THE UNIVERSITY OF MELBOURNE

Melbourne Institute Working Paper Series

Working Paper No. 34/07

What Drives Worker Flows?

Chew Lian Chua, Robert Dixon and G. C. Lim

MELBOURNE INSTITUTE
of Applied Economic and Social Research

What Drives Worker Flows?

Chew Lian Chua†, Robert Dixon# and G. C. Lim†

†Melbourne Institute of Applied Economic and Social Research, The University of Melbourne

#Department of Economics, The University Melbourne

Melbourne Institute Working Paper No. 34/07

ISSN 1328-4991 (Print)

ISSN 1447-5863 (Online)

ISBN 978-0-7340-3267-6

November 2007

Melbourne Institute of Applied Economic and Social Research
The University of Melbourne
Victoria 3010 Australia
***Telephone* (03) 8344 2100**
***Fax* (03) 8344 2111**
***Email* melb-inst@unimelb.edu.au**
***WWW Address* http://www.melbourneinstitute.com**

Abstract

This paper applies a multi-state latent factor intensity model to worker flows to obtain insights about the determinants of entry and exit rates pertaining to various labour market states. The analysis shows that one activity factor underpins the decision to move from employment and from unemployment and this result may be of special interest to policy makers concerned with understanding the rate of departures from the pool of both the employed and (especially) the unemployed. The paper also shows how to estimate a non-linear state space model using a Gibbs sampler that encompasses a Metropolis-Hastings algorithm as well as the auxiliary particle filter to estimate the latent process. The advantage of the approach is that it provides a parsimonious and efficient way to obtain key information about behaviour in labour markets.

Keywords: Bayesian statistics; Gibbs sampler; Metropolis-Hastings; Auxiliary particle filter; Worker flows; unemployment rate**.**

JEL classification: J60; E24; E32; C11

1 Introduction

Understanding the dynamics of labour markets has become increasingly important and there has been a steady build-up of empirical studies of worker flows.[1] The focus in these studies (as in the theoretical literature) is usually on the flows into and out of unemployment and especially the flows each way between unemployment and employment.[2]

Most studies appeal to a matching model (see Petrongolo and Pissarides (2001) for a survey of work along these lines) and relate these worker flows to the state of the labor market or the state of the business cycle as indicated by the unemployment rate or the level of job vacancies relative to the number unemployed. Alternatively, Hall (2005) and Shimer (2005)) rely on wage rigidities for an explanation of the fluctuations in the unemployment rate. The job-finding rate is treated as cyclical with the bargaining power of 'insiders' being dependent, at least in part, on the level of job vacancies relative to the number unemployed. While there has been much debate recently (concerning, inter alia, the relative importance of fluctuations in job separation and finding rates in driving the unemployment rate up and down), a common tread running through the literature is the role of business cycles in influencing flows between labour market states.

Recently, Hall (2003, 2005), Burgess and Turon (2005) and Dixon, Freebairn and Lim (2006) drew attention to unemployment 'entry' and 'exit' rates which are made up of combinations of various flows. The unemployment entry rate is defined as the size of the flow from employment to unemployment plus the size of the flow from not in the labour force to unemployment in any period, both measured relative to the number employed at the start of the period. The unemployment exit rate is defined as the size of the flow from

[1]The origin of empirical work in this area can be traced back to Singer (1939a and b). In recent times, works by Dale Mortensen (1970), Christopher Pissaridis (1986, 2000), Mortensen and Pissarides (1994), Robert Hall (2003, 2005 a,b,c) and Robert Shimer (2005), amongst others, have been seminal.

[2]Most authors study 'separation' and 'finding' rates. The separation rate is defined as the size of the flow from employment to unemployment in any period relative to the number employed at the start of the period. The finding rate is defined as the size of the flow from unemployment to employment in any period relative to the number unemployed at the start of the period respectively.

unemployment to employment plus the size of the flow from unemployment to not in the labour force in any period, both measured relative to the number unemployed at the start of the period. Again, it is common to relate these entry and exit rates to the state of the labor market or the state of the business cycle (see Burgess and Turon (2005) for an example). It has also been shown that the equilibrium unemployment rate implied by these entry and exit rates is highly correlated with the observed unemployment rate.

The aim of this paper is to extract the latent factor that determines the flow of labour between the three states in the labour force, namely, employed, unemployed and not in the labour force and in the process advance our understanding of the economic driver(s) of the various sets of exit and entry rates and the equilibrium rate of unemployment. The methodology utilizes all the information about flows in a consistent framework which is in contrast to existing studies which look at movements between unemployment and employment (and other states) independently of each other. It is important to model all the possible labour movements[3] in a coherent framework to allow for worker flows between states to be affected by general factors (i.e. common factors, such as the state of the economy) as well as by specific factors (such as the state one is transitioning from). The advantage of the latent common factor approach is that it identifies the representative activity index which in turn can be used in a parsimonous and efficient way to forecast behaviour in labour markets.

The paper has four related contributions. The first is to apply a new technique which utilises micro data about worker flows to understand macro aggregates. The second is to identify the common factor in the flows and relate it to suggested drivers of demand pressure in the labour market (for example, the vacancy to unemployment ratio). The third is to extend the analysis of the equilibrium unemployment rate implied by the flows to the case where the labour force is growing (in contrast to what is common in the literature where the labour force is held constant) and to use the estimated temporary equilibrium unemployment rate as a benchmark to evaluate ac-

[3]Unfortunately, it is not possible to obtain job-job flows data for Australia in a form amenable to time series analysis.

tual unemployment rate. The fourth is to estimate the model, which is in a nonlinear state space form, using a Gibbs sampler that encompasses a Metropolis-Hastings algorithm as well as a recent technique, the auxiliary particle filter, to estimate the latent process.

The paper is organized as follows. Section 2 sets out the multi state model of worker flows while section 3 gives a brief description of the econometric methodology. Results are presented and discussed in section 4 with concluding remarks in section 5.

2 Multi-state model of Worker Flows

At any point in time, an economic agent is in one or other of 3 states: (1) employed (E), (2) unemployed (U) and (3) not in the labour force (N) and the same agent can, in the next period, be in any one of the 3 states. This implies that there are 6 possible transitions or worker flows; from employed to unemployed (EU); from employed to out of the labour force (EN); from unemployed to employed (UE); from unemployed to out of labour force (UN); and from not in the labour force to employed (NE).

Let s represent the 6 possible state-to-state transitions such that $s = 1$ represents the movement EU, $s = 2$ represent EN, $s = 3$ is UE, and so on. Let L_t^ω denote the number of persons at time t in state ω ($\omega = [E, U, N]$). For example, L_t^E represents the number of persons who are employed at time t. Then define Λ_t^s as the transition probability:

$$\Lambda_t^s = \frac{\Omega_t^s}{L_{t-1}^\omega}, \tag{1}$$

where Ω_t^s is the number of persons experiencing a state s transition at t.

