Empirical topics in search and matching models of the labour market

by

Ashley Benedict Nanton

A Thesis Submitted in Fulfilment of the Requirements for the Degree of Doctor of Philosophy of Cardiff University

Economics Section of Cardiff Business School, Cardiff University

April 2014

Search and matching models such as those of Mortensen and Pissarides (1994) and Pissarides (2000) have come under criticism in recent years. Analysis of the model by Shimer (2005) and others has focussed in particular on the models' inability to generate sufficient volatility in variables such as the unemployment and vacancies rates, and the vacancy-unemployment ratio. Newer models have sought to ameliorate these empirical issues by changing the model – for example by adding wage rigidity or by amending the specification of the costs of search.

In Chapters 3 and 4 of this thesis, we re-address some of these issues using the method of indirect inference. The method allows us to formally test the hypothesis that data was generated by a particular model under a given set of parameter values. It therefore offers a statistically founded replacement for the somewhat arbitrary moment-by-moment comparisons found in much of the existing literature. We apply the method to Shimer's analysis of the Mortensen Pissarides model, and concur with his analysis that, under his chosen parameters, the model fails to fit the data. We also apply the method to the model used in Yashiv's (2006) paper, which argues using moment comparisons that the standard model can be improved by adding convex search costs. In contrast, we find that the augmented model is rejected under formal indirect-inference tests.

The aggregate search and matching literature has also generated an empirical debate about the relative importance of labour market flows, expressed in terms of the hazard rates of labour market transition faced by workers. Many studies decompose changes in steady-state unemployment in terms of the contributions of various hazard rates. This thesis also extends this literature so as to model the contributions of hazards for two distinct and contiguous geographical areas – those of Wales and the rest of the United Kingdom, using Labour-Force-Survey panel data. We find some evidence that in this regard, the UK hazards are weighted towards the hazards "out of" unemployment, whereas for Wales the hazards "into" and "out of" unemployment are of approximately equal importance. We also find however that the results are sensitive to whether or not the data are smoothed, and whether a steady-state is imposed.

Acknowledgements:

I would like to thank my supervisors, Professor Gerald Makepeace (Cardiff University) and Professor Patrick Minford (Cardiff University), for their help and support. Thanks also to Dr Panayiotis Pourpourides (Cardiff University) for valuable technical advice.
I also wish to thank Professor Eran Yashiv (Tel-Aviv University), for providing me with the technical appendix to his (2006) paper, and Professor Robert Shimer (Chicago University), whose Stata do-files I used to process gross flows data from the post-1976 U.S. Current Population Survey. Also, thanks to Professor Wouter Den Haan (London School of Economics), for making available on his website a long time series of the Conference Board's U.S. Help-Wanted Index.

Finally I would like to thank the ESRC for funding my PhD, and my parents for their support throughout.

Ashley Nanton

Disclosure:

The data analysis in Chapter 2 was based on data from the UK Labour Force Survey (LFS). The data has been made available by the Office of National Statistics (ONS) through the Data Archive. Neither the ONS not the Data Archive bear any responsibility for the analysis or interpretation of the data reported here.

Index of tables

Index of Figures

Introduction:

Understanding fluctuations in unemployment is a core task of both macroeconomics and microeconomics. It matters because employment is the primary means for the majority of the working-age population of accessing income and the resources they need to survive. Unemployment as defined and measured by modern labour market surveys represents a direct frustration of that aim. Various approaches have been used to analyse different aspects of unemployment. Search theory and the analysis of gross flows is one comparatively recent approach, which takes as its focus the transition of individuals between the different categories or states of employment, unemployment and inactivity. Within this broader category, work has spanned theoretical models to almost purely empirical applications. The former are attempts to understand the drivers of flows between labour market states, and the ways in which they interact. The latter examine which facts the models should aspire to reproduce. In this thesis, we use a relatively new empirical technique, simulation-based indirect inference and testing, to reappraise the performance of a version of an already well-known search theoretic model – the Mortensen Pissarides model. In keeping with the existing literature, we find the model a poor fit to US data, even when we estimate a set of "best-fitting" parameters. We also look at a variant of the model which has been purported to improve on the original model's performance by means of an augmented search cost function. We also conduct an empirical analysis of gross labour market flows for the country of Wales, which we compare with a similar analysis of the whole of the UK. To our knowledge this is the first attempt to perform an analysis of this type for this region.

How successful is search theory in modelling the aggregate labour market? The question has had a large literature lavished upon it in the last 20 years.
Central to the enterprise has been the contributions of Professors Mortensen and Pissarides, who provided a series of models which have become canonical. These original models are not well known because of their explanatory or predictive power. They are recognized as theoretical achievements, for bringing together search costs, labour demand and labour market flows in an internally coherent framework. However the empirical shortcomings of the models - even in terms of matching moments for plausible parameter values are well known. Essentially, the models

have trouble generating the appropriate level of business cycle volatility when their parameter values are chosen to be consistent with micro-econometric evidence. Some doubt has also arisen regarding the choice of model variables – which are the appropriate shocks, which variables can be modelled as exogenous and which are endogenous?
One could argue that the literature has responded to these doubts along two broad dimensions. Firstly, a large number of papers have attempted to add extensions to these basic models, in order to improve their empirical performances. Secondly, the interest in search and matching models has coincided with and perhaps caused an interest in the empirics of gross labour market flows. A source of debate has been the relative importance of inflows to unemployment, versus outflows from unemployment in determining the unemployment rate at business cycle frequencies. The issue is of relevance to aggregate search and matching theorists. One version of the Mortensen Pissarides model has an exogenous separation rate from employment, another has an endogenous separation rate, and both assumptions have been made in the recent literature. One of the aims of the empirical literature is to decide which sort of model is more appropriate.

This thesis makes a contribution to each of these strands of the literature. We note first that a large amount of the literature that has sought to improve upon the Mortensen Pissarides framework has followed a calibration-based methodology, in which the modeller chooses parameter values based on their consistency with micro-econometric evidence, in the hope that the model will produce simulated moments of endogenous variables that are the same as or close to those found in the data. The methodology is a controversial one, because it does not have a statistical foundation. No statistical metric is used to compare the model to the data, apart from "eyeball" comparisons between data and model of the modeller's chosen moments. In Chapter 3 of this thesis, we use the simulation based technique of indirect inference testing to test one of the basic aggregate search and matching models, under certain parameter values which have been suggested in the literature. This method of testing is statistical, in the sense that it uses the model's error to simulate the distribution of model outcomes, allowing in turn for a statistical test of the model's fit to the data. We also use the closely related method of indirect inference *estimation* to generate an ‘optimal’ set of parameters, and then examine the model’s fit to these data based on these.

One suggestion put forward for the improvement of the Mortensen Pissarides framework has been with regards to the way search costs are modelled. After some preliminary empirical investigation into this issue, Professor Eran Yashiv published a model (2006) with nonlinear search costs, which he claimed provided a closer fit to the data than the canonical aggregate search and matching models. The paper was based on the literature’s standard calibration methodology. In Chapter 4 of this thesis we reconstruct this model, so as to test it under Yashiv’s chosen parameter values, again using the indirect inference method. We reach different conclusions to those in Yashiv’s article.

This thesis also makes a contribution to the purely empirical strand of the literature. During the academic debate over the relative importance of flows for the evolution of unemployment, methods were developed for decomposing unemployment into hazard rates for the transition between labour force states faced by workers and the unemployed. In Chapter 2, we extend one such method used in a paper by Petrongolo and Pissarides (2008), for decomposing unemployment in order to analyse gross labour market flows in Wales, UK. The method allows for changes of state between countries – necessary in this case because of the close linkages of the Welsh economy with that of neighbouring England. The method used relies on the assumption of steady-state unemployment being a reasonable proxy for actual unemployment, and we also examine the validity of this assumption in Chapter 2.

Chapter 1: The Mortensen Pissarides framework and the aggregate labour market – a review essay

1.1 Introduction

The above is a brief overview of where this thesis sits in relation to existing literature. In the rest of this chapter we fill in the details. The next section, section 1.2 discusses two canonical versions of the aggregate search and matching model, credited to Mortensen and Pissarides. One version (here denoted version (a), with model equations labelled accordingly) contains exogenous separations of workers from employment to unemployment. This is the version of the model which we use in subsequent chapters for indirect inference testing and estimation. The other version (here denoted (b)) contains endogenous separations. Having explained the models, in section 1.3 we describe the empirical background to these models, their theoretical advantages, and their initial empirical performance. In section 1.4 we outline the findings of empirical literature which has sought to assess the relative importance of different labour market flows, and some of the questions it has thrown up. Crudely, this could be thought of as adjudicating between the models of type (a) and type (b) discussed in section 1.3 since one important issue is whether the separation rate plays a significant role in cyclical fluctuations. This section also is an overview of the literature that is relevant to our empirical exploration of gross flows in Wales, in Chapter 2. In section 1.5 we then outline the Shimer critique, which focuses on the model of type (a), and claims that the model does not generate enough labour market volatility to match the volatility seen in the data. Section 1.6 discusses ways in which the literature has responded. The last two sections aim to situate Yashiv's ideas regarding the modelling of search costs – and our response -

in the large pre-existing literature, in preparation for Chapters 3 and 4.

1.2 The canonical Mortensen Pissarides models

The models that are presented in this section are based on Pissarides (1986) and Mortensen and Pissarides (1994), but in terms of exposition we follow Chapters 1 and 2 of Pissarides (2000). The aim is to outline the way in which aggregate search and matching models work in order to provide the necessary background to the literature relevant to subsequent sections. Note that these models are in continuous time, whereas Chapters 3 and 4 use discrete time treatments.

The matching function

The canonical models begin with an economy with an exogenous labour force of size L of which uL denotes the number of unemployed workers and vL the number of job vacancies. The number of hires or matches of unemployed workers to vacancies at any instant is given by mL. In this version of the model, matches of workers who are already in employment to new jobs are ignored. There is a matching function that posits that there is a stable and increasing relationship between the stocks of vacant jobs, that is:

$$mL = m(uL, vL)$$

$$m_1(uL, vL) > 0, \qquad m_2(uL, vL) > 0 \tag{1.1}$$

Where $m(uL, vL)$ is a linear homogenous function and subscripts represent derivatives. The linear homogeneity of the matching function allows us to divide the matching function through by the labour force constant, so that it may be written in terms of rates out of the labour force: $m = m(u, v)$.
Furthermore, we may divide this again by the vacancy rate v, so as to write the following:

$$\frac{Lm(u,v)}{vL} = m\left(\frac{u}{v}, 1\right) = \ q(\theta) \tag{1.2}$$

Where $\theta \equiv \frac{v}{u}$ measures "market tightness," and where $q'(\theta) < 0$.

Equation (1.2) is the instantaneous probability of a firm filling a vacancy. Over the time interval dt, the probability is $q(\theta)dt$. The probability is decreasing in the number of vacancies. This is the model's version of "congestive externalities" on the firm's side of the market: the tighter the labour market, the more firms are chasing fewer unemployed workers.

The analogous probability for unemployed workers may be obtained by dividing the matching function expressed in rates by the unemployment rate. The result is conveniently expressed in terms of market-tightness and $q(\theta)$:

$$\frac{m(u,v)}{u} = \theta q(\theta) \tag{1.3}$$

Therefore, the probability that an unemployed worker finds a job over the time interval dt is given by $\theta q(\theta)dt$. This is *increasing* in $= \frac{v}{u}$.

Unemployment flows

For unemployment inflows, two different assumptions can be made. In the most basic form of the model, model (a), separation shocks hit with an exogenous Poisson probability δ. In this case the inflow rate to unemployment is given by the product of this parameter with the employment rate: $(1-u)\delta$. In model (b), separations are endogenous. One can suppose that firm and worker matches have idiosyncratic productivity indexed to the $[0,1]$ interval. When aggregate separation shocks hit with probability δ, matches have their productivity revalued. The $[0,1]$ interval is the support of the cumulative density function of idiosyncratic productivity which is denoted $G(.)$. An endogenous variable in the model is the reservation productivity level chosen by firms in response to economic conditions, $R \in [0,1]$. The separation rate into unemployment in this model is then given by $(1-u)\delta G(R)$. One can then write the equations for unemployment dynamics in each of the models – these are (1.4)(a) and (1.4)(b) respectively.

$$\dot{u} = (1-u)\delta - \theta q(\theta)u \tag{1.4)(a}$$

$$\dot{u} = (1-u)\delta G(R) - \theta q(\theta)u \tag{1.4)(b}$$

By setting the differential to zero these can be written as expressions for steady state unemployment:

$$u = \frac{\delta}{\delta + \theta q(\theta)} \tag{1.5)(a}$$

$$u = \frac{\delta G(R)}{\delta G(R) + \theta q(\theta)} \tag{1.5)(b}$$

Values of states

There are two types of agents in these models, and each type of agent can be in one of two states. Firms are either in possession of a vacant job, or a filled job. Workers are either employed or unemployed. For each of these four states, we can write down the return on the value of the state as a Bellman equation.

In the simple version of the model with an exogenous separation rate, the expected return on a filled job to a firm is given by:

$$rJ = \varrho - w - \delta J \tag{1.6)(a}$$

The benefit to a firm of a filled job is the output produced by the filled job (given by marginal productivity parameter ϱ, less wages w. There is also a probability δ of a separation shock, in which case the firm closes the job, and loses the job's current value J. r is the real interest rate. In the version with the endogenous rate, the value if the job is contingent on realised productivity $x \in [0,1]$. Also when a separation shock hits we need to allow for the possibility that the job will not be completely destroyed:

$$rJ(x) = \varrho x - w(x) + \delta \int_R^1 J(s)dG(s) - \delta J(x) \tag{1.6)(b}$$

Firms are risk neutral, and both filled and vacant jobs are treated as an asset. It is therefore the case that the rate of return on filled jobs and job vacancies is equal to the exogenous rate of return on capital, r.

The expected return on a vacancy in the model (a) is given by:

$$rV = -\varrho c + q(\theta)(J - V) \qquad (1.7)(\mathrm{a})$$

$-\varrho c$ represents the instantaneous search cost that the firm incurs while it has a vacancy. Note that this is some constant proportion, c, of a marginal occupied job's instantaneous output, ϱ.

Firms with vacancies fill them with a probability $q(\theta)$, this is therefore the probability that the firm's job changes state, from vacant to filled. When this occurs, firms gain the value of a filled job J and lose the value of a vacant job V. In model (b) the relevant equation is:

$$rV = -\varrho c + q(\theta)[J(1) - V] \qquad (1.7)(b)$$

The $J(1)$ arises from the simplifying assumption that new matches are assumed to have maximum idiosyncratic productivity. The justification for this is that firms will hire workers so as to be most productive with their current stock of capital[1].

The expected return to a worker from being in employment is equal to the wage received, plus the expected value of a change of state – which is in this case a change from the employed to the unemployed state. The probability of this occurrence is the probability that the firm suffers a negative shock. The return equation is therefore

$$rW = w + \delta(U - W) \qquad (1.8)(\mathrm{a})$$

In model (b):

$$rW(x) = w(x) + \delta \int_R^1 W(s)dG(s) + \delta G(R)U - \delta W(x) \qquad (1.8)(\mathrm{b})$$

Similarly, when the worker is unemployed, the expected return is the value of the unemployment benefit received, b, plus a term representing the expected return on finding a job. In model (a) this is:

[1] Capital is not modelled explicitly here, but one could argue that this line of reasoning makes it implicit.

$$rU = b + \theta q(\theta)(W - U) \tag{1.9)(a}$$

And in model (b):

$$rU = b + \theta q(\theta)[W(1) - U] \tag{1.9)(b}$$

Note again that when workers are newly matched in model (b), they receive the value of a job that pertains to maximum match productivity, $W(1)$.

Free entry

Both versions of the model assume that there is free entry of firms. This drives the value of a vacancy to zero in equilibrium. Using the condition that $V = 0$, equation (1.7)(a) solves to derive the equilibrium value of a filled job in model (a):

$$J = \frac{\varrho c}{q(\theta)} \tag{1.10)(a}$$

And in model (b):

$$J(1) = \frac{\varrho c}{q(\theta)} \tag{1.10)(b}$$

The value of a filled job is increasing in the costs of a vacancy: firms incur an instantaneous search costs of ϱc for vacancies, and $\frac{1}{q(\theta)}$ is the average duration of a vacant firm's search – in other words, the average duration over which the instantaneous search cost ϱc is incurred.

Equilibrium rents

By solving equations (1.8) and (1.9), it can be shown that the expected return on employment for a worker is at least as great as the expected return of a worker from unemployment, if it is assumed that $b \leq w$. (See Pissarides (2000), p14). Similarly, vacancies have a zero value in equilibrium by the free entry assumption, whereas filled jobs have a greater-than-zero value which is given by (1.10). The returns from being in a filled job, or owning a filled job, are therefore both greater to workers and firms respectively, than the respective values of being unemployed, or holding a vacant job. The implication is that there when firms and workers find a match, which puts both parties in their higher value state, there are rents that must be shared. This reasoning is used to find equilibrium wages in the model.

Wage bargaining

Wage determination is assumed to take place through Nash bargaining. Assuming that workers have bargaining power $\beta \in (0,1)$, and firms have power $1-\beta$, this implies that the wage bargained between the *ith* firm and worker pair, maximizes the joint surplus between them, i.e.

$$w_i = argmax(W_i - U)^{\beta}(J_i - V)^{1-\beta} \qquad (1.11)$$

Using this assumption, we can derive the wage equations for each version of the model:

$$w = (1-\beta)b + \beta\varrho(1 + c\theta) \qquad (1.12)(a)$$

In version (b), wages are match productivity contingent, hence:

$$w(x) = (1-\beta)b + \beta\varrho(x + c\theta) \qquad (1.12)(b)$$

The only endogenous variable in the wage equation is θ. A higher vacancy-unemployment ratio increases the bargaining power of workers, as it means that there are more firms trying to attract fewer workers, and workers are able to drive a harder bargain for the surplus. This effect is multiplied by workers' bargaining power,β, and the firms' costs of search ϱc. Greater productivity also increases the available surplus from which workers are able to bargain, hence the wage equation is also increasing in ϱ. Finally, workers demand a higher wage, the higher is the non-market wage (for example, unemployment benefit) available b.

Demand for labour

Equations (1.6)(a) and (1.10)(a) can be combined to derive what is effectively a demand for labour curve. Pissarides (2000) refers to it as a "job creation condition"

$$\varrho - w - \frac{(\delta + r)pc}{q(\theta)} = 0 \qquad (1.13)(\mathrm{a})$$

The equation looks like a simple zero marginal profit condition for a single job, however there is an additional term $-(\delta + r)\frac{\varrho c}{q(\theta)}$, which represents implicit and explicit flow search costs. Firms with vacancies incur an instantaneous search cost of ϱc over an average duration $\frac{1}{q(\theta)}$. A higher arrival rate of shocks, δ, increases the equilibrium search costs, because it means firms will close jobs more frequently and therefore will have to search more often. r represents the opportunity cost of putting resources into search.

Model (b) has a pair of equations governing labour demand, due to the presence of the extra endogenous variable, reservation productivity which is R:

$$(1-\beta)\frac{\varrho(1-R)}{r+\delta} = \frac{\varrho c}{q(\theta)} \qquad (1.13)(\mathrm{b})(i)$$

$$0 = R - \frac{b}{\varrho} - \frac{\beta}{1-\beta}c\theta + \frac{\delta}{r+\delta}\int_R^1 (s-R)dG(s) \qquad (1.13)(\mathrm{b})(ii)$$

$(1.13)(\mathrm{b})(i)$ is known as the job creation condition. It is useful to think about it as representing a locus of points in (θ, R) space, for which the relationship is a negative one. To see why, we can think about increasing θ from some positive initial value. The right hand side of $(1.13)(\mathrm{b})(i)$ is proportional to the expected search cost of filling a vacancy, as in $(1.13)(\mathrm{a})$. When the search costs are higher, the net expected benefit from vacancies or search (the left hand side) needs also to be higher. This can be achieved by lowering the reservation productivity R, since when R falls, firms are more tolerant of low realizations of productivity x, and when the firm is hit by a shock, there is a higher probability that the firm will keep the job open. Therefore, when R is lower, they will fire fewer workers, close down fewer jobs, and they will on average spend less time engaged in search for new workers. In other words, hired workers and jobs will last longer.

Equation $(1.13)(\mathrm{b})(ii)$, or the job destruction condition in contrast can be thought of as an upward sloping locus in (θ, R) space. It reflects the fact that a higher levels of market tightness θ, workers are able to bargain higher wages from firms (see $(1.12)(\mathrm{b})$). This means that firms then require jobs to have higher productivity in order for them to make sufficient profit for them to be worth keeping open. So as θ rises, firms require a higher reservation productivity R.

This completes our discussion of the basic forms of aggregate search models. Note what they have in common. Each model has a wage equation derived from the assumption of Nash bargaining, and each model has an equation governing unemployment dynamics, on which a steady state assumption may be imposed to give a solution for the equilibrium unemployment rate. Model (a) has one more equation which describes how vacancies are posted contingent on productivity wages, and separations. Model (b) has two such equations, since firms must not only determine how many vacancies to post, but the minimum level of idiosyncratic productivity that they are prepared to tolerate in any match.

1.3 The Mortensen Pissarides model, job reallocation and the choice of shocks

In this section I discuss the success of Mortensen Pissarides model with respect to its earlier goals – fitting stylized facts on job creation and destruction. Secondly, I discuss the fact that embedding the search and matching models into macroeconomic models has appeared to improve the fit of RBC models to the data along several dimensions – as found by Merz (1995) and Andolfatto (1996).

A literature that preceded the aggregate search and matching literature concerned the importance of shocks to the labour market, especially with relevance to the hiring behaviour of firms. The literature in some ways mirrored a similar debate in macroeconomics, regarding the relative importance of shocks to aggregate supply and aggregate demand-related factors. Supply-side shocks at the firm level have in some cases been ascribed to changes in the industrial structure of the economy – in terms of its relative demand for labour in different industrial sectors. The idea is that changing industrial structure creates frictional unemployment as workers move between expanding and contracting sectors. For example, Lilien (1982) found a positive empirical relationship between cross-sectional dispersion in U.S hiring growth rates at the 2-digit industry level, and layoffs, and in turn with unemployment. Lilien argued on the basis of exogeneity tests that causality runs from employment growth dispersion to unemployment, and that on this basis, sectoral reallocation can explain around 60% of the time variation in U.S unemployment over the period 1949-

1980 (p792). Abraham and Katz (1986) disputed Lilien's claimed importance of industry level shocks in favour of aggregate level shocks, by showing that countercyclical dispersion of employment growth at the industry level is quite compatible with aggregate shocks – on the condition that that slower growing industries are more cyclically sensitive. The authors presented evidence that this did indeed appear to be the case in the U.S economy. Instead of sectoral shocks, they argued for the importance of aggregate shocks, citing the aforementioned negative correlation between vacancy and unemployment rates at the aggregate level. Blanchard and Diamond (1989) used a structural VAR analysis to investigate the dynamics of the US unemployment, rate, vacancy rate and labour force between 1952 and 1988. They used sign and coefficient restrictions to identify the effects of cyclical and structural innovations (p59), and concluded that cyclical shocks were of much greater importance than structural shocks in explaining the variance of unemployment and vacancies over the short and medium run (p43). Their work thus echoed that of Abraham and Katz. In their (1992) analysis Davis and Haltiwanger examined the dynamics of employment growth using a panel of U.S manufacturing firms between 1973 and 1986. Defining the "job reallocation rate" as the sum of the rates of jobs created and destroyed during a particular year, they found, using a variance decomposition analysis that job reallocation is dominated by *idiosyncratic* changes in firm level employment growth, rather than sectoral changes, or shifts in the aggregate distribution of employment changes. Confusingly, the variances of job creation and job destruction rates were found to be *individually* dominated by aggregate and sectoral variation (pp853-854). Davis and Haltiwanger argued that their work suggested that models of employment dynamics should not rely on homogenous firms or homogenous sectoral responses to aggregate fluctuations (p859). The study provided more useful stylized facts: they found a large amount of job reallocation, in excess of the net reallocation rate – suggesting that one third to one-half of manufacturing job gains and losses are not related to aggregate labour demand (p834). They also found however that in the manufacturing sector there is a negative correlation between job reallocation and the net reallocation rate – as the latter is a good cyclical indicator this implies that job reallocation is countercyclical (p830). They argued that this shows that manufacturing job destruction is more volatile than job creation. The study was influential, and is cited by Pissarides (2000, p 63) for being one motivation for Mortensen and Pissarides (1994) – that is, model

(b) in section 1.2. At the end of their paper, Mortensen and Pissarides provided an illustrative simulation, the results of which showed that this model was able to generate both the negative correlation between job creation and job destruction, and a rate of job destruction more volatile than job creation (Mortensen and Pissarides p412). Cole and Rogerson (1999) examined Mortensen and Pissarides' (1994) model in more detail, studying in particular the ability of the model to match the stylized facts on job creation and job destruction noted above. Their approach was to note that for a simple version of the model with only two possible productivity states there were different structural versions of the model which all had the same reduced form. Therefore, they attempted to search for appropriate reduced form parameters which could reproduce those facts. As in Mortensen and Pissarides, Cole and Rogerson focused in particular on the negative correlation between job creation and job destruction and the fact that job destruction is more cyclically volatile than job creation – both facts reported by Davis and Haltiwanger. Cole and Rogerson claimed that under reduced form parameter values equivalent to the ones used by Mortensen and Pissarides the model was able to match at most one of those facts – but not both (p950). They argued that the model could fit both stylized facts simultaneously, if the steady-state unemployment hazard (that is – the hazard rate of an unemployed worker making a transition into employment) was reduced further below the value used by Mortensen and Pissarides, and indeed below the value that is consistent with official Bureau of Labour Statistics data. (p954). They argued that this might in practise be defensible, because the model abstracts from those workers on temporary quits and layoffs, who tend to have shorter durations thus pushing the BLS figure down, and because of evidence that there existed a subset of the officially economically inactive as classified by the BLS – that should have been rightfully be classified as unemployed. As these individuals tended to be less effective in searching, but should rightfully be counted as unemployed, their exclusion from the official BLS data means that the official unemployment hazard rate was again biased upwards. In summary, Cole and Rogerson showed that the reduced form of the Mortensen Pissarides model could be made to match the stylized facts on job creation and job destruction, for a certain restricted set of reduced form parameter values.

Aggregate search and matching models have themselves been used to improve the performance of RBC models according to calibration methodology. Merz (1995) is one effort in this direction. Merz added a matching function with variable search intensity and a search cost function on the part of unemployed workers to a macroeconomic model with productivity shocks. She found that search did improve the moment-matching capabilities of the model, on several dimensions. The labour market dynamics of the search model, for example, changed the contemporaneous correlation between employment and productivity relative to that found in the standard RBC model. With search frictions, employment reached its peak correlation with productivity at a lag –closer to the time series properties of the data (pp 282-283). Mertz also found that search costs dampened the volatility of wages, so that they were no longer of equal volatility to productivity, as in the standard RBC model with wages set according to the basic marginal product condition (pp281-282). This in turn generated a countercyclical labour share of output (as opposed to a constant share in the standard RBC model), also more in keeping with the data (p284). Finally, Merz noted that search frictions added persistence to unemployment and by extension to output. This contrasted well with pre-existing RBC models with no persistence, for example Hansen's (1985) indivisible labour model in which unemployment is determined period-by-period by lottery, for which the transition probabilities are independent of workers' labour market state. (p285) On the downside, the addition of search costs did result in insufficient output volatility (p285), and when variable search intensity was added to the model the observed negative correlation between vacancies and unemployment could not be re-created (pp284-295).

Andolfatto's (1996) paper is in a similar vein. Andolfatto stressed the ability of the RBC model, when combined with a search model to match the U.S data – in terms in particular of the volatility and covariance with output of employment and total hours. This he contrasted favourably with the basic RBC model, which for instance did not allow for labour inputs to vary via the employment rate at all. (p122). On the other hand, Andolfatto's hybrid model failed with respect to productivity and real wages – like the standard RBC model it displayed a near unit correlation between productivity and output, and between real wages and output respectively (p123).

The issues of the driving forces of job reallocation and real business cycles models are related – since all models mentioned in this section embody different views about the appropriate modelling of shocks to the labour market. The RBC framework has made a natural marriage with search theory, perhaps because both types of models have emphasised the importance of productivity shocks. It is worth mentioning that search theoretic labour market models have also been incorporated into the New Keynesian literature, see for example Krause and Lubik (2007), Barnichon (2007), who presents evidence that shocks in real-business cycle-style search models have their productivity shocks misidentified, conflating exogenous productivity movements with demand-driven endogenous shifts in work intensity (p3).

It is also the case that the Mortensen and Pissarides' own (1994) model evaluation exercise, the model evaluation exercise of Cole and Rogerson and the RBC hybrid models mentioned above are all based on calibration and pure moment-matching. We shall have more to say about that in Chapters 3 and 4 – but we note here that this is one of the ways in which this thesis diverges from the bulk of the existing model evaluation literature in this area.

1.4 Gross Flows

In section 1.2 we outlined two different versions of the aggregate search and matching model. In the first version, model (a), the rate at which employed workers made a transition to unemployment was given by a parameter δ. This meant that the only source of variation in the unemployment rate was variation in $m(u, v)$, or ultimately via variation in θ. Other models have allowed the separation shock arrival rate δ to vary, but as an exogenous variable only. In the second version, model (b), the total rate of separation was given by $\delta G(R)$, with δ being an exogenous parameter but R, and hence $G(R)$ being an endogenous variable.

The development of the aggregate search and matching literature has led to an interest in the question of which sort of model is more appropriate. How much of the cyclical volatility in unemployment is due to movements in the separation rate, and how much to outflows from unemployment? Economists have sought to settle the matter empirically as a basis for specifying search and matching models. In this section I review the evidence on this matter. Chapter 2 of this thesis draws directly on the methods used in this literature.

First however, I briefly consider the outflow from unemployment into employment. In contrast to inflows, the aggregate search and matching literature more or less universally accepts the aggregate matching function as an acceptable way of modelling hiring. In section 1.2 we postulated a matching function for use in the explanatory Mortensen-Pissarides type models of that section. The matching function posits a relationship between the stock of vacancies, the stock of unemployment, and the flow of hiring. We return to this now, but with a more specific functional form:

$$M_{t,t+1} = A_t U_t^{\alpha} V_t^{\beta}$$

$$\alpha \geq 0; \beta \geq 0$$

t refers to the beginning of a discrete period of time (typically a quarter or a month). The right-hand-side of the equation therefore refers to a set of labour market stocks at the beginning of period t. The flow of matches of the left-hand-side is a flow of matches that occurs during period t, up until the beginning of the new period $t+1$. This is the meaning of the double subscript.

This form has been widely adopted in the literature, including in the Mortensen Pissarides model. As this is a generalized Cobb-Douglas form, the returns to scale of the matching function are not restricted for $\alpha \geq 0; \beta \geq 0$, although the elasticity of substitution between unemployment and vacancies is restricted to be 1. The literature commonly estimates this relationship in logs using OLS, so that the estimating equation is linear. In many applications, the variables are also scaled as rates by the contemporaneous value of the labour force. It is also common to include a time trend. The literature produces a range of estimates, but the common finding is that α is in the range of 0.5-0.7. (Petrongolo and Pissarides (2000)). Estimates for α and β can be interpreted as "elasticities" of the stocks with respect to the flow of hires. The mechanism is simply that with more unemployed individuals, or more vacancies, or more of both, the rate of contact between them is higher - all other things being equal.

Most of the literature that estimates the matching function focuses on single-variable returns or returns-to-scale implied by the estimated parameters. If α and β are both less than 1, as implied by Petrongolo and Pissarides' study, then the standard interpretation given is that there are diminishing returns in terms of hiring to increasing either unemployment or vacancies. The implication is that there are not sufficient vacancies to meet the extra unemployed (or vice versa). Increasing vacancies and unemployment by the same proportion addresses the separate concept of returns to scale. The returns to scale are determined by the sum of the indices in the case presented above. If α and β sum to 1 exactly, then there are constant returns to scale in the matching function, $\alpha + \beta > 1$ implies increasing returns to scale and $\alpha + \beta < 1$ implies diminishing returns to scale. Coles and Smith (1996, p590) interpret increasing returns as being related to the density of the labour market – if vacancies and unemployment are scaled upwards by the same proportion, and all vacancies and unemployed individuals exist in the same area (that is, they are in the same market) – then increasing returns is what one would expect, since the rate of meeting will tend to be increasing in the density searching firms and workers. Petrongolo and Pissarides (2000) provide a meta-study on the returns-to-scale of the matching function literature. The results are not uniform across studies, although many are indeed consistent with constant returns (p425). Thus, many search-theoretic models have assumed constant returns to scale, including the Mortensen-Pissarides model. These findings matter, as the theoretical literature shows that differing qualitative returns to scale properties have different implications for the macroeconomy. (Diamond (1982).) However Barnichon (2013) shows that the standard matching function relationship of this type (with imposed constant returns to scale) breaks down after the 2007 for the US. Barnichon augments the standard matching function by including in (aggregate and labour-specific) match efficiency, time-varying labour market heterogeneities according to geography, industry and occupation. He shows that his specification provides a better fit to the post 1976 CPS labour market data, including the post great-recession period as far as 2012, owing to its capturing of time-varying effects upon dispersion and the composition of the labour market.

Blanchard and Diamond (1989) provide an early set of estimates of the U.S matching function, based on monthly data, for the period February 1968-December 1981. Across a range of specifications they find evidence of constant or increasing returns to scale in matching. They also highlight one potential econometric pitfall with matching function estimation. Because the regressors - vacancies and unemployment - are stocks which are reduced when regressand, hires, increases, any positive persistence in the deviations of hires around their predicted value (i.e. positive serial correlation in the error term) will be associated with subsequent periods of low unemployment and vacancies. This is simultaneity bias - OLS will over-attribute variation in hires to the error term and under-attribute it to variation in unemployment and vacancies so that the estimates will be depressed (p28). Recognising this, Blanchard and Diamond use the lagged values of unemployment and vacancies as instruments.

The residual is obviously also an important part of the matching function. In its most basic form, with just unemployment and vacancies and a time-trend, at least part of the residual can be attributed to 'mismatch'. The idea is that the residual may contain time-varying unobserved factors that affect hiring, such as search intensity, worker mobility, the propensity of workers to accept or reject job offers, which may in turn be influenced by the wage distribution and the benefit system. If data is available, these elements may be included in the matching function regression and their components identified.

Arguably the central concern of the aggregate class of aggregate search models to which the Mortensen-Pissarides model belongs is unemployment, of which hiring is only part of the story. Clearly, the inflows to unemployment (separations) as well as outflows (hires, or matches) must have some role in determining observed unemployment. The share of fluctuations in inflows and outflows in determining such fluctuations has been the subject of empirical debate in the literature, which we now consider.

Gross labour market flow equations can be expressed in terms of flow-levels. Consider equation (1.14). U_t, is the stock of the unemployed at time *t*. EU_t is the number of transitions from employment to unemployment during period *t,* and UE_t is the reverse flow, from unemployment to employment. Note that we have not included flows to or from inactivity (although these are in practise important), as the example given is merely illustrative:

$$U_{t+1} - U_t = EU_t - UE_t \tag{1.14}$$

The popularity of search theory has meant that analysis of equations such as (1.14) is no longer frequently used. Instead, let the labour force at time *t* be $L_t = E_t + U_t$, where E_t is the stock of employment at time t. Furthermore define $g^L_{t,t+1} = L_{t+1}/L_t$ to be the growth of the labour force between periods *t* and *t*+1 and define the unemployment rate and the employment rates respectively as $U_t/L_t = u_t$ and $E_t/L_t = 1 - u_t$. An equation for the evolution of the unemployment *rat*e can then be written as:

$$u_{t+1}g_{t,t+1} - u_t = (1 - u_t)\frac{EU_t}{E_t} - u_t\frac{UE_t}{U_t} \tag{1.15}$$

Search theoretic models tend to be written in terms of workers' probabilities of making labour market transitions between states, (for example, the probability of an employed worker transitioning from employment to unemployment). A simplifying assumption of many search models is to assume homogeneity of all the workers that are within a particular labour market state at a given time. This means, for instance, that $\frac{EU_t}{E_t}$ can be interpreted as the probability of any individual that is employed of making a transition from employment to unemployment during period *t* (the 'separation rate'), and $\frac{UE_t}{U_t}$ is analogously the probability of any individual that is unemployed of making a transition into employment (the 'job-finding rate').
Shimer (2005b) is recognised as a major contribution to the empirical literature on the relative influences of the separation and job-finding rates. His paper is notable for proposing a method of constructing estimates of the separation rate and job finding rate by matching individual households in the U.S Current Population Survey. His method has the advantage of correcting for time aggregation bias, a bias towards the under-reporting of labour market transitions when transition rates are calculated using stock data gathered at discrete intervals. If changes in monthly stock data are used to infer labour market transitions rates for example, then this will

fail to capture transitions that occurred within the month that were reversed within the same period. Shimer's method uses available data on short-duration unemployment to correct for this bias. He produces two distinct set of estimates, in which individuals can be in two and three states respectively. The two state analysis concerns unemployment and employment and the flows between them, and the three state analysis adds the possibility of inactivity and the extra requisite flows. Having used his new method to construct estimates for the job-finding and separation probabilities, Shimer begins by noting that setting the change in unemployment to zero in equation (1.15)[2] generates an estimated series for steady state unemployment which is highly correlated with the actual BLS unemployment rate between 1948 and 2004, with a coefficient of 0.99 (Shimer pp7-8). This suggests that U.S unemployment dynamics are sufficiently fast (that is, $\frac{EU_t}{E_t}$ and $\frac{UE_t}{U_t}$ are of sufficient magnitude), that U.S unemployment is never far from the steady state. Thus, Shimer argues that the convenience of looking at the effects of the transition rates on this measure of steady-state unemployment is justified. Thus, rather than modelling equations of type (1.4), one is justified in using the steady-state equivalent (1.5). (See Chapter 3).

Shimer's primary method in the 2-state case is to construct hypothetical unemployment rates, with only one or the other of the job-finding or separation probabilities allowed to vary through time. The other is held constant. In each case, Shimer then compares the resulting hypothetical series to the actual unemployment rate. He notes that the series in which the separation rate is held constant and the job finding rate is allowed to vary is more correlated with actual unemployment than in the opposite case. (The coefficients are 0.97 and 0.71 respectively). He takes this as evidence that the job finding rate makes a larger contribution than the separation rate to unemployment fluctuations (p8 and p33). Shimer also notes that the correlation between steady state unemployment with the job-finding probability held constant versus actual unemployment falls to 0.15 in the period following 1986 (p8). He infers that separations rates have become less important in the evolution of U.S unemployment since the mid-1980s.

Shimer follows a similar procedure for the three-state-system, which includes the

[2] Shimer's analysis also assumes that the labour force is constant, that is $g_{t,t+1} = 1$.

labour market flows between inactivity. Allowing for three states does not alter his conclusion that the flow from unemployment to employment explains more of the variation in unemployment than the reverse flow. The analysis upholds Shimer's 2-state findings, while also suggesting that the flow from unemployment to inactivity contributes to the increase in unemployment in low parts of the economic cycle, suggesting that workers prefer to retire from the labour force when the economy is doing well.

Shimer's (2005b) paper has become known for its emphasis on the job-finding rate – the "Outs" (from unemployment) - in explaining unemployment fluctuations. Hall's (2005) paper provides some additional evidence for this view.

His starting point is the JOLTS[3] survey, an official U.S survey of aggregate labour market transitions, time series for which begin in December 2000. The JOLTS survey confirms Shimer's finding about the recession of 2001 (which lasted according to the NBER from March 2001(Q1) to November 2001(Q4)), that there was no large increase in the separation rate compared to the trend level of separations. In fact, as Hall notes separations declined slowly from around 3.5% of employment to around 3% over the year 2001. This reflected a steady fall in the quit rate over the year, which slightly offset a gradual rise in involuntary separations over the same period. (See Hall pp104-105 and Figure 2.1 p105).   Yashiv (2007 p787) notes that separations figures recorded in JOLTS include job-to-job flows, as well as exits to employment and inactivity.

