are very similar to the FE results, suggesting that the RE model is not badly misspecified. Furthermore, we checked the appropriateness of the RE model using a linear probability model specification. Results from a Hausman test do not allow us to reject the null hypothesis that the FE and RE coefficients are the same. This gives us more confidence in our choice of the RE logit, as more of the information is used to estimate coefficients in the random effects model (Greene, 1997: 632-633).

5. DATA: THE PSLSD AND KIDS

The first data set we use is from the 1993 Project for Statistics on Living Standards and Development (PSLSD) household survey.¹⁶ The second is the KwaZulu-Natal Income Dynamics Study (KIDS). The KIDS is a 1998 re-survey of African and Indian households in KwaZulu-Natal, originally surveyed in the PSLSD.¹⁷ Conducted by researchers from the University of Natal, the University of Wisconsin, and the International Food Policy Research Institute, its primary purpose was to provide data to analyse poverty. It captured data on 1036 households using a questionnaire that mirrored the 1993 survey in most respects. In this section, we discuss the attraction of using panel data, as well as the potential problems involved. Some descriptive statistics are also provided.

(a) What panel data offers

¹⁶ Conducted by the South African Labour Development Research Unit (SALDRU).

¹⁷ May *et al.* (1999) discuss the details of the resurvey.

Panel data allows dynamic analyses of individual and household behaviour which is not possible with cross-sectional or time-series data alone. For example, using only the PSLSD 1993, we can observe how many South Africans were employed or unemployed in 1993, and what their characteristics were. We can make inferences based on these characteristics as to what factors are important for getting a job. However, we cannot say anything about whether those employed (or unemployed) in 1993 are the same ones employed (or unemployed) based on another cross section from 1998.

With panel data, we can start to pin down factors that are important for keeping an individual in a particular labour market state or moving them into other states. We can also temporally separate household formation in 1993 and labour market outcomes for individuals in these households in 1998. In this way, we have a better idea of which came first - the household or the employment status of the individual.

There may still be a lingering endogeneity that we cannot control for. Households may have formed in 1993 according to individual characteristics which also influence subsequent labour market outcomes. To make the point with a rather obvious example: marital status in 1993 may be a good predictor of employment in 1998, but jobless people married in 1993 were possibly better placed in the marriage market precisely because their future employment prospects were better than other individuals. Thus, outcomes in 1998 are dependent on household composition in 1993, but the household composition is in turn dependent on the individual characteristics of the members who choose to be in the household. Unless these individual characteristics are identified by the independent variables in our models, their effects will be swept up in the residual.

Another attraction of panel data is that with repeated observations on the same individual (and household), it is possible to control for unobserved heterogeneity, or time-

invariant characteristics, which cannot be measured or sufficiently proxied for, like the motivation to search for a job. In this way, the variance of parameter estimates can be reduced.

Despite the attraction of being able to identify household formation which precedes labour market outcomes and control for unobserved heterogeneity across individuals and households, panel data suffers from two major ailments: measurement error which generates attenuation bias and attrition which generates selection bias.

(b) Measurement error and attenuation bias

Measurement error plagues all household surveys (Deaton, 1998:26-32). Initial investigation of the age and education variables for the matched data files in the KIDS survey shows that there is room for potentially great mismeasurement. Generating age and education difference variables for the matched data set of working age individuals, we found that 9.42 per cent had age gaps outside of the acceptable 4 to 6 years. About 25 per cent gained more education than they could have, or lost some education in the time between 1993 and 1998.

Measurement error imparts a downward bias to estimates, termed attenuation bias. This problem can become more severe with panel data, where differences in variables are taken over time. Deaton (1998:99) shows that when a dependent and independent variable are measured with error, the probability limit of $\hat{\beta}$, where $\hat{\beta}$ is the coefficient on the differenced variables, becomes:

$$p \lim(\hat{\beta}) = \frac{\beta m_{\Delta}}{m_{\Delta} + \sigma_{\Delta}^2} = \beta \lambda$$

where m_{Δ} is the variance of the difference of the true x variable (the signal) and σ_{Δ}^2 is the variance of the difference of measurement error in x (the noise). λ is the measure of

attenuation bias, and will be less than one as long as measurement error is present.¹⁸

With repeated observations over time on individuals who have some mismeasured variables, the value of some variables may change only slowly or not at all: gender, or education for those over the age of 30. In this case m_{Δ} is small but the noise (σ^2_{Δ}) remains, and so attenuation bias increases (λ falls) (Deaton, 1998:108).

Deaton (1998:100) suggests procedures for dealing with measurement error. First, if it is possible to estimate the size of the measurement error, a value for $\hat{\beta}$ can be found. Thus, if you have some idea of how large or small λ is, you can adjust point estimates accordingly for attenuation bias. Second, in regression analysis, it may be possible to use instrumental variable techniques to obtain consistent estimates of parameters even with measurement error. Attenuation bias can also be corrected for if estimates of measurement error variances and covariances can be obtained.¹⁹

Poterba and Summers (1986) examine the extent of measurement error bias in estimating labour market transitions from panel data. Their evidence from the US Current Population Survey (CPS) Initial and Reinterview Surveys indicates that many spurious transitions between labour market states are reported, and that estimated flow rates between states are very sensitive to this measurement error (Poterba and Summers, 1986:1320). Since we examine transition rates between labour market states from 1993 to

¹⁸ Adding more accurately measured independent variables into the regression exacerbates the problem: these variables soak up some of the signal of the mismeasured variable, but do not reduce the noise, so attenuation bias increases (Deaton, 1998:99).

¹⁹ Hertz (2000) corrects returns to schooling coefficients for South Africa, by estimating the measurement error variance of education variables, using KIDS as the external validation data set.