Following Koopman, Lucas and Moneteiro (2007), we model λ_{kt}^s, the instantaneous probability of a person k experiencing a type s transition at time t given $t-1$ information as a proportional hazards specification:

$$\lambda_{kt}^s = R_{kt}^s \exp(\eta^s + \alpha^s \psi_t), \tag{2}$$

where R^s_{kt} is a dummy variable such that $R^s_{kt} = 1$ if k is 'at risk' of making a transition. The unknown parameters of the model are η^s, and α^s. The vector n^s represents the constant reference-level log-intensity of transition type s, and α^s measures the sensitivity of a type s transition to changes in an unobservable common dynamic latent factor[4] ψ_t.

The latent variable ψ_t accounts for unobserved dependence between the transition histories in a parsimonious way. We make the assumption that ψ_t follows an autoregressive process of order r

$$\psi_t = \rho(L)\psi_t + \varepsilon_t, \tag{3}$$

$$\varepsilon_t \sim N(0, \sigma^2), \tag{4}$$

where $\rho(L) = \rho_1 L + \rho_2 L^2 + \cdots + \rho_r L^r$ and all of the roots of $\rho(L)$ lie outside the unit circle. However we note that this assumption is not central to the methods we propose and that the model can easily be generalized to accommodate alternative functional forms. As α^s and σ are not simultaneously identified, we normalize the parameter space so that $\sigma = 1$.

The aggregate proportion Λ^s_t is then given as:

$$\Lambda^s_t = \lambda^s_{kt} \sum_{s=1}^{6} \Omega^s_{t-1}. \tag{5}$$

In other words, the product of the instantaneous probability of one person making a type s transition multiplied by all persons transiting is equal to the population probability of type s transition.

The advantage of the approach is that it allows the probability of transition to vary not only with general factors (i.e. common factors, such as the business cycle), but also with specific factors (such as the state one is transition from). The method may be viewed as a generalization of the regime-switching approach applied at a disaggregated level with the advantage that microeconomic characteristics (such as gender, location) can be

[4]In Koopman et. al., the model (2) is specified to model events in continuous time. However, employment data are reported monthly, and we have adapted the Koopman et.el model for discrete time analysis.

incorporated, if required, in a consistent manner.

3 Econometric methodology

Let Y_{kt}^s be a function that is equal to 1 when k experiences a transition event of type s at time t and zero otherwise. If we define

$$z_t = \{R_{1t}^1, ..., R_{Kt}^S, Y_{1t}^1, ..., Y_{Kt}^S\},$$

then the likelihood function of (θ, ψ_t, ρ) is

$$p(z_t|\theta, \psi_t,) = \prod_{k=1}^{K} \prod_{s=1}^{S} \exp\left(Y_{kt}^s \left(\eta^s + \alpha^s \psi_t\right) - R_{kt}^s \int_{t-1}^{t} \lambda_{kt}^s dt \right),$$

where $\theta = (\eta^1, ..., \eta^S, \alpha^1, ..., \alpha^S)$. In this case, $p(z_t|\theta, \psi_t)$ is the probability of survival for k in its current state at t, in other words, there is no transition event. When a transition event for k takes place at t (i.e. $Y_{kt}^s = 1$) this survival probability is multiplied by the hazard rate to yield the probability density of the transition event. The likelihood function for the whole sample period is then

$$p(z|\theta, \Psi, \rho) = \prod_{t=1}^{T} p(z_t|\theta, \psi_t) p(\psi_t|\psi_{t-1}, \theta, \rho),$$

where $\Psi = (\psi_1, ..., \psi_T)$.

In this paper, we adopt a Bayesian approach to estimate the model which require the specification of priors for the parameters.[5] We assume that each of the parameters is a priori independent and specify fairly non-informative

[5] This is in contrast to Koopman et. al. who estimate a similar model in the context of credit rating via a Monte Carlo maximum likelihood route which is a combination of importance sampling and the Kalman filter.

but proper priors for the parameters. The priors are

$$\begin{aligned}\eta^s &\sim N(\underline{\eta}^s, \Sigma_{\eta^s}),\\ \alpha^s &\sim N(0, \Sigma_{\alpha^s}),\\ \rho &\sim U(-1,1),\end{aligned}$$

where $\underline{\eta}^s = \ln\left(\sum_{t=1}^{T}\sum_{k=1}^{K} Y_{kt}^s\right) - \ln\left(\sum_{t=1}^{T}\sum_{k=1}^{K} R_{kt}^s\right)$ which is the MLE of $p(z|\eta^s) = \prod_{t=1}^{T}\prod_{k=1}^{K}\prod_{s=1}^{S} \exp\left((Y_{kt}^s - R_{kt}^s)\eta^s\right)$, the diagonal elements of Σ_{α^s} and Σ_{η^s} are set to 10000 and their off-diagonal elements to 0. Note that the specification of uniform prior for ρ with $(-1,1)$ is to ensure that ψ_t is stationary. To complete the Bayesian approach, these priors are combined with the likelihood function via Bayes Theorem to give the posterior distribution,

$$p(\theta, \Psi, \rho|z) \propto p(z|\theta, \Psi, \rho)p(\theta)I(\rho),$$

where $I(\rho)$ is an indicator function such that $I(\rho) = 1$ if $|\rho| < 1$, and 0 otherwise.

The aim of Bayesian analysis is to examine the marginal posterior distribution of each parameters. Since the marginal posterior distributions are analytically intractable we have to resort to sampling techniques for estimation. Furthermore, given the nonlinear state space form, standard MCMC algorithms such as the Gibbs sampler of Geman and Geman (1984) and the Metropolis-Hastings [6] algorithm of Metropolis et al. (1953) and Hastings (1970) cannot be applied directly, instead a version of MCMC that encompasses a Metropolis-Hastings algorithm and the auxiliary particle filter (APF) of Pitt and Shephard (1999) within a Gibbs sampler is used to produce a set of draws $(\eta^{s(h)}, \alpha^{s(h)}, \rho^{(h)}, \Psi^{(h)})$ from $p(\theta, \Psi, \rho|z)$. The MH algorithm is used to sample $p(\theta|\Psi, \rho, z)$ while the APF filter algorithm is used to generate the draws for $\psi_t, t = 1, 2..., T$. The steps for the sampling scheme are shown in the Appendix.

[6] A detailed survey of the Metropolis-Hastings algorithm is provided by Chib and Greenberg (1995).