Hall reports that the number of monthly U.S. separations as a share of employment is significantly larger than the standard deviation of the growth rate of employment, across a broad range of all industries and as a weighted average of most broad industrial categories of the U.S. economy. He interprets this to mean that a substantial amount of monthly separations over the period covered by the survey (at Hall's time of writing December 2000-October 2004) are associated with worker transitions that do *not* correspond to firms' changes in their employment levels, and can be interpreted simply as average gross flows for the industry.

In the same paper, Hall considers other sources of data on separations.   He reviews the CPS[4]-based estimates of gross flows constructed by Shimer, and notes that they differ from the JOLTS data in that the former do not include job-to-job flows.

---

[3] JOLTS stands for "Job Openings and Labor Turnover Survey".

[4] CPS stands for "Current Population Survey".

Although Shimer's CPS-based estimates show more counter-cyclicality in the separation rate in the period from mid-century to the end of the nineteen-eighties than they do in later decades, Hall conjectures that this might be due to the exclusion of job-to-job flows. Employer-to-employer transitions have been estimated to make up 40% of total separations (Fallick and Fleischman, p11). Therefore if the level of employment-employment separations is positively correlated with the business cycle, the former would tend to fall in recessions, potentially offsetting an increase in separations to unemployment. Hall concludes that: *"there is nothing in the CPS flows data to suggest that total separations rise in recessions*" (p111).
Hall mentions other evidence that can be used to examine the separation rate. Gottschalk and Moffitt's separation rate estimates from the Survey of Income and Program Participation (SIPP) are consistent in the level with JOLTS data, display no counter-cyclicality, but do display positive co-movement with employment *increases*. Hall also notes that the study on tenure by Jaeger and Stevens (1999) provides no evidence of counter-cyclical increases in short-tenure, which one would expect if separations were taking the burden of employment adjustments in recessions. Hall considers these short time series "tentative evidence" for the relative unimportance of the separation rate in cyclical employment adjustments.

Elsby, Michaels and Solon (2009) adopt a linear decomposition of log-changes in steady-state unemployment, as an alternative to Shimer's counterfactual composition method. They use the same dataset as Shimer, built from matched monthly CPS records. Their revised method of analysis alone gives an estimate of a 35:65 split between the separation and the job-finding hazard rate in explaining changes in log steady-state deviations in unemployment between 1948 and 2004. (Elsby, Michaels and Solon p11). They also consider several refinements to Shimer's method of construction of transition rates from the gross flows. The authors plot their estimates of the contributions (which are in logarithmic form) of the job-finding and separation hazards to U.S unemployment for all of the recessions between 1948-2004. Their analysis shows a decline in the importance of the separation rate relative to the job-finding probability, especially since the 1980s. However, although movements in the job-finding probability are still dominant, prior to the 1980s separations still played a considerable role.
Elsby, Michaels and Solon's analysis follows Hall's (2005) in decomposing inflow

rates to unemployment and outflow rates from employment according to whether the separation was forced or voluntary. One would expect more voluntary separations to be job-to-job flows than to be job-unemployment flows. Elsby, Michaels and Solon show Hall's conjecture to be broadly correct. That is - the log of the overall separation rate is a-cyclical, but that within this masks a pro-cyclical rate of separation among job leavers and a counter-cyclical rate of separation among job losers. Contrary to what one might think from looking at the cyclical behaviour of the aggregate separation rate, layoffs are an important cause of unemployment during recessions (Elsby, Michaels and Solon pp18-21).

Fujita and Ramey (2009) use a different approach to analyse the contribution of the different rates from the CPS-based gross flows dataset. They use a variance decomposition of steady state unemployment into components associated with the separation and job finding rates respectively. There is also an error term. To generate series on separations and job finding rates they follow a similar data treatment procedure to Shimer, using a two-state model that corrects for aggregation bias.

As a prelude to their variance decompositions, Fujita and Ramey examine the correlations with labour productivity and the unemployment rate, the latter taken as different measures of the economic cycle. They find the separation rate to be contemporaneously correlated with productivity, with a correlation coefficient of -0.58 when an HP-filter(1600) is used to de-trend the data (p420). Noting that unemployment is a lagging and counter-cyclical indicator, they also find that the separation rate leads unemployment with a peak correlation of 0.50. Thus they report substantial counter-cyclical variation in the separation rate, more or less contemporaneous with the economic cycle. There is also a strong correlation of the cycle with the job-finding rate, with a peak at around 3 quarters after that of the economic cycle and a correlation at 0.6. Thus according to Fujita and Ramey's analysis both the inflows and outflows from unemployment are cyclical variables.

Fujita and Ramey's analysis suggests that variation in the separation rate explains at least 40% of variation in the steady state unemployment rate. This finding is unchanged even when Shimer's original data set is used, indicating that the difference is due to Fujita and Ramey's and Shimer's methods of analysis, rather than data construction. (The full sample is from 1976Q1 in Fujita and Ramey's data and from 1967Q2 in that of Shimer). When the sample is restricted to post- 1985 data, the separation rate's share of the variance falls to 34% (27% in Shimer's data). [5] (See Fujita and Ramey p427).

Finally, Elsby, Michaels and Solon argue that the observed dominance of the job-finding-probability in most of these empirical gross-flows studies is in fact consistent with causality running from the separation rate to the job-finding probability. They note that the unemployment to employment flow is well-known to be *counter*cyclical when measured without normalization (simply in terms of numbers of workers), or when it is normalized by employment (Yashiv) or the working-aged population (Fallick and Fleischman (2004)). Conversely, the job-finding probability from unemployment is pro-cyclical. The unemployment and the separation rate from employment to unemployment are both countercyclical. For all of these facts to be true requires that unemployment rise by proportionately more than the gross flow out of unemployment in a recession, a change which can only occur through greater flow of separations into unemployment. The large plunge in the job-finding probability that occurs in recessions is thus entirely due to an increase in the denominator (i.e. unemployment). Elsby, Michaels and Solon note that this is potentially compatible with a burst of separations being the ultimate cause of the proximate rise in the job finding probability: the job-finding probability falls simply because of the rise in unemployment. Thus it could conceivably be the case that separations have a driving role in unemployment *despite* the apparent dominance of the contemporaneous job-finding probability in the variance of the changes of steady-state unemployment. Elsby, Michaels and Solon's argument is consistent with Fujita and Ramey's (2009)'s finding that the separation rate leads the unemployment rate in the economic cycle (p421).

To summarize, aggregate evidence for the U.S labour market suggests that the job-finding probability has the largest share of unemployment fluctuations over the post-war period. The separation probability has a smaller but definitely non-zero share for the whole of the post-war period, although its importance seems to have declined since the mid-1980s. However, the separation probability is composed of pro-cyclical quits and counter-cyclical layoffs – the effect of the whole separation rate on unemployment is dampened as the latter are counter-acted by the former in periods

[5] A smaller share of the variance of steady state unemployment is attributable to separations when Fujita and Ramey use a different duration-based dataset.

of recession. The findings suggest that neither model (a) nor model (b) is quite correct in terms of its modelling of separations. More accurate analysis would seek to model separations by type. In the next section we review another, perhaps more fundamental shortcoming.

1.5 Evaluating the Mortensen Pissarides model

In an influential (2005(a)) paper, Robert Shimer criticises the Mortensen-Pissarides model for failing to generate realistic labour market volatility in response to exogenous productivity fluctuations. He calibrates his own version of the model using parameter values based on microeconomic literature, and shows that the ratio of the volatility of the vacancy-unemployment ratio to the ratio of labour productivity in the data over the period 1951-2003 (at a quarterly frequency and with the HP-filtered trend removed) is a factor of ten times larger than in the model. (p39) Not only does the model fail to match this particular moment, but vacancies, unemployment and the job finding rate all display less volatility in the model than they should (p28 cf. p39). Shimer tries out a variety of combinations of shocks to the model, including using shocks to the separation rate. He finds that most of these lead to worse counterfactual model behaviour – for example separation shocks generate positive correlation between unemployment and vacancy rates, contrary to what is observed in the data (p40). His main result therefore uses productivity shocks alone. The explanation given for the erroneous model mechanism is that the Nash-bargained real wages respond too much to labour productivity, eroding profits and the incentive to job creation (p45). The discrepancy between the data and the model has become known as the “Shimer puzzle.”

Mortensen and Nagypál (2007) examine the Shimer puzzle closely. They note that the correlation between the log vacancy-unemployment ratio and log productivity has a value of 0.396 (p333) - far less than unity, implying that productivity cannot be the only cause of fluctuations. The authors argue that it is incoherent to expect a model with one driving force to explain more than the empirical elasticity of the vacancy unemployment ratio to productivity – a value they report as being 7.56. (p333) From this argument it can be surmised that the Shimer puzzle is still a puzzle – but the magnitude of the discrepancy between data and model is less severe than Shimer's paper suggests. Mortensen and Nagypál then examine the ideas which have been suggested as corrections to the Shimer puzzle. We discuss some of these in the following section.

1.6 Amending the Mortensen-Pissarides model

Many attempts to amend the Mortensen Pissarides model in light of the Shimer puzzle focus on the flexibility of wages implied by the model, though this is not the only approach that has been tried. Hagedorn and Manovskii (2008), for example, note that it is not only the cyclical volatility of wages that matters for the volatility of the labour market variables in the model, but also the level of profits. They argue that the volatility of the vacancy unemployment ratio is in fact decreasing in profits, so that attempting to increase the labour market volatility simply by reducing labour's bargaining power is doomed to failure (p1695). Instead, the authors show that large increases in model volatility can be generated by simultaneously reducing bargaining power, while increasing the value of the parameter representing unemployment benefit or the value of leisure. Since wages under Nash bargaining are increasing in the value of leisure, this counteracts the rise in profits. Mortensen and Nagypál are however sceptical of this solution, noting that it implies that worker's average benefit from employment versus not working is a mere 2.3% (p335). Costain and Reiter (2007) acknowledge Hagedorn and Manovksii's finding that the model can be made to produce more labour market volatility in response to productivity fluctuations by calibrating unemployment benefits such that firms' profit is very small. However they also show that not only does this produce large fluctuations in the vacancy-unemployment ratio as the data require, it also has the effect of greatly increasing the elasticity of unemployment with respect to changes in the benefit level. The authors use cross-country panel data on 19 OECD countries for the years 1960-1999 with unemployment as the regressand, and estimate a multiple regression coefficient of around 0.02 for the association of the benefit replacement rate with unemployment. They interpret their estimate as an elasticity of benefits on unemployment of roughly 2 (p1138). Costain and Reiter find that increasing the unemployment parameter to equal that of Hagedorn and Manovskii produces a model elasticity of 14.29 for the same figure. The Shimer puzzle is thus not resolved so easily.

That rigid wages in search models can resolve the Shimer puzzle was demonstrated by Shimer himself (2004). Continuing the theme, Hall (2004) shows that the Nash bargain solution for wages is in fact an arbitrary point within a broader set of wages that would be mutually agreeable to rational firms and workers. He suggests that wage stickiness can be modelled as a wage norm, so that wages change only when shocks are of sufficient magnitude that the prevailing wage is no longer part of the admissible bargaining set. In a separate contribution, Hall and Milgrom (2008) replace the Nash bargaining solution in the canonical Mortensen and Pissarides model with a bargaining game of alternating offers between the worker and the firm. The authors theorize that the threat point values (those of unemployment and a job vacancy) are unlikely to be the practically relevant threat points in actual wage negotiations. Rather than workers leaving bargaining altogether, Hall and Milgrom reason that it is more likely that in the face of an unacceptable initial offer workers will propose a counteroffer. The result is a delay which is somewhat costly for the firm and which causes workers to have to subsist on unemployment benefits for extra time. The worker reverts to unemployment only with a small, exogenously imposed probability – representing the total breakdown of negotiations. Hall and Milgrom admit to having scant evidence on actual wage negotiations (p 1655), however their model of wage bargaining has the desired effect of dampening cyclical wage fluctuations – relative to the canonical model with Nash bargaining. The reason is that the Nash bargaining solution contains an expression of the worker's full value of unemployment – this value is pro-cyclical via its continuation value. In Hall's model, the value of the unemployment in the wage solution is damped by the small exogenous probability of complete failure of wage negotiations with no counteroffer (p1660). This is another mechanism through which the desired wage rigidity can be incorporated into the Mortensen-Pissarides model.

Pissarides (2009) however expresses scepticism that simply imposing aggregate wage rigidity can resolve the Shimer puzzle. Pissarides creates a version of the aggregate matching model which differentiates between wages paid to new and existing matches. Theoretically, he shows that the wage relevant to the Shimer puzzle is that paid to *new* matches. The problem is that empirical evidence shows that wages tend to be flexible in new matches, and rigid and existing matches (Pissarides p1360). Therefore, Pissarides argues that the focus on wage rigidity in solving the Shimer puzzle is misleading. Instead, echoing Mortensen and Nagypál (pp337-338), he suggests that adding fixed firm hiring and firing costs may help to solve the Shimer puzzle.

Pissarides explains his reasoning in a model of canonical type (a) (above). To see the effects on the model, let H be the fixed hiring cost, incurred only after a match is formed. Equation $(1.7)(\mathrm{a})$ then becomes:

$$rV = -\varrho c + q(\theta)(J - H - V) \tag{1.16}$$

This in turn implies that the value of a job is given by:

$$J = \frac{\varrho c}{q(\theta)} + H \tag{1.17}$$

The job creation condition $(1.13)(\mathrm{a})$ and the wage equation under Nash bargaining $(1.12)(\mathrm{a})$ become:

$$\varrho - w - (r + \delta)\left[\frac{c}{q(\theta)} + H\right] = 0 \tag{1.18}$$

$$w = (1 - \beta)b + \beta(\varrho + \varrho c\theta + \theta q(\theta)H) \tag{1.19}$$

In (1.18), the search cost now contains a fixed component $(r + \delta)H$. In the wage equation there is an extra term $\beta\theta q(\theta)H$, which indicates the fact that hired workers save firms an expected hiring cost, for which they are compensated according to their bargaining power, β. (Pissarides (2009, p1364)). Without having specified any parameter values, equations (1.18) and (1.19) look as if they add volatility to the wage equation, and so might reduce firms' job creating incentives when there is a positive productivity shocks. However, Pissarides argues that by reallocating search costs away from vacancy costs c, towards fixed hiring costs, H, this version of the model is able to match the elasticity of θ with respect to ϱ, whereas the original version cannot. When the bulk of search costs do not depend upon the expected duration of the vacancy (in other words the inverse of $q(\theta)$, which is increasing in θ), there is less of a dampening effect on job creation when there is a positive productivity shock. Building on Pissarides' paper, Silva and Toledo (2013) find that a necessary condition for hiring costs to significantly raise the sensitivity of labour market variables with respect to productivity is that the hiring costs be *sunk* when wage bargaining occurs. If the hiring costs are not sunk then they raise the firm's threat point, and in equilibrium they are passed on as a discount in the wages of new entrants.

Silva and Toledo (2009) also explore the potential of hiring and firing costs to solve the Shimer puzzle using a model of endogenous job destruction, one of type (b) above. Crucially however, they distinguish between newly hired workers and 'incumbent' workers. It is assumed that newly hired workers are less productive than incumbent workers by a proportion of realized match-specific productivity. The assumption is founded on evidence from the Employer Opportunities Pilot Project, which reports large training costs born on average by firms *after* the recruitment of a new employee (Silva and Toledo pp79-80). In addition, the authors assume both an exogenous separation probability of matches common to new and incumbent workers, *and* a firing cost to the firm of firing incumbent workers only. The model dynamics are assumed to be driven by productivity.

Silva and Toledo show that under standard calibrated values these assumptions are sufficient to bring the moments of the model closer to the data in a number of ways. Vacancies, the job-finding probability and the vacancy-unemployment ratio are all increased by at least a factor of two relative to the standard model. (The data standard deviations are still however at least 3.25 times the amended model, so the puzzle is not fully resolved.) The volatility result works by reducing the match surplus relative to the model without hiring and firing costs, proportionately increasing the incentives for job creation when productivity shocks hit (p89). This is the mechanism identified by Hagedorn and Manovskii and by Costain and Reiter. However, the solution survives Costain and Reiter's critique in that the modification does not induce unrealistic sensitivity of unemployment with respect to the benefit

replacement ratio. Silva and Toledo argue that while the smaller surplus does make hiring more sensitive to benefit changes, the extra firing costs serve to reduce the sensitivity of separations to the same changes, neutralizing the overall effect on unemployment (pp90-91). Finally, the authors note that in contrast to Shimer's (2005)(a) observation, the introduction of endogenous separations to their model does not induce a positive correlation between vacancies and unemployment.
Again Silva and Toledo state that this is because of firing costs in their model, which induce firms to ensure that the burden of adjustments to their stock of labour fall on hires rather than separations (p92).

1.7 Conclusion

This chapter has discussed the evolution of the aggregate search theoretic literature over the last twenty years, with an emphasis on the Mortensen Pissarides model and the Shimer critique. We have explained the empirical controversy over the question of whether or not the separation rate is exogenous or endogenous, and shown how this relates to the theoretical literature with the use of two canonical models. Finally, we have explored the response to the Shimer critique, which consists of attempts to amend the canonical model so as to bring the simulated moments closer to the data.
The main approaches to this puzzle have been the introduction of wage rigidity, and exploration of alternative specifications for search costs – in particular fixed search costs, sunk when wages are bargained over, which crucially are independent of search duration. The latter approach has been argued to be more consistent with the data by Pissarides (2009).

The two broad areas surveyed overlap exactly with Chapters 2-4 of this thesis.
Hence the rest of this thesis proceeds as follows: In Chapter 2 we take the tools developed in the empirical gross flows literature, and adapt them for an analysis of a small and very open labour market – that of Wales. In line with the literature we measure the contribution of labour market hazards to unemployment on a quarterly basis since 1997. We also critically consider the robustness of the procedure.
In Chapters 3 and 4 we build on the large model evaluation literature with respect to the canonical Mortensen Pissarides model and its extensions. However, rather than using the calibration-based procedures which are popular in the literature, we turn instead to the simulation-based procedure of indirect inference, which in Chapter 3

we argue constitutes a more statistically founded method of model evaluation than calibration-based moment matching. Chapter 3 is a “test-bed” for the procedure, in which we use a canonical Mortensen-Pissarides model with static expectations, to see how the model performs. Here we do not expect the model to perform well – being cognizant of the Shimer critique. We also take the chance to use the closely related method of indirect inference estimation to attempt to find a better-fitting model. In Chapter 4, we turn to the literature which attempts to address the Shimer critique. We subject to rigorous testing Yashiv’s (2006) version of the model, which in the vein of Pissarides (2009) and Silva and Toledo (2009, 2013), attempts to address the model with the use of an amended search cost function.

Chapter 2: Do the Ins win in Wales? A hazard rate variance decomposition of Welsh unemployment using LFS panel data.

2.1 Introduction

As described in more detail in Chapter 1, the 2000s have seen a revival in the empirical study of gross labour-market flows, especially those of the United States. This has coincided with the rising popularity in the use of search-theoretic models to describe the aggregate U.S labour market. A central paper of this recent literature is that of Shimer (2005), who matched monthly observations on the labour-market status of U.S individuals in the U.S Current Population Survey (CPS) in order to create a set of quarterly aggregate gross flows for the period 1967-2004. Rather than simply computing gross flows in levels, Shimer expressed the flows in terms of hazard-rates. The idea was to summarize the data in a way that afforded comparison with the output of theoretical search models. In models in which the modelled individuals are sufficiently (ex-ante) homogenous, hazard rates can be interpreted as conditional transition probabilities for individuals of moving from one state to another[6], which makes them natural for modelling labour market flows in continuous time.

Shimer suggested that most of the variation in the U.S unemployment rate could be

[6] Conditional here means that it applies to individuals that have not yet made a transition.

attributed to cyclical changes in the probability of unemployed individuals finding a job, and that the probability of employed individuals becoming unemployed – the separation probability - was of little importance. If true, the finding could validate the modelling assumption of an exogenous separation rate, as used for example in Pissarides (2000, chapter 1), Yashiv (2006) and Chapters 3 and 4 of this thesis. Although it was acknowledged that Shimer's paper contributed novel ways of handling various data issues, many thought the conclusion to be driven by a combination of somewhat arbitrary data-handling procedures and a potentially misleading empirical method. (Gomes, 2009 p28). There followed several papers that built upon Shimer's study and sought to improve the empirical methods used, and which failed to uphold Shimer's central finding – changes in the U.S separation probability can explain at least 28% of the variation in U.S steady-state[7] unemployment, and more depending upon empirical and analytical choices. (See for example Fujita and Ramey, (2009) p427, and Elsby, Michaels and Solon (2007), p10).

There have been similar recent attempts to investigate the importance of different gross flows and hazard rates in the UK labour market. Most have had results similar to the more recent U.S analysis, finding that both job-finding rates and separation rates make important contributions to fluctuations in UK unemployment rate. Petrongolo and Pissarides (2008) decompose quarterly UK Labour-Force panel data and claimant-count data, and find that the inflow rate to unemployment explains a minimum of 20% and as much as 45% of movements in steady-state unemployment.[8] In a similar analysis which uses a range of methods, Gomes (2009) finds that the separation rate contributes around 40% of steady-state unemployment volatility at a quarterly frequency, although the contribution of the separation rate tends to be lower at monthly and weekly frequencies. Using a different sort of unemployment data - LFS recall-based micro-data that stretches back to 1975 - Elsby, Smith and Wadsworth (2011, p355) find a 64.1% - 28.2% split

[7] The analyses tend to maintain (with some justification) the assumption that U.S unemployment is well-approximated by steady-state unemployment – roughly speaking the unemployment rate that holds when the inflows and outflows to the unemployment are equal. See Chapter 1 for a broad discussion of the issue.

in the contribution of separations to job-finding hazard rates to the variance of changes in steady-state unemployment, with combined inactivity-related hazard rates contributing just under 10%. Smith (2011) provides a dynamic, non-steady state analysis of UK unemployment by hazard rate based on British Household Panel Survey (BHPS) data, in which she allows both current and past changes in hazard rates to affect observed unemployment, motivated by the UK's relatively small gross flows relative to those of the U.S. Her long sample allows her to calculate time-varying variance shares for inflows and outflows to unemployment. The results are broadly supportive of those of Elsby, Smith and Wadsworth, with lagged and contemporaneous inflows to unemployment accounting for over 100% of the variance of observed BHPS unemployment around the late 1980s and early 1990s, and around 40% since the late 1990s. She also finds however, that outflows from unemployment were comparatively more important than inflows during the mid-to-late 1990s, over the period of relative economic stability and declining unemployment (Smith p420).

The extension to the literature made here is to estimate the contributions of inflows and outflows to unemployment for Wales. We use Labour Force Survey panel data to obtain gross flows and hazard rates for the period 1997Q2-2010Q4. The analysis includes as far as possible cross-border flows between Wales and the rest of the UK. We use a steady-state framework in the style of Petrongolo and Pissarides, rather than a dynamic framework in the style of Smith, but we also examine how closely the steady-state approximates observed unemployment. This is to our knowledge the first attempt to apply this sort of flows analysis to regional data.

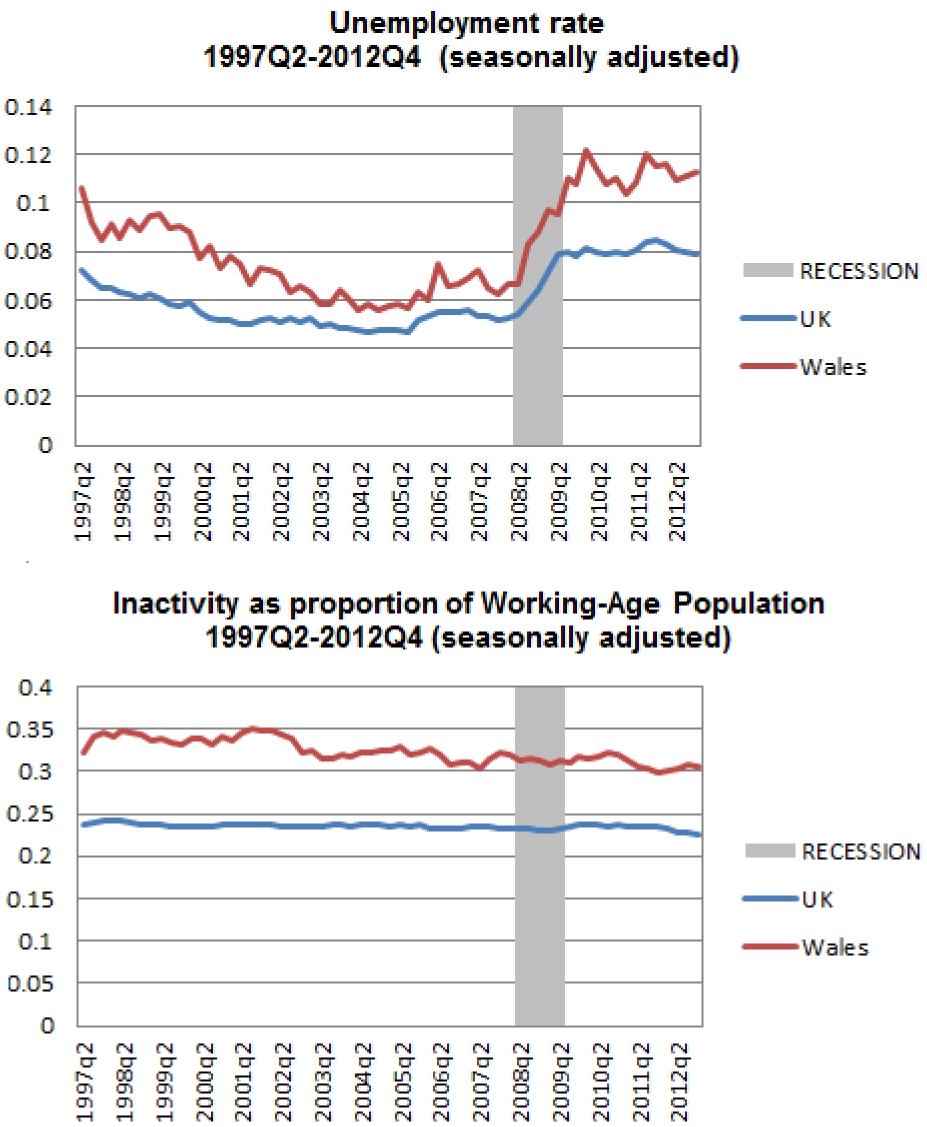

Figure 2.1: Unemployment and inactivity rates in Wales and the UK, 1997Q2-2012Q4 (seasonally adjusted and based on UK Labour Force Survey stock data).

Figure 2.1 shows the unemployment rates and inactivity rates for the UK and for Wales between 1997Q2 and 2012Q4. The graphs show that both are higher in Wales. Wales also suffered a sharper increase in unemployment during the recession that began in 2008 (shaded in grey in the figure). The graphs therefore suggest important differences between the labour markets in the two areas. This is confirmed by NOMIS data, which show, for example that the proportion of the population that is of working age (16-64) in Wales is persistently smaller than that of Great Britain (The UK minus Northern Ireland). The occupational structure of the

Welsh economy is also different, with persistently higher proportion of process, plant and machine operatives, caring, leisure and service occupations, and skilled trades, and a lower proportion of managerial, professional and technical occupations. Wales also has a different distribution of skills in its working-age-population compared to the whole of Great Britain. For example, the NOMIS data shows that Wales has a somewhat lower proportion of highly skilled workers than the rest of Great Britain (NVQ level 4 and above), and a higher proportion of individuals with no qualifications.[9] To the extent that these different worker characteristics are associated with different labour-market dynamics, we would expect Wales to exhibit a different pattern of gross flows in our analysis.

Section 2.2 describes the panel dataset, data-handling procedures, and provides summary statistics for the gross flows measures that we derive from it. We also describe how we define whether or not individuals in our dataset are associated with Wales. Section 2.3 explains how we obtain hazard rates from the gross-flow measures, and notes some of the methodological choices taken along the way. Section 2.4 reviews Petrongolo and Pissarides' (2008) variance decomposition of the changes in UK unemployment. As their analysis is the closest in the existing literature to our own, we redo their analysis with our own LFS panel dataset in order to provide UK estimates for our sample period which are comparable with our results for Wales, in terms of being estimated over the same sample period and in terms of being subject to the same data-handling procedures. Section 2.5 describes how we obtain steady-state estimates of the unemployment rates both for Wales and Outside-of-Wales. Section 2.6 analyses the correlation between steady-state unemployment rates (which are constructed using the hazard rates) and observed unemployment rates, for the UK, for Wales and for the area of the UK Outside-of-Wales. Section 2.7 contains the results of our variance decompositions for each of the geographical areas. In section 2.8 we discuss our results and section 2.9 concludes.

[9] http://www.nomisweb.co.uk/reports/lmp/gor/2013265930/report.aspx

2.2 Data

We use data from the 2-quarter Labour Force Survey panel, between 1997Q1 and 2010Q4. The Labour Force Survey (LFS) is a quarterly address-based survey conducted by the UK's official statistical agency, the Office of National Statistics (ONS). The survey is based upon five waves, meaning here that an address that has been selected for participation will be interviewed for five consecutive quarters. In any given quarter, survey respondents will be made up of responding addresses that are on their first of five interviews (wave 1), those that are on their second of five interviews (wave 2), etc, up to those wave 5 addresses that are on their fifth interview. When the address has been interviewed for five consecutive quarters, it exits the sample. Each quarter, new addresses must be selected to comprise the wave 1 part of the sample.

Attempts are made to interview each individual in a sampled household, although proxy responses are permitted in some cases. (Labour Force Survey, User Guide Volume 1 – LFS background and methodology p29.) Data are collected on many different sorts of variable. That which is most of interest here is of course data on labour-market status, and also geographical location. Quarterly LFS datasets are available in cross-sectional, two-quarter and five-quarter panel form. Each dataset includes a set of "person weights", which may be used to scale the desired observations up to the population level in each quarter. The weights in the two-quarter and five-quarter panels have the added feature of correcting for attrition bias.

We use the two-quarter panel datasets to obtain data on gross flows for the area of the UK Outside-of-Wales respectively. The panel datasets make this very easy, as each contains a ready-made variable[10] that indicates whether individuals at each sampled address made a transition between – or remained within – any of the labour market states of employment, unemployment and inactivity. The fact that changes in labour market status among individuals is already coded in the panel means that there is no need to match individuals based on other characteristics, as is the case

[10] The variable is called 'FLOW'. See *Labour Force Survey User Guide – LFS two quarter and five quarter longitudinal datasets-2011*; version 2.0 March 2012, p3 for more details.

for example, when constructing gross flows from the U.S Current Population Survey data.

Each of the quarterly panel datasets contains an estimated number of individuals in each of the initial labour-market states, which can be grossed up to an area population-level estimate using the panel weights. We use these panel data-based stocks, rather than the separate cross-sectional datasets which are also available, to ensure consistency between the stocks and flows.

It is clear that any analysis of gross flows between regions within an economy will require definitions of the economically relevant "boundary" between regions. How should "Wales" versus "not in Wales" be demarcated? The LFS dataset contains enough variables such that there is more than one possible answer. In the analysis that follows, we use individuals' location of residence[11] to assign individuals to the Wales or Outside-of-Wales categories, consistent with the standard geographical definitions of labour market states

Before discussing the gross flows, we briefly review summary statistics on the labour market *stocks* for the areas under consideration. These are displayed in table 2.1. The table shows that Wales contains roughly 4.7% of the UK's working age population, and 4.4% of the UK's labour force. Over the full sample from 1997Q2-2010Q4, the Welsh unemployment rate was on average 0.6 percentage points higher than that of the UK as a whole, however the difference is due to the period of economic weakness in the latter part of the sample that began in 2008Q2 – the difference in average unemployment between the earlier and later parts of the sample period was 3.4 percentage points in Wales but 2.6 percentage points for the whole of the UK. Table 2.1 also shows that over the full sample period, economic inactivity has been persistently higher in Wales than in the rest of the UK.

[11] The variables that identify location of residence in the 2-quarter panel dataset are called 'GOVTOR1' and 'GOVTOR2'.

Table 2.1: Summary statistics on labour markets			
		Working-age population	
	UK	*Outside-of-Wales*	*Wales*
Full sample 1997Q2-2010Q4 (T=55)	36,724,969	35,235,081	1,489,888
		Labour force	
	UK	*Outside-of-Wales*	*Wales*
Full sample 1997Q2-2010Q4 (T=55)	29,288,746	27,987,818	1,300,928
		Unemployment rate	
	UK	*Outside-of-Wales*	*Wales*
Full sample 1997Q2-2010Q4 (T=55)	0.058	0.057	0.063
1997Q2-2008Q1 (T=44)	0.054	0.054	0.060
2008Q2-2010Q4 *(T=11)*	0.076	0.072	0.076
		Inactivity/Labour force	
	UK	*Outside-of-Wales*	*Wales*
Full sample 1997Q2-2010Q4 (T=55)	0.271	0.267	0.347
1997Q2-2008Q1 (T=44)	0.276	0.268	0.351
2008Q2-2010Q4 *(T=17)*	0.266	0.263	0.327

Notes: Seasonally adjusted LFS data. T denotes sample size (number of quarters).

Table 2.2 shows summary statistics (means and standard deviations) for the levels of all the seasonally adjusted flows derived from the LFS 2-quarter panel dataset. Columns (i) and (ii) correspond to actual numbers of transitions, whereas (iii) and (iv) correspond to transitions expressed as a proportion of the relevant working-age population in the quarter. The flows ‘UE’, ‘EU’, ‘IU’, ‘UI’ ‘IE’, ‘EI’ and ‘EE’ pertain to the whole of the United Kingdom. For example, the first column of the row labelled ‘UE’ shows the estimated per quarter number of individuals in the United Kingdom that moved from the state of unemployment to that of employment over the sample period. The other flows contain suffixes to indicate whether the statistic includes individuals moving ‘from’ or ‘to’ a state associated with Wales (‘w’), or Outside-of-Wales (‘o’). For example, column (i) of the row labelled ‘U^wE^w’ contains the estimated per quarter number of individuals that moved from a state of being unemployed in Wales, to a state of being employed in Wales over the sample period.

In columns (iii) and (iv) the relevant working-age population in terms of which each flow is expressed is the area from which the flow originated. Flows that have the form ‘XY’ - in other words, flows estimated for the whole of the United Kingdom - are

expressed as a proportion of the working-age population of the United Kingdom. Flows that pertain to Wales and which therefore have the form 'X^wY^w' are expressed as a proportion of the working-age population of Wales. Flows that pertain to Outside-of-Wales and which therefore have the form 'X^oY^o' are expressed as a proportion of the working-age population Outside-of-Wales.
Note that each of the employment-employment flows is the largest of its type. An average of 1.9% of the working age population of the UK per quarter changed jobs over the sample period. In this context, EE means someone has changed jobs rather than remained in the same job thus placing the emphasis on employment flows. The employment-employment flows are presented only for context – we do not use them in the decomposition of unemployment that follows.

Table 2.2 makes clear that sample-size is an issue in estimating gross-flows for Wales using the 2-quarter LFS panel. Consider column (i). The LFS Longitudinal Panel Guide [12] stipulates a minimum publication threshold of 17,000, for any estimates derived from the 2-quarter dataset. At this level the sampling variation implies a standard error of around 20% (LFS Longitudinal Panel Guide, p5) – a level deemed unacceptable for official publication. Column (i) of table 2.2 shows that the gross flows estimates for the whole of the UK meet this criterion comfortably. This is also true for all flows of the form 'X^wY^w' and 'X^oY^o' (in other words, for all flows within Wales or within the Outside-of-Wales area), with the exception of separations from employment in Wales to unemployment in Wales ('E^wU^w'). For these flows, the sample sizes for the Outside-of-Wales area tend to be much greater than the sample sizes for inside of Wales, which means that the cross-quarter standard deviations of the flows as a percentage of the respective working-age populations in column (iv) tend to be greater for 'X^wY^w' estimates than for 'X^oY^o' estimates.

Ideally, we would like to include in our data analysis flows of the form 'X^oY^w' and 'X^wY^o.' However, the numbers of such quarterly flows are evidently so small that they are no such observations in the panel sample. It is highly unlikely that the number of individuals making such transitions is exactly zero in every quarter, but the point is that there are insufficiently many of such transitions to show up in the

[12] *Labour Force Survey User Guide – LFS two quarter and five quarter longitudinal datasets-2011*; version 2.0 March 2012.

LFS' limited panel sample. Hence they are not included in table 2.2, nor in further results.

Table 2.2 shows that, on average the flows from unemployment into employment were the second largest flows (second to employment-employment flows) over the sample period. Flows from employment to inactivity were almost as large. ('$E^w I^w$' is in fact estimated to be slightly bigger than '$U^w E^w$' but the difference is unlikely to be significant.) Flows from unemployment to inactivity tended to be smallest, at around 0.8% of the Working-age population for the UK a whole. Flows from employment to unemployment were somewhat bigger at around 0.9% for the UK.

Note that as a percentage of the working-age population, all gross flows in Wales are estimated to be greater than their counterparts for the area Outside-of-Wales and for the whole of the UK. We attribute this to the base-effect of the Wales working-age population. They also consistently more volatile, with more than double the standard deviation than their Outside-the-UK counterparts, for all flows with the exception of job-to-job flows ($E^W E^W$).

It would be useful to see whether we are able to reject the null hypotheses that each of the Wales and UK flows are equal, in each period or on average. However constructing confidence intervals for these flows based upon panel transitions (which are in turn based on weighted estimates of a smaller underlying sample) is not simple, and requires more information on the sampling process. We would also need to incorporate the sampling variability of the relevant working-age population in each area. We therefore leave the testing of this set of hypotheses for further work.

Table 2.2: Summary statistics for gross flows data, 1997Q2-2010Q4

Flow	*Mean level*	*Standard dev.*	*Mean, % of $Wpop$*	*Standard dev., % of $Wpop$*
	(i)	(ii)	(iii)	(iv)
UE	459,584	43,037	0.0125	0.0011
U^wE^w	21,636	4,874	0.0146	0.0034
U^oE^o	437,947	41,680	0.0124	0.0011
EU	338,782	50,012	0.0092	0.0013
E^wU^w	14,764	4,135	0.0099	0.0028
E^oU^o	324,017	48,404	0.0092	0.0013
IU	362,579	61,932	0.0098	0.0015
I^wU^w	19,364	5,508	0.0130	0.0035
I^oU^o	343,215	58,652	0.0097	0.0014
UI	288,811	47,475	0.0079	0.0012
U^wI^w	17,580	4,811	0.0118	0.0032
U^oI^o	271,230	45,712	0.0077	0.0012
IE	435,346	40,768	0.0119	0.0011
I^wE^w	20,959	5,238	0.0141	0.0035
I^oE^o	414,388	38,251	0.0118	0.0011
EI	454,624	37,028	0.0124	0.0011
E^wI^w	21,774	5,333	0.0146	0.0036
E^oI^o	432,850	34,888	0.0123	0.0010
EE	689,533	120,193	0.0189	0.0036
E^wE^w	29,337	6,704	0.0197	0.0046
E^oE^o	660,196	115,600	0.0188	0.0037
T	55	55	55	55

Notes: T denotes the sample size (number of quarters of data). "$Wpop$" stands for "Working-aged-Population". Seasonally adjusted data.

The gross flows expressed as a percentage of the relevant working-age population are displayed in figure 2.2 below. The graphs illustrate the fact that the gross flows for Wales are very much more volatile than those for the Outside-of-Wales and for the whole of the UK. They also show that as a share of the working-age-population, the mean level of gross flows in Wales tends to be somewhat higher than in the UK as a whole. A large increase in the flow into unemployment from employment is visible in all areas following the recession that began in 2008. Otherwise, the flows do not display a clear cyclical pattern.