1998 in our descriptive statistics below, we should be aware of the potential for measurement error effects on flow rates. However, we do not have a reinterview survey which can be used to cross-check responses in 1993 or 1998; thus we are unable to verify these transitions.

(c) Attrition and selection bias

The second serious problem with panel data is attrition and the selection bias it imparts to inferences made from the data. In any longitudinal survey, individuals attrit from the sample when they move and cannot be found, or when they refuse to be re-surveyed. These attritors are not likely to be random observations. When there is attrition, coefficient estimates are potentially affected by selection bias. Consider the following selection model:

$$y_{i,t+1} = x_{i,t} \beta_1 + \varepsilon_{i,t}$$

$$A_{i,t+1} = x_{i,t} \beta_2 + z_{i,t} \lambda + v_{i,t}$$

where $y_{i,t+1}$ is the labour market outcome of the individual in period $t+1$ based on characteristics $x_{i,t}$ observed in period t , some of which may be invariant over time, and $A_{i,t+1}$ is the function describing attritors between period t and $t+1$. $Y_{i,t+1}$ is only observed for the subset of the sample for which $A_{i,t+1}$ is less than zero i.e. for those who do not attrit. If there is any correlation between the error terms $\varepsilon_{i,t}$ and $v_{i,t}$ then inconsistent parameter estimates for β_1 will be generated. This is the problem of selection bias.

To minimise attrition, the KIDS survey attempted to track Core persons from 1993 households if these individuals had moved, and this procedure yielded a sample re-survey rate 84 per cent. May *et al.* (1999:10) state that this rate ‘seems quite good’, given the time span and the mobility of the South African population and compared with panel surveys in other developing countries.²⁰

²⁰ It should be noted that in the most recent re-release of the KIDS

In a paper which looks at the extent of attrition bias in the KIDS survey, and the effect this has on statistical inference, Maluccio (2000) finds that there are household and community characteristics which differ between attriting and non-attriting households in the sample.²¹ This suggests a systematic process which underlies the attrition. He argues that it is important not only to consider the level of attrition in the data set, but also the extent to which particular model-specific analyses are affected by attrition bias. His point is that selection bias is not necessarily the result of an attrited data set.

Selection bias may occur for reasons other than attrition. For example, we concentrate on 1998 labour market outcomes for the 1993 jobless, who are not likely to be a random sample of KZN inhabitants in this year. They have features which selected them out of employment in 1993.

One of the ways in which selection bias is usually dealt with is to use a Heckman selection model, in which the final outcome is modeled as the result of two sequential processes. The finally observed outcome is then dependent on the individual observation being included at the first and second level estimations.

It is questionable as to whether it is always necessary or relevant to take account of selection effects. We are actually interested in those who do not have jobs in 1993, and what happens to them over time, and so it makes sense to focus on them although they are not likely to be a representative group. It may also not be possible to isolate selection

data, a total of 146 households found to have falsified questionnaires were dropped from the sample. This paper was originally written using KIDS version 2, but is reformulated here for the most recent version 3 release.

²¹ Maluccio (2000) finds that wealthier households tend to move, but are traceable, whereas larger households are less likely to move and households not residing in the former Natal province are more likely to not be traced in 1998.

processes clearly enough to define them in the model set-up.²²

We undertake our analysis in descriptive mode and model outcomes and choices for a particular sub-sample of individuals, realising that our conclusions might not be readily generalisable to the South African population. We content ourselves with being aware of the potential selection problems and are cautious in our interpretation where necessary.

(d) Defining the subsample²³

For this analysis, we use the matched sample of working age Africans who have records in 1993 and 1998, shown in Table 1.²⁴ The 1993 working age population was restricted to (inclusive) ages: 16-60 for women and 16-64 for men.²⁵

Table 1. Number of observations

| | Working age individuals | Total households | Relevant sample of individuals | Households in sample |
|--------------------|-------------------------|------------------|--------------------------------|----------------------|
| KIDS matched files | 2708 | 846 | 1689 | 716 |

In constructing the relevant sample of jobless Africans – unemployed and NEA – who could transit into jobs by 1998, we omit individuals from 1993 who are in formal education in 1998. This is a cleaner comparison between the remaining

²² One of the conditions for setting up a selection model is that some independent variables used in the selection equation must be excluded from the substantive equation. Deaton (1998: 104) notes that this is a particularly difficult exercise and not likely to be theoretically sensible in any but rare cases.

²³ Details on how the full data set was obtained and how dependent and independent variables were constructed appear in STATA do-files, available from the author on request.

²⁴ The KIDS is restricted to Africans and Indians, but the Indian subsample is too small to provide any useful racial variation.

²⁵ Because we are looking at how these people have changed labour market states over time, we track them in 1998 at whatever ages they are recorded at. With no mismeasurement, these (inclusive) ages should be 20-65 for women and 20-69 for men.

individuals, who are not augmenting their human capital while being NEA. Exclusion of those in education is standard practice in the South African literature (see for example Bhorat and Leibbrandt (2001), Wittenberg (2001)).

To restrict the 1993 and 1998 samples in the relevant way, the dependent variable describing each individual's labour market status was constructed. Individuals were classified as employed, searching unemployed or NEA according to their responses to a number of questions around hours of work and the activities they were involved in. It was also possible to generate a classification breaking down the NEA group into the discouraged or non-searching unemployed, those in education, housewives, retired people and others, including the disabled.

A brief description of which individual and household level variables are used in section 6 is presented in Table A1. In the next section, we look at actual transition rates between labour market states for the 2708 matched data files in the KIDS survey.

*(e) Transition rates*²⁶

Transition rates were calculated by dividing the number of observations in state *i* in 1993 which are in state *j* in 1998 by the total number of observations in state *i* in 1993. The transitions for African men are given in Tables 2.1a and 2.1b and for African women in Tables 2.2a and 2.2b. Both tables indicate a fairly high degree of mobility into and out of employment.²⁷ About 70 per cent of all men and 60 per cent of all women who had jobs in 1993 were still employed in 1998.