4 Empirical Analysis

4.1 Description of Data

Data on gross flows between various labour market states in Australia has been published on a monthly basis by the Australian Bureau of Statistics since February 1980.[7] Measures of gross flows between any two months are compiled from data collected as part of the monthly Labour Force Survey (LFS) which currently includes around 60,000 households. Estimates of gross flows reflect the matching of responses by individuals in the any month's survey with responses by the same individuals in the previous month's survey (although there is sample rotation, around 7/8 of the sample is common across two successive months). These matched records are then "expanded up" to yield population estimates which, for various reasons, including non-responses, typically represent around 78% of the total civilian population aged 15 years and over.[8] This means that the balance of flows given in the published flows data will not be equal to the recorded changes in 'stocks' (such as the total number unemployed). Given the purpose of this paper, it is desirable to adjust the raw flows data so as to ensure that net flows and sums of rows and columns in the flow tables are equal to their stock counterparts. An iterative method has been applied to the published gross flows data to force the flow sums (including stay-puts) to be exactly equal to that of the labour force survey stocks data.[9] The most important feature

[7] The raw data on gross flows until March 2003 is taken from the tables of "Estimates of labour force status and gross changes (flows) derived from matched records" published in the ABS publication Labour Force: Australia, Cat No 6203.0. Raw data for March 2003 on is taken from the ABS data cube 6291.0.55.001 series GM. Where data was missing due to a new sample being rotated in, unpublished data was obtained from ABS microfiche and we have used that as the raw data for those periods. Detailed discussions of the Australian gross flows data and its limitations can be found in Foster (1981) and Dixon et. al. (2002).

[8] The reasons why the 'population represented by the matched records' is less than 100% of the total civilian population aged 15 years and over are explored in some detail in Dixon (et al) 2002.

[9] The approach entails an iterative method whereby rows and columns are scaled up or down until, (i) the sum of the entries across the rows of the 'new' flows table sum to the total number in each labour market state in the first of each pair of months as reported for Australia as a whole in the LFS, and (ii) that the implied unemployment

of the adjustment is that it forces the relative magnitude of the flows during the month to be consistent with the observed change in stock figures for the unemployment rate between months. Amongst other things, this means that when we enquire into the 'source' of changes in (say) the unemployment rate or we use our results to examine equilibrium unemployment rate (say), we can be sure that the sum of the (net) flows will be exactly equal to the changes in the stocks.

Table 1: Average monthly number of (off-diagonal) transitions for the whole sample period (thousands)

From	To		
	E	U	N
E	–	97.0	229.2
U	118.1	–	117.8
N	213.3	138.6	–

Over the whole of our sample period of 293 months, there were a total of 227,151.6 (thousand) transitions between states. The (monthly) mean number of transitions between the states E, U, and N (in '000s) are shown in Table (1) (the off-diagonal elements). One notable feature of the Australian labour market over the sample period (August 1979-November 2003) is that the inflow into unemployment ($EU + NU$) is about the same as the outflow from unemployment ($UE + UN$). This reflects the absence of any trend in the unemployment rate over the period. We also see that the inflow into the labour force ($NE + NU$) exceeds the outflow from the labour force ($EN + UN$), resulting in a rise in the labour force participation rate over the period. At the same time, the inflow into employment ($UE + NE$) exceeds the outflow from employment ($EU + EN$), resulting in a rise in the employment-population ratio over the period. On average then the growth

and participation rates in the rows of the 'new' flows table correspond exactly to those rates given for Australia as a whole in the LFS for the first of each pair of months. This is essentially the same as the "RAS" method which is commonly used in input-output analysis. More details can be found in Dixon et. al. (2006).

in the participation rate was accompanied by a rise in the employment to population ratio rather than a rise in the unemployment rate.

4.1.1 Sample Data Set

Our data set describes the flows between E, U and N for the population and the numbers are large (in the millions). Since it would not be feasible computationally to work with millions of observations, we will be tracking k representative agents. To obtain a parsimonious set of k, we performed two manipulations.

The first is to re-scale the flows from millions of persons to 1000 representative units $[1000 = (\sum_{\omega=1}^{3} L_t^{\omega})/\Phi_t]$ where Φ_t is the scaling factor. Applying this factor to all flows at time t will not alter any information about transitions at time t. However, it is likely that Φ_t will change over time along with changes in the working population.

The second manipulation is to take advantage of the fact that there is a core group that stays in the same state (EE, UU and NN) through out the sample period.

$$l_{t-1}^{\omega} = \overline{L^{\omega}} + i_{t-1}^{\omega}; \qquad \omega = E, U \text{ and } N,$$

where $l_{t-1}^{\omega} = (L_{t-1}^{\omega}/\Phi_{t-1})$, $\overline{L^{\omega}}$ is the core $(\overline{L}^{E} = 510, \overline{L}^{U} = 19, \overline{L}^{N} = 331)$ and i_{t-1}^{ω} is the residual. This implies that our sample data set contains 140 representative units (1000 less $\overline{L}^{E} + \overline{L}^{U} + \overline{L}^{N}$), who are likely to change states in the sample. Over the sample period of 293 months, there were 19905 transitions in the sample data set and the total number of transitions between the states E, U, and N are shown in Table (2).

The proportions for the population Λ_t^s are obtained by re-scaling the estimated $\widehat{\lambda}^s$ based on the reduced sample set of k representative agents as follows:

$$\Lambda_t^s = \widehat{\lambda}^s \left(\frac{i_{t-1}^{\omega}}{\overline{L}^{\omega} + i_{t-1}^{\omega}} \right) \left(\frac{\Phi_t}{\Phi_{t-1}} \right).$$

Table 2: Total number of transitions in the set of k for the whole sample period

From	To		
	E	U	N
E	–	2113	5029
U	2574	–	2561
N	4606	3022	–

4.2 Results

We estimated two models. The first model is without the latent variable Ψ, that is $\lambda^s_{kt} = R^s_{kt} \exp(\eta^s)$, while the second model includes the latent factor and is $\lambda^s_{kt} = R^s_{kt} \exp(\eta^s + \alpha^s \psi_t)$. In other words, the first model includes only state-specific factors while the second model allows for both common and state-specific factors.

Table (3) shows the posterior parameter estimates for the two model. A model comparison using the Bayes factor,[10] which is the ratio of marginal likelihood[11] of the two models, shows that the model with Ψ yields a better fit than the model without Ψ. In other words, each flow contains a fixed component captured in the estimate of η^s and a component which varies with the common latent factor $\alpha^s \psi_t$.

The results in the right hand column of Table (3) show that flows from employment to unemployment and from employment to out of the labour force (EU, EN) react positively to the common (latent) factor while the flows from unemployment (UE, UN) and from not in the labour force (NE, NU) react negatively to the common (latent) factor. In particular, following an increase in the latent variable, the negative signs on α^{UE} and α^{NE} suggest a decrease in flows to employment while the positive signs on α^{EU} and α^{EN} suggest an increase in flows out of employment. In other words, the latent

[10] In order to compute the Bayes factor, the marginal likelihood have to be computed first. In this paper, we estimate the marginal likelihood of each model using the harmonic-mean method of Gelfand and Dey (1994). Geweke (1999) offers an excellent exposition into their method.