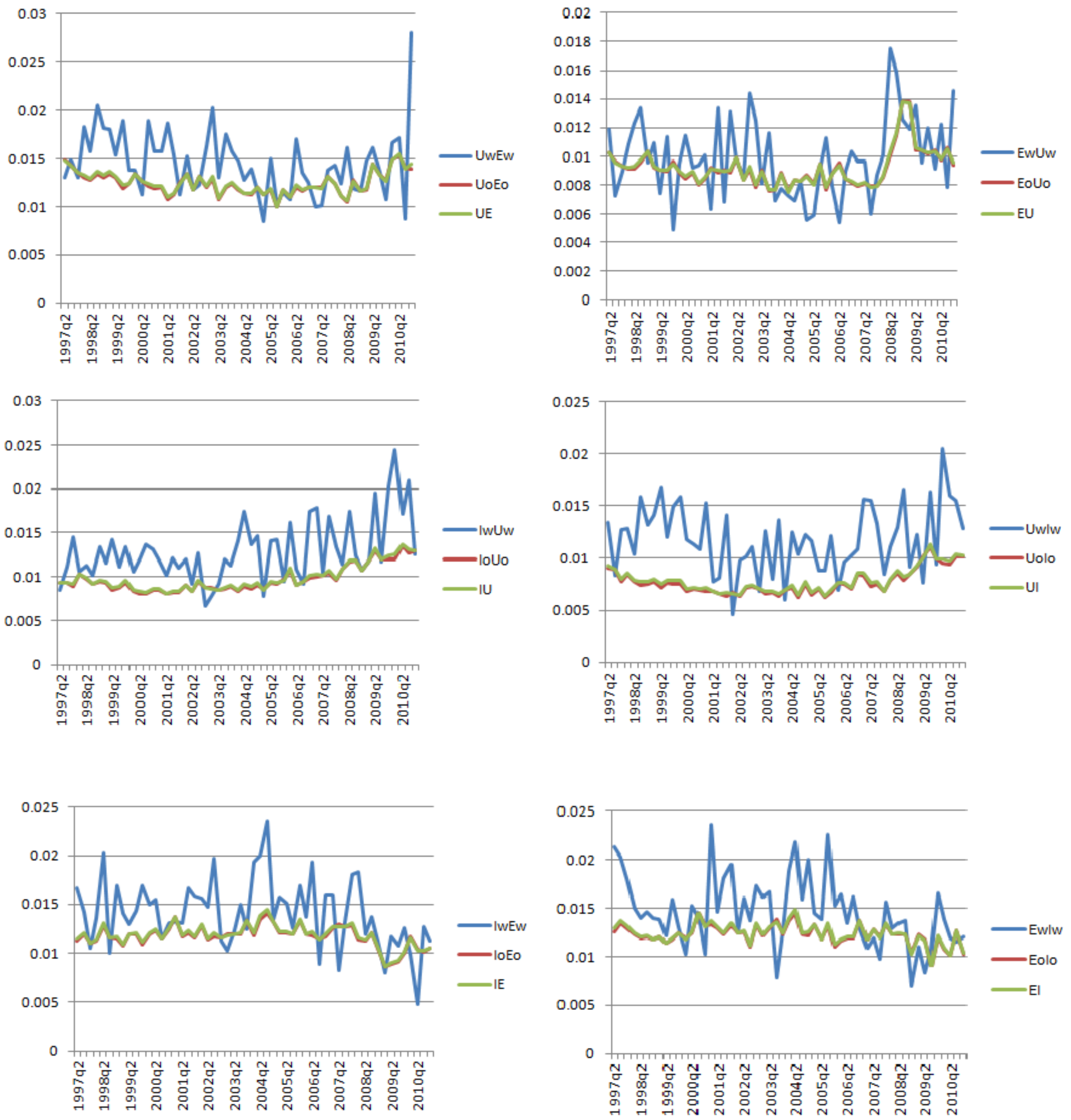

Figure 2.2: Gross flows for Wales, Outside Wales and the UK as a % of the relevant Working-Age population, seasonally adjusted

2.3 Calculating hazards rates from discrete gross flows: ignoring time aggregation bias

The gross flows literature on which this chapter is based obtains continuous time labour market measures from a discrete dataset. In this view, labour market flows change continuously according to the labour market's underlying hazard rates. The hazard rates are themselves assumed to be constant for the *discrete* period under examination (in our case, for a given quarter).

Table 2.2 summarizes the gross flows data that the LFS 2-quarter panel data allowed us to construct. How should one go about constructing hazard rates from this data?

The 'correct' way of constructing hazard rates in this way is to construct the hazard rates jointly. Unfortunately, this is only simple in model with two labour market states. We illustrate the point with a very simple model.

Consider a single country model in which workers can move only between employment and unemployment. Discrete quarters of time are indexed by $t = 1,2,,\ldots,T$. Time also elapses continuously between quarters, and is indexed by the intra-quarter variable ∈ $[0,1]$. Let $EU_t(\tau)$ be the cumulative number of workers that have made the transition between employment and unemployment since the beginning of period *t*, up to and including the instant $\tau \in [0,1]$ of period *t*. By definition, $EU_t(0) = 0$ since no time has elapsed and so no transitions have taken place. When $\tau = 1$ we have the total number of EU transitions for period *t*, this number can be denoted by $EU_t(1) = EU_t$. There is an opposite flow of workers from unemployment into employment, $UE_t(\tau)$, which obeys the same timing conventions. The stocks of unemployment and employment also vary continuously in the model and can therefore be written as $U_t(\tau)$ and $E_t(\tau)$ respectively.

It is assumed that within each period *t*, the flow from employment to unemployment $EU_t(\tau)$ is driven by a constant continuous time hazard rate s_t and the flow from unemployment to employment $UE_t(\tau)$ is driven by a constant continuous time hazard rate f_t. It is assumed that we are only able to observe any of the labour market stock or flow variables at discrete intervals – at the end of each of the discrete quarters when $\tau = 1$. The problem is to estimate the hazard rates in each period, s_t and f_t, from the discretely observed data. Note that this is a simplified version of the general question addressed by this chapter, the latter being complicated by the desire to model extra countries and states.

To proceed further, it is useful to review exactly what the assumption of a constant hazard rate entails, with reference to the rest of the model. Firstly, consider the magnitude $\frac{EU_t(\tau)}{E_t(0)}$ for some $\tau \in [0,1]$. This gives the proportion of individuals that were employed at the beginning of period *t* that have made the transition to unemployment by $\tau \in [0,1]$. The converse of this magnitude, $1 - \frac{EU_t(\tau)}{E_t(0)}$ is the proportion of those individuals that were employed at the beginning of the period who have *not* made the transition, and is called the *survivor function*. Finally, let $\eta_t^{EU}(\tau)$ be the instantaneous probability of an individual moving from employment to unemployment at time $\tau \in [0,1]$. $\eta_t^{EU}(\tau)$ is an unconditional transition probability.

Under these assumptions the proportion of those making transitions $\frac{EU_t(\tau)}{E_t(0)}$ and $\frac{UE_t(\tau)}{U_t(0)}$ satisfy the following pair of equations:

$$\frac{EU_t(\tau)}{E_t(0)} = \int_0^\tau \eta_t^{EU}(v)\,dv - \int_0^\tau \frac{\eta_t^{UE}(v)}{\left(1 - \frac{UE_t(v)}{U_t(0)}\right)} . \frac{EU_t(v)}{E_t(0)} . dv$$

$$\frac{UE_t(\tau)}{U_t(0)} = \int_0^\tau \eta_t^{UE}(v)\,dv - \int_0^\tau \frac{\eta_t^{EU}(v)}{\left(1 - \frac{EU_t(v)}{E_t(0)}\right)} . \frac{UE_t(v)}{U_t(0)} . dv$$

In general, the fact that the hazard rate of making the transition from employment to unemployment is assumed to be constant for period *t* means that the transition probability is a constant proportion of the survivor function in each period – that is:

$$s_t = \frac{\eta_t^{EU}(\tau)}{1 - \frac{EU_t(\tau)}{E_t(0)}}$$

$$t = 1, \ldots, T$$

Defining similar notation, there will be an equivalent expression for the job-finding hazard rate:

$$f_t = \frac{\eta_t^{UE}(\tau)}{1 - \frac{UE_t(\tau)}{U_t(0)}}$$

$$t = 1, \ldots, T$$

The pair of differential equations then becomes:

$$\frac{EU_t(\tau)}{E_t(0)} = \int_0^\tau s_t \left(1 - \frac{EU_t(v)}{E_t(0)}\right) dv - \int_0^\tau f_t . \frac{EU_t(v)}{E_t(0)} . dv \tag{2.1)(a}$$

$$\frac{UE_t(\tau)}{U_t(0)} = \int_0^\tau f_t \left(1 - \frac{UE_t(v)}{U_t(0)}\right) . dv - \int_0^\tau s_t \frac{UE_t(v\tau)}{U_t(0)} . dv \tag{2.1)(b}$$

$$t = 1, \ldots, T$$

Differentiating equations (2.1)(a) and (2.1)(b) gives a pair of differential equations in $\frac{\dot{EU_t}(\tau)}{E_t(0)}$ and $\frac{\dot{UE_t}(\tau)}{U_t(0)}$ respectively. By solving these equations and letting $\tau = 1$ then the hazard rates can be shown to satisfy the following pair of equations for each $t = 1, \ldots, T$:

$$\frac{EU_t}{E_t(0)} = \frac{s_t\left(1 - e^{-(s_t + f_t)}\right)}{s_t + f_t} \tag{2.2)(a}$$

$$\frac{UE_t}{U_t(0)} = \frac{f_t\left(1 - e^{-(s_t + f_t)}\right)}{s_t + f_t} \tag{2.2)(b}$$

$$t = 1, \ldots, T$$

$\frac{EU_t}{E_t(0)}$ and $\frac{UE_t}{U_t(0)}$ are observed in the data, so the equations can be inverted numerically and solved for s_t and f_t. Equations (2.2)(a) and (2.2)(b) shows that modelling time aggregation bias implies solutions for s_t and f_t that are mutually inter-dependent.

Note that these equations are equations (4) and (5) of Petrongolo and Pissarides (2008).

The idea behind these equations is as follows. Labour market surveys occur at discrete intervals. On a quarterly basis, a certain number of individuals will have a pattern of transition that is a single move from employment to unemployment, we can write this as $E-> U$. Clearly, we would like to construct our continuous time hazard rate s_t so as to take account of these individuals. However, there may be another group of individuals that have a pattern of transition $E-> U-> E$ which occurs during a single quarter. This means that at the two discrete quarterly interview dates that enclose one separation and one job-finding event, they are classified as employed. The survey design will not take into account the fact that they made a transition $E-> U$ and another $U-> E$ within the period between surveys, and so this transition will be incorrectly omitted from the measured EU and so from the calculated s_t. The longer the discrete interval between interview dates, all other things equal, the greater the chances of such unregistered transitions occurring. Thus this phenomenon is known as time aggregation bias. Unfortunately, when the number of states is greater than 2 it becomes more difficult to adjust for time-aggregation bias. Gomes (2009, Appendix) models the bias for a three-state system (including inactivity) using LFS data and a system of non-linear equations. Petrongolo and Pissarides (2008, p258 footnote 2), use LFS data and decompose UK unemployment using both a 2-state and 3-state model. Noting that adjusting their 2-state results for time aggregation bias does not substantially affect their results, they decline to make a similar adjustment to their 3-state results. As our model also contains flows between 3 states we follow Petrongolo and Pissarides, and compute our gross flows independently using the LFS panel data.

Consider the general flow for interval $[0, \tau]$ of period *t* from state X_t^j to state Y_t^k , $j, k \in \{o, w\}$. Abstracting from time aggregation bias implies the total number of transitions relative to the initial stock is simply the integral of the hazard:

$$\frac{X^j Y^k{}_t(\tau)}{X^j{}_t(0)} = \int_0^\tau \eta_t^{jk}(v)\, dv$$

Once more we assume the hazard rate is constant throughout the whole discrete time period. Let the constant hazard rate for flow $X^jY^k{}_t$ for period *t* be given be λ_t^{jk} Since by the definition of a hazard rate,

$$\lambda_t^{jk} = \frac{\eta_t^{jk}(\tau)}{\left(1 - \frac{X^jY^k{}_t(\tau)}{X^j{}_t(0)}\right)}$$

Where $\eta_t^{jk}(\tau)$ is the instantaneous transition probability and $\left(1 - \frac{X^jY^k{}_t(\tau)}{X^j{}_t(0)}\right)$ is the survivor function.

It is clear that when we abstract from time aggregation bias that:

$$\frac{X^jY^k{}_t(\tau)}{X_t^j(0)} = \int_0^\tau \lambda_t^{jk}\left(1 - \frac{X^jY^k{}_t(v)}{X_t^j(0)}\right)dv$$

$$t = 1, \ldots, T$$
$$j, k \in \{o, w\}$$

By solving this equation it can be shown that the relationship between the transition rate and the hazard rate is given by[13]:

Or:

$$\lambda_t^{jk} = -ln\left(1 - \frac{X^jY^k{}_t}{X_t^k(0)}\right)$$

$$t = 1, \ldots, T$$
$$k \in \{o, w\} \qquad (2.3)$$

Which when inverted gives:

[13] The formula is not original however most of the literature omits its derivation.

$$\frac{X^j Y^k{}_t}{X_t^j(0)} = 1 - e^{-\lambda_t^{jk}}$$

$$t = 1, \dots, T$$

$$k \in \{o, w\} \tag{2.3}$$

See Appendix 2.1 for details.

Equation (2.3) is the formula that we use to compute the hazard rates for our data. Table 2.3 gives the flows and the corresponding hazard rates, with summary statistics for the latter.

Data treatment in general, and smoothing in particular are important issues in gross-flows analysis. Yashiv (p12) discusses how different analyses have used different seasonal adjustment procedures. For simplicity we use a system of quarterly dummy variables to remove seasonal components from each of our series.[14]

Smoothing is more controversial. Gomes (2009) uses a four-quarter moving average to smooth his LFS-based UK gross flows, whereas other analysis (Petrongolo and Pissarides (2008) for example) give no mention of having applied extra smoothing to their data.

For the sake or robustness, in our analysis we provide results using both smoothed and unsmoothed data. For smoothing the data we use the fitted values of each series from a regression on a third-order polynomial in time. The order of polynomial is selected for convenience – we could of course have reduced the smoothing to an arbitrary degree by choosing a higher-order function. As we disaggregate the LFS data into two lower-level countries (Wales and Outside-of-Wales), the sample sizes on which our flow data are based are necessarily smaller than those available for the UK as a whole. This is especially true of our country of primary interest – Wales, as can be seen in table 2.2 – since Wales contains only around 4.4% of the UK's working-age population (table 2.1). Smoothing may mitigate the effects of sampling

[14] More specifically, we regress each series on a set of three seasonal dummy variables representing quarters. The de-seasonalised series we then take to be the constant plus the residual from each of these regressions. Some papers in the literature instead use an X-11 ARIMA, however Bell and Smith (2002, p23) try both methods and report little difference between them.

noise on changes in our gross flow estimates. On the other hand, it might also throw the baby out with the bathwater, discarding relevant time-variation in the hazard rates.

Table 2.3 displays summary statistics for both smoothed and unsmoothed hazard rates. The sample rates are calculated using the LFS gross flows data summarized in table 2.2, as well as LFS stock data for each of the relevant areas, as summarized in table 2.1. To this data we apply equation (2.3) to obtain each hazard rate. In the analysis that follows, we also experiment with smoothing various measures of unemployment. We cover both smoothed and unsmoothed analyses.

Table 2.3: Summary statistics for hazard rates data, 1997Q2-2010Q4

Flow	*Hazard*	*Hazard rates, unsmoothed* *Mean*	*Standard dev.*	*Hazard rates, smoothed* *Mean*	*Standard dev.*
		(i)	(ii)	(iii)	(iv)
UE	f_t	0.325	0.042	0.324	0.038
U^wE^w	f_t^{ww}	0.319	0.086	0.316	0.052
U^oE^o	f_t^{oo}	0.325	0.042	0.325	0.037
EU	s_t	0.013	0.002	0.013	0.001
E^wU^w	s_t^{ww}	0.016	0.004	0.016	0.002
E^oU^o	s_t^{oo}	0.012	0.002	0.012	0.001
IU	α_t	0.047	0.007	0.047	0.007
I^wU^w	α_t^{ww}	0.044	0.013	0.044	0.008
I^oU^o	α_t^{oo}	0.047	0.007	0.047	0.007
UI	b_t	0.189	0.015	0.188	0.009
U^wI^w	b_t^{ww}	0.248	0.071	0.245	0.021
U^oI^o	b_t^{oo}	0.186	0.015	0.185	0.009
IE	γ_t	0.057	0.005	0.057	0.004
I^wE^w	γ_t^{ww}	0.048	0.012	0.048	0.006
I^oE^o	γ_t^{oo}	0.057	0.005	0.057	0.004
EI	δ_t	0.017	0.001	0.017	0.001
E^wI^w	δ_t^{ww}	0.023	0.006	0.023	0.003
E^oI^o	δ_t^{oo}	0.017	0.001	0.017	0.001
T		55	55	55	55

Notes: T denotes the sample size (number of quarters of data). Seasonally adjusted data using quarterly seasonal dummies. Smoothing in columns (iii) and (iv) is done using a third-order polynomial in time.

The means of the hazard rates are uninformative with respect to their importance in explaining variation in unemployment, because their levels depend on the average size of their denominators. Hence, hazard rates calculated using flows out of employment are generally very small, especially relative to same-country flows out of unemployment. A comparison of columns (i) and (iii) shows that smoothing does not affect the means of the hazard rates, however, comparing columns (ii) and (iv) shows that it produces reasonably large reductions in the volatility of each of the flow series.

Table 2.3 shows that in general, the mean and standard deviation of the UK-wide hazards (XY) are very similar to the analogous hazards within the UK-Outside-of-Wales (of the form X^oY^o). This is entirely as one would expect given that the Outside-of-Wales part of the UK is far larger than Wales. By contrast, the E^wU^w hazards U^wI^w, E^wI^wappear to be somewhat greater on average than the analogous Outside-of-Wales and UK hazards, whereas the $\mathrm{U^wE^w}$, $\mathrm{I^wU^w}$ and $\mathrm{I^wE^w}$ hazards are on average smaller. Thus inflow hazards to unemployment in Wales are slightly weighted towards separations from employment, whereas inflows to employment, U^wE^w and I^wE^w hazards are on average, lower across the board than their UK values.

The smoothed and unsmoothed hazard rates are plotted in figures 2.3, 2.4 and 2.5, for the UK, Wales and Outside-of-Wales respectively.

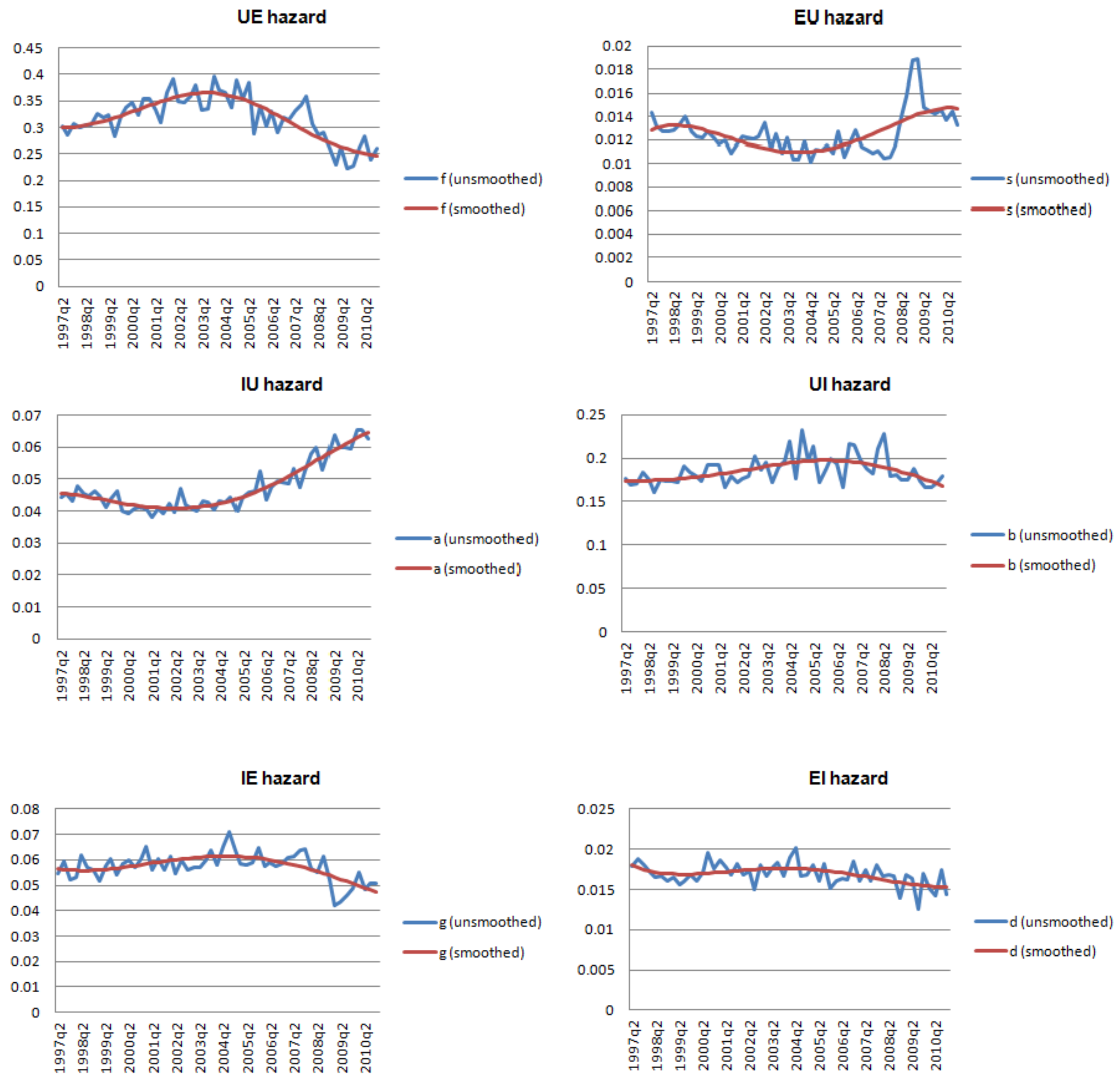

Figure 2.3: Hazard rates for UK Gross flows, seasonally adjusted

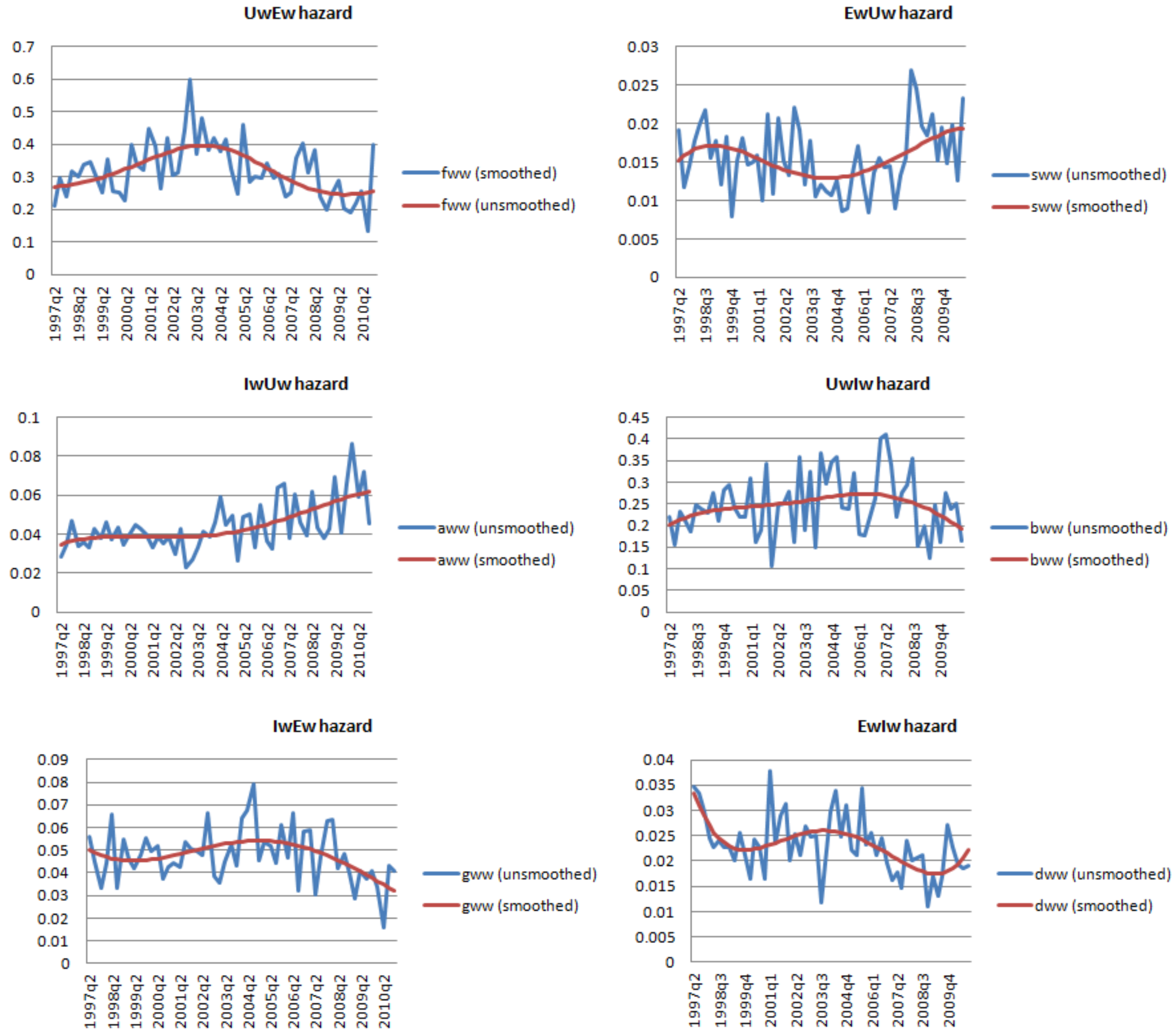

Figure 2.4: Hazard rates for Wales, seasonally adjusted

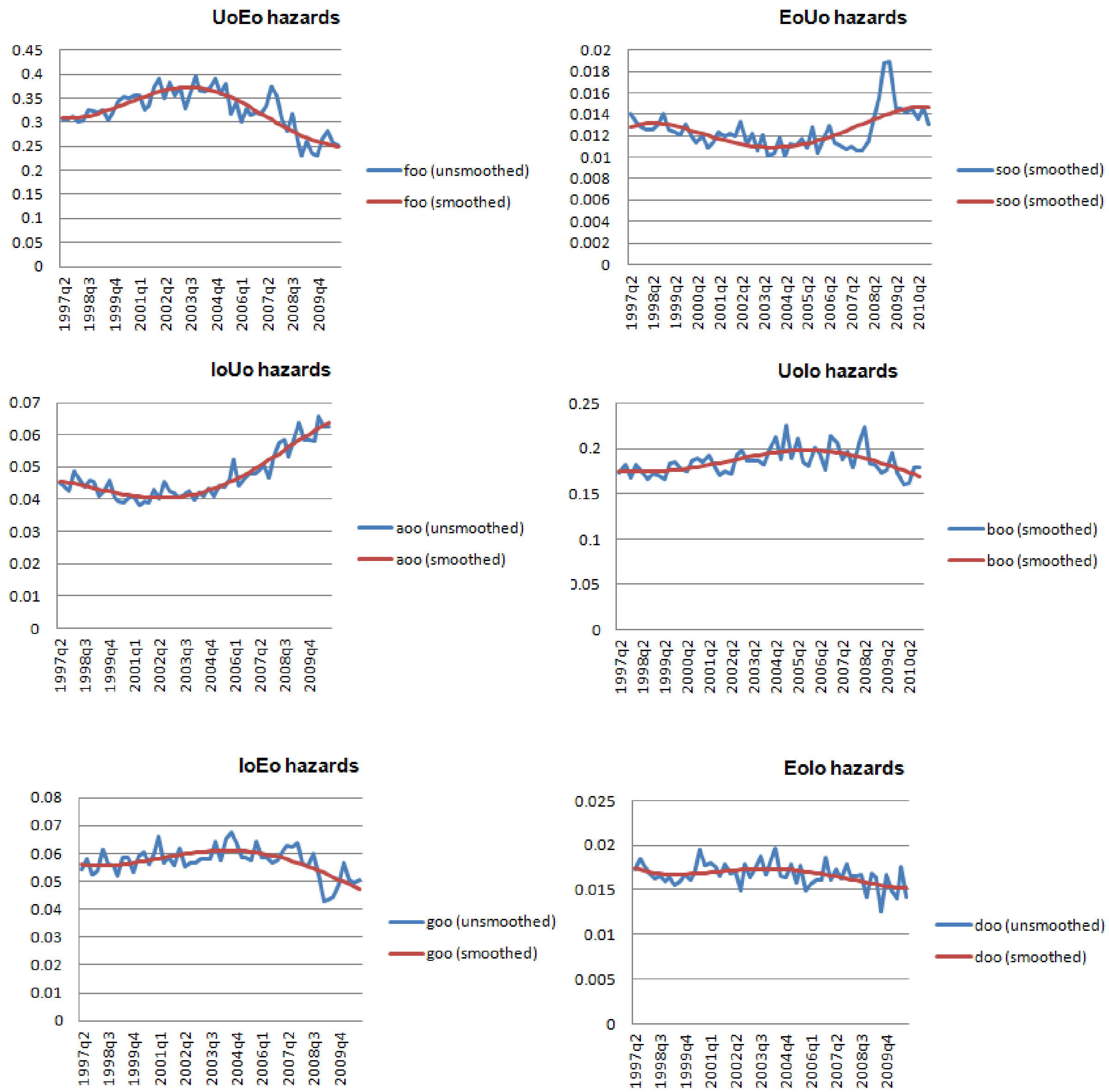

Figure 2.5: Hazard rates for Outside-of-Wales, seasonally adjusted

Table 2.4 gives the correlations of flows and hazard rates with the unemployment rates. The cyclicality of gross flows is of general interest in the gross flows literature and since the purpose of this analysis is to attempt to explain movements in unemployment, we compare the correlations of both our gross flows as shares of the relevant populations, and of our constructed hazard rates to unemployment. Unemployment is of course a *negatively* cyclical indicator.

Table 2.4: Cyclical correlations of flows and hazards (1997Q2-2010Q4)

	(i)		(ii)	(iii)
Flow:	*Correlation of flow as proportion of Working-age-population with unemployment rate*	*Hazard rate*	*Correlation of unsmoothed hazard rate with unsmoothed unemployment rate*	*Correlation of smoothed hazard rate with smoothed unemployment rate*
UK:	$\mathrm{corr}\left(\frac{YX}{Wpop}, u\right)$	λ	$\mathrm{corr}(\lambda, u)$	$\mathrm{corr}(\tilde{\lambda}, \tilde{u})$
Wales & Outside-of-Wales:	$\mathrm{corr}\left(\frac{X^jY^k}{Wpop}, u^j\right)$	λ^{jk}	$\mathrm{corr}(\lambda^{jk}, u^j)$	$\mathrm{corr}(\tilde{\lambda}^{jk}, \tilde{u}^j)$
UE	0.74	f_t	-0.82	-0.96
U^wE^w	0.44	f_t^{ww}	-0.45	-0.79
U^oE^o	0.74	f_t^{oo}	-0.82	-0.96
EU	0.59	s_t	0.64	0.97
E^wU^w	0.24	s_t^{ww}	0.32	0.95
E^oU^o	0.58	s_t^{oo}	0.63	0.97
IU	0.74	α_t	0.73	0.81
I^wU^w	0.19	α_t^{ww}	0.16	0.37
I^oU^o	0.76	α_t^{oo}	0.75	0.83
UI	0.90	b_t	-0.53	-0.74
U^wI^w	0.50	b_t^{ww}	-0.32	-0.89
U^oI^o	0.90	b_t^{oo}	-0.49	-0.70
IE	-0.72	γ_t	-0.73	-0.97
I^wE^w	-0.38	γ_t^{ww}	-0.43	-0.89
I^oE^o	-0.71	γ_t^{oo}	-0.73	-0.97
EI	-0.53	δ_t	-0.44	-0.78
E^wI^w	-0.23	δ_t^{ww}	-0.15	-0.27
E^oI^o	-0.53	δ_t^{oo}	-0.44	-0.83
T	55		55	55

Notes: T denotes the sample size (number of quarters of data). Seasonally adjusted data using quarterly dummy variables. Smoothing is done using a third-order polynomial in time. When correlations are reported as being "." it means that there are none of the relevant observations in the dataset.

Column (i) of Table 2.4 contains the correlation of each of the gross flows as a proportion of the relevant working-age population with the relevant unemployment rate. Flows of the type X^jY^j; $j, k \in \{o, w\}$ are expressed as a proportion of the working-age-population in country j, and the reported correlation is with the unemployment rate in country j. Flows for the whole of the UK (of type XY) are expressed as a proportion of the working-age population of the UK, and the reported correlation is with UK unemployment. In column (i), all flows and unemployment rates are seasonally adjusted using quarterly dummy variables, however we have not applied third-order polynomial smoothing to either the flows or the unemployment rates. In columns (ii) and (iii), seasonally adjusted flows are converted to hazard rates using equation (2.3) above, and the correlation with the appropriate unemployment rate (again that of country j for flows of the type X^jY^j; $j,k \in \{o,w\}$ and that of the whole of the UK for flows of type XY) are reported. In column (iii), the hazard rates and unemployment rates have been computed using third-order-polynomial smoothing, and in column (ii), no such smoothing has been used. Therefore, comparing columns (ii) and (iii) of table 2.4 gives a measure of the effects of our smoothing procedure on the cyclicality of the hazard rates. For each hazard rate, it is clear that the smoothing procedure increases the absolute magnitude of the correlation with unemployment – in other words it makes the cyclicality of each hazard rate more pronounced.

Table 2.4 shows that our data recreates a well-known stylized fact about the cyclicality of gross flows from unemployment to employment, and the hazard rate counterparts (f_t, f_t^{ww} and f_t^{oo}) to this flow. The hazard rates are strongly pro-cyclical whereas the flows are counter-cyclical. This has previously been noted in US data by Yashiv (2006 ,p17), and is also apparent in the analyses by Gomes (2009, table B p14) and Bell and Smith (2002, p33). Note that U^wE^w and f_t^{ww} appear to be less pro-cyclical than the measures for the UK and for Outside-of-Wales. The counter-cyclicality of the UE flow is what we would expect to find based on the theory of the matching function, which states that hiring is an increasing function of unemployment. Meanwhile the pro-cyclicality of f_t, f_t^{ww} and f_t^{oo} supports the familiar intuition that that jobs are easier to find for workers when the economy is strong. The reconciliation is of course that unemployment falls by proportionately more than the flows into employment during periods of economic expansion.
The flows and hazard rates from employment to unemployment are countercyclical, although more moderately so in Wales. Flows between inactivity and employment within countries are in general pro-cyclical, whereas those between inactivity and

unemployment are generally counter-cyclical (though note that the hazard rates corresponding to the flows from unemployment to inactivity are pro-cyclical).

2.4 The Petrongolo and Pissarides (2008) decomposition of UK unemployment

Petrongolo and Pissarides (2008) attempt to assign changes in steady-state unemployment to various UK hazard rates using LFS panel data. The method used is that developed and applied by Fujita and Ramey (2009), to decompose changes in steady-state U.S unemployment. [15]

The approach we take here to analyse changes in Welsh unemployment builds directly on these papers. For this reason, as well as the fact that it is useful for the purposes of our own analysis to have comparable results for the whole of the UK, we first reproduce Petrongolo and Pissarides' method of analysis here. As the sample we have at our disposal is different to that used by those authors[16], and because some of our data-treatment procedures differ (for example, the methods of seasonal adjustment) we then repeat their analysis using our later sample and our own data-treatment procedures. These results can then be compared back to the original results, and also to our results for Wales and Outside-out-Wales. The analytics for this part of the paper are however the same.

Petrongolo and Pissarides begin their three-state analysis with a pair of steady state equations for the UK labour market (Petrongolo and Pissarides, p258):

$$s_t E_t + \alpha_t I_t = (f_t + b_t) U_t$$

$$f_t U_t + \gamma_t I_t = (\delta_t + s_t) E_t \qquad (2.4)$$

s_t is the hazard rate corresponding to EU, α_t is the hazard rate corresponding to IU, f_t is the hazard rate corresponding to UE, b_t is the hazard rate corresponding to UI, γ_t is the hazard rate corresponding to IE and δ_t is the hazard rate corresponding to EI. The equations are two steady state conditions – the first of which says that the

[15] Gomes (2009) also analyses UK gross flows obtained from LFS panel data – however the variance decompositions that he reports are for a two-state model which excludes inactivity, rather than a three-state model as used in Petrongolo and Pissarides.

[16] Petrongolo and Pissarides' (2008) sample is 1993Q3-2003Q3. Our sample is 1997Q2-2010Q4.

total inflows to unemployment equals the total outflows, and the second of which says that the total inflows to employment equals the total outflows.

By defining the unemployment rate by $\frac{U_t}{U_t+E_t}$ the solution for steady-state unemployment in the UK is given by:

$$\bar{u}_t^{uk} = \frac{s_t + \frac{\alpha_t}{\alpha_t + \gamma_t}\delta_t}{s_t + \frac{\alpha_t}{\alpha_t + \gamma_t}\delta_t + f_t + \frac{\gamma_t}{\alpha_t + \gamma_t}b_t} \tag{2.5}$$

$\bar{u}_t^{uk}$ is the steady state unemployment rate for the UK under which the system of equations (2.4) holds. Note that if δ_t and b_t are both zero, one recovers the familiar steady state condition for unemployment that is used in the 2-state Mortensen-Pissarides model (Chapter 3).

Equation (2.5) says that steady-state unemployment is increasing in the separation hazard to unemployment, s_t and decreasing in the job-finding hazard to employment, f_t. One may also interpret $\frac{\alpha_t}{\alpha_t+\gamma_t}\delta_t$ as the hazard of making a transition from employment to inactivity via unemployment, and from inactivity to unemployment, since δ_t is the EI hazard and $\frac{\alpha_t}{\alpha_t+\gamma_t}$ is the hazard of making an IU transition, given that one transitions out of inactivity to some state. Similarly, $\frac{\gamma_t}{\alpha_t+\gamma_t}b_t$ is the hazard of making a transition to employment from unemployment via inactivity. (Smith (2011), p413).

Taking the first difference of equation (2.5) and manipulating, the change in steady-state unemployment can be decomposed as follows:

(2.6)

$$\begin{aligned}\Delta\bar{u}_t^{uk} = &\left(1-\bar{u}_t^{uk}\right)\bar{u}_{t-1}^{uk}\left[\frac{\Delta s_t}{s_{t-1}+\frac{\alpha_{t-1}}{\alpha_{t-1}+\gamma_{t-1}}\delta_{t-1}}\right] \\ &+\left(1-\bar{u}_t^{uk}\right)\bar{u}_{t-1}^{uk}\left[\frac{\Delta\left(\frac{\alpha}{\alpha+\gamma}\delta\right)_t}{s_{t-1}+\frac{\alpha_{t-1}}{\alpha_{t-1}+\gamma_{t-1}}\delta_{t-1}}\right] \\ &-\bar{u}_t^{uk}\left(1-\bar{u}_t^{uk}\right)\left[\frac{\Delta f_t}{s_{t-1}+\frac{\gamma_{t-1}}{\alpha_{t-1}+\gamma_{t-1}}b_{t-1}}\right] \\ &-\bar{u}_t^{uk}\left(1-\bar{u}_t^{uk}\right)\left[\frac{\Delta\left(\frac{\gamma}{\alpha+\gamma}\beta\right)_t}{s_{t-1}+\frac{\gamma_{t-1}}{\alpha_{t-1}+\gamma_{t-1}}b_{t-1}}\right]\end{aligned}$$

The idea is that equation (2.4) decomposes the first difference of unemployment into weighted changes in variables that are related to contemporaneous changes in the hazard rate counterparts to the gross flows. Clearly the first and third terms represent changes in the EU and UE hazard rates respectively. Petrongolo and Pissarides take the second term to represent changes in IU and the fourth term to represent changes in UI (see p258 and table 2 p260).