Table 2.1a. Transitions - African men

²⁶ These transition tables were calculated on the raw data from the matched data set. No weights were used to calculate the transition ratios.

²⁷ Within the transition rates matrices, elements on diagonal represent the degree of immobility for each state.

| Labour Market Status 1998 | | | | | |
|---------------------------|---------|---------|-------|-------|------|
| Labour | | Empl 98 | UE 98 | NEA98 | Tot |
| Market | Empl 93 | 273 | 18 | 93 | 385 |
| Status | UE 93 | 44 | 10 | 46 | 100 |
| 1993 | NEA 93 | 186 | 57 | 390 | 633 |
| | Total | 503 | 85 | 529 | 1118 |

Table 2.1b. Transition rates - African men

| Labour market status 1998 | | | | |
|---------------------------|---------|---------|-------|-------|
| Labour | | Empl 98 | UE 98 | NEA98 |
| Market | Empl 93 | 0.71 | 0.05 | 0.24 |
| Status | UE 93 | 0.44 | 0.10 | 0.46 |
| 1993 | NEA 93 | 0.29 | 0.09 | 0.62 |

Table 2.2a. Transitions - African women

| Labour Market Status 1998 | | | | | |
|---------------------------|---------|---------|-------|-------|------|
| Labour | | Empl 98 | UE 98 | NEA98 | Tot |
| Market | Empl 93 | 227 | 14 | 140 | 381 |
| Status | UE 93 | 38 | 3 | 37 | 78 |
| 1993 | NEA 93 | 274 | 53 | 804 | 1131 |
| | Total | 539 | 70 | 981 | 1590 |

Table 2.2b. Transition rates - African women

| Labour Market Status 1998 | | | | |
|---------------------------|---------|---------|-------|-------|
| Labour | | Empl 98 | UE 98 | NEA98 |
| Market | Empl 93 | 0.60 | 0.04 | 0.37 |
| Status | UE 93 | 0.49 | 0.04 | 0.47 |
| 1993 | NEA 93 | 0.24 | 0.05 | 0.71 |

Large parts of the non-working labour force also found jobs by 1998: over 44 per cent of all men and women searching for work were employed five years later, and almost 25 per cent of women and 30 per cent of men who were NEA 93 were working in 1998. However, men and women in the NEA 93 category were highly likely to remain in the same state in 1998. Less than 10 per cent of the NEA 93 transited into the category of searching unemployment by 1998.

There was also a substantial movement out of the searching unemployed category in 1998: almost half of African men and women transited either into employment or out of the labour force. This suggests that over time, individuals who search either get jobs or prefer to stop searching. In addition, there are very small numbers of people transiting into searching unemployment from the other labour market

states: this could indicate that in 1998, the value of being in search unemployment U is below the value of employment W and non-activity O for most individuals in the matched data set.

Table 3.1. Detailed labour market transitions - African men

| Labour market status 1993 | Labour market status 1998 | | | | | | |
|------------------------------|---------------------------|-----------------|----------------|-------------------|-----------|-----------|-----------------|
| | Empl. 98 | 98 Search UE | 98 Disc. UE | 98 House- work | 98 School | 98 Retire | 98 Other NEA |
| Empl. 93 | 0.71 | 0.05 | 0.16 | 0.01 | 0.01 | 0.06 | 0.02 |
| 93 Search UE | 0.44 | 0.10 | 0.37 | 0.06 | 0.01 | 0.01 | 0.01 |
| 93 Disc. UE | 0.39 | 0.12 | 0.35 | 0.04 | 0.03 | 0.03 | 0.02 |
| 93 Housework | 0.80 | 0.00 | 0.20 | 0.00 | 0.00 | 0.00 | 0.00 |
| 93 School | 0.25 | 0.08 | 0.32 | 0.01 | 0.31 | 0.00 | 0.03 |
| 93 Retire | 0.11 | 0.00 | 0.05 | 0.05 | 0.00 | 0.68 | 0.11 |
| 93 Other NEA | 0.19 | 0.04 | 0.33 | 0.11 | 0.06 | 0.13 | 0.15 |

Tables 3.1 and 3.2 show transition rates for the full spectrum of labour market states. They add more nuance to the picture of mobility described above and show where people move to when they transit out of search unemployment.

Men who lose or leave their jobs by 1998 move predominantly into non-searching unemployment, while women move into non-search and housework. This could indicate that the menu of acceptable alternative activities to searching unemployment is wider for women than men.

Table 3.2. Detailed labour market transitions - African women

| Labour market status 1993 | Labour market status 1998 | | | | | | |
|------------------------------|---------------------------|-----------------|----------------|-------------------|-----------|-----------|-----------------|
| | Empl. 98 | 98 Search UE | 98 Disc. UE | 98 House- work | 98 School | 98 Retire | 98 Other NEA |
| Empl. 93 | 0.60 | 0.04 | 0.14 | 0.11 | 0.02 | 0.09 | 0.01 |
| 93 Search UE | 0.49 | 0.04 | 0.35 | 0.09 | 0.04 | 0.00 | 0.00 |
| 93 Disc. UE | 0.31 | 0.06 | 0.44 | 0.13 | 0.04 | 0.01 | 0.01 |
| 93 Housework | 0.25 | 0.02 | 0.21 | 0.44 | 0.01 | 0.06 | 0.01 |
| 93 School | 0.18 | 0.07 | 0.34 | 0.05 | 0.35 | 0.00 | 0.01 |
| 93 Retire | 0.04 | 0.00 | 0.04 | 0.04 | 0.00 | 0.84 | 0.04 |
| 93 Other NEA | 0.28 | 0.03 | 0.21 | 0.18 | 0.09 | 0.11 | 0.09 |

The dynamics of those in school in 1993 are particularly interesting. About one fifth of women and one quarter of men transit from school into employment in 1998. The

alternatives for those who do not get or want jobs are either to remain in school (over 30 per cent of both genders in 1998) or to transit into the category of non-searching unemployment. Only a very small proportion of school leavers choose to start searching for work in 1998.