[11] The marginal likelihood is $p(z) = \int \int p(z|\theta, \rho) p(\theta, \rho) d\theta d\rho$.

factor is capturing an (inverse) business cycle or employment cycle activity variable.

The finding that the coefficients, α^{EU} and α^{EN} are positive while the coefficient α^{UN} is negative suggests an explanation for the negative correlation frequently commented on between changes in the unemployment rate and the participation rate. Our findings suggest that this negative correlation arises because flows between employment and not in the labour force (EN) and flows between employment and unemployment (EU) are positively correlated and not because flows between employment and unemployment (EU) and flows between unemployment and not in the labour force (UN) are positively correlated. This suggests that the 'discouraged-unemployed-worker' effect is not a credible explanation for the endogeneity of the participation rate.[12]

4.3 The latent variable

By construction, the latent activity variable Ψ is an autoregressive process with an estimated parameter of 0.9677 indicating high persistence. Figure 1 plots the path of Ψ against three key monthly labour variables, the unemployment rate (ur), the ratio of vacancies to number unemployed (v/u) and the employment to population ratio (er).[13] In addition, Table 4 presents the contemporaneous and lead-lag correlations between the latent variable Ψ and these key economic series. Since the sample period includes the period following the introduction of reforms to the industrial relations and wage setting systems in 1994, we have presented results for the pre and post reform subperiods.[14]

[12] This issue is discussed at greater length in the context of net flows between states in Dixon et. al. (2005).

[13] Source of data for the monthly unemployment rate is the Australian Bureau of Statistics while the job vacancy series is the ANZ job advertisement series obtained from Datastream

[14] We also examined the correlation of the latent variable with a number of other monthly activity variables, such as a leading index and interest rates, but these dynamic correlations were inferior to those reported here.

Table 3: Posterior parameter estimates and log marginal likelihoods

s	Without latent	With latent
η^{EU}	-6.2487	-6.2222
	(0.0210)	(0.0294)
η^{EN}	-5.3839	-5.3489
	(0.0143)	(0.0202)
η^{UE}	-5.3944	-5.3841
	(0.0199)	(0.0192)
η^{UN}	-5.4016	-5.3707
	(0.0189)	(0.0204)
η^{NE}	-5.4977	-5.4952
	(0.0147)	(0.0157)
η^{NU}	-5.9192	-5.9395
	(0.0178)	(0.0169)
α^{EU}	–	0.0902
		(0.0102)
α^{EN}	–	0.0713
		(0.0070)
α^{UE}	–	-0.0497
		(0.0069)
α^{UN}	–	-0.0606
		(0.0080)
α^{NE}	–	-0.0419
		(0.0050)
α^{NU}	–	-0.0378
		(0.0055)
ρ	–	0.9677
		(0.0064)
Log marginal likelihood	-130808	-130417

As shown in the Figures, the latent variable peaked during the two recessions (1981-83 and 1990-93) and has been declining since 2000, the approximate start of the recent long expansionary period. In other words, as surmised earlier, the latent variable is displaying countercyclical behaviour. The Figures show that Ψ is positively correlated with the procyclical nature of the unemployment rate and also positively correlated with the inverse vacancy/unemployment ratio and the inverse employment/population ratio.

According to the results in the Tables, the lead-lag relationship between the latent variable and the ur appear to be stronger post 1994, but the lead of v/u appear to have fallen in the latter period. However, the strongest correlation is with the employment population ratio. All of which is to suggest that the latent activity variable has time series properties similar to the employment population ratio. This has a useful implication - the employment population ratio may be used to forecast worker flows.

4.4 Sample Transition Probabilities

Figure 2 presents graphical comparisons of the observed transition probabilities (that is the size of the flow which occurred between any two states expressed as a proportion of the number in the originating state) with the equivalent probabilities calculated from the model for the case of the representative sample data set. The solid lines are the actual sample transition probabilities while the "thick" lines with upper and lower bounds are the estimated probabilities.[15] To mitigate the effects of starting values, the results are reported from January 1984.

As shown, the estimated proportions track actual sample probabilities associated with the movements EU and EN well, UE and UN reasonably well, but not flows out of N. In some ways this is not surprising given that we would expect flows out of N (that is flows into the labour force) to be far more dependent on sociological and demographic variables than flows out of E and U. However, the fact is that the model captures four (exits

[15] Since labour flows are subjected to considerable seasonal factors - a 12-month moving average have been applied to remove these effects.

from employment and exits from unemployments) of the six labour market flows well. This may be of special interest to policy makers concerned with forecasting the rate of departures from both the pool of the employed and (especially) the unemployed.

Table 4(a): Dynamic correlation: whole sample

	τ						
	-3	-2	-1	0	1	2	3
$\psi_t, ur_{t-\tau}$	0.6648	0.6741	0.6970	0.7141	0.7134	0.6924	0.6574
	(0.0168)	(0.0169)	(0.0170)	(0.0176)	(0.0185)	(0.0185)	(0.0181)
$\psi_t, (u/v)_{t-\tau}$	0.5671	0.5544	0.5370	0.5158	0.4917	0.4631	0.4310
	(0.0188)	(0.0195)	(0.0202)	(0.0209)	(0.0217)	(0.0224)	(0.0229)
$\psi_t, (1/er_{t-\tau})$	0.9475	0.9583	0.9653	0.9679	0.9666	0.9605	0.9503
	(0.0077)	(0.0064)	(0.0052)	(0.0043)	(0.0040)	(0.0042)	(0.0049)

(b): Dynamic correlation: pre 1994

	τ						
	-3	-2	-1	0	1	2	3
$\psi_t, ur_{t-\tau}$	0.6916	0.6919	0.7065	0.7181	0.7027	0.6760	0.6330
	(0.0209)	(0.0211)	(0.0214)	(0.0224)	(0.0236)	(0.0240)	(0.0243)
$\psi_t, (u/v)_{t-\tau}$	0.6340	0.6134	0.5872	0.5562	0.5184	0.4776	0.4313
	(0.0238)	(0.0250)	(0.0263)	(0.0276)	(0.0288)	(0.0299)	(0.0308)
$\psi_t, (1/er_{t-\tau})$	0.9392	0.9543	0.9638	0.9669	0.9644	0.9554	0.9398
	(0.0110)	(0.0087)	(0.0067)	(0.0053)	(0.0052)	(0.0062)	(0.0078)

(c): Dynamic correlation: post 1994

	τ						
	-3	-2	-1	0	1	2	3
$\psi_t, ur_{t-\tau}$	0.8243	0.8397	0.8781	0.8941	0.9011	0.8572	0.7791
	(0.0210)	(0.0198)	(0.0162)	(0.0140)	(0.0146)	(0.0171)	(0.0210)
$\psi_t, (u/v)_{t-\tau}$	0.6144	0.6137	0.6106	0.6079	0.5716	0.5299	0.4851
	(0.0376)	(0.0366)	(0.0356)	(0.0346)	(0.0366)	(0.0385)	(0.0407)
$\psi_t, (1/er_{t-\tau})$	0.8630	0.8865	0.9053	0.9189	0.9210	0.9175	0.9090
	(0.0223)	(0.0195)	(0.0169)	(0.0144)	(0.0137)	(0.0137)	(0.0144)