Collecting terms allows one to write the expression as a simple sum:

$$\Delta\bar{u}_t^{uk} = \pi_t^{EU} + \pi_t^{IU} + \pi_t^{UE} + \pi_t^{UI}$$

Where:

$$\pi_t^{EU} = \left(1-\bar{u}_t^{uk}\right)\bar{u}_{t-1}^{uk}\left[\frac{\Delta s_t}{s_{t-1}+\frac{\alpha_{t-1}}{\alpha_{t-1}+\gamma_{t-1}}\delta_{t-1}}\right] \qquad (2.7)$$

$$\pi_t^{IU} = \left(1-\bar{u}_t^{uk}\right)\bar{u}_{t-1}^{uk}\left[\frac{\Delta\left(\frac{\alpha}{\alpha+\gamma}\delta\right)_t}{s_{t-1}+\frac{\alpha_{t-1}}{\alpha_{t-1}+\gamma_{t-1}}\delta_{t-1}}\right]$$

$$\pi_t^{UE} = -\bar{u}_t^{uk}\left(1-\bar{u}_{t-1}^{uk}\right)\left[\frac{\Delta f_t}{s_{t-1}+\frac{\gamma_{t-1}}{\alpha_{t-1}+\gamma_{t-1}}b_{t-1}}\right]$$

$$\pi_t^{UI} = -\bar{u}_t^{uk}\left(1-\bar{u}_{t-1}^{uk}\right)\left[\frac{\Delta\left(\frac{\gamma}{\alpha+\gamma}b\right)_t}{s_{t-1}+\frac{\gamma_{t-1}}{\alpha_{t-1}+\gamma_{t-1}}b_{t-1}}\right]$$

The π_t^{XY}s are not themselves linear in the differences of hazards. The weights applied to them however, which consist of nonlinear combinations of the hazard levels (both contemporaneous and at first lag) result in a linear and exact decomposition of steady-state unemployment.

The sample variance of the change in steady state unemployment is then decomposed using a standard variance decomposition as follows:

$$\begin{aligned}\hat{\mathrm{var}}\left(\Delta\bar{u}_t^{uk}\right) &= \hat{\mathrm{var}}(\pi_t^{EU}) + \hat{\mathrm{cov}}(\pi_t^{EU},\pi_t^{IU}) + \hat{\mathrm{cov}}(\pi_t^{EU},\pi_t^{UE}) + \hat{\mathrm{cov}}(\pi_t^{EU},\pi_t^{UI}) \\ &+\hat{\mathrm{var}}(\pi_t^{IU}) + \hat{\mathrm{cov}}(\pi_t^{IU},\pi_t^{EU}) + \hat{\mathrm{cov}}(\pi_t^{IU},\pi_t^{UE}) + \hat{\mathrm{cov}}(\pi_t^{IU},\pi_t^{UI}) \\ &+\hat{\mathrm{var}}(\pi_t^{UE}) + \hat{\mathrm{cov}}(\pi_t^{UE},\pi_t^{EU}) + \hat{\mathrm{cov}}(\pi_t^{UE},\pi_t^{IU}) + \hat{\mathrm{cov}}(\pi_t^{UE},\pi_t^{UI}) \\ &+\hat{\mathrm{var}}(\pi_t^{UI}) + \hat{\mathrm{cov}}(\pi_t^{UI},\pi_t^{EU}) + \hat{\mathrm{cov}}(\pi_t^{UI},\pi_t^{IU}) + \hat{\mathrm{cov}}(\pi_t^{UI},\pi_t^{UE})\end{aligned} \tag{2.8}$$

Note that the variances and covariances are taken with respect to time.

Dividing through by $\hat{\mathrm{var}}\left(\Delta\bar{u}_t^{uk}\right)$ gives:

$$\begin{aligned}1 &= \frac{\hat{\mathrm{var}}(\pi_t^{EU}) + \hat{\mathrm{cov}}(\pi_t^{EU},\pi_t^{IU}) + \hat{\mathrm{cov}}(\pi_t^{EU},\pi_t^{UE}) + \hat{\mathrm{cov}}(\pi_t^{EU},\pi_t^{UI})}{\hat{\mathrm{var}}\left(\Delta\bar{u}_t^{uk}\right)} \\ &+ \frac{\hat{\mathrm{var}}(\pi_t^{IU}) + \hat{\mathrm{cov}}(\pi_t^{IU},\pi_t^{EU}) + \hat{\mathrm{cov}}(\pi_t^{IU},\pi_t^{UE}) + \hat{\mathrm{cov}}(\pi_t^{IU},\pi_t^{UI})}{\hat{\mathrm{var}}\left(\Delta\bar{u}_t^{uk}\right)} \\ &+ \frac{\hat{\mathrm{var}}(\pi_t^{UE}) + \hat{\mathrm{cov}}(\pi_t^{UE},\pi_t^{EU}) + \hat{\mathrm{cov}}(\pi_t^{UE},\pi_t^{IU}) + \hat{\mathrm{cov}}(\pi_t^{UE},\pi_t^{UI})}{\hat{\mathrm{var}}\left(\Delta\bar{u}_t^{uk}\right)}\end{aligned} \tag{2.9}$$

$$+\frac{\hat{\text{var}}(\pi_t^{UI})+\hat{\text{cov}}(\pi_t^{UI},\pi_t^{EU})+\hat{\text{cov}}(\pi_t^{UI},\pi_t^{IU})+\hat{\text{cov}}(\pi_t^{UI},\pi_t^{UE})}{\hat{\text{var}}\left(\Delta\bar{u}_t^{uk}\right)}$$

Equation (2.9) is an expression of the shares of the variance of the first-difference in steady state unemployment attributable to (weighted) changes in the hazard rate counterparts EU, IU, UE and UI. As is standard in the literature we label these shares $\hat{\beta}_{XY}$. They satisfy:

$$1=\hat{\beta}_{EU}+\hat{\beta}_{IU}+\hat{\beta}_{UE}+\hat{\beta}_{UI}$$

Where:

$$\hat{\beta}_{EU}=\frac{\hat{\text{var}}(\pi_t^{EU})+\hat{\text{cov}}(\pi_t^{EU},\pi_t^{IU})+\hat{\text{cov}}(\pi_t^{EU},\pi_t^{UE})+\hat{\text{cov}}(\pi_t^{EU},\pi_t^{UI})}{\hat{\text{var}}\left(\Delta\bar{u}_t^{uk}\right)}$$

$$\hat{\beta}_{IU}=\frac{\hat{\text{var}}(\pi_t^{IU})+\hat{\text{cov}}(\pi_t^{IU},\pi_t^{EU})+\hat{\text{cov}}(\pi_t^{IU},\pi_t^{UE})+\hat{\text{cov}}(\pi_t^{IU},\pi_t^{UI})}{\hat{\text{var}}\left(\Delta\bar{u}_t^{uk}\right)} \qquad (2.10)$$

$$\hat{\beta}_{UE}=\frac{\hat{\text{var}}(\pi_t^{UE})+\hat{\text{cov}}(\pi_t^{UE},\pi_t^{EU})+\hat{\text{cov}}(\pi_t^{UE},\pi_t^{IU})+\hat{\text{cov}}(\pi_t^{UE},\pi_t^{UI})}{\hat{\text{var}}\left(\Delta\bar{u}_t^{uk}\right)}$$

$$\hat{\beta}_{UI}=\frac{\hat{\text{var}}(\pi_t^{UI})+\hat{\text{cov}}(\pi_t^{UI},\pi_t^{EU})+\hat{\text{cov}}(\pi_t^{UI},\pi_t^{IU})+\hat{\text{cov}}(\pi_t^{UI},\pi_t^{UE})}{\hat{\text{var}}\left(\Delta\bar{u}_t^{uk}\right)}$$

The linearity property of covariances means that the betas above have a far more concise expression:

$$1=\hat{\beta}_{EU}+\hat{\beta}_{IU}+\hat{\beta}_{UE}+\hat{\beta}_{UI}$$

Where:

$$\hat{\beta}_{EU}=\frac{\hat{\text{cov}}\left(\pi_t^{EU},\Delta\bar{u}_t^{uk}\right)}{\hat{\text{var}}\left(\Delta\bar{u}_t^{uk}\right)}$$

$$\hat{\beta}_{IU}=\frac{\hat{\text{cov}}\left(\pi_t^{IU},\Delta\bar{u}_t^{uk}\right)}{\hat{\text{var}}\left(\Delta\bar{u}_t^{uk}\right)} \qquad (2.11)$$

$$\hat{\beta}_{UE}=\frac{\hat{\text{cov}}\left(\pi_t^{UE},\Delta\bar{u}_t^{uk}\right)}{\hat{\text{var}}\left(\Delta\bar{u}_t^{uk}\right)}$$

$$\hat{\beta}_{UI} = \frac{\hat{\text{cov}}\left(\pi_t^{UI}, \Delta\bar{u}_t^{uk}\right)}{\hat{\text{var}}\left(\Delta\bar{u}_t^{uk}\right)}$$

To understand the decomposition, it is useful to consider an ideal (and implausible) case in which the π_t^{XY} are entirely independent. In that case, the beta of each hazard is simply the ratio of the variance of the relevant π_t^{XY} to the total variance of the change in steady state unemployment. For example, $\hat{\beta}_{UE} = \frac{\hat{\text{var}}(\pi_t^{UE})}{\hat{\text{var}}(\Delta\bar{u}_t^{uk})}$. Now if we relax the assumption, and consider the case where there is some negative covariation between some of the π_t^{XY}. Suppose for example that $\hat{\text{cov}}(\pi_t^{UE}, \pi_t^{EU})$ is negative. In that case, $\frac{\hat{\text{var}}(\pi_t^{UE})}{\hat{\text{var}}(\Delta\bar{u}_t^{uk})}$ alone will over-estimate the contribution of the UE to the change in steady state unemployment because whenever there is an increase in π_t^{UE} there will be an offsetting decrease in π_t^{EU}. The magnitude of the offset will depend on the strength of the negative correlation. Similarly, positive covariation between the π_t^{XY} will mean that the variance ratio alone will be an under-estimate of the contribution of the role of the change in the XY hazard.
This discussion brings out an important feature of the method, which is that statistically it is no more causal than regression analysis. The method measures the strength of the association of transformations of the flows to an approximation to the change in the stock. In general, we understand flows to be the *proximate* causes of the change in stocks. The method can therefore do no more than measure the effect of variation in functions of these proximate causal factors on the change in the stock of unemployment. It will not uncover the driving force, or causal interrelationships between the flows.

This completes the analytic summary of method used by Petrongolo and Pissarides (2008) and Fujita and Ramey (2009).

We make one small extension to the method. Rather than decompose the variance of the first difference of *steady-state unemployment*, we decompose the variance of the first difference of observed unemployment, by adding an error term representing the change in observed unemployment that is not due to the change in steady state unemployment. The idea is to derive an extra beta term to indicate how closely changes in steady-state unemployment track changes in actual unemployment. Formally, the expression for the change in observed unemployment is given by:

$$\begin{aligned} \Delta u_t^{uk} &= \Delta \bar{u}_t^{uk} + \varepsilon_t \\ &= \pi_t^{EU} + \pi_t^{IU} + \pi_t^{UE} + \pi_t^{UI} + \varepsilon_t \end{aligned} \tag{2.12}$$

Where the πs have the same interpretation as before.

The variance of the change in observed unemployment is given by:

$$\begin{aligned} \hat{\text{var}}(\Delta u_t^{uk}) &= \hat{\text{var}}(\pi_t^{EU}) + \hat{\text{cov}}(\pi_t^{EU}, \pi_t^{IU}) + \hat{\text{cov}}(\pi_t^{EU}, \pi_t^{UE}) + \hat{\text{cov}}(\pi_t^{EU}, \pi_t^{UI}) \\ &+ \hat{\text{var}}(\pi_t^{IU}) + \hat{\text{cov}}(\pi_t^{IU}, \pi_t^{EU}) + \hat{\text{cov}}(\pi_t^{IU}, \pi_t^{UE}) + \hat{\text{cov}}(\pi_t^{IU}, \pi_t^{UI}) \\ &+ \hat{\text{var}}(\pi_t^{UE}) + \hat{\text{cov}}(\pi_t^{UE}, \pi_t^{EU}) + \hat{\text{cov}}(\pi_t^{UE}, \pi_t^{IU}) + \hat{\text{cov}}(\pi_t^{UE}, \pi_t^{UI}) \\ &+ \hat{\text{var}}(\pi_t^{UI}) + \hat{\text{cov}}(\pi_t^{UI}, \pi_t^{EU}) + \hat{\text{cov}}(\pi_t^{UI}, \pi_t^{IU}) + \hat{\text{cov}}(\pi_t^{UI}, \pi_t^{UE}) \\ &+ \hat{\text{var}}(\varepsilon_t) + 2\hat{\text{cov}}(\varepsilon_t, \pi_t^{EU}) + 2\hat{\text{cov}}(\varepsilon_t, \pi_t^{IU}) + 2\hat{\text{cov}}(\varepsilon_t, \pi_t^{UE}) + 2\hat{\text{cov}}(\varepsilon_t, \pi_t^{UI}) \end{aligned} \tag{2.13}$$

The betas now give the shares of each of the component flows in the change in observed unemployment. They satisfy:

$$1 = \hat{\beta}_{EU} + \hat{\beta}_{IU} + \hat{\beta}_{UE} + \hat{\beta}_{UI} + \hat{\beta}_{\varepsilon}$$

And are calculated as follows:

$$\begin{aligned} \hat{\beta}_{EU} &= \frac{\hat{\text{var}}(\pi_t^{EU}) + \hat{\text{cov}}(\pi_t^{EU}, \pi_t^{IU}) + \hat{\text{cov}}(\pi_t^{EU}, \pi_t^{UE}) + \hat{\text{cov}}(\pi_t^{EU}, \pi_t^{UI}) + \hat{\text{cov}}(\varepsilon_t, \pi_t^{EU})}{\hat{\text{var}}(\Delta u_t^{uk})} \\ \hat{\beta}_{IU} &= \frac{\hat{\text{var}}(\pi_t^{IU}) + \hat{\text{cov}}(\pi_t^{IU}, \pi_t^{EU}) + \hat{\text{cov}}(\pi_t^{IU}, \pi_t^{UE}) + \hat{\text{cov}}(\pi_t^{IU}, \pi_t^{UI}) + \hat{\text{cov}}(\varepsilon_t, \pi_t^{IU})}{\hat{\text{var}}(\Delta u_t^{uk})} \\ \hat{\beta}_{UE} &= \frac{\hat{\text{var}}(\pi_t^{UE}) + \hat{\text{cov}}(\pi_t^{UE}, \pi_t^{EU}) + \hat{\text{cov}}(\pi_t^{UE}, \pi_t^{IU}) + \hat{\text{cov}}(\pi_t^{UE}, \pi_t^{UI}) + \hat{\text{cov}}(\varepsilon_t, \pi_t^{UE})}{\hat{\text{var}}(\Delta u_t^{uk})} \\ \hat{\beta}_{UI} &= \frac{\hat{\text{var}}(\pi_t^{UI}) + \hat{\text{cov}}(\pi_t^{UI}, \pi_t^{EU}) + \hat{\text{cov}}(\pi_t^{UI}, \pi_t^{IU}) + \hat{\text{cov}}(\pi_t^{UI}, \pi_t^{UE}) + \hat{\text{cov}}(\varepsilon_t, \pi_t^{UI})}{\hat{\text{var}}(\Delta u_t^{uk})} \end{aligned} \tag{2.14}$$

$$\hat{\beta}_{\varepsilon} = \frac{\hat{\mathrm{var}}(\varepsilon_t) + \hat{\mathrm{cov}}(\varepsilon_t, \pi_t^{EU}) + \hat{\mathrm{cov}}(\varepsilon_t, \pi_t^{IU}) + \hat{\mathrm{cov}}(\varepsilon_t, \pi_t^{UE}) + \hat{\mathrm{cov}}(\varepsilon_t, \pi_t^{UI})}{\hat{\mathrm{var}}(\Delta u_t^{uk})}$$

Or, expressed in their concise form:

$$\hat{\beta}_{EU} = \frac{\hat{\mathrm{cov}}(\Delta u_t^{uk}, \pi_t^{EU})}{\hat{\mathrm{var}}(\Delta u_t^{uk})}$$

$$\hat{\beta}_{IU} = \frac{\hat{\mathrm{cov}}(\Delta u_t^{uk}, \pi_t^{IU})}{\hat{\mathrm{var}}(\Delta u_t^{uk})}$$

$$\hat{\beta}_{UE} = \frac{\hat{\mathrm{cov}}(\Delta u_t^{uk}, \pi_t^{UE})}{\hat{\mathrm{var}}(\Delta u_t^{uk})}$$

$$\hat{\beta}_{UI} = \frac{\hat{\mathrm{cov}}(\Delta u_t^{uk}, \pi_t^{UI})}{\hat{\mathrm{var}}(\Delta u_t^{uk})}$$

$$\hat{\beta}_{\varepsilon} = \frac{\hat{\mathrm{cov}}(\Delta u_t^{uk}, \varepsilon_t)}{\hat{\mathrm{var}}(\Delta u_t^{uk})}$$

$\hat{\beta}_{\varepsilon}$ is the component of the change in actual unemployment which goes unexplained in the analysis of Petrongolo and Pissarides (2008). Note that each beta contains a covariance with the error term– we want to know the extent to which each beta co-varies systematically with elements that are not explained by steady-state unemployment.

This completes the review of the analysis used on the UK as a whole. In the following section we adapt the methods presented here to decompose the variance of the change in unemployment for Wales.

2.5 Deriving steady-state labour force stocks for Wales and the rest of the UK

We derive the steady-state labour force stocks for the economically defined regions of Wales and Outside-of-Wales (the precise definitions of these were given above).

If it was the case that we could identify cross-border flows between Wales and outside of Wales in our dataset, this would increase the difficulty of the problem significantly. We would need to estimate the contribution of at least eight flows rather than four as in Petrongolo and Pissarides' analysis. As the number of cross-border changes of state appear to be small enough that no recorded changes appear in the dataset, we are restricted to the “within area” flows for Wales and Outside Wales. We therefore retain the Petrongolo and Pissarides decomposition for comparability’s sake, applying it on the within-area flows for each area.

The formulae are therefore familiar. The steady-state unemployment rate in each area is given by:

$$\bar{u}_t^j = \frac{s_t^{jj} + \frac{\alpha_t^{jj}}{\alpha_t^{jj} + \alpha_t^{jj}}\delta_t^{jj}}{s_t^{jj} + \frac{\alpha_t^{jj}}{\alpha_t^{jj} + \alpha_t^{jj}}\delta_t^{jj} + f_t^{jj} + \frac{\gamma_t^{jj}}{\alpha_t^{jj} + \gamma_t^{jj}}b_t^{jj}} \tag{2.15}$$

$$j \in \{o, w\}$$

Decomposing the change in each of the steady-state unemployment rates gives:

$$\begin{aligned} \Delta\bar{u}_t^j = {} & (1 - \bar{u}_t^j)\bar{u}_{t-1}^j \left[\frac{\Delta s_t^{jj}}{s_{t-1}^{jj} + \frac{\alpha_{t-1}^{jj}}{\alpha_{t-1}^{jj} + \gamma_{t-1}^{jj}}\delta_{t-1}^{jj}}\right] \\ & + (1 - \bar{u}_t^j)\bar{u}_{t-1}^j \left[\frac{\Delta\left(\frac{\alpha^{jj}}{\alpha^{jj} + \gamma^{jj}}\delta^{jj}\right)_t}{s_{t-1}^{jj} + \frac{\alpha_{t-1}^{jj}}{\alpha_{t-1}^{jj} + \gamma_{t-1}^{jj}}\delta_{t-1}^{jj}}\right] \\ & - \bar{u}_t^j(1 - \bar{u}_{t-1}^j)\left[\frac{\Delta f_t^{jj}}{s_{t-1}^{jj} + \frac{\gamma_{t-1}^{jj}}{\alpha_{t-1}^{jj} + \gamma_{t-1}^{jj}}b_{t-1}^{jj}}\right] \\ & - \bar{u}_t^j(1 - \bar{u}_{t-1}^j)\left[\frac{\Delta\left(\frac{\gamma^{jj}}{\alpha^{jj} + \gamma^{jj}}\beta^{jj}\right)_t}{s_{t-1}^{jj} + \frac{\gamma_{t-1}^{jj}}{\alpha_{t-1}^{jj} + \gamma_{t-1}^{jj}}b_{t-1}^{jj}}\right] \end{aligned} \tag{2.16}$$

$$j \in \{o, w\}$$

We can follow exactly the same logic as in section 2.4 to derive the betas for Wales and Outside of Wales respectively. That is, the change in steady-state unemployment for each time period and in each area is decomposed into a sum of linear components:

$$\Delta\bar{u}_t^j = \psi_t^{j,EU} + \psi_t^{j,IU} + \psi_t^{j,UE} + \psi_t^{j,UI}$$
$$j \in \{o, w\} \tag{2.17}$$

Where:

$$\psi_t^{j,EU} = (1-\bar{u}_t^j)\bar{u}_{t-1}^j \left[\frac{\Delta s_t^{jj}}{s_{t-1}^{jj} + \frac{\alpha_{t-1}^{jj}}{\alpha_{t-1}^{jj} + \gamma_{t-1}^{jj}} \delta_{t-1}^{jj}} \right]$$

$$\psi_t^{j,IU} = (1-\bar{u}_t^j)\bar{u}_{t-1}^j \left[\frac{\Delta \left(\frac{\alpha^{jj}}{\alpha^{jj} + \gamma^{jj}} \delta^{jj} \right)_t}{s_{t-1}^{jj} + \frac{\alpha_{t-1}^{jj}}{\alpha_{t-1}^{jj} + \gamma_{t-1}^{jj}} \delta_{t-1}^{jj}} \right] \tag{2.18}$$

$$\psi_t^{j,UE} = -\bar{u}_t^j(1-\bar{u}_{t-1}^j) \left[\frac{\Delta f_t^{jj}}{s_{t-1}^{jj} + \frac{\gamma_{t-1}^{jj}}{\alpha_{t-1}^{jj} + \gamma_{t-1}^{jj}} b_{t-1}^{jj}} \right]$$

$$\psi_t^{j,UI} = -\bar{u}_t^j(1-\bar{u}_{t-1}^j) \left[\frac{\Delta \left(\frac{\gamma^{jj}}{\alpha^{jj} + \gamma^{jj}} \beta^{jj} \right)_t}{s_{t-1}^{jj} + \frac{\gamma_{t-1}^{jj}}{\alpha_{t-1}^{jj} + \gamma_{t-1}^{jj}} b_{t-1}^{jj}} \right]$$

$$j \in \{o, w\}$$

As in the previous section, we may now calculate the shares of the variance of steady-state unemployment associated with each hazard component:

Where:

$$\hat{\beta}_{EU}^{j} = \frac{\hat{\text{cov}}\left(\psi_t^{j,EU}, \Delta\bar{u}_t^j\right)}{\hat{\text{var}}\left(\Delta\bar{u}_t^j\right)}$$

$$\hat{\beta}_{IU}^{j} = \frac{\hat{\text{cov}}\left(\psi_t^{j,IU}, \Delta\bar{u}_t^j\right)}{\hat{\text{var}}\left(\Delta\bar{u}_t^j\right)}$$

$$\hat{\beta}_{UE}^{j} = \frac{\hat{\text{cov}}\left(\psi_t^{j,UE}, \Delta\bar{u}_t^j\right)}{\hat{\text{var}}\left(\Delta\bar{u}_t^j\right)} \qquad (2.19)$$

$$\hat{\beta}_{UI}^{j} = \frac{\hat{\text{cov}}\left(\psi_t^{j,UI}, \Delta\bar{u}_t^j\right)}{\hat{\text{var}}\left(\Delta\bar{u}_t^j\right)}$$

$$1 = \hat{\beta}_{EU}^{j} + \hat{\beta}_{IU}^{j} + \hat{\beta}_{UE}^{j} + \hat{\beta}_{UI}^{j}$$

$$j \in \{o, w\}$$

Alternatively, we may as in the UK case, decompose the variance of observed rather than steady-state unemployment. The change in observed unemployment in each area is now the sum of the change in steady-state unemployment and an error term:

$$\Delta u_t^j = \Delta\bar{u}_t^j + \varepsilon_t^j \qquad (2.20)$$

$$j \in \{o, w\}$$

Decomposing observed unemployment in this way implies an extra beta for the error term. The betas are now:

$$\hat{\beta}_{EU}^{j} = \frac{\hat{\text{cov}}\left(\psi_t^{j,EU}, \Delta u_t^j\right)}{\hat{\text{var}}\left(\Delta u_t^j\right)}$$

$$\hat{\beta}_{IU}^{j} = \frac{\hat{\text{cov}}\left(\psi_t^{j,IU}, \Delta u_t^j\right)}{\hat{\text{var}}\left(\Delta u_t^j\right)}$$

$$\hat{\beta}_{UE}^{j} = \frac{\hat{\text{cov}}\left(\psi_t^{j,UE}, \Delta u_t^j\right)}{\hat{\text{var}}\left(\Delta u_t^j\right)}$$

$$\hat{\beta}_{UI}^{j} = \frac{\hat{\text{cov}}\left(\psi_t^{j,UI}, \Delta u_t^j\right)}{\hat{\text{var}}\left(\Delta u_t^j\right)} \tag{2.21}$$

$$\hat{\beta}_{\varepsilon}^{j} = \frac{\hat{\text{cov}}\left(\psi_t^{\varepsilon}, \Delta u_t^j\right)}{\hat{\text{var}}\left(\Delta u_t^j\right)}$$

$$1 = \hat{\beta}_{EU}^{j} + \hat{\beta}_{IU}^{j} + \hat{\beta}_{UE}^{j} + \hat{\beta}_{UI}^{j} + \hat{\beta}_{\varepsilon}^{j}$$

$$j \in \{o, w\}$$

2.6 The relationship between steady state and measures of observed unemployment rates

In the previous sections we discuss Petrongolo and Pissarides' (2008) method for decomposing the variance of UK unemployment rates into linear components which depend directly on changes in gross-flow hazard rates. The method involves the prior construction of an expression for the steady state unemployment rate. The expression is given in equation (2.5).

We then went on to construct steady state unemployment rates for Wales, and for the part of the UK that is Outside-of-Wales, in order to derive variance decompositions for the changes in these variables.

In all cases, we made clear that it is possible to decompose the variance of the change in the *observed* unemployment rate for each area, by adding a term that represents the difference between the change in the steady state and the change in the observed measure. Before moving on to the variance decomposition, we investigate how closely correlated the levels and first differences of the steady-state rates of unemployment are with the levels and first differences of observed rates.

Table 2.5: Correlations between measures of unemployment for the UK, Wales and Outside-of-Wales, 1997Q2-2010Q4.

Levels:

	UK:					Wales:					Outside-of-Wales:			
	(1)	(2)	(3)	(4)		(5)	(6)	(7)	(8)		(9)	(10)	(11)	(12)
	u	$\tilde{u}$	$\bar{u}$	$\bar{\bar{u}}$		u^w	$\widetilde{u^w}$	$\overline{u^w}$	$\overline{\overline{u^w}}$		u^o	$\widetilde{u^o}$	$\overline{u^o}$	$\overline{\overline{u^o}}$
u	1	0.93	0.90	0.92	u^w	1	0.91	0.75	0.94	u^o	1	0.93	0.91	0.92
$\tilde{u}$		1	0.84	0.99	$\widetilde{u^w}$		1	0.74	0.94	$\widetilde{u^o}$		1	0.83	0.99
$\bar{u}$			1	0.84	$\overline{u^w}$			1	0.79	$\overline{u^o}$			1	0.84
$\bar{\bar{u}}$				1	$\overline{\overline{u^w}}$				1	$\overline{\overline{u^o}}$				1

First differences:

	Δu	$\Delta\tilde{u}$	$\Delta\bar{u}$	$\Delta\bar{\bar{u}}$		Δu^w	$\Delta\widetilde{u^w}$	$\Delta\overline{u^w}$	$\Delta\overline{\overline{u^w}}$		Δu^o	$\Delta\widetilde{u^o}$	$\Delta\overline{u^o}$	$\Delta\overline{\overline{u^o}}$
Δu	1	0.27	0.19	0.28	Δu^w	1	0.20	0.31	0.19	Δu^o	1	0.27	0.41	0.28
$\Delta\tilde{u}$		1	0.08	0.95	$\Delta\widetilde{u^w}$		1	0.03	0.97	$\Delta\widetilde{u^o}$		1	0.10	0.94
$\Delta\bar{u}$			1	0.12	$\Delta\overline{u^w}$			1	0.03	$\Delta\overline{u^o}$			1	0.13
$\Delta\bar{\bar{u}}$				1	$\Delta\overline{\overline{u^w}}$				1	$\Delta\overline{\overline{u^o}}$				1

Notes: T=55 in levels and T=54 in first differences. Seasonally adjusted data using quarterly seasonal dummies. Smoothing is done using fitted values from a third-order polynomial in time. u is the observed, unsmoothed UK unemployment rate, $\tilde{u}$ is the observed, smoothed UK unemployment rate, $\bar{u}$ is the steady-state UK unemployment rate, uw is the observed, unsmoothed unemployment rate in Wales, $\bar{\bar{u}}$ is the steady-state UK unemployment rate constructed out of smoothed hazard rates, so that it is effectively a smoothed steady-state rate. $\widetilde{u^w}$ is the observed, smoothed unemployment rate in Wales, $\overline{u^w}$ is the steady-state unemployment in Wales,
$\overline{\overline{u^w}}$ is the steady-state Welsh unemployment rate constructed out of smoothed hazard rates. u^o is the observed, unsmoothed unemployment rate Outside-of-Wales, $\widetilde{u^o}$ is the observed, smoothed, unemployment rate Outside-of-Wales, $\overline{u^o}$ is the steady-state unemployment rate Outside-of-Wales, $\overline{\overline{u^o}}$ is the steady-state unemployment rate Outside-of-Wales constructed out of smoothed hazard rates.

We examine the correlation between three measures of the unemployment rate for each country in table 2.5. The variables u, u^w and u^o are observed unemployment rates for the UK, for Wales, and for Outside-of-Wales respectively. These we obtain from the cross-sectional LFS dataset. Each is seasonally adjusted using quarterly dummy variables.

$\tilde{u}$, $\widetilde{u^w}$ and $\widetilde{u^o}$ are observed unemployment rates for the UK, Wales, and for Outside-of-Wales. They differ from u, u^w and u^o respectively, only in the fact that we smooth them using a third-order polynomial. It makes sense to use at least one measure of unemployment which has had its high frequency components removed, to go along with the case in which the hazard rates are smoothed in the same way. The decompositions of changes in the u, u^w and u^o series are likely to have a greater variance share attributable to error than changes in the $\tilde{u}$, $\widetilde{u^w}$ and $\widetilde{u^o}$ series, because u, u^w and u^o contain high frequency fluctuations, whereas $\tilde{u}$, $\widetilde{u^w}$ and $\widetilde{u^o}$ and the smoothed hazard rates do not.

$\bar{u}$, $\overline{u^w}$ and $\overline{u^o}$ are the steady state unemployment rates for the UK, for Wales and for the area Outside-of-Wales respectively. They are constructed using the gross flow hazard rates, according to equations (2.5) and (2.15). The hazard rates that make up these measures of steady-state unemployment are not smoothed in this case, so that the steady-state series are not smoothed.

Finally,$\bar{\tilde{u}}$, $\overline{\widetilde{u^w}}$ and $\overline{\widetilde{u^o}}$ are the smoothed steady state unemployment rates for the UK, for Wales and for the area Outside-of-Wales respectively. They are smoothed because they are constructed out of smoothed hazard rates, in contrast to $\bar{u}$, $\overline{u^w}$ and $\overline{u^o}$.

The first part of table (2.5) shows that there is a high correlation in levels between the different measures of unemployment for all the geographical areas under examination. In the UK, and the Outside of Wales region the correlation between all measures of unemployment (whether steady-state or observed, smoothed or unsmoothed) in levels are all above 0.83. For Wales all correlations are above 0.74. Thus, in *levels*, steady-state unemployment is a reasonably good approximation to observed unemployment whether the data are smoothed or unsmoothed.

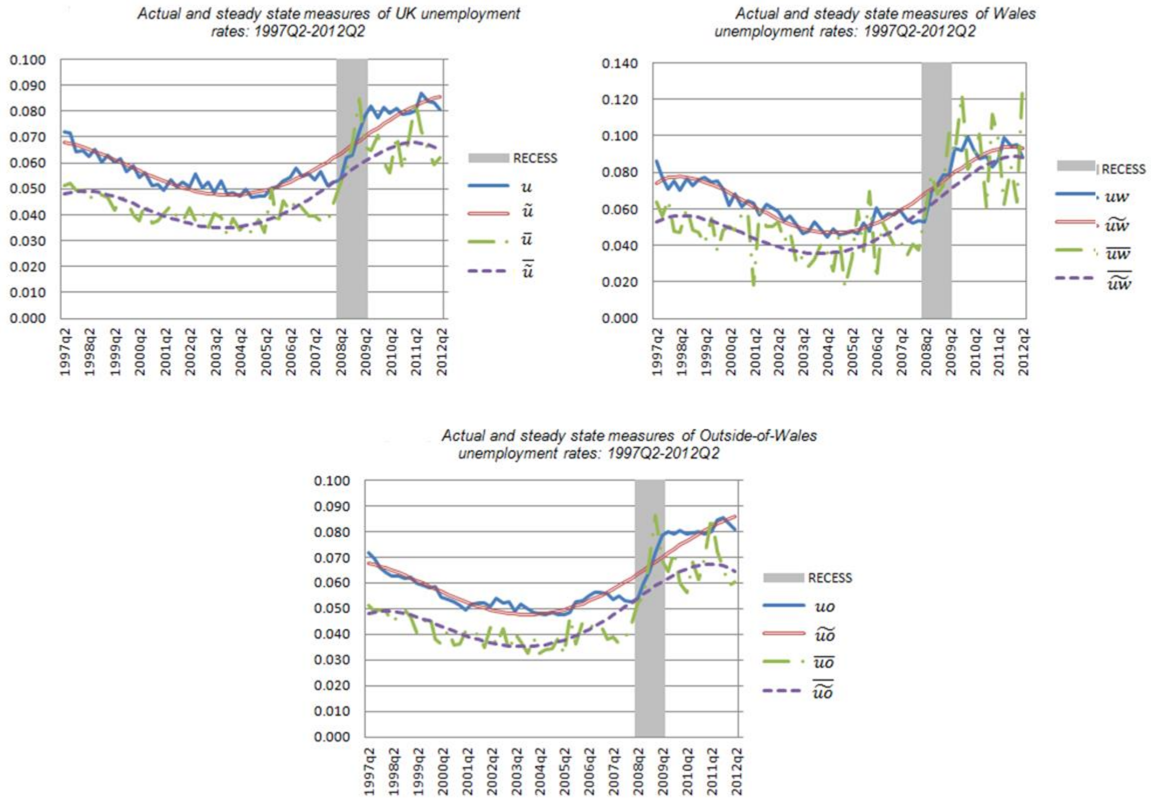

Figure 2.6: Observed and steady state unemployment, smoothed and unsmoothed, for the UK, Wales and Outside the UK.

The graphs in figure 2.6 show the levels of each of the four measures of unemployment for the three areas. The official period of recession which is 2008Q2-2009Q2 is shaded in grey. The charts show that for each geographical area, the unsmoothed and smoothed observed unemployment rates track each other closely in levels. However, from 2007 the unsmoothed unemployment rate series undergoes more fluctuations and deviates more from the smoothed series, especially with the onset of the 2008 recession.

In general, as unemployment stocks have to be the result of prior flows, we would expect steady-state unemployment to be a leading indicator of actual unemployment. When steady state unemployment is below observed unemployment, observed unemployment should be falling, and vice versa. In our analysis, this is true of the large increase in unemployment, during or following the 2008 recession (in which the upwards-spike in steady-state unemployment does indeed proceed the rise in actual unemployment), and during the period 1998-2003, in which observed unemployment was on a falling trend. However, throughout most of the sample steady-state unemployment bias does appear to be consistently lower than it should be, especially in periods of fairly stable observed unemployment. Comparison of non-seasonally adjusted and seasonally adjusted data (reported above) suggests that some, but not all of this discrepancy is due to seasonal adjustment (since removing seasonal components using dummy variables has an effect on the level of the data). It is not clear what explains the rest of the discrepancy, although the reason may relate to the effects of the LFS panel weights. However, the variance decomposition analysis that we perform is not in levels but in first differences. We therefore do not explore further the discrepancies between the different measures of the levels of unemployment since they are not critical to our analysis. It is however, relevant to consider whether the high correlations between the levels of the different measures of the unemployment rate in each geographical area are also present in the first differences of the series. The relevant correlations are given in the second part of table 2.5.

The table shows that the correlation between the unsmoothed first differences of steady-state measures of unemployment ($\Delta\bar{u}, \Delta\bar{u}^{w}, \Delta\bar{u}^{o}$, respectively) and the unsmoothed first differences of observed measures of unemployment ($\Delta u, \Delta u^{w}, \Delta u^{o}$, respectively) are in general much weaker. The greatest at 0.41 is for the Outside-of-Wales area. The value for Wales is 0.31 and 0.19 for the whole of the UK. This is because differencing series tends to exacerbate the effect of idiosyncratic noise present in the data. However, the correlations between changes in smoothed measures of observed unemployment ($\Delta\tilde{u}, \Delta\tilde{u}^{w}, \Delta\tilde{u}^{o}$, respectively) and the changes in smoothed measures of steady-state unemployment ($\Delta\tilde{\bar{u}}, \Delta\tilde{\bar{u}}^{w}, \Delta\tilde{\bar{u}}^{o}$, respectively) are very high (above 0.90 for all areas). The reason is that the smoothing removes idiosyncratic movements from both steady-state and observed unemployment series. What is left is a common trend which is highly correlated between series.

2.7 Variance decomposition results:

Table 2.6 reports the results of our variance decompositions for the United Kingdom. The figures in the table are the relevant "betas" which are hazard shares that correspond to the "flows" indicated in the second column of the same table.[17]

[17] In the discussion that follows, for the sake of brevity we often refer to the betas as "hazard shares".

The first column labelled P&P contains Petrongolo and Pissarides' results for comparison purposes. For our analysis, we present a range of results, which differ according to whether we decompose steady-state unemployment (columns (1) and (2)) or observed unemployment (columns (3) and (4)). Within these categories, we differentiate between analyses in which we smooth the unemployment and hazard-rate data using a 3rd-order-polynomial (columns (1) and (3)) and analyses in which we perform no such smoothing (columns (2) and (4))[18]. Note that in the steady-state analyses of columns (1) and (2), the decomposition of the unemployment rate is constructed out of the hazard rate according to equations (2.5) and (2.6), so that using smoothed hazards means automatically that the steady-state unemployment will also be smoothed[19].