It is important to note that the KIDS presents a set of snapshots on the same individuals from which we can infer something about the dynamics of labour force attachment. It is not possible to say whether the employed in 1993 who are still employed in 1998 are in the same jobs, or whether they changed jobs or experienced a spell of unemployment in the five year period.

(f) Additional summary statistics

Recall that the subpopulation whose outcomes we are interested in modelling is the set of jobless (searchers and NEA) individuals in 1993 who are resident household members in 1993. We track their employment status in 1998, although do not impose the condition that they should remain household members in 1998. As previous work on South Africa has suggested, one of the most powerful ways an individual can ensure persistent unemployment is to remain within a rural household.

Of the individuals who are jobless and household members in 1993, 32 per cent are no longer resident household members in 1998 although they are still attached to the household in some way. Table 4 shows that from the sample of jobless individuals in 1993, almost 45 per cent of the people who do become employed by 1998 are no longer resident household members.

Table 4. 1998 Household residency and labour market status for 1993 jobless sample

| | Resident HH member in 1998 | Non-resident HH member in 1998 | Total |
|-----------------|-------------------------------|-----------------------------------|-------|
| Jobless in 1998 | 841 | 306 | 1147 |
| Working in 1998 | 301 | 241 | 542 |
| Total | 1142 | 547 | 1689 |

Thus, not only is there a fair amount of labour market

mobility in our sample, there also seems to be movement out of households. We constructed means of the independent variables for the subpopulation of jobless men and women in 1993, and Table A2 presents these results. Of those who become employed from the pool of jobless, a much larger proportion were searching in 1993: about 20 per cent of the working 1998 men were searching in 1993 and about 12 per cent of the working 1998 women. For many of the other individual and household level variables, there is little observable difference between the means for those working by 1998 and those who are still jobless. One point is worth mentioning with respect to the jobless in 1998: over 50 per cent of this group live with migrants, and more of them (compared to those men who work, and all African women) live with pensioners. One interpretation of this is that the household might function more as a safety net for those without jobs than as a source of finances for successful search activities. African men with a higher b might be able to reduce search efforts as $O > U$.

Our results in section 6 show that some household composition variables do affect the success of search activities, and that there are additional unobserved household effects which help to explain why some individuals find work and others do not.

6. RESULTS

The results of the logit estimations for the matched sample are presented in columns (1) to (3) in Table 5. Coefficients are interpreted as the odds ratio of being employed in 1998, given the set of independent variables. An odds ratio greater than one implies that the variable increases the odds of the individual being in employment; an odds ratio less than one implies that the variable decreases the odds of the individual being in employment. The results suggest that:

(a) *Search effort matters*

Search efforts may indicate individual motivation to find work or perhaps the presence of resources for financing search. In Table 5 column (3), it is clear that search activities in 1993 have a large and significant impact on the probability of obtaining a job in 1998. Women who are looking for work in 1993 have a more than 150 per cent increase in their baseline odds of finding a job compared to non-searchers, while men have a 60 per cent increase in their odds of employment.

The finding that the activity of search has such a large positive impact on search success may seem somewhat obvious, but it is crucial to observe this variable having the effect it does, particularly since our transition matrices indicated that people also transit out of the NEA and non-searching unemployed states into jobs by 1998. It suggests that search does matter for getting a job – that sitting around waiting for a job to arrive, as in the case of discouraged workers, reduces the chances of finding work relative to searchers. Investment in search was apparently a worthwhile investment in the KZN labour market in 1993, particularly for women.

(b) Residing in an urban area reduces search success

The coefficient on the urban dummy for the matched sample indicates that individuals in urban areas have reduced chances of finding employment. Compared to the baseline odds of employment in rural areas, individuals in urban households have an almost 20 per cent lower chance of finding work by 1998. Since it was indicated in section 5 that a fair number of the jobless sub-sample were no longer living in the same household as 1993, but were still attached as members, it may be the case that individuals from rural households in 1993 moved to urban areas and so were able to find jobs. Those observed as jobless in urban areas in 1993 were initially less employable than these rural movers and so were not successful by 1998. However, this effect – while statistically

insignificant – is indeed surprising, and may simply be an artefact of the geographical region in which the sample was taken.

(c) Pensioners and household size matter

Having a larger proportion of men and women of pensionable age in the household reduces search success probability for men, probably through raising the actual income part of b . In a ten-person household, replacing one adult female with one male or female pensioner reduces the odds ratio of employment probability for men by about 99 . This is a large and statistically significant effect. It is consistent with the theoretical possibility that $O > U$ even though W might be larger than O : men have a lower incentive to go out and be successful searchers. This finding supports the story of reduced adult male labour supply in response to pensions in the household told by Bertrand et al. (2003).²⁸

Female search success is affected negatively by having old men in the household, possibly through a combination of raising b and increasing the opportunity costs of leaving the household to enter the labour market. These older men are associated with almost a 99 per cent reduction in the odds of employment for women, although the coefficient is only significant at the 10 per cent level. This negative pensioner effect on the chances of finding work is reversed for women, when there are female pensioners in the household. One extra female pensioner in 1993 increases the odds of a women finding a job by 1998 by almost 430 per cent. Thus, the presence of a male or female pensioner does not affect b in the same way for women.