Figure 1(a): Ψ and the unemployment rate

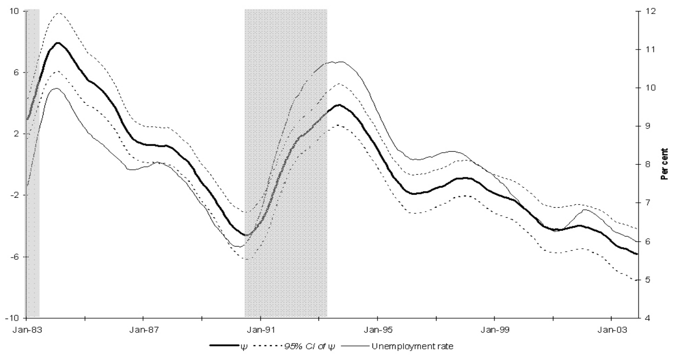


Figure 1(b) Ψ and the inverse (vacancy/unemployment) ratio

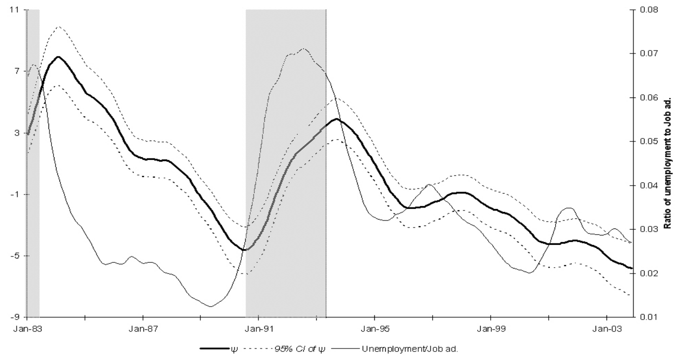
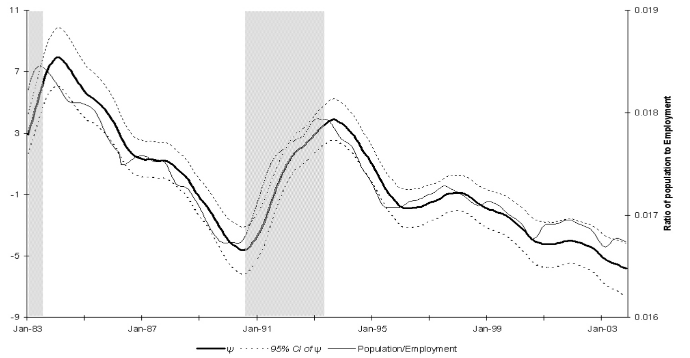


Figure 1(c): Ψ and the inverse (employment/population) ratio



5 Dynamics of the Unemployment rate with a growing Labour force

The unemployment rate is defined as the ratio of the number unemployed (U) to the total labour force ($LF = E + U$). Allowing for both U and LF to vary over time, the change in the unemployment rate can be computed as:

$$\begin{aligned} \Delta\left(\frac{U}{LF}\right)_t &= \frac{U_t}{LF_t} - \frac{U_{t-1}}{LF_{t-1}} \\ &= \frac{\Delta U_t}{LF_t} - \left(\frac{U_{t-1}}{LF_t}\right)\left(\frac{\Delta LF_t}{LF_{t-1}}\right), \end{aligned} \tag{6}$$

where Δ represents a discrete change operator. The above equation may also be written as:

$$\Delta\left(\frac{U}{LF}\right)_t = \frac{(\Delta U_t) - U_{t-1}(\Delta LF_t/LF_{t-1})}{LF_t}.$$

The two terms in the numerator on the RHS may be given a rather interesting interpretation. The first term, is simply the balance of inflows and outflows over any period and is equal to the observed (i.e. the actual) change in the number unemployed over the period. The second term, measures the extent to which the number unemployed can change when there is a growing labour force without changing the unemployment rate.. Put another way, for the unemployment rate to be constant over time, we require the rate of growth in unemployment to equal the rate of growth in the labour force, i.e.: $\Delta U_t/U_{t-1} = \Delta LF_t/LF_{t-1}$ or $\Delta U_t = U_{t-1}\Delta LF_t/LF_{t-1}$. Clearly, if the first term in the numerator (i.e., the actual change) exceeds the second (i.e., the change consistent with the unemployment rate remaining constant) the unemployment rate will rise. Only if the first term is exactly equal to the second will the unemployment rate be constant. In fact, even when (ΔU_t) equals zero, the unemployment rate can rise or fall depending on the rate of growth of the labour force. This should not be surprising. If the Labour Force is (say) rising over time then the number unemployed must rise at the same rate to keep the ratio between the two (this is the unemployment rate, (U/LF)) constant. However, for the number unemployed to rise over time there must be a net inflow into unemployment, that is (ΔU_t) must be positive, not zero.

Put another way, for the unemployment rate to remain constant between successive periods,[16] we require:

$$\left(\frac{U}{LF}\right)_t^* = \frac{\Delta U_t}{\Delta LF_t}. \tag{7}$$

Changes in the number unemployed over time (ΔU) reflect the balance between two flows, the inflow into unemployment and the outflow from unemployment:

$$\Delta U_t = (EU_t + NU_t) - (UN_t + UE_t). \tag{8}$$

Similarly, changes in the labour force reflect the flows in and out of the labour force:

$$\Delta LF_t = (NE_t + NU_t) - (EN_t + UN_t). \tag{9}$$

Taken together these three expressions imply that the equilibrium unemployment rate will equal:

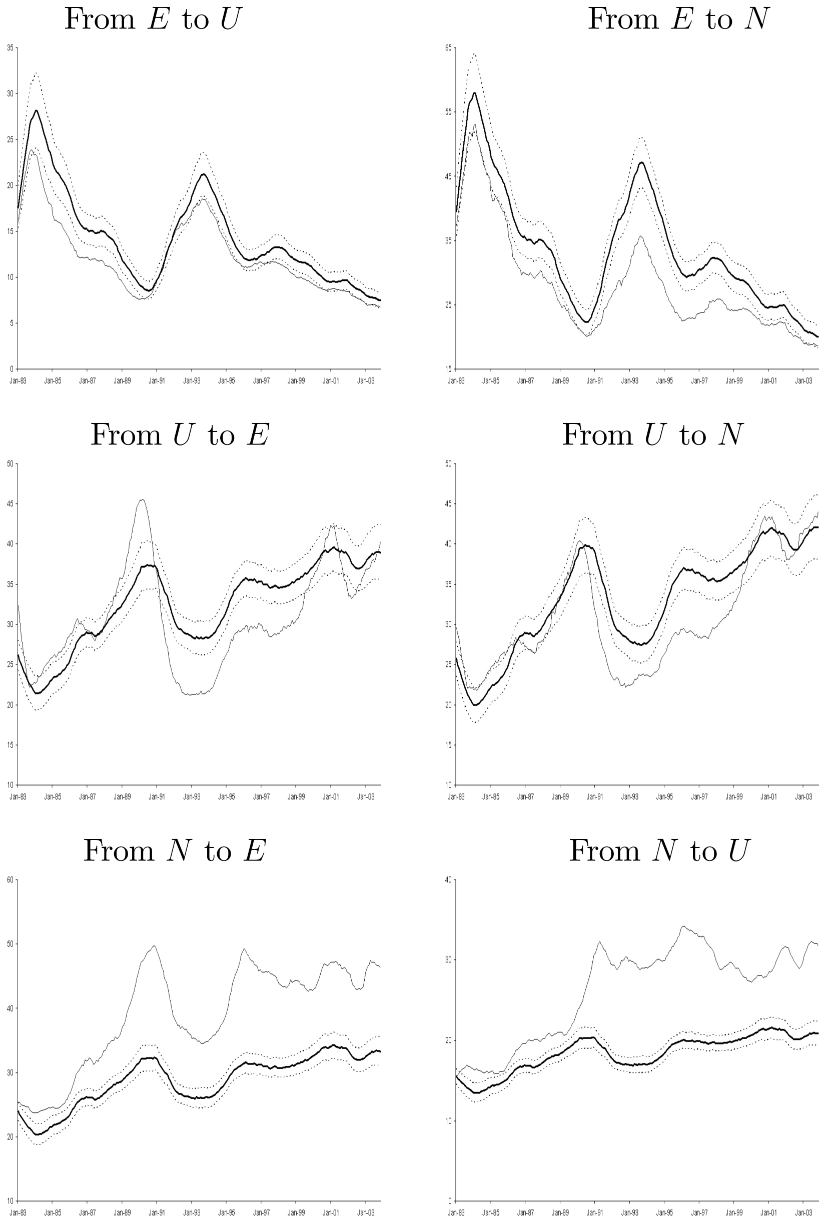
$$\left(\frac{U}{LF}\right)_t^* = \frac{(EU_t + NU_t) - (UN_t + UE_t)}{(NE_t + NU_t) - (EN_t + UN_t)}. \tag{10}$$

Although flows occur between three labour market states (employed, unemployed and not in the labour force), it is useful to model unemployment and especially unemployment dynamics in a parsimonious fashion with the aid of only a single entry rate to unemployment and a single exit rate from unemployment. Applying the concepts of 'entry' and 'exit' rates in and out unemployment (U), and in and out of the labour force $(LF = E + U)$ respectively as :[17]

[16] Hall calls this the "stochastic equilibrium unemployment rate" (Hall, 2003, p 148 and 2005a, p 399).

[17] To some extent these definitions are arbitrary. For example, the entry rate into unemployment could be defined as $(EU + NU)/(E + N)$ and the entry rate into the labour force could be defined as $(NE + NU)/(N)$. We have chosen definitions which allow us to derive the most straightforward expression for the equilibrium unemployment rate.

Figure 2: Graphs of posterior Λ_t^s compared to sample proportions



$$\begin{aligned}
enu_t &= \frac{EU_t + NU_t}{E_{t-1}} = \Lambda_t^{eu} + \Lambda_t^{nu}\frac{N_{t-1}}{E_{t-1}}, \\
exu_t &= \frac{UN_t + UE_t}{U_{t-1}} = \Lambda_t^{un} + \Lambda_t^{ue}, \\
\\
enlf_t &= \frac{NE_t + NU_t}{LF_{t-1}} = \left(\Lambda_t^{ne} + \Lambda_t^{nu}\right)\frac{N_{t-1}}{LF_{t-1}}, \\
exlf_t &= \frac{EN_t + UN_t}{LF_{t-1}} = \Lambda_t^{en}\frac{E_{t-1}}{LF_{t-1}} + \Lambda_t^{un}\frac{U_{t-1}}{LF_{t-1}}.
\end{aligned}$$

Given these definitions and noting that $E_{t-1} = LF_{t-1} - U_{t-1}$, we may write equation (10) as:

$$\left(\frac{U}{LF}\right)_t^* = \frac{enu_t * (LF_{t-1} - U_{t-1}) - exu_t * U_{t-1}}{enlf_t * LF_{t-1} - exlf_t * LF_{t-1}},$$

which may be rearranged as:

$$\left(\frac{U}{LF}\right)_t^* = \frac{enu_t}{(enu_t + exu_t) + (enlf_t - exlf_t)}. \tag{11}$$

Note in passing that if the labour force is constant, equation (10) may be written as:

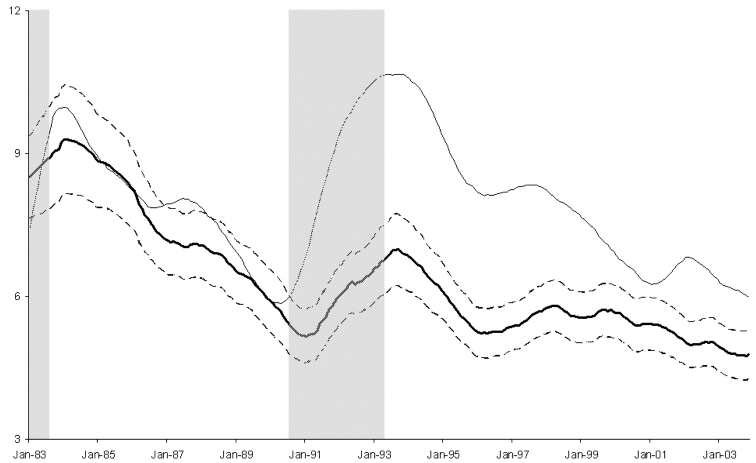
$$\left(\frac{U}{LF}\right)_t^* = \frac{enu_t}{(enu_t + exu_t)}, \tag{12}$$

which is the expression to be found in Hall (2005a) and Burgess and Turon (2005).

All this is reasonably straightforward, although it may not be obvious that, for given entry and exit rates, the equilibrium unemployment rate will be lower when the case where the labour force is growing (equation 10) than in the case where the labour force is constant i.e., when $(enlf_t - exlf_t)$ is zero (equation 12). The intuition goes something like this: in equilibrium the unemployment rate will be constant. This means that the rate of growth in the number unemployed must equal the rate of growth in the labour force and this means that ΔU has to be positive in equilibrium, not zero as would be the case if the labour force was not growing. So $(EU_t + NU_t)$ has to be greater than $(UN_t + UE_t)$ and, if the unemployment rate is to be constant, $(EU_t + NU_t)$ has to be greater than $(UN_t + UE_t)$ by an amount that matches the growth in the labour force multiplied by the number unemployed at the start of the period $U_{t-1}(\Delta LF_t/LF_{t-1})$. In other words, for given entry and exit rates, the only way in which inflow can rise relative to outflow is if employment rises relative to unemployment, i.e. if the unemployment rate falls.