Column (1) of table 2.6 shows the decomposition of the change in steady-state unemployment according to equation (2.10). The steady-state unemployment rate and its changes are constructed out of smoothed hazards. These results are most similar to those of Petrongolo and Pissarides. In our analysis, the share attributable to the hazard rate of moving from unemployment to employment (UE), at 0.43, is around 0.07 greater than in Petrongolo and Pissarides' analysis (0.36), with a roughly corresponding relative reduction in the hazard rate of moving from employment to unemployment (our figure is 0.07 lower than Petrongolo and Pissarides' EU hazard share of 0.35). The shares attributed to each of the inactivity hazard shares (IU and UI) in column (1) are similar to those of Petrongolo and Pissarides. In column (2) of table 2.6 we do the same steady-state decomposition without smoothing the hazard rates. In this case the IU hazard share, at 0.21, is far more important relative to the EU and UE hazard shares, gaining at the expense of the UI hazard share (which falls from 0.18 in column (1) to 0.08 in column (2)). Interestingly, the EU and UE hazard shares are relatively unaffected by the lack of smoothing. The results suggest that the volatility in the IU share is disproportionately affected by smoothing.

[18] This pattern of analyses with respect to the table columns is the same for our Wales and Outside-of-Wales analyses, and for each sample length analysed, i.e. in tables 2.7, 2.8, 2.10, 2.11 and 2.12, as well as in table 2.6.

[19] This is also true of the steady-state unemployment rates for Wales and for Outside-of-Wales.

Table 2.6: UK Variance decompositions 1997Q2-2010Q4

		P&P	(1)	(2)	(3)		(4)	
Steady state unemp?		Yes	Yes	Yes	No		No	
Smoothed unemp?		-	Yes	No	Yes		No	
Smoothed hazards?		-	Yes	No	Yes		No	
Sample		-	Full	Full	Full	*Of which*	Full	*Of which*
Beta:	*Flow:*							
$\hat{\beta}_{EU}$	EU	0.35	0.28	0.28	0.18	*0.21*	0.26	*-0.25*
$\hat{\beta}_{IU}$	IU	0.13	0.12	0.21	0.11	*0.16*	0.16	*-0.11*
$\hat{\beta}_{UE}$	UE	0.36	0.43	0.43	0.32	*0.41*	-0.04	*0.90*
$\hat{\beta}_{UI}$	UI	0.15	0.18	0.08	0.15	*0.22*	-0.01	*0.45*
$\hat{\beta}_{\varepsilon}$	Error				0.24		0.63	
T		41	55	55	55		55	

Notes: T denotes the sample size (number of quarters of data). Note that the reported numbers are the "Beta" as listed in the leftmost column. The column entitled "Flow" is included simply to remind the reader to which flow the reported betas correspond. Seasonally adjusted data using quarterly seasonal dummies. Smoothing is done using a third-order polynomial in time. "Of which" refers to the column directly to the left. It is the ratio of the row's entry to the total non-error share in column. For example, 0.21=0.18/(0.18+0.11+0.32+0.15). P&P refers to Petrongolo and Pissarides' (2008) results for the period 1993Q3-2003Q3 which also use the Labour Force Survey panel.

Cognizant of the correlations between changes in observed and steady state unemployment in table 2.5 (0.999 with smoothed data but only 0.75 in the unsmoothed analysis) we turn to the decomposition of observed unemployment. Column (3) shows that changes in steady-state unemployment explain only 76% of the time-variation in observed unemployment, when both observed unemployment and the hazards that determine steady-state unemployment are subject to smoothing. However, in the column directly to the right of column (3), (labelled "*Of which*"), we recalculate the hazard shares excluding the error term, so that they are expressed as a proportion of steady-state unemployment only. From this column it can be seen that the results are very similar to those in column (1), except with slightly more emphasis on the inactivity-related (that is UI and IU - related) hazard shares. This is broadly what we would expect with both the hazard-rate data and variable to be explained (either observed or steady state unemployment) smoothed in the same way in each analysis.

In columns (4) we relax the aggressive smoothing of column (3), both on observed unemployment and on the hazard rates that constitute steady-state unemployment.

The results are striking - suggesting that, without smoothing, just under 2/3rds of the change in observed unemployment is not attributable to change in the steady-state. Also in column (4), it can be seen that both of the outflow hazard shares from unemployment (the UE and UI hazard shares) make small negative contributions (-4% and-1% respectively) to the variance of the change in observed unemployment. Moreover, because the corresponding *steady state* analysis in column (2) contains no negative shares, we can infer that the negative shares of components in column (4) arise from negative correlation between those components and the error term. This implies that in our sample, when the change in unemployment increases due to non-steady state factors there is a tendency for the outflow hazard rates to fall. Column (4) also suggests a relatively important role for changes in inflow hazard shares (that is $EU$and IU hazard shares) in explaining changes in the observed unemployment rate.

Table 2.7: Wales Variance decompositions 1997Q2-2010Q4

		(1)	(2)	(3)		(4)	
Steady state unemp?		Yes	Yes	No		No	
Smoothed unemp?		Yes	No	Yes		No	
Smoothed hazards?		Yes	No	Yes		No	
Sample		Full	Full	Full	*Of which*	Full	*Of which*
Beta:	*Flow:*						
$\hat{\beta}_{EU}^{ww}$	E^wU^w	0.22	0.27	0.20	*0.26*	0.08	*0.15*
$\hat{\beta}_{IU}^{ww}$	I^wU^w	0.27	0.33	0.18	*0.23*	0.07	*0.14*
$\hat{\beta}_{UE}^{ww}.$	U^wE^w	0.23	0.25	0.16	*0.21*	0.17	*0.31*
$\hat{\beta}_{UI}^{ww}$	U^wI^w	0.28	0.15	0.22	*0.29*	0.21	*0.40*
$\hat{\beta}_{\varepsilon}^{ww}$	Error			0.24		0.46	
T		55	55	55		55	

Notes: T denotes the sample size (number of quarters of data). Note that the reported numbers are the "Beta" as listed in the leftmost column. The column entitled "Flow" is included simply to remind the reader to which flow the reported beta correspond. Seasonally adjusted data using quarterly seasonal dummies. Smoothing is done using a third-order polynomial in time. When shares are reported as being "." it means that there are none of the relevant observations in the dataset. "Of which" refers to the column directly to the left. It is the ratio of the row's entry to the total non-error share in column. For example, 0.26=0.20/(0.20+0.18+0.16+0.22).

Tables 2.7 and 2.8 repeat the full-sample variance decomposition analyses for Wales and for Outside-of-Wales respectively. The most notable difference between the results for the UK and Wales for the smoothed steady-state decompositions is the relative importance of the inactivity (that is, U^wI^w and I^wU^w) hazard shares. The

I^wU^w hazard share at 27% and the U^wI^w hazard share at 28% is greater than either of the E^wU^w or U^wE^w contributions (22% and 23% respectively). As for the UK analysis, when smoothing is relaxed in column (2), the I^wU^w share increases (to 33%) and the U^wI^w share falls (to 15%), whereas the E^wU^w and U^wE^w hazard shares are relatively unchanged. Another similarity with the UK analysis is that the column (3) results in table 2.7 (smoothed data but decomposing the observed rather than steady-state unemployment in Wales) are very similar to the column (1) results in table 2.7 (smoothed data and steady-state unemployment in Wales). The chief difference is that there is an error term, relating changes in steady-state unemployment in Wales to changes in observed unemployment in Wales. In column (3) we can see that the error comprises 24% of the variance of changes in observed unemployment.

Column (4) of table 2.7 decomposes changes in observed unemployment in Wales, without any data smoothing. We can see that this changes the results greatly. The error term representing non-steady state changes in observed unemployment comprises 46% of the variance of observed unemployment in the sample. In contrast to the results for the UK, the U^wE^w and U^wI^w hazard shares (Outflows) appear to matter more than the inflow hazard shares, with shares of 31% and 41% of steady-state unemployment respectively (or 17% and 21% of observed unemployment).

As we would expect given that the Outside-of-Wales region comprises around 96% of the working-age-population of the UK, the results in table 2.8 are mostly very similar to the results for the whole UK in table 2.6. The results are different in column (4), however. The error term takes a smaller share in the Outside-of-Wales area than in the whole of the UK (25% rather than 63%), and the outflow hazard shares take positive shares of the variance of observed unemployment. The shares are in fact strikingly similar to the steady-state shares using unsmoothed data in column (2), once allowance is made for the contribution of the error term.

Table 2.8: Outside-of-Wales Variance decompositions 1997Q2-2010Q4							
		(1)	(2)	(3)		(4)	
Steady state unemp?		Yes	Yes	No		No	
Smoothed unemp?		Yes	No	Yes		No	
Smoothed hazards?		Yes	No	Yes		No	
Sample		Full	Full	Full	*Of which*	Full	*Of which*
Beta:	*Flow:*						
$\hat{\beta}^{oo}_{EU}$	E^oU^o	0.28	0.35	0.18	*0.24*	0.27	*0.36*
$\hat{\beta}^{oo}_{IU}$	I^oU^o	0.11	0.23	0.10	*0.13*	0.14	*0.18*
$\hat{\beta}^{oo}_{UE}$	U^oE^o	0.44	0.38	0.33	*0.43*	0.26	*0.35*
$\hat{\beta}^{oo}_{UI}$	U^oI^o	0.17	0.05	0.15	*0.20*	0.08	*0.10*
$\hat{\beta}^{oo}_{\varepsilon}$	Error			0.24		0.25	
T		55	55	55		55	

Notes: T denotes the sample size (number of quarters of data). Note that the reported numbers are the "Beta" as listed in the leftmost column. The column entitled "Flow" is included simply to remind the reader to which flow the reported beta correspond. Seasonally adjusted data using quarterly seasonal dummies. Smoothing is done using a third-order polynomial in time. When shares are reported as being "." it means that there are none of the relevant observations in the dataset. We use "~" to denote shares that are very close to zero, but for which there are observations. "Of which" refers to the column directly to the left. It is the ratio of the row's entry to the total non-error share in column. For example, 0.24=0.18/(0.18+0.10+0.33+0.15).

A significant event of our full sample is the period including and following the recession that began in 2008Q2. We are interested in the extent to which the large increase in unemployment that occurred at this time is driving our variance decomposition results. To investigate this further, we split our sample into a period before the recession that spans the years from 1997Q2 to 2008Q1 ($T = 44$) and a second smaller sample that covers the period 2008Q2-2010Q4. We then analyse the correlations between observed unsmoothed, observed smoothed, and steady-state unemployment rates in each of the three areas for the earlier subsample. We do not use the subsample that runs from 2008Q2-2010Q4, because it contains only 11 observations. Results for the pre-recession subsample are presented in table 2.9. The table may be compared with table 2.5, which contains the same correlations but for the full 1997Q2-2010Q4 sample.

Table 2.9: Correlations between measures of unemployment for the UK, Wales and Outside-of-Wales, 1997Q2-2008Q1.

Levels:

	UK:					Wales:					Outside-of-Wales:			
	(1)	(2)	(3)	(4)		(5)	(6)	(7)	(8)		(9)	(10)	(11)	(12)
	u	$\tilde{u}$	$\bar{u}$	$\bar{\tilde{u}}$		u^w	$\widetilde{u^w}$	$\overline{u^w}$	$\overline{\widetilde{u^w}}$		u^o	$\widetilde{u^o}$	$\overline{u^o}$	$\overline{\widetilde{u^o}}$
u	1	0.85	0.78	0.76	u^w	1	0.85	0.58	0.99	u^o	1	0.84	0.84	0.75
$\tilde{u}$		1	0.70	0.97	$\widetilde{u^w}$		1	0.55	0.99	$\widetilde{u^o}$		1	0.72	0.97
$\bar{u}$			1	0.62	$\overline{u^w}$			1	0.53	$\overline{u^o}$			1	0.65
$\bar{\tilde{u}}$				1	$\overline{\widetilde{u^w}}$				1	$\overline{\widetilde{u^o}}$				1

First differences:

	Δu	$\Delta\tilde{u}$	$\Delta\bar{u}$	$\Delta\bar{\tilde{u}}$		Δu^w	$\Delta\widetilde{u^w}$	$\Delta\overline{u^w}$	$\Delta\overline{\widetilde{u^w}}$		Δu^o	$\Delta\widetilde{u^o}$	$\Delta\overline{u^o}$	$\Delta\overline{\widetilde{u^o}}$
Δu	1	0.26	0.12	0.21	Δu^w	1	0.18	0.14	0.16	Δu^o	1	0.27	0.41	0.21
$\Delta\tilde{u}$		1	0.16	0.967	$\Delta\widetilde{u^w}$		1	0.09	0.940	$\Delta\widetilde{u^o}$		1	0.19	0.956
$\Delta\bar{u}$			1	0.15	$\Delta\overline{u^w}$			1	0.10	$\Delta\overline{u^o}$			1	0.17
$\Delta\bar{\tilde{u}}$				1	$\Delta\overline{\widetilde{u^w}}$				1	$\Delta\overline{\widetilde{u^o}}$				1

Notes: *T=43*. Seasonally adjusted data using quarterly seasonal dummies. Smoothing is done using a third-order polynomial in time. u is the observed, unsmoothed UK unemployment rate, $\tilde{u}$ is the observed, smoothed UK unemployment rate, $\bar{u}$ is the steady-state UK unemployment rate, uw is the observed, unsmoothed unemployment rate in Wales, $\bar{\tilde{u}}$ is the steady-state UK unemployment rate constructed out of smoothed hazard rates, so that it is effectively a smoothed steady-state rate. $\widetilde{u^w}$ is the observed, smoothed unemployment rate in Wales, $\overline{u^w}$ is the steady-state unemployment in Wales, $\overline{\widetilde{u^w}}$ is the steady-state Welsh unemployment rate constructed out of smoothed hazard rates. u^o is the observed, unsmoothed unemployment rate Outside-of-Wales, $\widetilde{u^o}$ is the observed, smoothed, unemployment rate Outside-of-Wales, $\overline{u^o}$ is the steady-state unemployment rate Outside-of-Wales, $\overline{\widetilde{u^o}}$ is the steady-state unemployment rate Outside-of-Wales constructed out of smoothed hazard rates.

Comparing the first rows of tables 2.5 and 2.9 we see that in most cases, excluding the quarters from 2008Q2 to 2010Q4 which cover the recession weakens the correlation between the unemployment level series. The exception to this is the correlation between the smoothed steady state unemployment series for Wales, $\overline{\widetilde{u^w}}$ and the smoothed and unsmoothed observed series for the same area, u^w and $\widetilde{u^w}$. A similar effect is observed in first-differences when we move from the longer to the shorter sample comparing the change in unsmoothed observed and unsmoothed steady-state unemployment, for Wales and for the whole of the UK. The effect for Wales is particularly stark: $\Delta\overline{u^w}$ and Δu^whave a correlation of 0.31 in the full sample, and just 0.14 in the shorter sample. However, shortening the sample has very little effect on the correlation between changes in the *smoothed* observed and steady-state series, in all cases.

The variance decomposition results for the UK in the shorter sample in table 2.10 are very similar to those for the whole 1997Q2-2010Q4 period in table 2.6. This is especially true of columns (1)-(3) in each table, the smoothed analyses and the unsmoothed steady-state analyses suggest similar relative roles for each type of hazard. There are a few exceptions however. Firstly, note that in the analysis of observed unemployment, with smoothed hazards and smoothed observed unemployment, the error term takes a smaller share over the shorter period than in the full sample (0.24 in the full sample and 0.07 in the short sample – see columns (3) of tables 2.6 and 2.10 respectively). Secondly, consider column (4) of table 2.10, which is the analysis without any smoothing and without the imposition of steady-state unemployment. The negative contributions of the UE and UI hazard shares are present in both samples, however they are more pronounced in the pre-recession subsample. The magnitudes of the EU hazard is the same in each sample. However, the error share in the1997Q2-2008Q1 is much greater than in the full sample (95%, rather than 63%).

Table 2.10: UK variance decompositions 1997Q2-2008Q1

		P&P	(1)	(2)	(3)		(4)	
Steady state unemp?		Yes	Yes	Yes	No		No	
Smoothed unemp?		-	Yes	No	Yes		No	
Smoothed hazards?		-	Yes	No	Yes		No	
Sample		-	Short	Short	Short	*Of which*	Short	*Of which*
Betas:	*Flow:*							
$\hat{\beta}_s$	EU	0.35	0.33	0.26	0.27	*0.30*	0.26	*5.54*
$\hat{\beta}_{i0}$	IU	0.13	0.10	0.18	0.11	*0.12*	0.05	*1.12*
$\hat{\beta}_f$	UE	0.36	0.45	0.46	0.42	*0.46*	-0.14	*-3.01*
$\hat{\beta}_{i1}$	UI	0.15	0.12	0.10	0.11	*0.12*	-0.12	*-2.65*
$\hat{\beta}_\varepsilon$	Error				0.07		0.95	
T		41	43	43	43		43	

Notes: T denotes the sample size (number of quarters of data). Note that the reported numbers are the "Beta" as listed in the leftmost column. The column entitled "Flow" is included simply to remind the reader to which flow the reported beta correspond. Seasonally adjusted data using quarterly seasonal dummies. Smoothing is done using a third-order polynomial in time. "Of which" refers to the column directly to the left. It is the ratio of the row's entry to the total non-error share in column. For example, 0.30=0.27/(0.27+0.11+0.45+0.11). P&P refers to Petrongolo and Pissarides' (2008) results for the period 1993Q3-2003Q3 which also use the Labour Force Survey panel.

We now consider the pre-recession results for Wales and Outside-of-Wales (tables 2.11 and 2.12). Firstly, note the smoothed steady-state and non-steady-state results for Wales in columns (1) and (3) of 2.11. Shortening the sample so as to exclude the recessionary period serves to greatly reduce the variance shares of inactivity-related flows. In the non-steady-state analysis, for example, the contribution of the $I^w U^w$ hazard share actually becomes negative in the short sample (-0.08), from a hazard share of 0.10 in the full sample. The shares attributable to the $E^w U^w$ hazard are increased by 5% in the smoothed steady-state case column (1) and 10% in the smoothed non-steady-state case (column (3)). The shares attributable to $U^w E^w$ are increased even more by shortening the sample, by 13 percentage points in the steady-state case and 27 percentage points in the non-steady-state analysis. This contrasts with the *unsmoothed* steady-state analysis, in which the $I^w U^w$ hazard share increases slightly in the shortened sample to 0.36 and the $U^w I^w$ hazard share is 7 percentage points lower at 0.08.

Table 2.11: Wales variance decompositions 1997Q2-2008Q1

		(1)	(2)	(3)		(4)	
Steady state unemp?		Yes	Yes	No		No	
Smoothed unemp?		Yes	No	Yes		No	
Smoothed hazards?		Yes	No	Yes		No	
Sample		Short	Short	Short	*Of which*	Short	*Of which*
Betas:	*Flow:*						
$\hat{\beta}_{EU}^{ww}$	$E^w U^w$	0.27	0.34	0.30	*0.38*	-0.18	*-1.16*
$\hat{\beta}_{IU}^{ww}$	$I^w U^w$	0.01	0.36	-0.08	*-0.10*	0.03	*0.16*
$\hat{\beta}_{UE}^{ww}$	$U^w E^w$	0.57	0.22	0.43	*0.56*	0.24	*1.55*
$\hat{\beta}_{UI}^{ww}$	$U^w I^w$	0.15	0.08	0.12	*0.16*	0.07	*0.44*
$\hat{\beta}_{\varepsilon}^{ww}$	Error			0.22		0.84	
T		43	43	43		43	

Notes: T denotes the sample size (number of quarters of data). Note that the reported numbers are the “Beta” as listed in the leftmost column. The column entitled “Flow” is included simply to remind the reader to which flow the reported beta correspond. Seasonally adjusted data using quarterly seasonal dummies. Smoothing is done using a third-order polynomial in time. When shares are reported as being “.” it means that there are none of the relevant observations in the dataset. We use “~” to denote shares that are very close to zero, but for which there are observations. “Of which” refers to the column directly to the left. It is the ratio of the row’s entry to the total non-error share in column. For example, 0.38=0.30/(0.30-0.08+0.43+0.12).

In the non-steady-state analysis of column (3), moving to a shorter sample does not have a large effect on the share of the error term – it falls by just 2 percentage points. However, as already pointed out, the distribution of the explained volatility is very different, falling more heavily on the unemployment-related flows in the shorter sample. Finally we turn to column (4) of table 2.11, which is the analysis for observed unemployment without any smoothing. Interestingly, the hazard shares for movements out of unemployment (U^wE^w and E^wI^w) are almost the same, whereas the E^wU^w hazard appears to contribute negatively in the shorter sample (compared to 0.27 in table 2.8), and the share attributable to I^wU^w is almost zero (compared to 0.14 in table 2.8). Instead, the share of the error term is much greater in the early sample, at 0.84 (compared with 0.25 in the full sample).

Table 2.12: Outside-of-Wales variance decompositions 1997Q2-2008Q1

		(1)	(2)	(3)		(4)	
Steady state unemp?		Yes	Yes	No		No	
Smoothed unemp?		Yes	No	Yes		No	
Smoothed hazards?		Yes	No	Yes		No	
Sample		Short	Short	Short	*Of which*	Short	*Of which*
Betas:	*Flow:*						
$\hat{\beta}^{oo}_{EU}$	E^oU^o	0.33	0.37	0.27	*0.29*	0.30	*0.55*
$\hat{\beta}^{oo}_{IU}$	I^oU^o	0.10	0.22	0.12	*0.13*	0.05	*0.09*
$\hat{\beta}^{oo}_{UE}$	U^oE^o	0.45	0.36	0.42	*0.46*	0.18	*0.34*
$\hat{\beta}^{oo}_{UI}$	U^oI^o	0.12	0.04	0.11	*0.12*	0.01	*0.02*
$\hat{\beta}^{oo}_{\varepsilon}$	Error			0.09		0.46	
T		43	43	43		43	

Notes: T denotes the sample size (number of quarters of data). Note that the reported numbers are the "Beta" as listed in the leftmost column. The column entitled "Flow" is included simply to remind the reader to which flow the reported beta correspond. Seasonally adjusted data using quarterly seasonal dummies. Smoothing is done using a third-order polynomial in time. When shares are reported as being "." it means that there are none of the relevant observations in the dataset. We use "~" to denote shares that are very close to zero, but for which there are observations. . "Of which" refers to the column directly to the left. It is the ratio of the row's entry to the total non-error share in column. For example, 0.29=0.27/(0.27+0.12+0.42+0.11).

The results for the 1997Q2-2008Q1 sample for the area Outside-of-Wales are given in table 2.12. We note that for the smoothed results (columns (1) and (3)) these are for the most-part, very similar to those in table 2.10, which cover the whole of the UK for the same period. The unsmoothed analysis results are however somewhat different, for example with a much larger share attributed to the EU hazard in the Outside-of-Wales region in the steady-state analysis (column (2)). The results also differ greatly from those of the whole of the UK for the unsmoothed, non-steady analysis in column (4).

As with the results for the whole of the UK, shortening the sample does not greatly affect the shares for the Outside-of-Wales region, again with the exception of column (4).

2.8 Discussion

Do the “ins” (that is, the EU and IU hazards), or the “outs” (UE and UI hazards) - “win” in the sense of dominating the first difference of unemployment? Our answer for the whole of the UK depends on whether or not the data are smoothed. The results for the whole of the UK appear to suggest that the “ins” and “outs” are of approximately equal importance, if the steady-state is imposed and the data are not smoothed (column (2) of table 2.6). This result is in line with those of Petrongolo and Pissarides, who find that the total shares of the “ins” and “outs” are approximately equal. However, when the data are smoothed (columns (1) and (3) of table 2.6), the “outs” gain in importance relative to the “ins”.
Column (4) of table 2.6 (in which we neither smooth the series nor impose the steady-state unemployment) appears to overturn this result. Here, the contribution of the “ins” is positive whereas that of the “outs” is slightly negative, and 63% of the variation in observed unemployment is attributed to the error term. Thus the results for the UK analysis with the least restrictive assumptions and data-treatment appear to contradict the assertion that the “ins” and “outs” have approximately equal hazard shares.

This is one reason for caution in interpreting these results. Looking at table 2.8 gives further reasons to be cautious. In the smoothed steady-state and non-steady state analysis (columns (1) and (3)), the results for Outside-of-Wales are almost exactly the same for those of the UK. This is as it should be, given that the Outside-of-Wales area is 95.5% of the population of the UK. But in column (2), where we impose the steady state but do not smooth the data, the results are surprisingly different. The share of the “ins” is approximately 9 percentage points higher than that for the whole of the UK: 58% against 43% for the “outs”. Now it is possible that the “ins win” Outside of Wales, and the “outs win” in Wales, giving a “draw” (roughly 50% share for ins and outs) in the UK. However for this explanation to work we would need the “outs” to dominate in column (2) of table 2.7, which is not the case. In fact, the results for column (2) in table 2.7 show that the “ins” win by slightly more in Wales. Therefore, a more likely explanation is that the unsmoothed series for Outside of Wales and for Wales are excessively noisy. A worrying implication is that, in reducing the sample size for the hazard rates in each period when we break the UK down into geographical sub-regions, we are reducing the scope for answering the question about the relative importance of hazard shares with the given data set. In other words, if we had estimated confidence intervals in the analyses of observed unemployment using the error term, the confidence intervals for the Outside-of-Wales analysis would wide enough to be consistent with the fact that the “ins win” *or* that the “outs win.” The fact that the results for “outs” differ so much for the UK compared with Outside-of-Wales area in the non-steady state, unsmoothed analysis (column (4) results in table 2.6 and 2.8) underline these concerns.

In the shorter, 1997Q2-2008Q1 sample, our results for the steady-state analysis without smoothing (column (2)) are tipped in favour of the “outs” for the UK, and the “ins” for Outside-of-Wales (tables 2.10 and 2.12). Thus the fragility of our results is once again evident. This is also the case in column (4) of those tables (the non-steady state, non-smoothed results), where the “outs” have a negative share of the variance for the whole of the UK and a positive share for the area Outside-of-Wales. The results for the smoothed data (steady-state and non-steady-state) are however qualitatively similar to those in the larger sample: the Outs appear to dominate. Given our concerns about the sample size for the Outside-of-Wales analysis, these concerns are even greater for our results for Wales, for which the number of each type of transition in our dataset is inevitably much smaller. Note that in both the longer and shorter samples (tables 2.7 and 2.11), the unsmoothed results (columns (2) and (4)) suggest that the “ins” win in Wales. In the analysis with smoothed data (steady state and non-steady-state – columns (1) and (3)), the longer sample suggests that “ins” and “outs” contribute equally, whereas the shorter sample suggests the “outs” are more important.

What then are we to make of these results which conflict across specifications? We argue that these results do not allow us to draw a conclusive inference about whether or not the “ins” or the “outs” win, and that we should be equally cautious about attributing relative importance to each of the four individual hazard shares.

The analyses to which we feel it is proper to attach most weight – namely the column (4) analyses in which we neither smooth the data and attempt to explain observed rather than steady-state unemployment, contains large error terms and fails to display similar results for the whole UK and for Outside-of-Wales. This is an important failure. We feel that it is important to take account of the fact that non-steady state unemployment deviates from steady-state unemployment, which is why we attach more weight to this analysis than the more plausible column (2) results (which come very close to those of Petrongolo and Pissarides, despite the difference in our sample period – see column (2) of table 2.6), which are unsmoothed yet do not contain an error term.

We are less inclined to put weight on our smoothed results. The reason is that we cannot know how much noise and how much legitimate variation it removes from the data. It is the case that the 3rd order polynomial smoother is very "aggressive" form of smoothing, and it might be possible in future work to examine the results of smoothing that removes less of the original variation. One could experiment, for example, with moving average smoothers of different lengths or higher-order polynomials.

Unfortunately, it seems that the main objective of our analysis, to ascertain the relative importance of the gross-flows in Wales, is hamstrung by the sample size in the LFS panel. The Welsh hazard rate series we construct are far more volatile in than those for the larger regions, making the precision of our results weak. This is, we argue, the reason for the wide variation in our estimated Wales hazard-rate shares between specifications. In particular, the smoothed results tell us that the "ins" and "outs" are approximately of the same importance, whereas the unsmoothed results are themselves inconsistent, depending in turn on whether or not the steady-state is imposed.

Finally, we note that our choosing to decompose observed unemployment in addition to steady-state unemployment allows us to ascertain the magnitude of departures of changes in actual unemployment from changes in steady-state unemployment, in a departure from the otherwise comparable method of Petrongolo and Pissarides. Although the magnitude of this estimated error-share varied between specifications, in general it was at least of 24% of the variation in observed unemployment (with smoothed data in the full sample – column (3) of tables 2.6-2.8), and sometimes as great as 95%(column (4) of table 2.10). This suggests that there is scope for further dynamic modelling of unemployment in relation to the hazard rates, for example of the type undertaken by Smith (2011), may also be fruitful at a regional level.

2.9 Conclusion

We have constructed gross flows for Wales based upon the Labour Force Survey panel dataset. This builds upon similar work that has been done for the whole of the UK, in Gomes (2009), Bell and Smith (2002) and Petrongolo and Pissarides (2008).

Our results are shown to be heavily contingent on the precise assumptions of the analysis – whether we decompose steady-state or observed unemployment, and whether the measure of unemployment we choose to decompose is smoothed or unsmoothed. In the UK and Outside-of-Wales, the smoothed analyses suggests that exists from unemployment dominate, in particular the hazard shares from unemployment to employment. This is true both in the pre-recession sample period and in the full sample. These results are dominated by UE and EU flows in particular.

The unsmoothed analyses give different results, depending on whether it is steady-state unemployment or observed unemployment that is decomposed, the length of the sample period and whether the whole of the UK is analysed, or just the area Outside-of-Wales.

The smoothed results for Wales have more weight on the “ins”, so that the hazard shared between the “ins” and “outs” are approximately equal for the smoothed analysis. For these results there is also more of an emphasis on inactivity related hazard shares in the full sample. The unsmoothed steady-state results are fairly consistent with the smoothed result, however when we relax the steady-state assumption with the unsmoothed data we get anomalous results.

A potential criticism of our method is that we did not deeply explore the non-steady state analysis, in the sense of providing a comprehensive account of out-of-steady state unemployment dynamics using our error term. Smith (2011) suggests that unemployment dynamics are a far slower in the UK than in the US, so that there is a smaller volume of labour turnover and it takes unemployment longer to converge to its steady state value in the absence of intervening disturbances. (Smith, p417.)

Table 2.2 and 2.3 provide no reason to think that unemployment dynamics are any faster in Wales than for the rest of the UK. This suggests that if a dynamic analysis is appropriate for the UK, then it is just as appropriate for Wales. We leave for further work the issue of how to reconcile a dynamic analysis with multiple states and gross labour market flows between different countries.

Appendix 2.1: Deriving the relationship between the hazard rate and the transition rate

Note that the formula is not original, but it is hardly ever included in the literature. (Indeed we were forced to derive it ourselves). We include it here for completeness as it is part of the method:

We show that given $\frac{X^jY^k{}_t(\tau)}{X_t^j(0)} = \int_0^\tau \lambda_t^{jk}\left(1 - \frac{X^jY^k{}_t(v)}{X_t^j(0)}\right)dv$ for $t = 1,\ldots,T$ and $j,k \in \{o,w\}$, it follows that:

$$\lambda_t^{jk} = -ln\left(1 - \frac{X^jY^k{}_t}{X_t^k(0)}\right)$$

Proof:

$$\frac{X^jY^k{}_t(\tau)}{X_t^j(0)} = \int_0^\tau \lambda_t^{jk}\left(1 - \frac{X^jY^k{}_t(v)}{X_t^j(0)}\right)dv \ = \int_0^\tau \lambda_t^{jk}\,dv - \lambda_t^{jk}\int_0^\tau \frac{X^jY^k{}_t(v)}{X_t^j(0)}dv$$

$$= \lambda_t^{jk}\int_0^\tau dv - \lambda_t^{jk}\int_0^\tau \frac{X^jY^k{}_t(v)}{X_t^j(0)}dv = \lambda_t^{jk}[\tau - 0] - \lambda_t^{jk}\int_0^\tau \frac{X^jY^k{}_t(v)}{X_t^j(0)}dv$$

$$\frac{X^jY^k{}_t(\tau)}{X_t^j(0)} = \lambda_t^{jk}\tau - \lambda_t^{jk}\int_0^\tau \frac{X^jY^k{}_t(v)}{X_t^j(0)}dv$$

Differentiating with respect to τ:

$$\frac{X^j\dot{Y}^k{}_t(\tau)}{X_t^j(0)} = \lambda_t^{jk} - \lambda_t^{jk}\frac{\partial}{\partial\tau}\left(\int_0^\tau \frac{X^jY^k{}_t(v)}{X_t^j(0)}dv\right)$$

By the Leibniz rule,

$$\frac{\partial}{\partial\tau}\left(\int_0^\tau \frac{X^jY^k{}_t(v)}{X_t^j(0)}dv\right) = \int_0^\tau \frac{\partial}{\partial\tau}\left(\frac{X^jY^k{}_t(v)}{X_t^j(0)}\right)dv + \frac{X^jY^k{}_t(\tau)}{X_t^j(0)}\frac{\partial\tau}{\partial\tau}$$

$$= 0 + \frac{X^jY^k{}_t(\tau)}{X_t^j(0)} = \frac{X^jY^k{}_t(\tau)}{X_t^j(0)}$$

Hence:

$$\frac{X^j\dot{Y}^k{}_t(\tau)}{X_t^j(0)} + \lambda_t^{jk}\frac{X^jY^k{}_t(\tau)}{X_t^j(0)} = \lambda_t^{jk}$$

Multiply through by $e^{-\int_0^\tau -\lambda_t^{jk}d\tau} = e^{\lambda_t^{jk}\tau}$

$$\frac{X^j\dot{Y}^k{}_t(\tau)}{X_t^j(0)}e^{\lambda_t^{jk}\tau} + \lambda_t^{jk}\frac{X^jY^k{}_t(\tau)}{X_t^j(0)}e^{\lambda_t^{jk}\tau} = e^{\lambda_t^{jk}\tau}\lambda_t^{jk}$$

$$X^j\dot{Y}^k{}_t(\tau)e^{\lambda_t^{jk}\tau} + \lambda_t^{jk}X^jY^k{}_t(\tau)e^{\lambda_t^{jk}\tau} = X_t^j(0)e^{\lambda_t^{jk}\tau}\lambda_t^{jk}$$

The right hand side is equivalent to $\frac{d}{d\tau}\left(X^jY^k{}_t(\tau)e^{\lambda_t^{jk}\tau}\right)$ so we have:

$$\frac{d}{d\tau}\left(X^jY^k{}_t(\tau)e^{\lambda_t^{jk}\tau}\right) = X_t^j(0)e^{\lambda_t^{jk}\tau}\lambda_t^{jk}$$

Integrate with respect to τ over the interval $[0,1]$:

$$\frac{X^jY^k{}_t(\tau)}{X_t^j(0)}e^{\lambda_t^{jk}\tau} = \int_0^1 e^{\lambda_t^{jk}\tau}\lambda_t^{jk}\,d\tau = \lambda_t^{jk}\int_0^1 e^{\lambda_t^{jk}\tau}\,d\tau = \frac{\lambda_t^{jk}}{\lambda_t^{jk}}\left(e^{\lambda_t^{jk}\tau} - 1\right) = \left(e^{\lambda_t^{jk}\tau} - 1\right)$$

$$\frac{X^jY^k{}_t(\tau)}{X_t^j(0)}e^{\lambda_t^{jk}\tau} = \left(e^{\lambda_t^{jk}\tau} - 1\right)$$

Let $\tau = 1$:

$$\frac{X^jY^k{}_t}{X_t^j(0)}e^{\lambda_t^{jk}} = \left(e^{\lambda_t^{jk}} - 1\right)$$

Multiplying through by $e^{-\lambda_t^{jk}}$ gives:

$$\frac{X^jY^k{}_t}{X_t^j(0)} = 1 - e^{-\lambda_t^{jk}}$$

Or:

$$\lambda_t^{jk} = -ln\left(1 - \frac{X^jY^k{}_t}{X_t^j(0)}\right)$$

As required.

Chapter 3: Testing and estimating the basic Mortensen-Pissarides model using indirect inference

3.1 Indirect Inference

3.2 A discrete-time Mortensen-Pissarides model

3.3 Data

3.4 Calibration

3.5 Model testing using Indirect Inference with Shimer's (2005) parameters

3.6 Indirect inference estimation of the model

3.7 Discussion

Chapter 1 outlines the Mortensen-Pissarides model of the labour market and the associated Shimer puzzle, which is the failure of the model to match moments in aggregate US data such as the standard deviations of vacancies, unemployment, market tightness (the vacancy/unemployment ratio) when the model is subject to shocks to productivity. The usual explanation given for the Shimer puzzle is that wages absorb too much of changes in productivity, so that productivity shocks result in insufficiently strong incentives for job creation.

Shimer (2005(a)) is the original formulation of the Shimer puzzle. Here Shimer calibrates key parameters of the Mortensen - Pissarides model, and comparing moments of endogenous variables with those in the data, finds important discrepancies. Importantly, Shimer's moment comparison method is simply based on inspection of the distance between moments from the model and the data. The moment comparisons are done independently, without reference to the joint distribution of parameters implied by the model. In this chapter, we work with a similar version of the Mortensen-Pissarides model. Rather than following Shimer's methodology of comparing moments, we *use indirect inference evaluation* to evaluate the model under Shimer's suggested parameters.

Indirect inference offers a statistically-founded way of evaluating the performance of a model subject to shocks. It takes a set of primitive model parameters, and simulates – using a bootstrap procedure - a distribution of model outcomes which can be compared to the data in the form of a statistical hypothesis test[20]. Crucially, the statistical test is based upon the joint distribution of data moments. We therefore argue that in the context of the Mortensen-Pissarides model, the use of indirect-inference offers an advantage over Shimer's moment - by - moment comparison method, while at the same time retaining the focus on endogenous-variable moments that form the basis of the Shimer critique.

Exploring the implementation of the method is therefore the contribution of this chapter. Our intention here is primarily to see what conclusions about the basic Mortensen-Pissarides model we can draw using this method. We are interested in whether the statistical testing procedure of indirect inference reaches the same conclusion as Shimer's calibration procedure – namely that the Mortensen Pissarides model fits the data poorly due to wages absorbing productivity shocks.

Indirect inference can also be used for estimation of primitive parameters. This is an extension of the model evaluation procedure - one can search for primitive parameters that give the best fit of the model in terms of moments of interest, by searching for the minimal test statistic. This method is particularly useful in the context of models that do not fit the data under calibrated parameters. We argue that indirect inference estimation is therefore useful for our purposes, given that the Shimer puzzle suggests that the basic form of the Mortensen-Pissarides model is likely to be rejected under Shimer's parameters, and given that the literature that claims that choosing new parameters can improve the model's performance (Mortensen and Nagypál (2007), Hagerdorn and Manovskii (2008)).

In what follows, we find that the Mortensen-Pissarides model under Shimer's calibration is indeed rejected using indirect inference evaluation. Turning to indirect inference estimation, we are able to estimate and report a set of parameters that do fit the data using a statistical test based on the standard deviations of the endogenous variables, but that do not fit the data using a test based on VAR coefficients.

[20] The hypothesis is that the observed data were generated by the model under the given set of primitive parameters. See the extended discussion below for more details.

We begin by summarizing the method of indirect inference both for testing and estimating applications in section 3.1. Section 3.2 summarizes the main features of the model, and we provide more detail in Appendix 3.1. Section 3.3 provides details of all of the data used in this chapter, both in terms of sources and in terms of summary statistics. Section 3.4 explains how our model is calibrated so that in section 3.5 we can evaluate it under parameters equivalent to those of Shimer (2005). In section 3.5 we estimate the model under indirect inference and comment on our findings. We discuss and interpret the results in section 3.6.