²⁸ Although the presence of pension income should be captured in the log per capita income variable in the regression, as mentioned above, the income variable is very poorly measured in the 1993 Saldrú survey. This means that the coefficient on the proportion of pension-age adults in the household may contain some of the pension income effect.

Table 5. Predicting probability of an unemployed 1993 individual being employed by 1998 logit results (odds ratios).

| Independent variables | (1) Simple | | (2) Fixed effects | | (3) Random effects | |
|------------------------------------|----------------------------------|-----------------------------|----------------------------|----------------------------------|----------------------------------|----------------------------|
| | Women | Men | Women | Men | Women | Men |
| Gender (base = female) | | 10.2735 (22.8729) | | 1.932 (6.3131) | | 10.0255 (24.1459) |
| Age 98 | 1.2106* (0.0492) | 1.2355* (0.0855) | 1.267* (0.108) | 1.3327* (0.1236) | 1.2541* (0.0801) | 1.2694* (0.0908) |
| Age ² 98 | 0.9974* (0.0005) | 0.9971* (0.0008) | 0.9967* (0.0011) | 0.9962* (0.0012) | 0.9969* (0.0008) | 0.9968* (0.0009) |
| Married in 93 (base=unmarried) | 1.0536 (0.2285) | 1.5372 (0.4685) | 1.027 (0.3311) | 2.0211 (0.8561) | 1.0444 (0.2365) | 1.7479 (0.6087) |
| Urban 93 (base=rural) | 0.7658 (0.2372) | 0.7572 (0.1654) | | 0.9105 (0.427) | 0.7877 (0.1942) | 0.7982 (0.2199) |
| LMS 93 ¹ (base=NEA) | 2.1516* (0.5559) | 1.452 (0.4021) | 2.5507 (1.0664) | 2.0308 ^{^^} (0.8456) | 2.5266* (0.7823) | 1.5971 (0.4588) |
| <i>Education 98 (base=none)</i> | | | | | | |
| Grade1-Std1 | 2.5343 (0.9784) | 0.961 (0.496) | 1.4907 (0.9035) | 0.7032 (0.591) | 2.6986 (1.1357) | 0.9822 (0.5885) |
| Std 2 - 5 | 2.0418* (0.5349) | 0.8033 (0.2873) | 1.427 (0.6776) | 0.6885 (0.4238) | 2.1612 (0.739) | 0.8344 (0.3725) |
| Std 6 - 9 | 2.5903* (0.7509) | 1.7613 (0.7852) | 1.4035 (0.7738) | 1.8107 (1.1797) | 2.7377* (1.0503) | 1.9983 (0.9384) |
| Matric | 3.7413* (1.3207) | 3.5494* (1.5033) | 1.7352 (1.0554) | 3.0196 (2.13) | 4.099* (1.7605) | 4.2507* (2.1888) |
| Postmatric | 21.6627* (11.5915) | 10.8208* (8.6577) | 3.5275 (4.1235) | 4.5183 (5.4377) | 24.7142* (18.2224) | 15.5243 (16.504) |
| <i>HH variables 93²</i> | | | | | | |
| Proportion infants | 3.7878 (4.04) | 0.2364 (0.3641) | | 0.0454 (0.1168) | 4.7758 (5.0776) | 0.3066 (0.4715) |
| Proportion kids | 1.4324 (1.2452) | 0.2191 (0.3303) | | 1.5328 (4.181) | 1.4505 (1.5496) | 0.2802 (0.4145) |
| Proportion m. youth | 2.4951 (2.3918) | 0.3773 (0.5408) | | 0.4733 (1.3275) | 2.9006 (3.2893) | 0.5274 (0.7928) |
| Proportion f. youth | 0.5543 (0.5162) | 0.2812 (0.4419) | | 0.9684 (2.5715) | 0.5939 (0.6524) | 0.4032 (0.6186) |
| Proportion m. adults | 2.4849 (3.357) | 0.185 (0.3547) | | 0.0028 (0.0102) | 3.1363 (4.4823) | 0.1713 (0.2908) |
| Proportion old men | 0.0247 ^{^^} (0.0529) | 0.0028 (0.0068) | | 0.5149 (2.585) | 0.0113 ^{^^} (0.0262) | 0.0007 (0.0021) |
| Proportion old women | 4.3594 (5.1681) | 0.0224* (0.0312) | | 0.0163 (0.0543) | 5.2689 (7.4761) | 0.0152 (0.0263) |
| Independent variables | (1) Simple | | (2) Fixed effects | | (3) Random effects | |
| | Women | Men | Women | Men | Women | Men |

| | | | | | |
|-----------------------|--------------------------|----------------------------|---------------------------------|----------------------------|--------------------|
| HH variables 93 | | | | | |
| LnHH income (p.c.) | 0.9055 (0.1043) | 0.8786 (0.1137) | 0.9118 (0.2035) | 0.8934 (0.0968) | 0.8731 (0.1239) |
| Average adult edn | 0.9856 (0.0342) | 0.9643 (0.048) | 0.9538 (0.0926) | 0.9862 (0.0427) | 0.9603 (0.0551) |
| Total adults working | 1.1238 (0.1143) | 0.8314 (0.1101) | 0.9021 (0.1951) | 1.1294 (0.1311) | 0.8533 (0.1177) |
| Migrant in HH 93 | 0.8915 (0.1634) | 0.4529* (0.0873) | 0.5461 (0.2118) | 0.8527 (0.1694) | 0.9192 (0.1885) |
| Household size | 0.935 (0.0264) | 1.002 (0.0357) | 1.092 ^{^^} (0.0581) | 0.9249* (0.0252) | 0.997 (0.0342) |
| Sample n | 1689 | | 806 | 1689 | |
| Pseudo-R ² | | | 0.1507 | | |

Notes: Standard errors in brackets.