Figure 3 plots these "equilibrium rates" using the estimated exit and entry rates generated by the single factor model. As shown, actual unemployment rates were greater than predicted during the recession years (1990-93) and during the post Asian crisis period (1998).

Figure 3: Posterior equilibrium unemployment rate and actual unemployment rate



6 Concluding Remarks

Understanding the determinants of worker flows contributes to our understanding of the dynamics of the unemployment rate. In this paper we have applied a new technique which utilises micro data about worker flows to understand macro aggregates. In particular, we have identified the common factor in the flows. The results show that one activity factor underpins the decision to move from employment and from unemployment, and that this latent variable is highly correlated with a lagged employment to population ratio variable. This implies that we can forecast worker flows from employment and unemployment well and this may be of special interest to policy makers concerned with understanding the rate of departures from the pool of both the employed and (especially) the unemployed. Furthermore, extending the analysis of the equilibrium unemployment rate implied by the flows to the case where the labour force is growing (in contrast to what is common in the literature where the labour force is held constant) we provide estimates of the temporary equilibrium unemployment rate and use it to benchmark the actual unemployment rate.

Finally, the model, which is in a nonlinear state space form, has been estimated using a Gibbs sampler that encompasses a Metropolis-Hastings algorithm as well as a recent technique, the auxiliary particle filter, to estimate the latent process. This development allows us to extend the application of the multi-state latent factor intensity model to the case of yet another factor, as well as for the inclusion of micro demographic characteristics that are crucial to understanding the decision to participate. The advantage of the approach is that it provides a parsimonious and efficient way to obtain key information about behaviour in labour markets.

7 Appendix: Gibbs Sampler with Metropolis-Hastings and Auxiliary Particle Filter

The steps in the MCMC algorithm are shown below. Note that, since the transition density $q(\theta^{*(j)}, \theta^{(j)})$ in the Metropolis-Hastings step is a random walk, the acceptance probability at each step depends only on the ratio of the product of the likelihood and the prior between the potential candidate and the current candidate.

Algorithm 1 : Gibbs sampler

1. Choose an arbitrary starting point for $(\theta^{(j)}, \rho^{(j)})$ and set $j = 0$.
2. Given $(\theta^{(j)}, \rho^{(j)})$, generate $\psi_t^{(j+1)}, t = 1, ..., T$ using the APF algorithm.
3. Given $\psi_t^{(j+1)}$, compute $\rho^{(j+1)}$.
4. Repeat Step 2 if $\left|\rho^{(j+1)}\right| > 1$.
5. Given $(\psi_t^{(j+1)}, \rho^{(j+1)})$, generate $\theta^{(j+1)}$ using the MH algorithm
6. Set $j = j + 1$ and return to Step 2.

Algorithm 2: Metropolis-Hastings step

1. Given $\theta^{(j)}$, generate a candidate $\theta^{*(j)}$ from a random walk transition density $q(\theta^{*(j)}, \theta^{(j)})$.
2. Calculate the acceptance probability

$$\alpha(\theta^{(j)}, \theta^{*(j)}) = \min\left[\frac{p(\theta^{*(j)})p(z|\theta^{*(j)}, \Psi^{(j+1)}, \rho^{(j+1)})}{p(\theta^{(j)})p(z|\theta^{(j)}, \Psi^{(j+1)}, \rho^{(j+1)})}, 1\right].$$

3. Generate an independent random variable u from $U(0, 1)$.
4. Set $\theta^{(j)} = \theta^{*(j)}$ if $u < \alpha(\theta^{(j)}, \theta^{*(j)})$ or else $\theta^{(j)} = \theta^{(j)}$.
5. Repeat Steps 1 to 4 N times. Note that this is to allow for burn-in.
6. Set $\theta^{(j+1)} = \theta^{(j)}$.

Algorithm 3 : Auxciliary particle filter step

1. Given $(\theta^{(j)}, \rho^{(j)})$ obtain G draws of ψ_0 from $N(\psi_{-1}^{(g)}, 1)$ and set $t = 1$. Note that $\psi_{-1}^{(g)}$ are assumed to be zeros.

2. Given $\psi_{t-1}^{(g)}$ predict $\psi_t^{(g)}$ from

$$\psi_t^{(g)} \sim N(\rho^{(j)}\psi_{t-1}{}^{(g)}, 1) \qquad g = 1, ..., G.$$

3. Estimate $p(z_t|\theta^{(j)}, \mathcal{F}_{t-1}) = \int p(z_t|\theta^{(j)}, \mathcal{F}_{t-1})p(\psi_t|\mathcal{F}_{t-1})d\psi_t$ through a simple averaging

$$\widehat{p}(z_t|\theta^{(j)}, \mathcal{F}_{t-1}) = \frac{1}{G}\sum_{g=1}^{G} p(z_t|\theta^{(j)}, \psi_t{}^{(g)}),$$

   where $\mathcal{F}_{t-1}$ denotes the history of observations up to time $t-1$

4. Given $\psi_t^{(g)}$

   (a) compute the expectation of

$$\widehat{\psi}_t^{*(g)} = \rho^{(j)}\psi_{t-1}^{(g)} \qquad g = 1, ..., G.$$

   (b) compute weight

$$w_g = p(z_t|\theta^{(j)}, \widehat{\psi}_t^{*(g)}) \qquad g = 1, ..., G.$$

5. Normalise w_g

$$\pi_g = \frac{w_g}{\sum_{l=1}^{G} w_l} \qquad g = 1, ..., G.$$

6. Construct CDF for π_k

$$c_g = c_{g-1} + \pi_g \quad g = 1, ..., G.$$

7. Draw $\widehat{\psi}_t^{*(k_1)}$ and $\psi_{t-1}^{(k_1)}$ as follows

(a) Draw u from $U(0,1)$.

(b) Starting from c_1 check for the first c_g that is greater than u.

(c) Select the associated $\widehat{\psi}_t^{*(g)}$ and $\psi_{t-1}^{(g)}$.

(d) Set $\widehat{\psi}_t^{*(k_1)} = \widehat{\psi}_t^{*(g)}$ and $\psi_{t-1}^{(k_1)} = \psi_{t-1}^{(g)}$.

8. Repeat Step 7 R times to obtain $\{\widehat{\psi}_t^{(k_1)}, ..., \widehat{\psi}_t^{(k_R)}\}$ and $\{\psi_{t-1}^{(k_1)}, ..., \psi_{t-1}^{(k_R)}\}$.