3.1 Indirect inference

Consider an economic model in structural form:

$$\mathbf{A}\mathbf{y_t} = \mathbf{B}\mathbf{x_t} + \mathbf{C}E_t\mathbf{y_{t+1}} + \mathbf{e_t}$$

$$t = 1, \dots, T$$

We assume that there are r equations, and that vector $\mathbf{y_t}$ contains r endogenous variables. There are n exogenous variables in $\mathbf{x_t}$, and all variables are assumed to be stationary. t is a time subscript, and E_t is the expectations operator.

The elements of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are structural parameters, since the equation is in structural form. These elements will be combinations of primitive parameters $\boldsymbol{\theta}$, each of which has a theoretical economic interpretation. Let all the primitives of $\boldsymbol{\theta}$ in $\mathbf{A}$ be $\boldsymbol{\theta}^{\mathbf{A}}$, let those in $\mathbf{B}$ be $\boldsymbol{\theta}^{\mathbf{B}}$ and those in $\mathbf{C}$ be $\boldsymbol{\theta}^{\mathbf{C}}$. The model can then be written more fully as:

$$\mathbf{A}(\boldsymbol{\theta}^{\mathbf{A}})\mathbf{y_t} = \mathbf{B}(\boldsymbol{\theta}^{\mathbf{B}})\mathbf{x_t} + \mathbf{C}(\boldsymbol{\theta}^{\mathbf{C}})E_t\mathbf{y_{t+1}} + \mathbf{e_t}$$

$$t = 1, \dots, T$$

Indirect inference is a simulation-based method.[21]   It has two broad uses with respect to a model like the one above. (i) It can be used to evaluate the model against the data for a particular choice of $\boldsymbol{\theta}$ and for a particular choice of model metric. We denote the particular set of primitive parameters under consideration by $\widehat{\boldsymbol{\theta}}$ and assume that as in the case of the general values these can be grouped as $\widehat{\boldsymbol{\theta}}^{\mathbf{A}}$,

[21] An important general reference for the use of indirect inference in econometrics is Gourieroux and Monfort (1993).

$\widehat{\boldsymbol{\theta}}^{\mathbf{B}}$ and $\widehat{\boldsymbol{\theta}}^{\mathbf{C}}$, according to the structural matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ in which they make an appearance. (ii) It can be used to estimate the parameters of $\boldsymbol{\theta}$, finding a set of values $\boldsymbol{\theta}^*$ that brings the performance of the model closest to the data based on a pre-specified model metric. In this chapter we use both applications. Estimation, (ii), can however be easily described with reference to the testing procedure, (i), so we focus first on the latter. The method is very difficult to describe piecemeal, so to ease our exposition we first present a schematic summary of indirect inference testing. It begins by assuming that we have a model in the structural form above, and we wish to test the fit of the model against real world data for a particular vector of primitive parameters, $\widehat{\boldsymbol{\theta}}$.

Schema for testing

(1) Choose a *K*-parameter auxiliary model. The parameters of the auxiliary model should be the means by which we wish to compare the model against the data. The model should be simple to estimate. We first estimate the auxiliary model on our data, which should correspond to the real-world observations on the r endogenous variables of our structural model contained in $\mathbf{y_t}$. At the end of this step we will have one set of *K* auxiliary-model parameters. We denote the kth such parameter estimate by $\hat{\phi}_k$, with $k = 1,..,K$. In the applications that follow in this and the following chapters, three different auxiliary models will be used. The first consists simply of a vector of r standard deviations of the endogenous variables, so that $K = r$. The second is a set of VAR(1) coefficients estimated on the endogenous variables only and with no constant terms, in which case $K = r^2$. The third auxiliary model combines both standard deviations and VAR coefficients, which implies that $K = r + r^2$.

(2) Using the structural model, compute the *r* structural errors in $\hat{\mathbf{e}}_\mathbf{t}$ under $\widehat{\boldsymbol{\theta}}$ using

$$\hat{\mathbf{e}}_\mathbf{t} = \mathbf{A}(\widehat{\boldsymbol{\theta}}^\mathbf{A})\mathbf{y_t} - \mathbf{B}(\widehat{\boldsymbol{\theta}}^\mathbf{B})\mathbf{x_t} - \mathbf{C}(\widehat{\boldsymbol{\theta}}^\mathbf{C})E_t\mathbf{y_{t+1}}$$

using actual data for $\mathbf{y_t}$, $\mathbf{x_t}$ and if necessary, using a VAR to obtain estimated data values for $E_t\mathbf{y_{t+1}}$.

(3) Estimate autoregressive (here illustrated as AR(1)s[22]) processes for each of the *r* structural errors in $\mathbf{e_t}$ so as to obtain serially independent residuals for each. (The residuals do not have to be normal – see Meenagh, Minford and Wickens (2008), p6).

$$e_{1t} = \hat{\rho}_1 e_{1t-1} + \hat{\varepsilon}_{1t}$$
$$\dots$$
$$e_{rt} = \hat{\rho}_r e_{rt-1} + \hat{\varepsilon}_{rt}$$

Estimate autoregressive processes for the exogenous variables as well, and obtain the residuals which are also white noise.

$$x_{1t} = \hat{\varpi}_1 x_{1t-1} + \hat{\zeta}_{1t}$$
$$\dots$$
$$x_{nt} = \hat{\varpi}_n x_{nt-1} + \hat{\zeta}_{nt}$$

And the end of step (3) one will have parameters $\hat{\rho}_1, \dots, \hat{\rho}_r; \hat{\varpi}_1, \dots, \hat{\varpi}_n$ and residuals $\hat{\varepsilon}_{1t}, \dots, \hat{\varepsilon}_{rt}; \hat{\zeta}_{1t}, \dots, \hat{\zeta}_{nt}\ \forall t$ in the notation above.

(4) Sample the residuals in order to create 2000 vector-bootstrapped samples of innovations. This means using each of the estimated univariate processes to generate 2000 simulated samples of structural errors, and 2000 samples of exogenous variables. In the notation above, this is the stage in which we create, for $j = 1, \dots, 2000$:

$$\hat{\hat{e}}_{1t}^{j} = \hat{\rho}_1 \hat{\hat{e}}_{1t-1}^{j} + \hat{\hat{\varepsilon}}_{1t}^{j}$$
$$\dots$$
$$\hat{\hat{e}}_{rt}^{j} = \hat{\rho}_r \hat{\hat{e}}_{rt-1}^{j} + \hat{\hat{\varepsilon}}_{rt}^{j}$$
$$\hat{\hat{x}}_{1t}^{j} = \hat{\varpi}_1 \hat{\hat{x}}_{1t-1}^{j} + \hat{\hat{\zeta}}_{1t}^{j}$$
$$\dots$$
$$\hat{\hat{x}}_{nt}^{j} = \hat{\varpi}_n \hat{\hat{x}}_{nt-1}^{j} + \hat{\hat{\zeta}}_{nt}^{j}$$

(5) Solve the model under $\hat{\boldsymbol{\theta}}$ for $\mathbf{y_t}$, in terms of the structural errors and exogenous variables. Use the simulated data and the model solution to generate 2000 samples of pseudo-data for the endogenous variables.

(6) Estimate the chosen K-parameter auxiliary model on each of the 2000 simulated samples. We denote the kth auxiliary parameter estimate on the

[22] In what follows in this chapter we use AR(1) processes to generate innovations to our structural error terms for all variables. We check each series for serial independence. For all of our variables but one (wages), we fail to reject the null of serial independent residuals after the original series has been fitted to an AR(1) process. Wages are found to have serially independent residuals when they are fitted to an AR(2). To check the robustness of our results, we run indirect inference tests again using AR(2) processes to extract the residuals. The results are not substantially affected however.

pseudo-data from the $\hat{\phi}_k^j(\hat{\boldsymbol{\theta}})$ for $j = 1, \ldots, 2000$ and $k = 1, \ldots, K$. Compute the mean of each parameter across the simulations: $\bar{\phi}_k(\hat{\boldsymbol{\theta}}) = \frac{1}{2000}\sum_j^{2000} \hat{\phi}_k^j(\hat{\boldsymbol{\theta}})$ for $k = 1, \ldots, K$.

(7) Compute the Wald statistic. For this let:

$$\mathbf{v} = [\hat{\phi}_1 - \bar{\phi}_1(\hat{\boldsymbol{\theta}}) \quad \ldots \quad \hat{\phi}_k - \bar{\phi}_k(\hat{\boldsymbol{\theta}}) \quad \ldots \quad \hat{\phi}_K - \bar{\phi}_K(\hat{\boldsymbol{\theta}})]$$

where $\hat{\phi}_1, \ldots, \hat{\phi}_k, \ldots, \hat{\phi}_K$ come from step (1) and $\bar{\phi}_1(\hat{\boldsymbol{\theta}}), \ldots, \bar{\phi}_k(\hat{\boldsymbol{\theta}}), \ldots, \bar{\phi}_K(\hat{\boldsymbol{\theta}})$ come from step (6). Also let:

$$\mathbf{W}(\hat{\boldsymbol{\theta}}) = \left[E\left(\hat{\phi}_i - \bar{\phi}_j(\hat{\boldsymbol{\theta}})\right)\left(\hat{\phi}_i - \bar{\phi}_j(\hat{\boldsymbol{\theta}})\right)\right]^{-\mathbf{1}}; \quad i, j \in \{1, \ldots, K\}$$

In other words, $\mathbf{W}(\hat{\boldsymbol{\theta}})$ is the $K \times K$ inverse of the variance-covariance matrix of $\mathbf{v}$.

The Wald statistic is given by $\mathbf{v}\mathbf{W}(\hat{\boldsymbol{\theta}})\mathbf{v}'$.

(8) Compute the distribution of the Wald under the null as follows: $WALD_0^j = \mathbf{v}_\mathbf{0}^\mathbf{j}\mathbf{W}(\hat{\boldsymbol{\theta}})\mathbf{v}_\mathbf{0}^{\mathbf{j}\,\prime}$ for $j = 1, \ldots 2000$, where:

$$\mathbf{v}_\mathbf{0}^\mathbf{j} = [\hat{\phi}_1^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_1(\hat{\boldsymbol{\theta}}) \quad \ldots \quad \hat{\phi}_k^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_k(\hat{\boldsymbol{\theta}}) \quad \ldots \quad \hat{\phi}_K^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_K(\hat{\boldsymbol{\theta}})]$$

Note that under the null hypothesis, the Wald has an asymptotically chi-squared (*K)* distribution, although one does not need to rely on its asymptotic properties. For inference one instead may use the finite-sample simulated distribution that can be constructed by ordering the 2000 values of $WALD_0^j$ by increasing magnitude.

(9) The Wald statistic from step (7) can then be compared to the distribution of Wald under the null hypothesis from the previous step. The model under $\hat{\boldsymbol{\theta}}$ is rejected if the Wald statistic lies in the tail outside of 95% of the values of the finite sample Wald distribution constructed in step (8). A p-value for the Wald can also be constructed, by comparing the value of the Wald statistic to the percentile of the closest value of the Wald under the null hypothesis that the data was generated by the model.

(10) (If required). Compute the t-statistic equivalent of the Wald. This can be easier to interpret than the Wald's p-value, and is defined so that under the null hypothesis that the true data was generated by the model, the 95th

percentile of the Wald distribution has a value of 1.645. Let $WALD$ be the Wald statistic from step (7) and let $WALD_{95}$ be the 95th percentile value of the Wald distribution under the null. K is the number of auxiliary model parameters. The formula for the t-statistic (see Meenagh and Le (2013)) is then given by:

$$t = 1.645 \times \left(\frac{\sqrt{2 \times WALD} - \sqrt{2 \times (K-1)}}{\sqrt{2 \times WALD_{95}} - \sqrt{2 \times (K-1)}} \right)$$

Having provided an outline we now clarify the method.

Step (1) concerns the choice of auxiliary model, the estimated parameters of which are first used to represent the true data, and which is later estimated repeatedly on the pseudo-data. In economic applications the auxiliary model is typically a pure statistical model with little economic content such as volatilities of endogenous variables, VAR coefficients, impulse response functions or some combinations of these. The particular choice should be informed by the objectives of the modelling exercise. This means that the choice is essentially ad-hoc, although it is the case that the more parameters in the auxiliary model, the greater the power of indirect inference to reject a mis-specified model. (This result is described in Le, Meenagh, Minford and Wickens (2012), see below for more details on this). One could arguably count this ad-hocness as a strength or a weakness. In its favour, it means that the method is adaptable – the modeller can focus on the moments that are most of interest, and is free to exclude others from consideration. Of course, such ad-hocness adds the ambiguity of interpretation – what are we to make of a model that fits the data well in terms of standard deviations, but only three out of six VAR coefficients?

In step (2), the structural errors are calculated for each model equation and each time period. The equation in step (2) makes clear that they are simply the difference between the data-values for the endogenous variables, and the model's predicted values given the data for the exogenous variables and under the particular choice of $\hat{\boldsymbol{\theta}}$ that are being tested. We use the structural errors to create bootstrapped samples of shocks, but as is made clear in step (3), we do not bootstrap the structural errors themselves but their estimated white-noise innovations from a univariate time-series process. We also estimate univariate processes on the exogenous variables, and obtain white-noise innovations for these. Thus with r structural equations and n exogenous variables, step (3) requires that we estimate $r+n$ univariate time series equations:

$$\begin{gathered} e_{1t} = \hat{\rho}_1 e_{1t-1} + \hat{\varepsilon}_{1t} \\ \dots \\ e_{rt} = \hat{\rho}_r e_{rt-1} + \hat{\varepsilon}_{rt} \end{gathered}$$

$$\begin{gathered} x_{1t} = \widehat{\varpi}_1 x_{1t-1} + \hat{\zeta}_{1t} \\ \dots \\ x_{nt} = \widehat{\varpi}_n x_{nt-1} + \hat{\zeta}_{nt} \end{gathered}$$

The estimated residual innovations to each of the estimated AR(1)s $\hat{\varepsilon}_{1t}, \dots, \hat{\varepsilon}_{rt}; \hat{\zeta}_{1t}, \dots, \hat{\zeta}_{nt}$ should be white-noise – mean zero, stationary, and with all the temporal correlation removed - otherwise one must reconsider the time-series model used. We assume here that univariate AR(1)s are sufficient to produce white noise innovations for each equation.

In step (4) we draw bootstrap samples of the innovations. The bootstrapping procedure used is a *vector* bootstrap. This means that rather than drawing samples of innovations individually, we do the following. Suppose that our sample of real-world data is of length T. This means that estimating univariate AR(1) processes on the structural errors and innovations will result in $(r+n)$ vectors of fitted innovations each of which has dimension $(T-1)$. To create a single simulated bootstrap sample of the same length, we draw $T-1$ integers at random and with replacement from the discrete uniform distribution of the integers between 1 and $T-1$. The drawn integers represent the time index of the estimated innovations $\hat{\varepsilon}_{1t}, \dots, \hat{\varepsilon}_{rt}; \hat{\zeta}_{1t}, \dots, \hat{\zeta}_{nt}$ in the original sample. For example, suppose we are making draws for our jth of 2000 simulated samples. If the first integer we draw at random is 6, the values of the first set of innovations in the bootstrap sample are $\hat{\varepsilon}_{16}, \dots, \hat{\varepsilon}_{r6}; \hat{\zeta}_{16}, \dots, \hat{\zeta}_{n6}$. In this way the drawn bootstrap samples preserve any intra-temporal correlation that may exist in the original structural innovations. We denote estimated parameter values (including the original time-series residuals) with a single-hat and bootstrapped residuals and simulated values with a double-hat. Under this notation, in our example, we would therefore set:

$$\left[\hat{\hat{\varepsilon}}_{11}^{j}, \dots, \hat{\hat{\varepsilon}}_{r1}^{j}; \hat{\hat{\zeta}}_{11}^{j}, \dots, \hat{\hat{\zeta}}_{n1}^{j}\right] = \left[\hat{\varepsilon}_{16}, \dots, \hat{\varepsilon}_{r6}; \hat{\zeta}_{16}, \dots, \hat{\zeta}_{n6}\right]$$

To create simulated structural errors for each bootstrap sample, one then simply uses the estimated parameter set $\hat{\rho}_1, \dots, \hat{\rho}_r; \hat{\varpi}_1, \dots, \hat{\varpi}_n$ and an initial drawing of $e_{10}, \dots, e_{r0}; \zeta_{10}, \dots, \zeta_{n0}$ as follows.

$$\begin{gathered}
\hat{\hat{e}}_{1t}^{j} = \hat{\rho}_1 \hat{\hat{e}}_{1t-1}^{j} + \hat{\hat{\varepsilon}}_{1t}^{j} \\
\dots \\
\hat{\hat{e}}_{rt}^{j} = \hat{\rho}_r \hat{\hat{e}}_{rt-1}^{j} + \hat{\hat{\varepsilon}}_{rt}^{j}
\end{gathered}$$

$$\begin{gathered}
\hat{\hat{x}}_{1t}^{j} = \hat{\varpi}_1 \hat{\hat{x}}_{1t-1}^{j} + \hat{\hat{\zeta}}_{1t}^{j} \\
\dots \\
\hat{\hat{x}}_{nt}^{j} = \hat{\varpi}_n \hat{\hat{x}}_{nt-1}^{j} + \hat{\hat{\zeta}}_{nt}^{j}
\end{gathered}$$

We create 2000 bootstrap samples so $j = 1, \dots, 2000$.

Steps (5) and (6) complete the process of generating the 2000 samples of pseudo-data. This requires us to find the solution to the structural model, also known as the reduced form. This can be done with conventional methods – the solution procedure will depend on the particular way that expectations are specified and other features of the model. We may then simulate the endogenous variables by applying the pseudo-samples of $\hat{\hat{e}}_{1t}^{j}, \dots, \hat{\hat{e}}_{rt}^{j}$ and $\hat{\hat{x}}_{1t}^{j}, \dots, \hat{\hat{x}}_{nt}^{j}$. The equation for generating the pseudo-samples of endogenous variables is given below:

$$\begin{gathered}
\hat{\hat{\mathbf{y}}}_{t}^{j} = \mathbf{\Upsilon_0}(\hat{\boldsymbol{\theta}}) \hat{\hat{\mathbf{x}}}_{t}^{j} + \mathbf{\Upsilon_1}(\hat{\boldsymbol{\theta}}) \hat{\hat{\mathbf{e}}}_{t}^{j} \\
t = 1, \dots, T \\
j = 1, \dots, 2000
\end{gathered}$$

where $\mathbf{\Upsilon_0}(\hat{\boldsymbol{\theta}})$ and $\mathbf{\Upsilon_1}(\hat{\boldsymbol{\theta}})$ are respectively $r \times n$ and $r \times r$ matrices of reduced form parameters, which are the solution to the structural model under $\hat{\boldsymbol{\theta}}$. These will depend on the structural matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$, and hence ultimately on the primitive parameters $\hat{\boldsymbol{\theta}}$.

Step (6) requires estimating the auxiliary model 2000 times upon the pseudo data, and generating 2000 sets of estimated auxiliary model parameters. The result is a

K-dimensional distribution of features of interest under the assumption that the model is the true data-generating mechanism. From step (1), we also now have a single K-dimensional vector of auxiliary model parameters from the data. It should be clear that we have the necessary ingredients for a statistical test (an empirical distribution under a null hypothesis, and a high-dimensional version of a comparable test-statistic), except that there are too many dimensions. To reduce the dimensions down from K to 1, we use steps (7) to construct a single Wald-statistic based on the step (1) parameters, and we use step (8) to construct a (univariate) distribution of the Wald under the null. We discuss the Wald statistic and the distribution of the Wald in more detail below.

In steps (9) and (10) we conduct the actual inference procedure. In step (9) we use the p-value of the Wald statistic, which can be obtained by looking at the percentile at which it lies in the Wald distribution. If the Wald-statistic lies outside the maximum point of the simulated Wald distribution however, it is more informative to calculate the equivalent t-statistic which is given by the formula in step (10). The model is rejected at the 5% level of significance if the p-value is less than 0.05 or if the t-statistic is greater than 1.645.

More about the Wald statistic

We now use a particular example of an auxiliary model to understand the construction of the Wald statistic in more detail. The discussion is an algebraic version of the presentation given in Minford, Meenagh and Wickens (2009). Consider again the model:

$$\mathbf{A}(\boldsymbol{\theta}^{\mathbf{A}})\mathbf{y_t} = \mathbf{B}(\boldsymbol{\theta}^{\mathbf{B}})\mathbf{x_t} + \mathbf{C}(\boldsymbol{\theta}^{\mathbf{C}})E_t\mathbf{y_{t+1}} + \mathbf{e_t}$$

$$t = 1, \ldots, T$$

Suppose more specifically that there are two endogenous variables so that $\mathbf{A}(\boldsymbol{\theta}^{\mathbf{A}})$ is a 2×2 matrix and $\mathbf{y_t} = [y_{1t} \quad y_{2t}]'$ is a vector of dimension 2. Suppose that our auxiliary model is a VAR(1) of the endogenous variables, which for simplicity is orthogonal, in other words:

$$\mathbf{y_t} = \begin{bmatrix} \phi_1 & 0 \\ 0 & \phi_2 \end{bmatrix} \mathbf{y_{t-1}} + \boldsymbol{\epsilon}_\mathbf{t}$$

Let $\hat{\phi}_k : k = 1,2$ be the coefficients obtained from running the VAR on the true (observed) data. Let $\hat{\phi}_k^j(\widehat{\boldsymbol{\theta}}) : k = 1,2$ be the respective VAR coefficients obtained by estimating a VAR on the jth sample that was simulated by the model under $\widehat{\boldsymbol{\theta}}$**,** and let $\bar{\phi}_k(\widehat{\boldsymbol{\theta}}) : k = 1,2$ denote the respective *means* of these model-generated VAR coefficients. The Wald statistic will then be calculated as follows:

$$WALD = \mathbf{v}'\mathbf{W}(\widehat{\boldsymbol{\theta}})\mathbf{v}$$

where $\mathbf{v}$ is a vector of deviations of the data-estimated VAR coefficients from the mean of the model generated coefficients, in other words:

$$\mathbf{v} = \left[\hat{\phi}_1 - \bar{\phi}_1(\widehat{\boldsymbol{\theta}}) \quad \hat{\phi}_2 - \bar{\phi}_2(\widehat{\boldsymbol{\theta}})\right]'$$

$\mathbf{W}(\widehat{\boldsymbol{\theta}})$ is a weight matrix created out of the inverse of variance-covariance matrix, denoted as follows:

$$\mathbf{W}(\hat{\boldsymbol{\theta}}) = \begin{bmatrix} \mathrm{var}\left(\hat{\phi}_1^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_1(\hat{\boldsymbol{\theta}})\right) & \mathrm{cov}\left(\left(\hat{\phi}_1^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_1(\hat{\boldsymbol{\theta}})\right)\left(\hat{\phi}_2^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_2(\hat{\boldsymbol{\theta}})\right)\right) \\ \mathrm{cov}\left(\left(\left(\widehat{\phi_1^j}(\hat{\boldsymbol{\theta}}) - \bar{\phi}_1(\hat{\boldsymbol{\theta}})\right)\left(\hat{\phi}_2^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_2(\hat{\boldsymbol{\theta}})\right)\right)\right) & \mathrm{var}\left(\hat{\phi}_2^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_2(\hat{\boldsymbol{\theta}})\right) \end{bmatrix}^{-1}$$

In this simple case, by working through the algebra to obtain $\mathbf{v}'\mathbf{W}(\widehat{\boldsymbol{\theta}})\mathbf{v}$, the Wald statistic can be shown to be equivalent to the following:

$$WALD = \left(\hat{\phi}_1 - \bar{\phi}_1(\hat{\boldsymbol{\theta}})\right)^2 \left[\frac{1}{\mathrm{var}\left(\hat{\phi}_1^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_1(\hat{\boldsymbol{\theta}})\right) - \frac{\mathrm{cov}^2\left(\left(\left(\hat{\phi}_1^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_1(\hat{\boldsymbol{\theta}})\right)\left(\hat{\phi}_2^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_2(\hat{\boldsymbol{\theta}})\right)\right)\right)}{\mathrm{var}\left(\hat{\phi}_2^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_2(\hat{\boldsymbol{\theta}})\right)}}\right] +$$

$$\left(\hat{\phi}_2 - \bar{\phi}_2(\hat{\boldsymbol{\theta}})\right)^2 \left[\frac{1}{\mathrm{var}\left(\hat{\phi}_2^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_2(\hat{\boldsymbol{\theta}})\right) - \frac{\mathrm{cov}^2\left(\left(\hat{\phi}_1^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_1(\hat{\boldsymbol{\theta}})\right)\left(\hat{\phi}_2^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_2(\hat{\boldsymbol{\theta}})\right)\right)}{\mathrm{var}\left(\hat{\phi}_1^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_1(\hat{\boldsymbol{\theta}})\right)}}\right]$$

$$- \left(\hat{\phi}_1 - \bar{\phi}_1(\hat{\boldsymbol{\theta}})\right)\left(\hat{\phi}_2 - \bar{\phi}_2(\hat{\boldsymbol{\theta}})\right)\left[\frac{\operatorname{cov}\left(\left(\hat{\phi}_1^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_1(\hat{\boldsymbol{\theta}})\right)\left(\hat{\phi}_2^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_2(\hat{\boldsymbol{\theta}})\right)\right)}{\operatorname{var}\left(\hat{\phi}_1^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_1(\hat{\boldsymbol{\theta}})\right)\operatorname{var}\left(\hat{\phi}_2^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_2(\hat{\boldsymbol{\theta}})\right) - \operatorname{cov}^2\left(\left(\hat{\phi}_1^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_1(\hat{\boldsymbol{\theta}})\right)\left(\hat{\phi}_2^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_2(\hat{\boldsymbol{\theta}})\right)\right)}\right]$$

In other words, the Wald statistic is a weighted, linear combination of squared deviations of the data-generated auxiliary parameters from the mean of the simulated outcomes under $\hat{\boldsymbol{\theta}}$, plus a cross product term. The weights come from the inverse variance-covariance matrix $\mathbf{W}(\hat{\boldsymbol{\theta}})$, a substitute for the de-meaned distribution of auxiliary parameters of the model under $\hat{\boldsymbol{\theta}}$.

One can see from the equation above that the Wald criterion uses auxiliary parameter variation implied by the model under $\hat{\boldsymbol{\theta}}$ to infer the extent to which deviation of the mean of these parameters from those generated by the data should be 'penalised.' For example, if $\operatorname{var}\left(\hat{\phi}_1^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_1(\hat{\boldsymbol{\theta}})\right)$ is relatively large, deviations in $\hat{\phi}_1$ around $\bar{\phi}_1(\hat{\boldsymbol{\theta}})$ will be weighted less heavily than if $\operatorname{var}\left(\hat{\phi}_1^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_1(\hat{\boldsymbol{\theta}})\right)$ is relatively small. When there is a relatively large covariance between de-meaned $\hat{\phi}_1^j(\hat{\boldsymbol{\theta}})$ and de-meaned $\hat{\phi}_2^j(\hat{\boldsymbol{\theta}})$ (i.e. when it is clear that the joint distribution of $\left(\hat{\phi}_1^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_1(\hat{\boldsymbol{\theta}})\right)$ and $\left(\hat{\phi}_2^j(\hat{\boldsymbol{\theta}}) - \bar{\phi}_2(\hat{\boldsymbol{\theta}})\right)$ is far from linearly independent), values of $\left(\hat{\phi}_k - \bar{\phi}_k(\hat{\boldsymbol{\theta}})\right)^2$ will be penalised more heavily than they would if the auxiliary model parameters were assumed to be independent. (This would be the case if $\mathbf{W}(\hat{\boldsymbol{\theta}})$ were assumed to be a diagonal matrix). Another way of saying this is that use of the Wald statistic with $\mathbf{W}(\hat{\boldsymbol{\theta}})$ models the joint distribution of auxiliary model parameters. For a numerical illustration of this point see Le, Minford and Wickens (2010).

Indirect inference estimation for a given choice of auxiliary model:

Estimating the model means finding the vector of primitive parameter values $\hat{\boldsymbol{\theta}}$ within the parameter space of the model that minimizes the Wald statistic. In practise this means running through steps (2)-(7) repeatedly, for different candidate values of $\hat{\boldsymbol{\theta}}$ in the parameter space. An obvious (and slow) way to do this is via grid search. Another way is to use a different sort of stochastic search algorithm. For example, Meenagh and Le (2013) recommend the use of the simulated annealing algorithm, and this is also the search algorithm which we use in this chapter's estimation. Under this algorithm, an initial choice of parameter vector $\boldsymbol{\theta}$ is chosen, and the Wald at that point is evaluated by running through steps (2)-(7) above. The algorithm then moves randomly to try a new point in the parameter space. When a new point in the parameter space is found to have a smaller Wald than any point preceding it in the sequence, it is chosen to be the current point from which the search for the minimum proceeds. The algorithm can also move to points which have a larger Wald, although the probability of this happening decreases with the number of points at which Wald statistics have previously been evaluated[23]. Eventually, after a certain number of best points are found, the search is once again widened, by increasing the acceptance probability. There are many different available stopping rules for the algorithm. In this chapter we set the maximum number of iterations to 500.

Note that due to the random nature of the bootstrap, there is stochastic volatility in the Wald. Running through steps (2)-(7) for the *same* $\widehat{\boldsymbol{\theta}}$ will not give an identical Wald each time. The volatility of the Wald can however be reduced by increasing the number of simulated samples drawn for each test.

One issue which remains relatively unexplored in the literature is the small sample efficiency properties of indirect inference estimators. This relates to the more practical issue of what an adequate sample size might be to provide sufficiently precise estimates. While we are unable to provide a benchmark sample size with

[23] In fact, the acceptance probability is not directly a function of the number of evaluated Wald statistics. Instead the number of evaluated Wald statistics is negatively related to an intermediate variable, known as the temperature. The temperature starts at some high pre-specified maximum value and as the number of evaluated Wald statistics increases the value of the temperature falls, according to a pre-specified schedule. The schedule determines the specific rate of descent of the temperature with respect to the number of evaluated Wald statistics. The acceptance probability decreases as the temperature decreases, so that with more evaluations, the algorithm becomes decreasingly likely to accept points in the parameter space of $\boldsymbol{\theta}$ which have a greater Wald than the current point. An added complication is that the acceptance probability is also decreasing in the distance between the current candidate point and the point with the minimum Wald that has been found in the parameter space thus far. Eventually, after a certain number of candidate points have been accepted, the temperature and hence the search probability are reset to their maximum values, and the algorithm begins another iteration. The algorithm is designed this way to reduce the chance that it will end up at a local rather than a global minimum.
See http://www.mathworks.co.uk/help/gads/how-simulated-annealing-works.html

regards to efficiency, we note that our 114 observations for the period 1976Q1-2003Q3 compares favourably with the 98 quarters used in the estimation results of Meenagh, Minford and Wickens (2009, p450).

Why use indirect inference?

At some stage of econometric model evaluation, data must be somehow used in conjunction with the model. One broad approach concerns itself with matching data moments – a canonical reference is Kydland and Prescott (1982). Another approach posits that the model in question is a restricted version of a more general a-theoretical model, and consists of likelihood ratio tests of the two representations of the data. Most research on the aggregate search and matching model falls clearly within the first group. (See Chapter 1.)

Le, Minford and Wickens (2010) describe the first approach in detail, labelling it the Puzzles Methodology. They associate this methodology with five modelling tendencies. These are: (i) The modeller has a free choice of shocks, (ii) Shocks are scaled arbitrarily to match moments of endogenous variables, (iii) Required values for other parameters are obtained from VAR estimation, (v) Particular moments of endogenous variables are compared with the data, (iv) but only one-by-one – the joint distributions of the moments are ignored. (Le, Minford and Wickens 2010 (p))[24]. Le, Minford and Wickens argue that these tendencies are without clear justification, and that indirect inference provides a means of replacing them with something more well-founded. Broadly, the way in which structural errors are derived under indirect inference deals with (i) and (ii), the block-bootstrap deals with (iii), (iv) and (v). In the case of (i) and (ii), shocks pertain to each of the structural equations and the scale of the errors is determined by the error inherent in the structural equations. Errors are therefore no longer at the modeller's discretion. In the case of (iii), (iv) and (v), indirect inference replaces the moment-by-moment comparison with a statistical test of the moments of interest. The advantages of this as opposed to looking at only the differences between data and model moments should be obvious from the discussion of the concept of the distribution of the Wald above – without some reference to the stochastic volatility implied by the model the "closeness" of the

[24] The authors take pains to point out that not every practitioner of the puzzles methodology adopts all five of these tendencies.

data moments to those implied by a particular model is not meaningful. Thus, we agree that model evaluation by indirect inference is superior to mere moment matching.

Still, one could reasonably ask why one should bother with indirect inference when likelihood ratio tests are a well-established means of econometric model evaluation. Firstly, many models have likelihoods which are difficult to model. Indirect inference requires only that the model of interest may be simulated, so that for some models the latter may be much easier to implement than the former. However, many models in current use including the popular Dynamic Stochastic General Equilibrium models of macroeconomics and the partial equilibrium models used in this thesis have representations as VAR or VARMA time series models, so this argument does not apply. Instead, Le, Meenagh, Minford and Wickens (2012) show that model evaluation by maximum likelihood (they use the term 'direct method') and by indirect inference evaluation are based on fundamentally different criteria. The former is a test of the model's in-sample predictive power (p13), which is not the same as the testing whether the moments of interest in the data (as represented by the parameters of the auxiliary model) could have been generated by the model.[25] If indirect inference had a lower power than the direct method of inference then this would weigh in favour the latter – however the authors provide evidence that this is not the case. In fact, they use a Monte-Carlo analysis to show that indirect inference has a higher rejection rate among subtly mis-specified models than direct inference (p13) (although both methods of evaluation have good power properties for more dramatic mis-specification – in the sense that as the null model diverges from the true model the rejection rates of both tests tend to one-hundred percent.) Furthermore, they show that in contrast to direct inference, indirect inference has the feature that one can increase the power of the test by adding more parameters to the auxiliary model (pp17-18). In other words, the higher the dimension of the joint distribution of simulated auxiliary model parameters, the harder the test that the

[25] In fact, Le, Meenagh, Minford and Wickens (2012, pp14-15) show using Monte-Carlo simulations that the Wald statistics arising from Indirect Inference testing and Likelihood Ratio statistics arising from Likelihood tests on a given model are essentially uncorrelated for a given model. That is, the model's ability to reproduce moments of a chosen auxiliary model is entirely distinct from its forecasting ability.

model faces under indirect inference.[26]

We have outlined some broad justifications for indirect inference from the literature. The question remains – why use indirect inference for the *particular* purpose of this paper - re-evaluating the Mortensen Pissarides model and one of its extensions?

The literature reviewed in Chapter 1 that evaluates the original Mortensen Pissarides model and that attempts to solve the Shimer puzzle adopts the puzzles methodology. Of the tendencies outlined above, (i), (iv) and (v) are especially prevalent in this literature. Shocks are generally assumed to be shocks to productivity, and sometimes to the separations rate. This corresponds to tendency (i). Furthermore, shocks are sometimes excluded for producing data moments which differ widely from the data.[27] The entire Shimer puzzle itself is conceived around points (iv) and (v). The discrepancy between the standard deviation of various labour market variable moments (such as vacancies, unemployment and the vacancy unemployment ratio) and that of labour productivity is established by a series of moment-by-moment comparisons. The joint distribution of the moments of interest are typically not considered.

In this thesis, we take it for granted that matching the moments is a desirable means of model evaluation for models of this type. In other words, we are fundamentally interested in the moments of the model (and their distribution) rather than the model's property as an in-sample predictor. Based on the discussion in Le, Meenagh, Minford and Wickens (2012), this obviates the use of the direct inference method. Indirect inference provides us with a *statistical* inference procedure based on the moments that are of concern to search theorists.

[26] Note that an issue that indirect inference alone does not resolve occurs where the model of interest may is observationally equivalent to another model. In fact, this problem may be may be exacerbated by the use of linear structural models using aggregate variables –many primitive models may result in the same structural form. Note however, that our models concerns a set of variables which are fairly specific to search theoretic studies of the aggregate labour market. Therefore we do not consider the issue to be a severe problem in the material presented in this thesis.

[27] See Shimer (2005), who abandons the separation rate because its inclusion violates the negative correlation between the vacancy and unemployment rates observed in the data.

3.2 A discrete-time Mortensen-Pissarides model

We turn now to the version of the version of the model that we wish to test and estimate. The model is a simplified version of that used by Yashiv (2006).[28] The model is in discrete time, in contrast to most of the search and matching literature. Expectations are assumed to be static. In Chapter 4 we perform the same analysis on the full version of the Yashiv model, which uses rational expectations, a more complicated vacancy cost function, and a modified Nash bargaining condition.

This version of the model starts from the premise of a representative, profit-maximizing firm that must choose a quantity of labour to operate a fixed capital stock. The firm cannot simply employ as much labour as it likes directly however, it must open vacancies to attract unemployed workers. Hiring takes place via the matching function, which implies that it depends both on the stock of vacancies maintained by the firm and on the level of unemployment. The firm pays a cost per vacancy maintained, which is also proportional to average labour productivity. As in the standard Mortensen-Pissarides framework, real wages are determined by Nash bargaining. Also consistent with the original Mortensen Pissarides model, unemployment is modeled as steady-state unemployment, which balances exogenous separations and endogenous hires. The model seeks to determine either the vacancy-unemployment ratio or the job-finding probability, real wages (or compensation paid to labour) and the unemployment rate. These are the endogenous variables. The exogenous variables are labour productivity, here proxied by GDP per worker, and the separation rate into unemployment.

[28] At this point it is useful to clarify the relationship between the model of this chapter, and that of the Chapter 4, especially in terms of what they owe to Professor Yashiv's work. Chapter 4 contains a version of the model in Yashiv's (2006). The model is an amended version of the standard Mortensen-Pissarides model, which has been converted to discrete time. The model's equations are explicitly derived from a model of a profit maximizing firm treating the decision to hire labour as an investment decision, as well as the standard search-theoretic Bellman equations. The model differs from the canonical model in that it makes use of a non-standard search cost function, allows wages to respond to employment, and has a non-standard choice of variables. It is also a log-linearized model. The model used here in Chapter 3 is essentially the Chapter 4 model, but with all of the non-standard assumptions and variable choices removed. This brings the model back to the canonical Mortensen Pissarides model with exogenous separation rate shocks as well as the usual productivity shocks. We therefore refer to the model in this chapter as the "Mortensen-Pissarides model". The presentation does still owe something to Yashiv however – since we follow his method of framing the analysis as an investment-in-labour decision for the firm.

The full derivation of this model is written in Appendix 1 to this chapter. Here we summarize the model's main parts. The representative firm maximizes the following stream of discounted expected future profits, with respect to N_{t+1}[29] and V_t:

$$\max \ E_0 \sum_{t=1}^{\infty} \left(\prod_{j=0}^{t} \beta_j \right) [F_t - W_t N_t - \Gamma_t(V_t)] \qquad (3.1)(a)$$

Subject to:

$$N_{t+1} - (1 - \delta_{t,t+1}) N_t - Q_{t,t+1} V_t = 0 \qquad (3.1)(b)$$

In this chapter's version of the model, the search cost function is given by:

$$\Gamma_t(V_t) = c V_t \left(\frac{F_t}{N_t} \right) \qquad (3.1)(c)$$

Here F_t is period t output, $W_t$is the real wage and N_t is the number of workers employed. $\Gamma_t(V_t)$ is the level of search costs, specified as a function of the stock of vacancies posted by the firm, V_t. (3.1)(c) shows that this total cost is proportional to the firm's output-per-worker, $\frac{F_t}{N_t}$, and a constant of proportionality c, as well as the number of vacancies themselves. (3.1)(b) is the equation for the evolution of employment, which depends positively upon those that were employed last period that did *not* separate into unemployment $(1 - \delta_{t,t+1})N_t$, and negatively upon the number of vacancies that are become matches during period t, $Q_{t,t+1}V_t$. (See below for a more precise definition of $Q_{t,t+1}$).