Significance at 1% level is (bold with *), 5% (bold), 10% (^{^^})

Sigma_u is an estimate of u_i , the household effect

\square is the proportion of the residual due to the household effect

1 LMS 93 indicates that the person was searching unemployed in 1993

2 Excluded household composition category is adult female

We begin to see how the composition of the household directly impacts on the individual, and on outcomes of search success in the labour market, and how it matters differently for men and women.

Household size has a significant negative effect on female labour market success, perhaps because larger households require more housework and so raise O>U. The effect is also negative for men, possibly indicating that with more people in the household, resources for investing in effective search activities are limited. However, the coefficient on household size is only statistically significant for women, and since it is so close to one, the economic significance of this channel in affecting the probability of finding work is questionable.²⁹

(d) *The effect of labour market contacts on search success is uncertain*

²⁹ As my anonymous referee point out, this lack of economic significance is particularly stark, if one considers that a one unit change in household size is about a three standard deviation increase in this variable for women not working in 1998. See Table A.2 for household size averages.

The evidence for households having positive measurable effects on labour market outcomes via contacts is not convincing. Migrants and the total number of working adults in the household in 1993 have a negative effect on employment probability by 1998 for men, but neither of these coefficients are statistically different from zero in the RE logit. These types of households may support inactivity by increasing b rather than increasing search activity and success directly through raising P ; however, we cannot tell whether this is the case with this data. At the very least, if people are relying on labour market links outside of the house, these links do not seem to be very successful at delivering the information or the jobs.

(e) Unmeasurable households effects have a large impact on search outcomes

The contribution of unmeasurable household effects over time to explaining search success and failure is large. Twenty per cent of the unexplained residual in our employability estimations for the matched sample is accounted for by unobserved household effects. Even after controlling for household composition, living in a particular household with a particular set of people has significant unobserved effects on employment probability over time. These effects may be emanating from the household providing labour market information, or from the culture of the household, or some combination of the two.

From the five main results extracted here, we have tried to indicate how search models are relevant for understanding individual labour market transitions in South Africa, and that such models cannot sensibly omit the context of the household. Furthermore, this context needs to be incorporated in a way that takes account of potential endogenous effects.

7. CONCLUSIONS

We have applied a search theory of labour supply choices and outcomes to a sub-set of South African individuals in KwaZulu-Natal and shown how households mediate between these individuals and the labour market.

At the most obvious level, our results confirm the validity of using the search framework in the South African labour market. Controlling for household level heterogeneity using a random effects logit on panel data, we find that search effort matters for search success of men and women. That search has a positive effect on employment probability supports the notion that labour markets are characterised by frictions of time, space and information. In South Africa, overcoming these frictions is not a trivial task for the individual, and the household plays both a supporting and constraining role in individual labour supply choices.

The four different household themes identified in the South African literature have all been brought to bear on our interpretation of the results. We find that households with more pensioners (especially male) and working adults may act as support nets rather than as resource-providers for successful search activity. We also see that the composition of the household has a different influence on female labour supply choices and outcomes, compared to men, supporting the theme that households are productive units requiring different inputs from members.

The direct evidence from our sample for household contacts facilitating search activity and success is not strong. However, some of these information effects as well as other unobservable effects of motivation and search effort are presumably captured in the residual attributable to household effects.

Finally, despite the dangers of using panel data highlighted in section 5 and the fact that the gap between the surveys is fairly long, at five years, this study has been methodologically useful. Separating out household formation from labour market outcomes and labour supply choices has

previously not been done convincingly on South African cross section data alone. We have been able to show how households influence subsequent labour market outcomes for individuals. Further research would obviously benefit from longer panel studies conducted at more regular intervals. New data from the South African Labour Force Surveys should obviously help, in this regard. The ability to construct labour market transition histories for South Africans would considerably add to our understanding of the interactions between individuals and their household contexts in the labour market.

APPENDIX

Table A1: Independent variables used in econometric exercise

| Variable | Description | Hypothesised impact on components of value functions of search model ¹ |
|---------------------------------------|--|---|
| Gender | Male=1, female=0 | All components |
| Age, Age ² | Based on 1998 age | rW and P |
| Marital Status 1993 | Married=1, otherwise=0 | P (the effect could differ for men and women) |
| Urban93 | Urban/metro=1, rural=0 | P, c |
| 1993 labour market status | Searching unemployed=1, NEA=0 | P (Gives some indication of the prior motivation of the individual who is willing to search for a job in 1993, compared to others who are NEA) |
| Education group 1998 | Presumably, the education level of the individual in 1993 matters; however, any additional education between 1993 and 1998 makes an individual more employable (thus rW increases). For this reason, the most recent education data from 1998 is used. | rW, P |
| Household composition variables 1993 | Generated as the proportion of individuals in the household belonging to the following groups: infants (0-6), children (7-15), male & female youths (separately) (16-29), male & female adults (separately) (31-59) and male & female old people (separately) (over 60). | b and c (for men and women differently) |
| Log per capita income 1993 | Based on household size including resident household members. | only b ¹ |
| Average adult education in HH 1993 | Based on resident HH adults. | P (through assistance they might provide in finding and exploiting labour market information). |
| Num. Working age adults in HH in 1993 | Women age 16-60, men aged 16-64 inclusive; HH resident members only | P (potentially direct labour market links). |
| Migrant worker 1993 | Migrant=1 if at least one migrant attached to HH, otherwise=0. The variable does not specify whether migrant is employed or not. | P, c (if people are in distant labour markets, the HH itself has access to more information about job opportunities, whether or not the migrant has a job). |
| Household size 1993 | Includes only those members resident at least 15 days in the month prior to the interview. | a, b, c (differently for men and women). |