9. For each k_l

(a) simulate

$$\psi_t^{*(l)} \sim N(\rho^{(j)}\psi_{t-1}^{(k_l)}, 1) \qquad l = 1, ..., R.$$

(b) compute their associated weight as shown below

$$w_l^* = \frac{p(z_i|\theta^{(j)}, \psi_t^{*(l)})}{p(z_i|\theta^{(j)}, \widehat{\psi}_t^{*(k_l)})} \qquad l = 1, .., R.$$

10. Normalise w_l^*

$$\pi_l^* = \frac{w_l^*}{\sum_{k=1}^{R} w_l^*} \qquad l = 1, ..., R.$$

11. Construct CDF for π_l^*

$$c_l^* = c_{l-1}^* + \pi_l^* \qquad l = 1, ..., R.$$

12. Draw $\psi_t^{(1)}$ as follows

(a) Draw u^* from $U(0,1)$.

(b) Starting from c_1^* check for the first c_l^* that is greater than u^*.

(c) Select the associated $\psi_t^{*(l)}$ and set $\psi_t^{(1)} = \psi_t^{*(l)}$.
Note that $\psi_t^{(1)}$ is deemed to be from $p(\psi_t|\theta^{(j)}, \mathcal{F}_t)$.

13. Repeat Steps 9 to 12 G times to obtain $\{\psi_t^{(1)}, ..., \psi_t^{(G)}\}$.

14. Set $t = t + 1$ and return to Step 2.

15. Compute $\psi_t^{(j+1)} = G^{-1} \sum_{g=1}^{G} \psi_t^{(g)}$ for $t = 1, ..., T$. $\psi_t^{(j+1)}$ is the mean of the sample draw which is the minimum mean estimator of the model given $\theta^{(j)}$.

The number of simulated draws is 10000 and the first 1000 are discarded. The simulation is computationally demanding and require a substantial amount of time. This is because (a) for each set of $\eta_s^{(h)}, \alpha_s^{(h)}$, M draws are required to produce $\Psi^{(h)}$ in the auxiliary particle filter and (b) for each set of $\rho^{(h)}, \Psi^{(h)}$, we allow the first 200 draws of $\eta_s^{(h)}, \alpha_s^{(h)}$ as burn-in in the Metropolis-Hastings steps. Hence, to produce a 10000 draws of $(\eta_s^{(h)}, \alpha_s^{(h)}, \rho^{(h)}, \Psi^{(h)})$, the total number of draws involved is roughly $2000000 \times M$.

References

[1] Burgess, S. and Turon, H. (2005), 'Unemployment Dynamics in Britain', *Economic Journal*, 115, 423-48.

[2] Chib S. and Greensberg, E. (1995), 'Understanding the Metropolis-Hastings Algorithm', *The American Statistician*, 49, 327-335.

[3] Dixon, R., Lim, G. and Thomson, J. (2002), 'Australian Gross Flows Data: The Labour Force Survey and the Size of the Population Represented by the Matched Sample', *Australian Journal of Labour Economics*, 5, 1-21.

[4] Dixon, R., Freebairn, J. and G C Lim (2005), 'An Examination of Net Flows in the Australian Labour Market', *Australian Journal of Labour Economics*, 8(1): pp. 25-42

[5] Dixon, R., Freebairn, J. and Lim, G. (2006), 'Time-Varying Equilibrium Rates of Unemployment: An Analysis with Australian Data', Melbourne Institute Working Paper Series wp2006-11, Melbourne Institute of Applied Economic and Social Research, The University of Melbourne.

[6] Foster, W. (1981), 'Gross Flows in the Australian Labour Market', *Australian Economic Review*, 14, 57-64.

[7] Gelfand, A.E. and Dey, D.K. (1994), 'Bayesian Model Choice: Asymptotics and Exact Calculations', *Journal of the Royal Statistical Society Series* B, 56, 501-514.

[8] Geweke, J. (1999), 'Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communciation', *Econometric Reviews* 18, 1-73.

[9] Geman, S, and Geman D. (1984), 'Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, 721-741.

[10] Hall, R. (2003), 'Modern Theory of Unemployment Fluctuations: Empirics and Policy Applications', *American Economic Review*, 93, 145-150.

[11] Hall, R. (2005a), 'Employment Efficiency and Sticky Wages: Evidence from Flows in the Labor Market', *Review of Economics and Statistics*, 87, 397-407.

[12] Hall, R. (2005b), 'Job Loss, Job Finding, and Unemployment in the U.S. Economy Over the Past Fifty Years', *NBER Macroeconomics Annual*, , 101-137.

[13] Hall, R. (2005c), 'Employment Fluctuations with Equilibrium Wage Stickiness', *American Economic Review*, 95, 50-65.

[14] Hasting, W.K. (1970), 'Monte Carlo Sampling Methods Using Markov Chains and Their Applications', *Biometrika*, 57, 97-109.

[15] Koopman,S.J., Lucas, A. and Monteiro, A.A. (forthcoming), 'The multi-state latent factor intensity model for credit rating transitions', *Journal of Econometrics.*

[16] Metropolis, N., Rosenbluth, A.W., Rosenbulth, M.N., Teller, A.H and Teller, E. (1953), 'Equations of State Calculations by Fast Computing Machines', *Journal of Chemical Physics*, 21, 1087-1092.

[17] Mortensen, Dale T. (1970), 'Job Search, the Duration of Unemployment, and the Phillips Curve', *American Economic Review*, Vol. 60, No. 5, December, pp. 847-862

[18] Mortensen, Dale,T. and Christopher A. Pissarides (1994), 'Job Creation and Job Destruction in the Theory of Unemployment', *Review of Economic Studies*, Vol. 61, No. 3, July, pp. 397-415

[19] Pitt, M. and Shephard, N. (1999), 'Filtering via simulation: auxiliary particle filter," *Journal of the American Statistical Association* 94, 590-599.

[20] Petrongolo, P. and Pissarides, C. (2001), 'Looking into the Black Box: A Survey of the Matching Function", *Journal of Economic Literature*, 39, 390-431.

[21] Pissarides, C. (1986), 'Unemployment and Vacancies in Britain", *Economic Policy*, 3, 499-559.

[22] Pissarides, C. (2000), *Equilibrium Unemployment Theory*, Second edition, Cambridge, Mass, MIT Press.

[23] Shimer, R. (2005). 'The Cyclical Behavior of Equilibrium Unemployment and Vacancies," *American Economic Review*, vol. 95, 25-49.

[24] Singer, H. (1939a). 'The process of unemployment in the Depressed Areas (1935-1938)", *Review of Economic Studies*, vol. 6, 177-88

[25] Singer, H. (1939b). 'Regional labour markets and the process of unemployment", *Review of Economic Studies*, vol. 7, 42-58