As is standard in the search literature, four Bellman equations give the values of the two possible states that the firm's job positions can take (that is, vacancies or filled jobs), and of the two labour force states that workers may occupy. Workers may be unemployed or employed, and are otherwise homogenous. They enjoy utility from either state, but would rather be employed since it affords them more utility. (See Pissarides (2000).) J_t^F represents the value to the firm of a filled job, and J_t^V the value of a vacancy. J_t^U is the value to a worker of being unemployed, and J_t^W is the value of employment. The four Bellman equations can then be written:

[29] Note that N_t is predetermined.

$$J_t^F = \frac{\partial F_t}{\partial N_t} - W_t + E_t\beta_t\left[\delta_{t,t+1}J_{t+1}^V + \left(1-\delta_{t,t+1}\right)J_{t+1}^F\right] \tag{3.2)(a}$$

$$J_t^V = -\frac{\partial \Gamma_t}{\partial V_t} + E_t\beta_t\left[Q_{t,t+1}J_{t+1}^F + \left(1-Q_{t,t+1}\right)J_{t+1}^V\right] \tag{3.2)(b}$$

$$J_t^W = W_t + E_t\beta_t\left[\left(1-\delta_{t,t+1}\right)J_{t+1}^W + \delta_{t,t+1}J_{t+1}^U\right] \tag{3.2)(c}$$

$$J_t^U = b_t + E_t\beta_t\left[P_{t,t+1}J_{t+1}^W + \left(1-P_{t,t+1}\right)J_{t+1}^U\right] \tag{3.2)(d}$$

$\frac{\partial F_t}{\partial N_t}$ is the firm's marginal product of labour and $W_t$is the real wage, to be determined by Nash bargaining. $\delta_{t,t+1}$ is the separation rate of workers from employment into unemployment during period *t*. β_t is a discount factor (common both to workers and firms). In this version of the model it will be treated as a constant. It is more standard in the search and matching literature to write it as an interest rate r, so that $\beta = \frac{1}{1+r}$ or $r = \frac{1-\beta}{\beta}$.

The matching function is assumed to be Cobb-Douglas and is given by:

$$M_{t,t+1} = \mu U_t^\sigma V_t^{1-\sigma} \tag{3.3)(a}$$

The parameter σ is the elasticity of hiring with respect to unemployment, restricted here to lie between 0 and 1. μ is a scale parameter. The matching function can be more usefully written with $M_{t,t+1}$, U_t and V_t expressed in lower-case as proportions of the labour force (the sum of employment and unemployment).

$$m_{t,t+1} = \mu u_t^\sigma v_t^{1-\sigma} \tag{3.3)(b}$$

The variables $Q_{t,t+1}$ and $P_{t,t+1}$ which both appear in the Bellman equations are the job-filling rate for firms and the job-finding rate for workers, respectively. They can be derived from the matching function as follows:

$$Q_{t,t+1} = \mu\frac{u_t^\sigma v_t^{1-\sigma}}{v_t} = \mu\frac{u_t^\sigma}{v_t^\sigma} = \mu\left(\frac{u_t}{v_t}\right)^\sigma = \mu\left(\frac{v_t}{u_t}\right)^{-\sigma} = \mu{\theta_t}^{-\sigma} \tag{3.4)(a}$$

$$P_{t,t+1} = \mu \frac{u_t^{\sigma} v_t^{1-\sigma}}{u_t} = \mu \frac{u_t^{\sigma} v_t^{1-\sigma}}{u_t^{\sigma}} = \mu \left(\frac{v_t}{u_t}\right)^{1-\sigma} = \mu {\theta_t}^{1-\sigma} \qquad (3.4)(\mathrm{b})$$

θ_t is the ratio of vacancies to unemployment (or their equivalent rates, since both are expressed here as a proportion of the labour force). This is the crucial market-tightness variable, and we make it endogenous for the purposes of one of the versions of the model that we test in this chapter. Equation (3.4)(b) shows that the job finding rate $P_{t,t+1}$ is a positive monotonic transformation of θ_t for $\mu \geq 0$ and $\sigma \leq 1$.

In this chapter we estimate and test two versions of the model, using θ_t and $P_{t,t+1}$as measures of market tightness respectively. As can be seen from our data summary (below), our measure of θ_t is based on the U.S help-wanted index, an aggregate index of job-advertising. The same index has been used in previous empirical work on aggregate search and matching models, see for example Shimer (2005). However, Yashiv (2006) argues that this measure of vacancies conflates vacancies that are eventually filled by individuals outside of employment with those filled by individuals moving directly between jobs. Underlying his argument is the fact the matching function is embedded within the Mortensen-Pissarides model, which is only supposed to explain hires from outside of employment. There is a general argument for including job-to-job vacancies in this matching function since there is a possibility of them crowding-each other out, but Yashiv provides evidence that the cyclical behaviour of the two (ex-post) types of vacancies are different, leading him to conclude that they should not be treated in the same way. In other words, the Mortensen-Pissarides model does not model job-to-job flows, but Help-Wanted Index data inextricably contains job-to-job flows. Rather than using the Help-Wanted Index, and pretending that all vacancies are subsequently filled by individuals outside of employment, one should instead not use the vacancies data.
However, we would like to contrast our results to those of Shimer (2005), who uses the Help-Wanted Index. Therefore, our first version of the model uses the Help-Wanted Index to construct θ_t, and the caveat that an important variable may be mis-measured should be kept in mind. Our second version substitutes the job-finding probability of the unemployed $P_{t,t+1}$ as the measure of market tightness, using equation (3.4)(b). Note that this transformation relies on the assumption of a

constant-returns-to-scale matching function. The transformation is useful because we can construct job-finding-probability *data* without using the Help-Wanted Index: it is the ratio of the hiring rate from unemployment (which we estimate directly from matched CPS data) to the unemployment rate. Further details are given in the data section below.

We note that our approach does not entirely solve the problem, since job-to-job flows can affect the job finding probabilities of workers that transition between unemployment and employment. Our measure of $P_{t,t+1}$ is at least based upon the theoretically relevant population, which we argue is an advantage over θ_t. To resolve this issue completely, our model would need to take explicit account of job-to-job flows directly, however so far this has not been a major concern of the Shimer-puzzle literature and we have not yet extended our results in this direction.

The first equation of the structural model is given by:

$$(1-\alpha)\frac{F_t}{N_t} - W_t = \frac{c}{\mu}\left[r + \delta_{t,t+1}\right]\frac{F_t}{N_t}\theta_t{}^{\sigma} \qquad (3.5)(a)$$

Equation (3.5)(a) modifies the standard neoclassical result for a perfectly competitive firm, that the marginal product of labour equals the real wage. A Cobb-Douglas production function is assumed, such that $F(N_t, L_t) = K_t^{\alpha} L_t^{1-\alpha}$. This implies that the marginal product of labour is given by $(1-\alpha)\frac{F_t}{N_t}$. The term on the right hand side represents the cost of search to the firm. It is increasing in the search cost parameter c and in the vacancy-unemployment ratio θ_t. The real interest rate r appears as a measure of the opportunity cost to the firm of devoting resources to search activity. The term also contains the separation rate $\delta_{t,t+1}$, since a higher separation rate serves to increase the amount of search the firm needs to take in order to maintain its desired level of employment. The equation is derived in Appendix A.3.1.1

Real wages are assumed to be set by Nash bargaining. The result is equation (3.5)(b). The equation serves to allocate the surplus from matching between firms and workers. Details of the derivation are in Appendix A.3.1.2.

$$W_t = (1-\xi)b + \xi \frac{F_t}{N_t}\big((1-\alpha) + c\theta_t\big) \qquad (3.5)(\mathrm{b})$$

The final equation of the model is an expression for steady-state unemployment.

$$u_t = \frac{\delta_{t,t+1} + (G^L - 1)}{\left[(G^L-1) + \delta_{t,t+1} + \mu\theta_t^{1-\sigma}\right]} \qquad (3.5)(c)$$

Expressing these in terms of the job-finding probability for workers, rather than market tightness gives:

$$(1-\alpha)\frac{F_t}{N_t} - W_t = \frac{c}{\mu}\left[r + \delta_{t,t+1}\right]\frac{F_t}{N_t}\mu^{\frac{\sigma}{\sigma-1}}\left(P_{t,t+1}\right)^{\frac{\sigma}{1-\sigma}} \qquad (3.6)(\mathrm{a})$$

$$W_t = (1-\xi)b + \xi\frac{F_t}{N_t}\left((1-\alpha) + c\mu^{\frac{1}{\sigma-1}}\left(P_{t,t+1}\right)^{\frac{1}{1-\sigma}}\right) \qquad (3.6)(\mathrm{b})$$

$$u_t = \frac{\delta_{t,t+1} + (G^L - 1)}{\left[(G^L-1) + \delta_{t,t+1} + P_{t,t+1}\right]} \qquad (3.6)(\mathrm{c})$$

Following Shimer (2005(a)), we log-linearize the variables around their HP-filtered trend. We use a smoothing parameter equal to 1600 – this is the standard for quarterly data (Cogley (2006), p6), rather than Shimer's value of 10^5. We adopt the lower-case ϱ_t as a short-hand for the average productivity of labour F_t/N_t. We use the notation $\tilde{x}_t$ to denote the log deviation of a variable x_t from its steady state value x.

The log-linearized equations in terms of $\tilde{\theta}_t$ are given by:

$$\left[\frac{c}{\mu}\varrho\theta^{\sigma}(r+\delta)\sigma\right]\tilde{\theta}_t + [W]\widetilde{W}_t \qquad (3.7)(\mathrm{a})$$

$$= \left[(1-\alpha)\varrho - \frac{c}{\mu}\varrho\theta^{\sigma}(r+\delta)\right]\tilde{\varrho}_t - \left[\frac{c}{\mu}\varrho\theta^{\sigma}\delta\right]\tilde{\delta}_{t,t+1}$$

$$-[\xi\varrho c\theta]\tilde{\theta}_t + [W]\widetilde{W}_t = \left[\xi\varrho\big((1-\alpha)+\theta c\big)\right]\tilde{\varrho}_t \tag{3.7)(b}$$

$$\left[\frac{(1-\sigma)\mu\theta^{1-\sigma}}{(G^L-1)+\delta+\mu\theta^{1-\sigma}}\right]\tilde{\theta}_t + \tilde{u}_t \approx \left[\frac{\delta\mu\theta^{1-\sigma}}{\big(\delta+(G^L-1)\big)\big((G^L-1)+\delta+\mu\theta^{1-\sigma}\big)}\right]\tilde{\delta}_{t,t+1} \tag{3.7)(c}$$

In terms of $\tilde{P}_{t,t+1}$ they are given by:

$$\left[c\mu^{\frac{1}{\sigma-1}}\varrho P^{\frac{\sigma}{1-\sigma}}(r+\delta)\frac{\sigma}{1-\sigma}\right]\tilde{P}_{t,t+1} + W\widetilde{W}_t = \left[(1-\alpha)\varrho - c\mu^{\frac{1}{\sigma-1}}\varrho P^{\frac{\sigma}{1-\sigma}}(r+\delta)\right]\tilde{\varrho}_t - \left[c\mu^{\frac{1}{\sigma-1}}\varrho P^{\frac{\sigma}{1-\sigma}}\delta\right]\tilde{\delta}_{t,t+1} \tag{3.8)(a}$$

$$-\left[\xi\varrho c\mu^{\frac{1}{\sigma-1}}P^{\frac{1}{1-\sigma}}\frac{1}{1-\sigma}\right]\tilde{P}_{t,t+1} + W\widetilde{W}_t = \left[\xi\varrho(1-\alpha)+\xi\varrho c\mu^{\frac{1}{\sigma-1}}P^{\frac{1}{1-\sigma}}\right]\tilde{\varrho}_t \tag{3.8)(b}$$

$$\tilde{P}_{t,t+1}\left[\frac{P}{(G^L-1)+\delta+P}\right] + \tilde{u}_t = \tilde{\delta}_{t,t+1}\left[\frac{\delta P}{\big(\delta+(G^L-1)\big)(\delta+(G^L-1)+P\ )}\right] \tag{3.8)(c}$$

3.3 Data

Table 3.1 contains details of the variables used in our analysis of the model, in terms of both their construction and their sources. The unemployment rate, the employment rate, and the labour force are the standard BLS[30] series used for aggregate labour market analysis of U.S data - Yashiv (2006) uses the same data for these variables. We also follow Yashiv in constructing the job-finding-probability as the hiring rate divided by the unemployment rate, and in using the Index of Real Compensation as the measure of U.S. real wages. We differ from Yashiv somewhat in the construction of gross flows data, meaning the quarterly flows of matching and separations and their associated rates. The underlying data is from the same monthly Current Population Survey, from which individuals must be matched in consecutive months in order to estimate the number of labour market transitions. The monthly estimates can then be time-aggregated to produce quarterly figures. The flows that Yashiv derives from this data are matched by Bleakly et. al ((1999), see Yashiv (2006), p934). For reasons of access, ours are matched using a set of Stata programs written by Robert Shimer, which are available at his website.[31] The variables that are at least partially based on matched CPS data are $\delta_{t,t+1}, m_{t,t+1}$ and $P_{t,t+1}$.

We follow Shimer (2005) in using output per worker as a measure of average labour productivity, although our variable definitions are not precisely the same as Shimer's. The use of Conference Board's Help-Wanted-Index as a proxy for vacancies is widespread in the literature (see, for example Shimer (2005), Blanchard and Diamond (1989)), when estimates are required for any period before 2002.

[30] BLS stands for" Bureau of Labor statistics".

[31] See https://sites.google.com/site/robertshimer/research/flows, under the section "Gross Worker Flows" for Shimer's Stata programs. We downloaded the monthly CPS files from http://www.nber.org/data/cps_basic.html.

Table 3.1: Data definitions and sources.

Variable	*Variable definition in model*	*Source*
Endogenous variables:		
θ_t	Vacancy-unemployment ratio. $\theta_t = \frac{v_t}{u_t} = \frac{V_t}{U_t}$	Vacancy data calculated by Wouter Den Haan. For vacancies he uses the Conference Board's index of newspaper advertising, the Help-Wanted Index. This is a seasonally adjusted monthly series, which Den Haan converts into a quarterly series by averaging. To keep the ratio consistent, he does the same averaging procedure with seasonally adjusted unemployment. The resulting series is a quarterly series for market tightness. We downloaded it from his website http://www.wouterdenhaan.com/data.htm. The spreadsheet is the one that accompanies his paper Anticipated Growth and Business Cycles in Matching Models (2009).
$P_{t,t+1}$	Job-finding probability among workers. $P_{t,t+1} = \frac{m_{t,t+1}}{u_t}$	The job finding-probability for workers is calculated as the ratio of our series for the hiring rate $m_{t,t+1}$ to our series for the unemployment rate u_t.
W_t	Real wages paid to workers.	We create an index of real compensation by dividing the quarterly seasonally adjusted BEA series COE (National Income: Compensation of Employees: Paid), by the BEA Implicit Price Deflator for GDP (series GDPDEF). Both series were downloaded from http://research.stlouisfed.org/.
u_t	Unemployment rate expressed as a proportion of the labour force.	Unemployment is BLS series LNS13000000Q divided our series for L_t. The former is quarterly seasonally adjusted U.S unemployment, and was downloaded from http://www.bls.gov/cps/. For details of L_t see below in this table.
Exogenous variables:		
ϱ_t	Average product of labour.	We created an index of the average product of labour by dividing the quarterly seasonally adjusted BEA series for real U.S GDP (series GDPC1) by quarterly seasonally adjusted U.S employment (BLS series LNS12000000Q). The former was downloaded from http://research.stlouisfed.org and the latter from http://www.bls.gov/cps/.
$\delta_{t,t+1}$	Separation probability of workers from employment into unemployment.	We create estimates of quarterly flows by matching individuals in consecutive months using the U.S Current Population survey, and observing and recording any changes in their labour market status between months. To do this we adapted code which has been made public by Robert Shimer. See the main text for details. This method gives a quarterly seasonally adjusted series of estimates for movements from employment to unemployment. To generate the separation

		probability $\delta_{t,t+1}$we divide by N_t.

Other variables (used as means or in the calculation of exogenous and endogenous variables):

$m_{t,t+1}$		We create estimates of quarterly flows by matching individuals in consecutive months using the U.S Current Population survey, and observing and recording any changes in their labour market status between months. To do this we adapted code which has been made public by Robert Shimer. This method gives a quarterly seasonally adjusted series of estimates for movements from unemployment to employment, $M_{t,t+1}$. To get the rate $m_{t,t+1}$ we divide by the labour force L_t (see below).
L_t	Labour force. Note that the unemployment rate u_t, and where required the vacancy rate v_t and matching rate m_t are all expressed as proportions of the labour force.	For the labour force we use BLS series LNS11000000Q. This is the quarterly seasonally adjusted U.S civilian labour force. The data was downloaded from http://www.bls.gov/cps/.
n_t	Employment rate expressed as a proportion of the labour force.	Employment is BLS series LNS12000000Q divided our series for L_t. The former is quarterly seasonally adjusted U.S unemployment, and was downloaded from http://www.bls.gov/cps/. For details of $L_t$see above in this table.
$G^L_{t,t+1}$	Quarterly (gross) growth rate of the labour force. $G^L_{t,t+1} = \frac{L_{t+1}}{L_t}$ Note that we only actually use the mean of this value for the sample period, so that G^L takes the role of a parameter.	Calculated as the gross quarter on quarter growth of L_t.

Table 3.2 presents summary statistics for each of the variables used in the version of model that is estimated and tested here. We use deviations from the HP- filtered trend for the model, the level means in the second column are included for illustrative purposes, and because we calibrate the steady-state values of variables (which enter the linearized version of the model, below) for the most part using sample means.

Table 3.2: Data moments: Means of levels and standard deviations of log deviations from HP(1600)-filtered trend, 1976Q1-2003Q4

Variable x_t	*Mean* $\bar{x}_t$	*Log deviation* $\tilde{x}_t = ln(x_t) - ln(x)$	*Standard deviation* $\sigma(\tilde{x}_t)$
Endogenous variables:			
θ_t	1.08	$\tilde{\theta}_t$	0.22
$P_{t,t+1}$	0.54	$\tilde{P}_{t,t+1}$	0.07
W_t	1,616	$\tilde{W}_t$	0.01
u_t	0.063	$\tilde{u}_t$	0.10
Exogenous variables:			
ϱ_t	68,836	$\tilde{\varrho}_t$	0.01
$\delta_{t,t+1}$	0.031	$\tilde{\delta}_{t,t+1}$	0.07
Other variables:			
$m_{t,t+1}$	0.033	$\tilde{m}_{t,t+1}$	0.06
n_t	0.94		
$G^L_{t,t+1}$	1.0039		

Notes: U.S. data. See table 3.1 for sources.

The fourth column of table 3.2 gives the data standard deviations of the de-trended HP filtered variables for the 1976Q1-2003Q4 sample period. The idea of using the filter is to take out the trend from the data, and hence isolate cyclical movements in the variables of interest. We are aware of the disadvantages of the procedure – the use of the filter means that the trend is essentially estimated using an a-theoretical time series method, and adds a degree of ad-hockness into the analysis. There is also the possibility of introducing spurious cyclicality into the data, under the presence of a random-walk with drift in the original series (see Cogley (2006) p8). However, in using the filter we are following much of the literature, including for example Shimer (2005(a)) and Hagerdorn and Manovksii (2008).

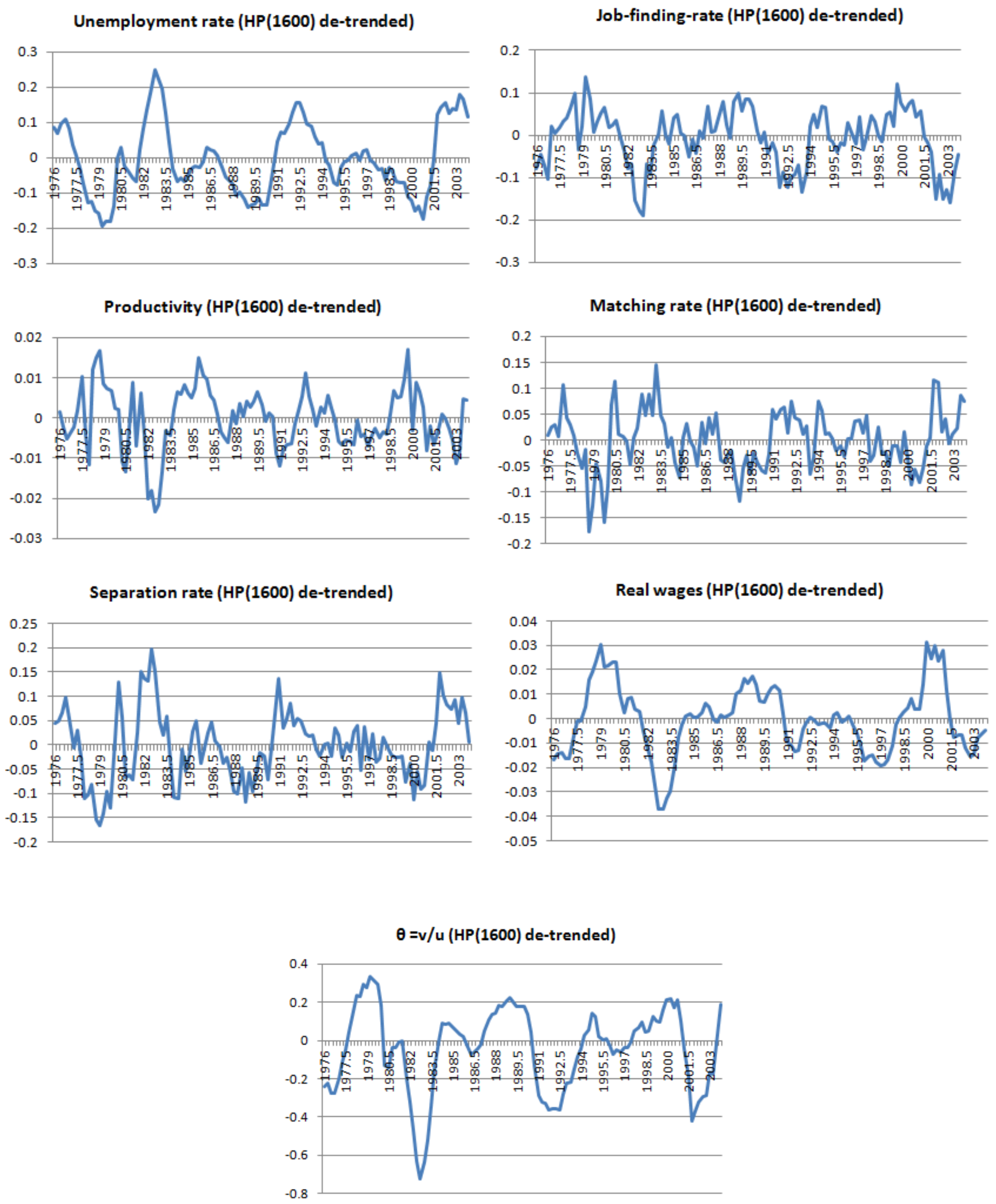

Figure 3.1 Graphs of HP(1600)-filtered variables for 1976Q1-2003Q4 sample

The standard deviations are important as they are some of the most basic stylized facts that the model needs to come close to recreating in order to fit the data. By far the most volatile variable is the vacancy-unemployment ratio, next is the unemployment rate which is less than half as volatile. The alternative measure of market tightness, the job-finding probability is less than one-third as volatile as the vacancy-unemployment ratio. Least volatile are the average productivity of labour and real wages.

An examination of the signs of the correlations in table 3.3 confirms that the de-trended variables have the cyclicality that one would expect. The vacancy unemployment ratio $\tilde{\theta}_t$ and the job-finding probability $\tilde{P}_{t,t+1}$ are both highly negatively correlated with the unemployment rate, and more modestly positively correlated with productivity (output-per-worker). Our measure of real wages is also highly pro-cyclical. As one might expect, the separation rate is positively correlated with unemployment and negatively correlated with all of the pro-cyclical variables in both samples.

The only variable which surprises in terms of correlations is the counter-cyclicality of the matching rate $\widetilde{m}_{t,t+1}$. This is however consistent with previous empirical literature, for example Yashiv (2006, p925) argues that this is because during recessions, the effect of the increase in unemployment tends to outweigh the effect of the fall in the job-finding probability. Note that the matching rate $m_{t,t+1}$does not play a direct role in the either version of the model presented here, we include information on its cyclical properties because we use it to construct the job-finding probability $P_{t,t+1}$.

Table 3.3: Correlations between variables in log-deviations from HP(1600)-filtered trend, 1976Q1-2003Q4

	$\tilde{\theta}_t$	$\tilde{P}_{t,t+1}$	$\widetilde{W}_t$	$\tilde{u}_t$	$\tilde{\varrho}_t$	$\tilde{\delta}_{t,t+1}$	$\widetilde{m}_{t,t+1}$
$\tilde{\theta}_t$	1						
$\tilde{P}_{t,t+1}$	0.78	1					
$\widetilde{W}_t$	0.79	0.66	1				
$\tilde{u}_t$	-0.92	-0.85	-0.82	1			
$\tilde{\varrho}_t$	0.57	0.37	0.53	-0.47	1		
$\tilde{\delta}_{t,t+1}$	-0.80	-0.59	-0.65	0.77	-0.56	1	
$\widetilde{m}_{t,t+1}$	-0.71	-0.31	-0.67	0.76	-0.40	0.68	1

Notes: U.S. data. See table 3.1 for sources.

Table 3.4 contains univariate AR(1) correlation coefficients for each of the relevant variables. All variables are moderately to highly persistent. The coefficients are similar in each of the two samples.

Table 3.4: AR(1) autocorrelation parameters for variables in log-deviations from HP(1600) filtered trend, 1976Q1-2003Q4

$\tilde{\theta}_t$	0.928
$\tilde{P}_{t,t+1}$	0.740
$\widetilde{W}_t$	0.919
$\tilde{u}_t$	0.933
$\tilde{\varrho}_t$	0.664
$\tilde{\delta}_{t,t+1}$	0.727
$\widetilde{m}_{t,t+1}$	0.583

Notes: U.S. data. See table 3.1 for sources.

3.4 Calibration

We have already described the method of indirect inference in detail above. Details of the procedure for implementing this model are in Appendix 2. Here in the main text we give details of the calibrated model parameters that we use.

An examination of the log-linearized equations of the model reveals that simulation of the model requires the assumption of steady state values for most of the variables. There are also three parameters, the matching function scale parameter μ, the capital-elasticity of output, α, and the real interest rate, r, the values of which we fix throughout the analysis.

We set $\alpha = 0.32$, the same value as that used by Yashiv (2006), whose model is the closest cousin of that used here. For r we use a quarterly rate of 0.012, which is the same as that used by Shimer (2005). We set μ to the value which reconciles the mean job-finding probability in the data between 1976Q1 and 2003Q4 (0.538), to the same sample counterpart for the mean vacancy unemployment ratio θ_t (1.076), when the value for matching elasticity of unemployment is set to Yashiv's (2006) baseline calibration value of 0.4. As the job finding probability is given by $P_{t,t+1} = \mu\theta_t{}^{1-\sigma}$, our calibrated figure is given by:

$$\mu = exp\big(ln(0.538) - (1 - 0.4) \times ln(1.076)\big) = 0.5150$$

The calibration and steady-state values are summarized in table 3.5.

Table 3.5: Calibrated values for indirect inference estimation and model testing

Steady state value/parameter	Calibrated value (short sample)	Reasoning/source:
θ	1.08	sample mean
P	0.54	sample mean
W	0.58p	see text
ϱ	68,836	sample mean
δ	0.0314	sample mean
m	0.033	sample mean
G^L	1.0039	sample mean
μ	0.5150	see text
α	0.32	Yashiv(2006)
r	0.012	Shimer (2005)

The steady state real wage W is calibrated as follows. The mean labour share of GDP for our sample period is 0.58. Normalizing the price level to 1 gives $0.58 = \frac{WN}{F}$ or $W = 0.58\frac{F}{N} = 0.58\varrho$. ϱ is set to the mean output per worker for the sample period, although an examination of the model equations' condition shows that this level is arbitrary; ϱ or W enter every term of the linearized job-creation conditions and wage equations, so that what matters is not the absolute level of either but their relative values.

The search cost parameter c in the model acts as a multiplier on the vacancy–employment ratio. To see this, note the firm's total per-period profit is given by:

$$F_t - W_tN_t - cV_t\left(\frac{F_t}{N_t}\right)$$

Dividing through by output gives:

$$1 - \frac{W_t N_t}{F_t} - c\left(\frac{V_t}{N_t}\right)$$

Note that $\frac{W_t N_t}{F_t}$ is the labour share s_t if the price level is normalized to 1. Dividing both total vacancies and employment by the labour force gives:

$$1 - s_t - c\left(\frac{v_t}{n_t}\right)$$

The mean value of the employment rate n_t for our sample is 0.94. The mean vacancy rate can be inferred from the implied mean vacancy-unemployment ratio which is 0.0680.

$$v = \theta u = 1.08 \times 0.063 = 0.068$$

Shimer (2005) uses a value of vacancy costs which is equivalent to 0.213 of output per worker. As the Shimer model is like the model presented here, based on a search cost function that is linear in vacancies, in terms of our model this implies that:

$$c = 0.213$$

Based on our mean values of the vacancy and employment rates, this implies that vacancies are a very small share of output, around 1.5%, since:

$$0.213 \times \left(\frac{v}{n}\right) = 0.213 \times \left(\frac{0.068}{0.94}\right) = 0.0154$$

Below where we estimate the model, we impose the upper bound on c as being 7, in order to help the estimation algorithm by limiting the area over which it needs to search. Our reasoning is that a value of c greater than 7 would be implausible. Based on a mean vacancy rate of 0.068 this would imply a total vacancy cost equal to over half of output, a magnitude that seems too large:

$$7 \times \left(\frac{v}{n}\right) = 7 \times \left(\frac{0.068}{0.94}\right) = 0.5064$$

We set the number of bootstrap samples that go into the calculation of each Wald at 2000. We chose this value based on a rough trade-off between statistical efficiency and computing time. The initial number of model replications was 1000, but we found that doubling this figure produced a roughly 60% reduction in the Wald's statistic's volatility for an acceptable increase in time taken. Computational efficiency gains dropped off sharply after that.

In our sample of data which runs from 1976Q1 to 2003Q4 we have 112 quarterly observations. We therefore set the simulated samples of data to be the same length.

3.5 Model testing using Indirect Inference with Shimer's (2005) parameters

Table 3.6 shows that under Shimer's parameters, our version of the model fails to fit the data either with an auxiliary model based either on volatilities of VAR coefficients, when both θ_t and $P_{t,t+1}$ are used as measures of market tightness. Tables 3.7 to 3.9 reveal why.

Table 3.6: Model evaluation results for Shimer parameterisation 1976Q1-2003Q4.

Estimation criterion and Model:	*Wald*	*t-stat.*	*P-value*
(A) Standard deviations, $\tilde{\theta}_t$	*** 24.875	4.17	0.0005
(B) Standard deviations, $\tilde{P}_{t,t+1}$	*** 29.502	4.71	0.0000
(C) VAR coefficients, $\tilde{\theta}_t$	*** 51.003	4.48	0.0005
(D) VAR coefficients, $\tilde{P}_{t,t+1}$	*** 31.172	3.06	0.0025
(E) Standard deviations & VAR coefficients, $\tilde{\theta}_t$	*** 76.275	5.55	0.0000
(F) Standard deviations & VAR coefficients, $\tilde{P}_{t,t+1}$	*** 54.278	4.36	0.0000

Notes: Model distribution data is generated by 2000 replications of each model using Shimer's (2005) parameters: *c=0.213, σ=0.72, ξ=0.72.* *** Indicates statistically significant at the 1% level. The column labelled 't-stat' is a t-statistic which may be compared to a critical value of 1.645. This corresponds to a 5% level of significance.

Table 3.7 shows the individual standard deviations of the endogenous variables in the data. It also shows various characteristics of the distribution of standard deviations for the same variables across 2000 simulations of the model under Shimer's parameters. The column entitled "In or Out?" says whether the standard deviation in the data lies inside or outside the 5%-95% range of simulated standard deviations from the model. Model (A) is the version of the model with the vacancy-unemployment ratio as its measure of market tightness, and model (B) is the version with the job-finding-probability as its measure of market tightness. The table shows that the main points of failure in terms of volatilities are wages and the unemployment rate. Simulated wages and unemployment are both insufficiently volatile relative to the data, for both models. Conversely, in terms of the volatilities of market tightness, both models perform reasonably well, in the sense that the data is within the 95% range of model-generated outcomes. Table 3.6 shows that neither model (A) nor (B) fit the data overall under Shimer's parameters.

Table 3.7: Distribution of auxiliary model parameters (endogenous variable volatilities) under Shimer's (2005) parameters. 1976Q1-2003Q4.

Model:	Parameter:	Data:	In or out?	Model distribution:			
				Mean:	5%	50%	95%
(A)	$std(\tilde{\theta}_t)$	0.219	IN	0.171	0.118	0.166	0.239
	$std(\widetilde{W}_t)$	0.0141	OUT	0.0108	0.0083	0.0107	0.0138
	$std(\tilde{u}_t)$	0.101	OUT	0.063	0.047	0.062	0.084
(B)	$std(\tilde{P}_{t,t+1})$	0.069	IN	0.063	0.051	0.062	0.075
	$std(\widetilde{W}_t)$	0.0141	OUT	0.0099	0.0076	0.0098	0.0126
	$std(\tilde{u}_t)$	0.101	OUT	0.062	0.050	0.062	0.076

Notes: $std(\tilde{x}_t)$ refers to the standard deviation of variable $\tilde{x}_t$. Model distribution data is generated by 2000 replications of each model using Shimer's (2005) parameters: *c=0.213, σ=0.72, ξ=0.72.*

Table 3.8 displays similar results for the dynamics of the model under Shimer's parameters. The auxiliary model is a VAR(1) on the three endogenous variables only. Here and in the remainder of this chapter we use the notation "$\mathrm{var}(Y_1, Y_2)$" to refer to the estimate of VAR(1) coefficient that associates variable Y_1 with the lag of variable Y_2. Model (C) is the version of the model the vacancy-unemployment ratio as its measure of market tightness, and model (D) is the version with the job-finding-probability as its measure of market tightness. Note that the model with θ_t as a measure of market tightness and the model with $P_{t,t+1}$ as a measure of market tightness naturally have different VAR coefficients in the data. The largest differences between the coefficients are (as one would probably expect), between the VAR coefficients for the dynamics of the market tightness variable, but there are also considerable differences between the estimated coefficients for unemployment on lagged wages and between the estimated coefficients for the joint unemployment persistence.

In terms of joint persistence, both versions of the model are able to fit the data on wages, in the sense that VAR coefficient of wages on its lag value from the VAR estimated on the actual data in both cases falls within each model's 95% bounds. In addition, model (C) fits the joint persistence of unemployment. Neither model captures the joint persistence of market tightness. Model (D) fits all of the VAR coefficients for real wages and unemployment, with the exception of unemployment's joint persistence. Model (C) fails to fit the VAR coefficients for real wages and for unemployment with respect to lagged market tightness Table 3.6 shows that models (C) and (D) are both rejected overall, although model (D) with the job-finding probability as its measure of market tightness fits marginally better.

Table 3.8. Distribution of auxiliary model parameters (VAR coefficients) under Shimer's (2005) parameters. 1976Q1-2003Q4.

Model:	Parameter:	Data:	In or out?	Model distribution:			
				Mean:	5%	50%	95%
(C)	$\mathrm{var}(\tilde{\theta}_t, \tilde{\theta}_t)$	1.317	OUT	0.886	0.717	0.890	1.038
	$\mathrm{var}(\tilde{\theta}_t, \widetilde{W}_t)$	-1.097	IN	-0.321	-1.715	-0.327	1.012
	$\mathrm{var}(\tilde{\theta}_t, \tilde{u}_t)$	0.819	OUT	-0.003	-0.424	-0.010	0.454
	$\mathrm{var}(\widetilde{W}_t, \tilde{\theta}_t)$	0.017	OUT	-0.005	-0.019	-0.004	0.00874
	$\mathrm{var}(\widetilde{W}_t, \widetilde{W}_t)$	0.778	IN	0.800	0.675	0.807	0.900
	$\mathrm{var}(\widetilde{W}_t, \tilde{u}_t)$	0.012	IN	0.004	-0.032	0.004	0.041
	$\mathrm{var}(\tilde{u}_t, \tilde{\theta}_t)$	-0.149	OUT	-0.043	-0.119	-0.042	0.029
	$\mathrm{var}(\tilde{u}_t, \widetilde{W}_t)$	0.461	IN	0.112	-0.498	0.113	0.711
	$\mathrm{var}(\tilde{u}_t, \tilde{u}_t)$	0.687	IN	0.725	0.506	0.735	0.918
(D)	$\mathrm{var}(\tilde{P}_t, \tilde{P}_t)$	0.300	OUT	0.680	0.521	0.686	0.826
	$\mathrm{var}(\tilde{P}_t, \widetilde{W}_t)$	-0.253	IN	0.116	-0.819	0.119	1.022
	$\mathrm{var}(\tilde{P}_t, \tilde{u}_t)$	-0.387	OUT	0.013	-0.159	0.013	0.187
	$\mathrm{var}(\widetilde{W}_t, \tilde{P}_t)$	0.0162	IN	0.009	-0.013	0.009	0.032
	$\mathrm{var}(\widetilde{W}_t, \widetilde{W}_t)$	0.819	IN	0.768	0.629	0.775	0.880
	$\mathrm{var}(\widetilde{W}_t, \tilde{u}_t)$	-0.008	IN	0.004	-0.021	0.004	0.028
	$\mathrm{var}(\tilde{u}_t, \tilde{P}_t)$	-0.165	IN	-0.148	-0.286	-0.148	-0.011
	$\mathrm{var}(\tilde{u}_t, \widetilde{W}_t)$	0.095	IN	-0.047	-0.876	-0.057	0.768
	$\mathrm{var}(\tilde{u}_t, \tilde{u}_t)$	0.849	OUT	0.670	0.504	0.673	0.821

Notes: $\mathrm{var}(\tilde{y}_t, \tilde{x}_t)$ refers to the VAR(1) coefficient that associates the variable $\tilde{y}_t$ with the first lag of $\tilde{x}_t$. Model distribution data is generated by 2000 replications of each model using Shimer's (2005) parameters:
c=0.213, σ=0.72, ξ=0.72.

Finally, table 3.9 displays results for a version of the model that again uses the equivalent of Shimer's parameters, but which also combines volatilities and VAR coefficients in the Wald criterion. This is not really necessary as the results are essentially the same. Model (E) uses the vacancy-unemployment ratio as its measure of market tightness and model (F) uses the job-finding probability as its measure of market tightness. Table 3.9 shows that, coefficient for coefficient, the version of the model with $P_{t,t+1}$ is superior, with the exception of the joint persistence of unemployment. This is reflected in the somewhat smaller Wald coefficient for model (F) compared to that of model (E) (see table 3.6), although both models are rejected overall.

Table 3.9: Distribution of auxiliary model parameters (endogenous variable volatilities and VAR coefficients) under Shimer's (2005) parameters. 1976Q1-2003Q4.