¹. There were several missing values for household income. All missing values were set to zero & a dummy variable was included to mark these missing values. This dummy was

always insignificant. Note also that the KIDS did not collect wage & income data in 1998 very well; thus there are many imputed values for household income in 1998

Table A2. Means of independent variables for the jobless 1993, by 1998 labour market status

| Independent variables | African women | | African men | |
|--|---------------|----------------|-------------|----------------|
| | Working 98 | Not working 98 | Working 98 | Not working 98 |
| Age in 98 | 33.4872 | 36.5335 | 31.0743 | 34.0021 |
| | 0.4376 | 0.4204 | 0.6771 | 0.7704 |
| Proportion married 93 * | 0.3911 | 0.4333 | 0.1913 | 0.1956 |
| | 0.0336 | 0.0282 | 0.0275 | 0.0245 |
| Proportion in urban areas * | 0.2110 | 0.1895 | 0.2732 | 0.2677 |
| | 0.0641 | 0.0566 | 0.0730 | 0.0695 |
| Proportion searching unemployed 93 * | 0.1218 | 0.0495 | 0.1913 | 0.1379 |
| | 0.0203 | 0.0104 | 0.0294 | 0.0195 |
| Ave yrs of education, 1998 | 8.2431 | 6.7125 | 8.6939 | 7.1155 |
| | 0.3240 | 0.3177 | 0.3675 | 0.3066 |
| Prop infants 93 (age<=6)~ | 0.1800 | 0.1718 | 0.1253 | 0.1268 |
| | 0.0085 | 0.0090 | 0.0088 | 0.0088 |
| Prop kids 93 (6<age<=15)~ | 0.2558 | 0.2573 | 0.2265 | 0.2223 |
| | 0.0108 | 0.0095 | 0.0128 | 0.0099 |
| Prop m. youth 93 (15<age<=30)~ | 0.1105 | 0.1078 | 0.2329 | 0.2149 |
| | 0.0076 | 0.0083 | 0.0124 | 0.0111 |
| Prop f. youth 93 (15<age<=30)~ | 0.1960 | 0.2029 | 0.1373 | 0.1412 |
| | 0.0088 | 0.0064 | 0.0087 | 0.0071 |
| Prop m. adult 93 (30<age<=59)~ | 0.0717 | 0.0639 | 0.0956 | 0.0989 |
| | 0.0071 | 0.0057 | 0.0104 | 0.0088 |
| Prop f. adult 93 (30<age<=59)~ | 0.1264 | 0.1299 | 0.1228 | 0.1107 |
| | 0.0070 | 0.0045 | 0.0075 | 0.0052 |
| Prop old male 93 (60+)~ | 0.0138 | 0.0228 | 0.0158 | 0.0282 |
| | 0.0024 | 0.0026 | 0.0033 | 0.0036 |
| Prop old female 93 (60+)~ | 0.0458 | 0.0435 | 0.0437 | 0.0571 |
| | 0.0044 | 0.0037 | 0.0053 | 0.0047 |
| Log of HH income per capita 93 | 4.4592 | 4.4579 | 4.3448 | 4.4807 |
| | 0.0843 | 0.0681 | 0.1161 | 0.0730 |
| Ave adult edn in HH 93 | 6.5112 | 6.0542 | 6.6233 | 6.3403 |
| | 0.2662 | 0.2817 | 0.2485 | 0.2584 |
| Total # working adults in HH 93 | 0.8040 | 0.7547 | 0.8603 | 0.8842 |
| | 0.0673 | 0.0769 | 0.0877 | 0.0843 |
| Total # adults in HH 93 | 4.5231 | 4.9930 | 5.4652 | 5.5735 |
| | 0.1544 | 0.1897 | 0.2032 | 0.1876 |
| Prop of working adults in HH 93 | 0.1760 | 0.1512 | 0.1686 | 0.1623 |
| | 0.0154 | 0.0152 | 0.0168 | 0.0162 |
| Prop living with migrant in HH * (at least one in 93) | 0.4458 | 0.5124 | 0.3438 | 0.5092 |
| | 0.0466 | 0.0444 | 0.0450 | 0.0411 |
| Ave Hhsize 93 | 8.2519 | 9.0302 | 8.8490 | 9.0233 |
| | 0.2703 | 0.3302 | 0.3270 | 0.3093 |

n (number of jobless working age people in 1993)

1689

Notes: This Table displays survey-weighted means, with standard errors beneath.

*proportion of relevant category (employed/not in 1998) with this feature eg: married ~proportion of hh members of that type (e.g. infants) living in hhs with individuals in the relevant category (employed/not in 1998)

REFERENCES

- BECKER, G. (1965). 'A theory of the allocation of time', *Economic Journal*, 75: 493-517.
 BERTRAND, M., MULLAINATHAN, S. AND MILLER, S. (2003). 'Public Policy and Extended Families: Evidence from Pensions in South Africa', *World Bank Economic Review*, 17(1): 27-50.

- BHORAT, H., M. LEIBBRANDT, M. MAZIYA, S. VAN DER BERG AND I. WOOLARD (2001). *Fighting Poverty: Labour markets and inequality in South Africa*. Cape Town: UCT Press.
- BHORAT, H. AND M. LEIBBRANDT (2001). 'Modelling vulnerability and low earnings in the South African labour market', Chapter 4 in Bhorat, H. et. al. (Eds.), *Fighting Poverty: Labour markets and inequality in South Africa* (2001). Cape Town: UCT Press.
- CARTER, M. AND J. MAY (2000) 'One kind of freedom: Poverty dynamics in post-apartheid South Africa', Working paper, Department of Agricultural and Applied Economics, University of Wisconsin, Madison.
- CICHELO, P., G. FIELDS AND M. LEIBBRANDT (2000). 'Which African workers have gotten ahead and by how much? The story of KwaZulu-Natal, South Africa, 1993-1998', Paper presented at the Trade and Industrial Secretariat (TIPS) Annual Forum. <http://www.tips.org.za/profile/activities/policy/2000papers.htm>
- DEATON, A. (1998). *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*. Baltimore: Johns Hopkins University Press.
- DEVINE, T.J. AND N.M. KIEFER (1991). *Empirical Labour Economics: The Search Approach*. Oxford: Oxford University Press.
- DINKELMAN, T. (2002). 'Investigating the impact of households on search choice and outcomes in the KwaZulu-Natal labour market 1993-1998' Unpublished dissertation submitted in partial fulfilment of the MComm degree, University of the Witwatersrand, Johannesburg.
- _____ and F. PIROUZ (2002). 'Individual, household and regional determinants of labour force attachment in South Africa: Evidence from the 1997 October Household Survey', *South African Journal of Economics*, June, 70:5.
- GREENE, W. (1997). *Econometric Analysis* (3e), New Jersey: Prentice-Hall.
- HERTZ, T. (2000). 'Upward Bias in the Estimated Return to Education: Evidence from South Africa', Doctoral paper, University of Massachusetts at Amherst.
- INTERNATIONAL LABOUR ORGANISATION (ILO). (1996). *Restructuring the Labour Market: The South African Challenge*, Geneva: ILO Publications Office.
- KESWELL, M. (2000). 'Labour market dynamics in South Africa: Evidence from KwaZulu-Natal province', Paper presented at the TIPS Annual Forum, September. Available at <http://www.tips.org.za/profile/activities/policy/2000papers.htm>
- KINGDON, G. AND J. KNIGHT (2000). 'Are searching and non-searching unemployment distinct states when unemployment is high? The case of South Africa', Centre for the Study of African Economies (CSAE) Working Paper 2, Oxford University.
- KLASEN, S. AND I. WOOLARD (1998). 'Unemployment, household formation, poverty and nutrition in South Africa', Paper presented at Workshop on Unemployment at the University of Pretoria, April 1998, mimeographed: University of Munich.
- _____ (2000). 'Surviving unemployment without state support: Unemployment and household formation in South Africa', IZA Discussion Paper No. 237, Bonn.
- LEIBBRANDT, M. AND I. WOOLARD (2000). 'An analysis of income inequality and income mobility of African households in KwaZulu-Natal, 1993 and 1998', Paper presented at the 5th Annual Conference on Econometric Modelling for Africa, Univ of the Witwatersrand, July 2000.
- LIPPMAN, S. A. AND J. MCCALL (1976A). 'The economics of job search: a survey', *Economic Inquiry*, June, 14: 155-189, part 1.
- _____ (1976a). 'The economics of job search: a survey', *Economic Inquiry*, Sept, 14: 347-366, part 2.
- MALUCCIO, K., D. THOMAS AND L. HADDAD (1999). 'The mobility of adults in post-apartheid South Africa: Levels, determinants and consequences', Paper prepared for PAA meetings, March 1999.
- _____ (2000). 'Attrition in the KwaZulu-Natal Income Dynamics Study, 1993-1998', FCND Discussion Paper No. 96, International Food Policy Research Institute (IFPRI).
- MAY, J., M. CARTER, L. HADDAD AND J. MALUCCIO (1999) 'KwaZulu-Natal income dynamics study (KIDS) 1993-1998: A longitudinal household data set for South African policy analysis',

- Centre for Social and Development Studies Working Paper Number 21, *Durban: University of Natal.*
- MORTENSEN, D. (1986). 'Job search and labour market analysis', in O. Ashenfelter and R. Layard (Eds.), *Handbook of Labour Economics*, Vol 2., Ch 15: 849-918, *Amsterdam: North Holland.*
- ____ AND C. A. PISSARIDES (1999). 'New developments in search theory', in O. Ashenfelter and D. Card (Eds.) *Handbook of Labour Economics*, vol 3. Ch 39:2567-2627. *Amsterdam, New York: Elsevier.*
- PISSARIDES, C. (2000). *Equilibrium Unemployment Theory*, *Cambridge, MA: MIT Press.*
- POTERBA, J. AND L. SUMMERS (1986). 'Reporting Errors and Labour Market Dynamics', *Econometrica*, 54(6): 1319-1338.
- SOUTH AFRICAN LABOUR DEVELOPMENT RESEARCH UNIT (SALDRU) (1993) 'Project for Living Standards and Development (PSLSD)', *University of Cape Town, School of Economics. Coding manual, survey questionnaire and 1993 data.*
- STATA CORP. (2001). *Stata Statistical Software Release 7.0. College Station, TX: Stata Corporation.*
- STIGLER, G. (1962). 'Information in the labour market', *Journal of Political Economy*, October supplement, 5: 94-105, pt 2.
- WITTENBERG, M. (1999). 'Job search and household structure in an era of mass unemployment: a semi-parametric analysis of the South African labour market', *Econometric Research Southern Africa (ERSA) Working Paper No.3, University of the Witwatersrand, Johannesburg.*
- ____ (2001). 'Spatial dimensions of unemployment', *Paper presented at the Development Policy Research Unit (DPRU)/Friedrich Ebert Stiftung (FES) Conference on Labour Markets and Poverty in South Africa, November. Available at <http://www.uct.ac.za/depts/dpru/fes01.htm>*
- UNIVERSITY OF NATAL, UNIVERSITY OF WISCONSIN AND INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE (IFPRI). (2002). 'KwaZulu-Natal Income Dynamics Study (KIDS) – version 3', *Overview of data files, data set.*
- ____ (1998). 'KwaZulu-Natal Income Dynamics Study (KIDS)', *Coding manual, survey questionnaire, data set.*