Model:	Parameter:	Data:	In or out?	Model distribution: Mean:	5%	50%	95%
(E)	$std(\tilde{\theta}_t)$	0.219	IN	0.171	0.119	0.166	0.235
	$std(\widetilde{W}_t)$	0.0141	OUT	0.011	0.008	0.011	0.01404
	$std(\tilde{u}_t)$	0.101	OUT	0.063	0.047	0.062	0.083
	$\mathrm{var}(\tilde{\theta}_t, \tilde{\theta}_t)$	1.317	OUT	0.889	0.711	0.898	1.037
	$\mathrm{var}(\tilde{\theta}_t, \widetilde{W}_t)$	-1.097	IN	-0.289	-1.660	-0.276	1.028
	$\mathrm{var}(\tilde{\theta}_t, \tilde{u}_t)$	0.819	OUT	0.010	-0.415	0.004	0.455
	$\mathrm{var}(\widetilde{W}_t, \tilde{\theta}_t)$	0.017	OUT	-0.004	-0.018	-0.004	0.009001
	$\mathrm{var}(\widetilde{W}_t, \widetilde{W}_t)$	0.778	IN	0.800	0.671	0.807	0.901
	$\mathrm{var}(\widetilde{W}_t, \tilde{u}_t)$	0.012	IN	0.006	-0.031	0.006	0.042
	$\mathrm{var}(\tilde{u}_t, \tilde{\theta}_t)$	-0.149	OUT	-0.046	-0.119	-0.045	0.026
	$\mathrm{var}(\tilde{u}_t, \widetilde{W}_t)$	0.461	IN	0.103	-0.538	0.099	0.730
	$\mathrm{var}(\tilde{u}_t, \tilde{u}_t)$	0.687	IN	0.717	0.501	0.725	0.905
(F)	$std(\tilde{P}_t)$	0.0686	IN	0.063	0.051	0.062	0.075
	$std(\widetilde{W}_t)$	0.0141	OUT	0.010	0.008	0.010	0.012
	$std(\tilde{u}_t)$	0.101	OUT	0.062	0.050	0.062	0.076
	$\mathrm{var}(\tilde{P}_t, \tilde{P}_t)$	0.300	OUT	0.680	0.519	0.688	0.818
	$\mathrm{var}(\tilde{P}_t, \widetilde{W}_t)$	-0.253	IN	0.113	-0.804	0.104	1.033
	$\mathrm{var}(\tilde{P}_t, \tilde{u}_t)$	-0.387	OUT	0.010	-0.163	0.009	0.185
	$\mathrm{var}(\widetilde{W}_t, \tilde{P}_t)$	0.016	IN	0.010	-0.010	0.010	0.030
	$\mathrm{var}(\widetilde{W}_t, \widetilde{W}_t)$	0.819	IN	0.769	0.640	0.774	0.882
	$\mathrm{var}(\widetilde{W}_t, \tilde{u}_t)$	-0.008	IN	0.004	-0.019	0.004	0.028
	$\mathrm{var}(\tilde{u}_t, \tilde{P}_t)$	-0.165	IN	-0.147	-0.291	-0.142	-0.015
	$\mathrm{var}(\tilde{u}_t, \widetilde{W}_t)$	0.095	IN	-0.066	-0.912	-0.069	0.766
	$\mathrm{var}(\tilde{u}_t, \tilde{u}_t)$	0.849	OUT	0.671	0.509	0.676	0.820

Notes: $std(\tilde{x}_t)$ refers to the standard deviation of variable $\tilde{x}_t$. $\mathrm{var}(\tilde{y}_t, \tilde{x}_t)$ refers to the VAR(1) coefficient that associates the variable $\tilde{y}_t$ with the first lag of $\tilde{x}_t$. Model distribution data is generated by 2000 replications of each model using Shimer's (2005) parameters. *c=0.213, σ=0.72, ξ=0.72.*

3.6 Indirect inference estimation of the model

In this section we report the results for estimating the model by minimizing the Wald statistic. We perform three different sets of estimation, with three different types of auxiliary models.

As discussed in section 3.1, we attempt to use the simulated annealing algorithm to minimize the Wald. This turns out not to be straightforward however. We find that setting a stopping rule of a maximum of 500 iterations per estimation sometimes results in different estimated parameters for the same model. There appears to be two problems. Firstly, the simulated annealing does not always visit the areas of the parameter space where the minimal Wald is to be found. Secondly, even when it does, there is a possibility for the particular draw of the Wald statistic to not be especially low – recall that the Wald is subject to stochastic variation because of the bootstrapping procedure.

To resolve the issue, we estimate each model several times, compare the Wald statistics and obtain a set of candidate plausible values. For each of these sets of optimal parameters, we run indirect inference *testing* on each 50 times, and take the average of the Wald statistics, in order to remove the effect of out-lying Walds based on out-lying bootstrap samples. We then search manually around the minimum points using indirect inference testing for smaller values of the Wald, averaging in the same way. In many cases we find that parameter estimates that minimize the Wald are at or near the corner solutions, as can been below. We vary the parameters values in the neighbourhood of the resulting values, to ascertain that we are close to or at the minimum.

In all, the minimization algorithm used can be described less as pure simulated annealing, than a combination of simulated annealing and manual search.

Table 3.10 shows the estimated parameter values for each version of the model and each of three Wald criterions used for the short sample.

Table 3.10: Estimation results for indirect inference, 1976Q1-2003Q4.

Estimation criterion and Model:	$c*$	$\sigma*$	$\xi*$	*Wald*	*t-stat.*	*P-value*
(A) Standard deviations, $\tilde{\theta}_t$	7.00	0.03	1.00	0.688	-0.67	0.875
(B) Standard deviations, $\tilde{P}_{t,t+1}$	7.00	0.31	0.00	3.698	0.59	0.257
(C) VAR coefficients, $\tilde{\theta}_t$	0.4	0.65	0.04	*** 35.244	3.20	0.003
(D) VAR coefficients, $\tilde{P}_{t,t+1}$	6.88	0.01	0.02	*** 24.433	2.38	0.009
(E) Standard deviations & VAR coefficients, $\tilde{\theta}_t$	4.8	0.58	0.01	*** 45.834	3.54	0.003
(F) Standard deviations & VAR coefficients, $\tilde{P}_{t,t+1}$	7.00	0.02	0.02	** 27.822	2.22	0.049

Estimation predominantly based on simulated annealing. (See main text for details). Estimation restrictions imposed are [0,0,0] (lower) and [7,1,1] upper for the model for $[\theta, W, u]'$, and [0,0,0] (lower) and [7,0.999,1] for the model for $[P, W, u]'$. The column labelled 't-stat' is a t-statistic which may be compared to a critical value of 1.645. This corresponds to a 5% level of significance.

Table 3.10 shows that the indirect inference estimation procedure is able to find parameter values for which the model is able to fit the data based on the standard deviations of the three endogenous variables (models (A) and (B)), but not based on a full set of nine VAR(1) coefficients (models (C) and (D)). Of course, when the auxiliary model combines volatilities with VAR coefficients, the presence of the VAR coefficients again creates problems for the fit of the model, hence models (E) and (F) are also fail to fit the data.

Table 3.11 shows the distribution of the model volatilities of endogenous variables, when the auxiliary model used to estimate the model are the volatilities themselves. Model (A) is the model in which market tightness is the vacancy unemployment ratio θ_t and model (B) is the one in which market tightness is given by $P_{t,t+1}$. Note that all of the data standard deviations lie within the 95% range of model outcomes in each.

The value of 7 for the estimate of the vacancy cost parameter *c* for both estimations (A) and (B) is the upper bound that we imposed on the estimation on the grounds of plausibility. Similarly, the value for the elasticity of hiring with respect to unemployment in model (A) is close to its lower estimation bound of 0. The labour bargaining-power parameter ξ is close to its upper bound at 1.

In model (B), where the measure of market tightness is the job-finding probability $P_{t,t+1}$ rather than the vacancy-unemployment ratio θ_t, the result is reversed for the labour bargaining power – it takes on a value of zero rather than one. The elasticity of unemployment with respect to hiring σ takes a value within the [0,1] interval – of 0.31. This is far below conventional estimates which tend to be in the [0.5, 0.7] range (Petrongolo and Pissarides, 2001). A value of 7 for the search cost parameter seems high – taking the implied mean vacancy rate to be 0.066 as discussed above, this would imply that vacancies costs of 0.066x7=46.2% of output per worker.

It is important to remember however that these estimated parameter values cannot be given a straightforward comparative-static interpretation solely in terms of the exogenous variables. The reason is that the parameters also enter the matrix that converts the bootstrapped structural errors into reduced form errors in the model simulations. The volatilities of the endogenous variables come both from the exogenous variables and from the bootstrapped shocks to the endogenous variables, the latter are converted to reduced form shocks in the model solution and so the results for the volatility estimation become difficult to interpret. The estimation algorithm simply fits the volatilities of the endogenous variables as closely as possible – it does not care about the source of the volatility.

Table 3.11: Distribution of auxiliary model parameters (endogenous variable volatilities) under estimated parameters.

Model:	Parameter:	Data:	In or out?	Model distribution:			
				Mean:	5%	50%	95%
(A)	$std(\tilde{\theta}_t)$	0.219	IN	0.188	0.126	0.183	0.266
	$std(\widetilde{W}_t)$	0.0141	IN	0.0140	0.0109	0.0139	0.0176
	$std(\tilde{u}_t)$	0.101	IN	0.089	0.062	0.087	0.125
(B)	$std(\tilde{P}_t)$	0.069	IN	0.068	0.057	0.068	0.082
	$std(\widetilde{W}_t)$	0.0141	IN	0.0115	0.0081	0.0113	0.0160
	$std(\tilde{u}_t)$	0.101	IN	0.089	0.070	0.088	0.110

Notes: $std(\tilde{x}_t)$ refers to the standard deviation of variable $\tilde{x}_t$. Model distribution data is generated by 2000 replications of each model using the estimated parameters as estimated by simulated annealing (see table 3.10). For model (A) these are *c*=7, σ*=0.03, ξ*=1* and for model (B) these are *c*=7, σ*=0.31, ξ*=0.00.*

Table 3.12 shows the parameter-by-parameter results for models (C) and (D). Both models manages to fit the dynamics of unemployment with respect to lagged wages, and also the joint persistence of unemployment. Otherwise, the only other VAR parameter that (C) and (D) both mange to fit is that of market tightness with respect to lagged real wages. Model (D), which has $P_{t,t+1}$as its measure of market tightness, in fact manages to fit all of the VAR parameters with respect to unemployment. It does not fit any on any of the dynamics of wages, except for with respect to the lagged job-finding probability. Model (C) meanwhile fits the dynamics of wages with respect to both unemployment and for wages' joint persistence. In terms of market tightness, model (C) fits only on lagged wages, and model (D) fits both this and the joint persistence of $P_{t,t+1}$.

Table 3.12: Distribution of auxiliary model parameters (VAR coefficients) under estimated parameters.

Model:	Parameter:	Data:	In or out?	Model distribution:			
				Mean:	5%	50%	95%
(C)	$\mathrm{var}(\tilde{\theta}_t, \tilde{\theta}_t)$	1.317	OUT	0.865	0.639	0.869	1.086
	$\mathrm{var}(\tilde{\theta}_t, \widetilde{W}_t)$	-1.097	IN	-1.184	-3.262	-1.145	0.733
	$\mathrm{var}(\tilde{\theta}_t, \tilde{u}_t)$	0.819	OUT	0.110	-0.384	0.100	0.650
	$\mathrm{var}(\widetilde{W}_t, \tilde{\theta}_t)$	0.017	OUT	0.000	-0.014	0.000	0.013
	$\mathrm{var}(\widetilde{W}_t, \widetilde{W}_t)$	0.778	IN	0.831	0.696	0.840	0.933
	$\mathrm{var}(\widetilde{W}_t, \tilde{u}_t)$	0.012	IN	0.003	-0.029	0.003	0.033
	$\mathrm{var}(\tilde{u}_t, \tilde{\theta}_t)$	-0.149	OUT	-0.028	-0.126	-0.028	0.068
	$\mathrm{var}(\tilde{u}_t, \widetilde{W}_t)$	0.461	IN	0.386	-0.484	0.361	1.348
	$\mathrm{var}(\tilde{u}_t, \tilde{u}_t)$	0.687	IN	0.720	0.482	0.728	0.934
(D)	$\mathrm{var}(\tilde{P}_t, \tilde{P}_t)$	0.300	IN	0.466	0.290	0.469	0.625
	$\mathrm{var}(\tilde{P}_t, \widetilde{W}_t)$	-0.253	IN	0.747	-0.397	0.737	1.878
	$\mathrm{var}(\tilde{P}_t, \tilde{u}_t)$	-0.387	OUT	-0.176	-0.372	-0.171	0.010
	$\mathrm{var}(\widetilde{W}_t, \tilde{P}_t)$	0.0162	IN	-0.013	-0.035	-0.013	0.009
	$\mathrm{var}(\widetilde{W}_t, \widetilde{W}_t)$	0.819	OUT	0.666	0.516	0.670	0.801
	$\mathrm{var}(\widetilde{W}_t, \tilde{u}_t)$	-0.008	OUT	-0.039	-0.064	-0.039	-0.016
	$\mathrm{var}(\tilde{u}_t, \tilde{P}_t)$	-0.165	IN	-0.013	-0.173	-0.011	0.151
	$\mathrm{var}(\tilde{u}_t, \widetilde{W}_t)$	0.095	IN	-0.773	-1.917	-0.761	0.319
	$\mathrm{var}(\tilde{u}_t, \tilde{u}_t)$	0.849	IN	0.778	0.594	0.779	0.946

Notes: $\mathrm{var}(\tilde{y}_t, \tilde{x}_t)$ refers to the VAR(1) coefficient that associates the variable $\tilde{y}_t$ with the first lag of $\tilde{x}_t$. Model distribution data is generated by 2000 replications of each model using the estimated parameters as estimated by simulated annealing (see table 3.10). For model (C) these are *c*=0.40, σ*=0.65, ξ*=0.04* and for model (D) these are *c*=6.88, σ*=0.01, ξ*=0.02.*

Table 3.13 shows parameter by parameter results for models (E) and (F). The results are naturally somewhat different to those for models (A) to (D), because the Wald criterion now includes VAR(1) coefficients and volatilities together. This implies a different set of estimated optimal parameters for each measure of market tightness. For example, if the results for model (E) and model (A) are compared, it is apparent that the inclusion of VAR coefficients in the Wald criterion implies that the model's optimal parameters no longer fit the volatility of the vacancy-unemployment ratio nor unemployment, as the model volatility is too low in both cases. On the other hand, model (E) fits all of the VAR coefficients that model (C) fits – that is the joint persistences of real wages and unemployment, the coefficients for market tightness and unemployment with respect to lagged real wages, and that of real wages with respect to unemployment.

In model (F) (compared with model (B)) it is the volatility of real wages that no longer fits the data once the VAR coefficients are added to the auxiliary model. In this case, there is insufficient volatility for model generated real wages to fit the data within the model's 95% bounds. In terms of VAR coefficients, the only results that are different for model (D) compared with model (F) is the job-finding probability, which is too persistent relative to the data in (F), the version with both volatilities and VAR coefficients.

Table 3.13: Distribution of auxiliary model parameters (standard deviations and VAR coefficients) under estimated parameters.

Model:	Parameter:	Data:	In or out?	Model distribution: Mean:	5%	50%	95%
(E)	$std(\tilde{\theta}_t)$	0.219	OUT	0.161	0.118	0.158	0.216
	$std(\widetilde{W}_t)$	0.0141	IN	0.011	0.008	0.011	0.015
	$std(\tilde{u}_t)$	0.101	OUT	0.071	0.054	0.069	0.093
	$\mathrm{var}(\tilde{\theta}_t, \tilde{\theta}_t)$	1.317	OUT	0.833	0.617	0.843	1.023
	$\mathrm{var}(\tilde{\theta}_t, \widetilde{W}_t)$	-1.097	IN	0.077	-1.640	0.070	1.844
	$\mathrm{var}(\tilde{\theta}_t, \tilde{u}_t)$	0.819	OUT	0.001	-0.437	-0.001	0.450
	$\mathrm{var}(\widetilde{W}_t, \tilde{\theta}_t)$	0.017	OUT	-0.004	-0.018	-0.003	0.010
	$\mathrm{var}(\widetilde{W}_t, \widetilde{W}_t)$	0.778	IN	0.870	0.731	0.877	0.979
	$\mathrm{var}(\widetilde{W}_t, \tilde{u}_t)$	0.012	IN	-0.002	-0.034	-0.001	0.029
	$\mathrm{var}(\tilde{u}_t, \tilde{\theta}_t)$	-0.149	OUT	-0.007	-0.104	-0.008	0.097
	$\mathrm{var}(\tilde{u}_t, \widetilde{W}_t)$	0.461	IN	-0.061	-0.918	-0.072	0.793
	$\mathrm{var}(\tilde{u}_t, \tilde{u}_t)$	0.687	IN	0.787	0.564	0.794	0.993
(F)	$std(\tilde{P}_t)$	0.0686	IN	0.065	0.053	0.065	0.080
	$std(\widetilde{W}_t)$	0.0141	OUT	0.0109	0.0083	0.0109	0.0139
	$std(\tilde{u}_t)$	0.101	IN	0.081	0.062	0.081	0.104
	$\mathrm{var}(\tilde{P}_t, \tilde{P}_t)$	0.300	OUT	0.477	0.311	0.478	0.636
	$\mathrm{var}(\tilde{P}_t, \widetilde{W}_t)$	-0.253	IN	0.683	-0.472	0.675	1.860
	$\mathrm{var}(\tilde{P}_t, \tilde{u}_t)$	-0.387	OUT	-0.191	-0.384	-0.193	0.003
	$\mathrm{var}(\widetilde{W}_t, \tilde{P}_t)$	0.016	OUT	-0.013	-0.035	-0.012	0.009
	$\mathrm{var}(\widetilde{W}_t, \widetilde{W}_t)$	0.819	OUT	0.655	0.500	0.659	0.807
	$\mathrm{var}(\widetilde{W}_t, \tilde{u}_t)$	-0.008	OUT	-0.040	-0.065	-0.040	-0.013
	$\mathrm{var}(\tilde{u}_t, \tilde{P}_t)$	-0.165	IN	-0.022	-0.188	-0.019	0.145
	$\mathrm{var}(\tilde{u}_t, \widetilde{W}_t)$	0.095	IN	-0.724	-1.911	-0.723	0.388
	$\mathrm{var}(\tilde{u}_t, \tilde{u}_t)$	0.849	IN	0.789	0.601	0.792	0.963

Notes: $std(\tilde{x}_t)$ refers to the standard deviation of variable $\tilde{x}_t$. $\mathrm{var}(\tilde{y}_t, \tilde{x}_t)$ refers to the VAR(1) coefficient that associates the variable $\tilde{y}_t$ with the first lag of $\tilde{x}_t$. Model distribution data is generated by 2000 replications of each model using the estimated parameters as estimated by simulated annealing (see table 3.10). For model (E) these are *c*=4.8, σ*=0.58 ξ*=0.01* and for model (F) these are *c*=7, σ*=0.02, ξ*=0.02.*

3.7 Discussion

It is not surprising that the model under Shimer's parameters is rejected by our tests in table 3.7. Recall from Chapter 1 that the point of Shimer's (2005) paper was to document the fact that the model implied insufficient volatility in market tightness (the vacancy-unemployment ratio θ_t, as measured by the Help-Wanted index), unemployment and vacancies relative to productivity. Are our testing results under Shimer's parameters consistent with the Shimer puzzle? Model (A) in table 3.7 would superficially suggest not. The one coefficient for which the data lies in the model bounds is the volatility of the vacancy-unemployment ratio θ_t – the mean standard deviation of the vacancy unemployment ratio generated by our model is 0.171, compared with 0.219 in the data. The maximum volatility that that Shimer's model is able to generate in θ_t for any combination of productivity and/or separation rate shocks is 0.037 (compared to a volatility of 0.382 in his sample – note that there is a discrepancy between Shimer's sample period and ours, and he also uses a difference HP-filter with a different smoothing parameter.[32]) It therefore appears that the lack of volatility is less of an issue in our model. Furthermore, when we simulate the model under Shimer's parameters, the 95% distribution of the standard deviation of real wages lies *below* rather than above the standard deviation in the data. There is however, a very important difference between our model and Shimer's that makes this comparison misleading. Our model contains an important extra source of volatility which comes from the bootstrapped shocks to the structural errors. Shimer's (2005) model only includes the responses of the endogenous variables to productivity and/or the separation rate. It is however possible for us to decompose the variances of the endogenous variables in our model, and to thus examine the amount of variation that is attributable to the exogenous variables and to the endogenous variables' shocks in each. We present the results of this exercise in the table 3.14 below, under Shimer's parameters and for both measures of market tightness.

[32] We use the standard quarterly HP-filter with a smoothing parameter of 1600. This is the standard for macroeconomic data (Cogley (2006), p6). Shimer uses a smoothing parameter of 10^5 (Shimer (2005) p27). The latter represents an "extremely low frequency trend".

Table 3.14 Variance decompositions for endogenous variables under Shimer's (2005) parameters.

	Model with θ_t as measure of market tightness			Model with $P_{t,t+1}$ as measure of market tightness		
Shares attributable to:	θ_t	W_t	u_t	$P_{t,t+1}$	W_t	u_t
exogenous	0.00093	0.65	0.86	0.00057	0.80	0.90
structural	0.97	1.38	0.81	0.995	1.41	0.67

Notes: The exogenous variance share is obtained by averaging the variance generated by productivity and the separation rate across 2000 samples, and taking the ratio of this to the total volatility across 2000 samples. The structural volatility and covariance shares are the model variation attributable to the simulated structural shocks. All results are generated under Shimer's (2005) parameters: *c=0.213, σ=0.72, ξ=0.72.* Omitted is the error share, which is necessarily negative in the case that the sum of the rows exceeds 1.

Table 3.14 shows that for both measures of market tightness, the vast majority of the volatility comes from the shocks to the endogenous variables rather than from the exogenous variables. Conversely there is still plenty of model variation in real wages that is attributable to variation in productivity and the separation rate. It therefore seems that the lack of volatility in market tightness with respect to productivity fluctuations is still present in our model – it is simply masked by the treating the structural errors as shocks and using them to generate aggregate volatility. This volatility is omitted by Shimer's (2005) stochastic simulations – this is we believe the reason for the discrepancy. It is also the case that Shimer does not compare the standard deviation of simulated wages against its counterpart in the data – his assertion that wages are too volatile relates to the model mechanism that generates volatility in market tightness. We therefore argue that our results are broadly consistent with the Shimer puzzle.

To summarize our estimation results, it is clear from table 3.10 that the estimated model can fit the data based on volatilities of the endogenous variables, but not based on dynamics (that is – VAR coefficients). It must be noted that the parameter estimates in the volatility-based estimates (models (A) and (B)) are at odds with the values conventionally used in the search and matching literature. While this is not in itself a problem, the fact that four out the six estimated parameters for models (A) and (B) are on the corners of the estimation bounds implies that the model could fit the data more closely by violating the theoretical (and in the case of the upper bound of the vacancy cost parameter, pragmatic) constraints imposed upon it.

So far in this discussion we have focussed on the results of our model in terms of the standard deviation of the endogenous variables, as most of the closely related literature has been preoccupied with the relative volatilities of market tightness, unemployment, vacancies and productivity rather than the dynamics. Fitting the data based on VAR coefficients is a stringent test and it is not surprising that Shimer's parameters in our model perform poorly. Neither is it surprising that the estimation procedure fails to find any set of parameters in the permitted parameter space that can fit the model to the data based on dynamics. The failure of propagation in the basic Mortensen-Pissarides model has been already noted in the literature, by for example Fujita and Ramey (2007).

It must be noted that our modelling of shocks is somewhat different to the original Mortensen-Pissarides (1994) model, as well as that of Shimer (2005). The original formulation of the model uses a three state Markov process as an aggregate shock. These shocks are to (in our notation) ϱ (interpreted in that model as "price", but since more commonly given the interpretation of aggregate productivity, including in Shimer (2005(a))). There is also a shock to *idiosyncratic* job productivity in the original paper, assumed to be uniform on the [-1,1] interval. The original Mortensen Pissarides model is of type (b) (in the categorisation of chapter 1). Shimer's version of the Mortensen Pissarides model is however of type (a). This means that separations are exogenous, and are not modelled as a response to changes in the idiosyncratic productivity of jobs as in the type (b) model. Aggregate shocks are modelled as a Poisson process on a grid. Shimer tries out a variety of different shock variables, including to the separation rate and to workers' bargaining power, but finds that shocks to ϱ alone allow the endogenous variables of the model to match the moments of the data better, for the reasons discussed in Chapter 1. Yashiv (2006) adds shocks to the separation rate and to the discount rate.

The point here is that the relationship between the shocks in the theoretical and the empirical model is weak. The literature has not taken a uniform stance on which shocks are included – Shimer tries a variety of specifications and uses that which results in the best fit. This may be because Shimer, and the associated Shimer puzzle literature, takes a model which was originally calibrated to quarterly data for the U.S manufacturing sector alone, for the short 17 year period 1972-1988, and applied to aggregate US data for a much longer 53 year period, 1951-2003. Later work has combines the basic search and matching model with (for example), monetary shocks (see Barnichon (2010.) It is true however that productivity shocks are included in most Mortensen-Pissarides type models.

Our model abandons the discrete-grid approach to shocks, used by Mortensen and Pissarides, Shimer and others. The reason is that our procedure requires us to bootstrap the innovations to the exogenous variables. Discretising productivity can needlessly add approximation error if there are not enough steps, therefore we find it simplest to keep productivity on the continuous scale in which it appears in the data. The same applies to the separation-rate shocks.

Productivity shocks are common to almost all Mortensen-Pissarides type models, and are an essential part of the Shimer puzzle. Separation rate shocks are more controversial (Shimer rejects their use as he finds they lead to a positive correlation between vacancies and unemployment). However, we have followed Yashiv (2006) in including them. (Apart from modelling endogenous separations, which would take us too far from the Shimer framework, we would otherwise have to keep separations constant, which we feel would be worse, especially given the results of the more recent empirical gross flows literature, which shows that separations do play a role in the evolution of unemployment, see for example, Fujita and Ramey (2009)).

Finally, we note that the structural shocks which we introduce as part of our indirect inference procedure are a departure from previous work. These shocks contain the effects of unobserved factors, which lie outside of the purview of the theoretical model. The theory therefore does not have anything to say about the expected magnitude or source of these shocks, and the fact that they are large suggests that these omitted factors are important.

In our version of the model we take non-wage income or home production, to be a parameter, as in most of the related literature. The reason for this is that it is very difficult to come up with an adequate summary measure of a highly contingent benefit system for the whole U.S economy which is suitable for use in an aggregate model. Similarly, creating a time-series for the value of home-production would hardly be straightforward. In our model, non-wage income enters the solution to the Nash-bargain wage equation, however it vanishes when we linearize the model. This is not the case in most of the search and matching literature, which does not tend to linearize its models before simulating them. The fact that our model lacks the parameter prevents us from making full comparisons to certain other models and testing their full set of parameterisations. Shimer's (2005) model includes a non-wage income parameter, so does the model of Hagedorn and Manovskii (2008). The latter stress an interaction between the level of non-labour income and bargaining power in the determination of the response of market-tightness to productivity changes in the Mortensen-Pissarides model. Obviously, our version of the model is unable to capture this effect. One solution, proposed by Pissarides (2000) and used by Yashiv (2006) would be to make the parameter proportional to wages – the parameter then becomes a replacement ratio. Again, the following chapter extends the model in this direction.

This chapter attempted to estimate and test a very simple version of the Mortensen-Pissarides model using indirect inference. One major simplification was the imposition of static expectations – this is contrary to the usual assumption in this literature (including that in Yashiv's (2006) paper) which is that expectations are rational. Another is the fact that the model used here ignores non-labour income and unemployment benefits as mentioned previously. It is clear that we need to deal with these issues before coming to definitive conclusions about the model using indirect inference. A separate issue is that the basic Mortensen-Pissarides model has *already* been seen to be an inadequate description of the labour market – from the point of view of the Shimer puzzle. As discussed in Chapter 1, a large literature has emerged to try to generate volatility closer to that seen in the data. Examples include wage-rigidity (see Shimer (2004) and Hall (2005)), and the specification of the search cost function (Yashiv (2006)). The model here addresses none of these innovations, so that even if our reproduction of the Mortensen-Pissarides model had been entirely faithful, there is a limit to what we could expect to see in terms of empirical performance. In Chapter 4 we attempt to address the problems simultaneously, by testing Yashiv's (2006) full model, again using the method of indirect inference as a tool of exploration.

Appendix 3.1: Mathematics of the model

The model starts with a representative firm, that chooses N_{t+1} and V_t to solve the following maximization problem:

$$\max \; E_0 \sum_{t=1}^{\infty} \left(\prod_{j=0}^{t} \beta_j \right) [F_t - W_t N_t - \Gamma_t(V_t)] \qquad (3.1)(a)$$

The choice is subject to the evolution of employment equation which is a binding constraint in each period:

$$N_{t+1} - \left(1 - \delta_{t,t+1}\right) N_t - Q_{t,t+1} V_t = 0 \qquad (3.1)(b)$$

The first order conditions are as follows:

$$FOC(V_t): \qquad \frac{\partial \Gamma_t(V_t,)}{\partial V_t} = c\frac{F_t}{N_t} = Q_{t,t+1}\Lambda_t$$

$$FOC(N_{t+1}): \qquad \Lambda_t = E_t \beta_t \left[\frac{\partial F_{t+1}}{\partial N_{t+1}} - W_{t+1} \right] + E_t \left(1 - \delta_{t+1,t+2}\right) \beta_t \Lambda_{t+1}$$

$$FOC(\Lambda_t): \qquad N_{t+1} - \left(1 - \delta_{t,t+1}\right) N_t - Q_{t,t+1} V_t = 0$$

A transversality condition is also necessary.

Note that in this model we treat the discount factor β_t as a parameter, so that $\beta_t = \beta$ for all t. The same goes for the discount rate r, since $\beta = \frac{1}{1+r}$.

A.3.1.1 Deriving the job-creation conditions (3.5)(a) and (3.6)(a):

Beginning using the Bellman equation for the firm's marginal value of a vacancy:

$$J_t^V = -\frac{\partial \Gamma_t}{\partial V_t} + E_t \beta_t \left[Q_{t,t+1} J_{t+1}^F + \left(1 - Q_{t,t+1}\right) J_{t+1}^V \right] \qquad (3.2)(b)$$

The model is predicated on the assumption that the firms open vacancies until the marginal value of a vacancy is driven to zero, in other words, $J_t^V = 0$ for all *t*. This results in the condition:

$$0 = -\frac{\partial \Gamma_t}{\partial V_t} + E_t \beta \left[Q_{t,t+1} J_{t+1}^F \right]$$

Furthermore, in this version of the model we use static expectations, with the implication that $E_t J_{t+1}^F = J_t^F$.

$$0 = -\frac{\partial \Gamma_t}{\partial V_t} + \beta Q_{t,t+1} J_t^F$$

This implies that the value to the firm of a marginal filled job is:

$$J_t^F = \frac{c\frac{F_t}{N_t}}{\beta Q_{t,t+1}}$$

To derive the job creation condition, we substitute this and the $J_t^V = 0$ condition into the Bellman equation that gives value of the marginal value of a filled job to the firm, which is

$$J_t^F = \frac{\partial F_t}{\partial N_t} - W_t + E_t \beta_t \left[\delta_{t,t+1} J_{t+1}^V + \left(1 - \delta_{t,t+1}\right) J_{t+1}^F\right] \qquad (3.2)(a)$$

Imposing the static expectations condition once again, and rearranging can be shown to give the result:

$$0 = \frac{\partial F_t}{\partial N_t} - W_t + \left[\frac{\beta - 1}{\beta} - \delta_{t,t+1}\right] \frac{c\frac{F_t}{N_t}}{Q_{t,t+1}}$$

Since $r = \frac{1-\beta}{\beta}$,

$$\frac{\partial F_t}{\partial N_t} - W_t - \left[r + \delta_{t,t+1}\right] \frac{c\frac{F_t}{N_t}}{Q_{t,t+1}} = 0$$

The matching function implies that the firm's job filling probability is given by $Q_{t,t+1} = \mu {\theta_t}^{-\sigma}$. Furthermore, the assumption of a Cobb-Douglas production function with a labour elasticity of output $(1 - \alpha)$means that the marginal product of labour is $(1 - \alpha)$ times the average product of labour. This gives the final form of the job creation condition, expressed in terms of the vacancy-unemployment ratio θ_t.

$$(1 - \alpha)\frac{F_t}{N_t} - W_t = \frac{c}{\mu}\left[r + \delta_{t,t+1}\right] \frac{F_t}{N_t} {\theta_t}^{\sigma} \qquad (3.5)(a)$$

Since by the matching function $P_{t,t+1} = \mu \theta_t^{1-\sigma}$, the vacancy-unemployment ratio may also be re-written in terms of the worker's job finding probability. Inverting the relationship gives $\theta_t = \mu^{\frac{1}{\sigma-1}} \left(P_{t,t+1}\right)^{\frac{1}{1-\sigma}}$, and so:

$$(1 - \alpha)\frac{F_t}{N_t} - W_t = \frac{c}{\mu}\left[r + \delta_{t,t+1}\right] \frac{F_t}{N_t} \mu^{\frac{\sigma}{\sigma-1}} \left(P_{t,t+1}\right)^{\frac{\sigma}{1-\sigma}}$$

Collecting the terms in μ gives:

$$(1 - \alpha)\frac{F_t}{N_t} - W_t = c\left[r + \delta_{t,t+1}\right] \frac{F_t}{N_t} \mu^{\frac{1}{\sigma-1}} \left(P_{t,t+1}\right)^{\frac{\sigma}{1-\sigma}} \qquad (3.6)(a)$$

This is the second form of the job finding probability that we use in the testing and estimation of the model.

A.3.1.2. Deriving the wage equation under the assumption of Nash bargaining (3.5)(b) and (3.6)(b):

The bargained wage is assumed to maximize the "Nash product". The bargaining power of the worker is given by the parameter ξ which can lie anywhere in the closed interval between zero and one. The bargaining power of the firm is given by$1-\xi$. The wage then satisfies:

$$W_t = \arg\max\ (J_t^F - J_t^V)^{1-\xi}(J_t^W - J_t^U)^{\xi}$$

The solution to the problem can be shown to satisfy:

$$J_t^F = J_t^V + (1-\xi)(J_t^F + J_t^W - J_t^V - J_t^U)$$

Since it is assumed here that vacancies are opened until their marginal value $J_t^V = 0$ at all times, this simplifies to:

$$(J_t^W - J_t^U) = \left[\frac{\xi}{1-\xi}\right] J_t^F$$

To derive the wage equation, both of the Bellman equations for workers are required, both that for the state of being employed and for the state of unemployment:

$$J_t^W = W_t + E_t\beta\left[\left(1-\delta_{t,t+1}\right)J_{t+1}^W + \delta_{t,t+1}J_{t+1}^U\right]$$

$$J_t^U = b_t + E_t\beta\left[P_{t,t+1}J_{t+1}^W + \left(1-P_{t,t+1}\right)J_{t+1}^U\right]$$

Taking the difference and simplifying gives:

$$(J_t^W - J_t^U) = W_t - b_t + E_t\beta\left(1-\delta_{t,t+1}-P_{t,t+1}\right)(J_{t+1}^W - J_{t+1}^U)$$

Substituting $\left[\frac{\xi}{1-\xi}\right] J_t^F$ for $(J_t^W - J_t^U)$ gives:

$$\left[\frac{\xi}{1-\xi}\right] J_t^F = W_t - b_t + E_t\beta\left(1-\delta_{t,t+1}-P_{t,t+1}\right)\left[\frac{\xi}{1-\xi}\right] J_{t+1}^F$$

In the place of J_t^F we substitute the firm's Bellman's equation for the marginal value of filled job, in which the condition $J_t^V = 0$ for all *t* is assumed to hold, which is given by:

$$J_t^F = \frac{\partial F_t}{\partial N_t} - W_t + E_t\beta\left[\left(1-\delta_{t,t+1}\right)J_{t+1}^F\right]$$

Substituting and rearranging gives:

$$\xi\frac{\partial F_t}{\partial N_t} = W_t - (1-\xi)b_t - P_{t,t+1}\xi E_t\beta J_{t+1}^F$$

Finally, the assumption of static expectations allows us to write:

$$E_t J_{t+1}^F = J_t^F = \frac{c\frac{F_t}{N_t}}{\beta Q_{t,t+1}}$$

So that:

$$\xi \frac{\partial F_t}{\partial N_t} = W_t - (1-\xi)b_t - P_{t,t+1}\xi \frac{c\frac{F_t}{N_t}}{Q_{t,t+1}}$$

Substituting average for marginal productivity, and solving for W_t gives:

$$W_t = (1-\xi)b_t + \xi \frac{F_t}{N_t}\left((1-\alpha) + c\frac{P_{t,t+1}}{Q_{t,t+1}}\right)$$

The matching function allows us to re-express the ratio $\frac{P_{t,t+1}}{Q_{t,t+1}}$ as:

$$\frac{P_{t,t+1}}{Q_{t,t+1}} = \frac{\mu {\theta_t}^{1-\sigma}}{\mu {\theta_t}^{-\sigma}} = \theta_t$$

$$W_t = (1-\xi)b_t + \xi \frac{F_t}{N_t}\big((1-\alpha) + c\theta_t\big) \qquad (3.5)(b)$$

Without any data on b_t we assume it to be a constant, $b_t = b$. This gives the wage equation under Nash-bargaining:

$$W_t = (1-\xi)b + \xi \frac{F_t}{N_t}\big((1-\alpha) + c\theta_t\big)$$

Alternatively it may be expressed in terms of the job finding probability $P_{t,t+1}$:

$$W_t = (1-\xi)b + \xi \frac{F_t}{N_t}\left((1-\alpha) + c\mu^{\frac{1}{\sigma-1}}\big(P_{t,t+1}\big)^{\frac{1}{1-\sigma}}\right) \qquad (3.6)(b)$$

A.3.1.3. Deriving steady state unemployment (3.5)(c) and (3.6)(c):

Under the assumption that all labour force entrants are initially unemployed, the constraint can be written in terms of unemployment:

$$U_{t+1} - U_t = \delta_{t,t+1}(L_t - U_t) - M_{t,t+1} + L_{t+1} - L_t$$

The model is however expressed in terms of rates. We use lower case to denote rates out of the labour force L_t. G_t^L denotes the gross rate of growth of the labour force. This gives:

$$u_{t+1}G_t^L - u_t = \delta_{t,t+1}(1-u_t) - m_{t,t+1} + (G_t^L - 1)$$

We assume labour force growth to be constant, which parameterises G_t^L such that $G_t^L = G^L$.

$$u_{t+1}G^L - u_t = \delta_{t,t+1}(1-u_t) - m_{t,t+1} + (G^L - 1)$$

By the matching function (see the main text) $m_{t,t+1} = P_{t,t+1}u_t$. Therefore:

$$u_{t+1}G^L - u_t = \delta_{t,t+1}(1-u_t) - P_{t,t+1}u_t + (G^L - 1)$$

Finally we assume that unemployment is always in its steady state in each period. This allows us to substitute u_t for u_{t+1}.

$$u_tG^L - u_t = \delta_{t,t+1}(1-u_t) - P_{t,t+1}u_t + (G^L - 1)$$

Under these assumptions, steady state unemployment can be written as:

$$u_t = \frac{\delta_{t,t+1} + (G^L - 1)}{\left[(G^L - 1) + \delta_{t,t+1} + P_{t,t+1}\right]} \tag{3.6)(c}$$

Alternatively, since by the matching function $P_{t,t+1} = \mu\theta_t^{1-\sigma}$, unemployment can be written in terms of market-tightness as:

$$u_t = \frac{\delta_{t,t+1} + (G^L - 1)}{\left[(G^L - 1) + \delta_{t,t+1} + \mu\theta_t^{1-\sigma}\right]} \tag{3.5)(c}$$

Appendix 3.2: Applying the method of Indirect Inference testing

There are two versions of the simple Yashiv-Mortensen-Pissarides model, the first of which has endogenous variables θ_t, W_t and u_t and the second of which contains $P_{t,t+1}$, W_t and u_t. Note that the exogenous variables, p_t and $\delta_{t,t+1}$, are the same in both versions of the model. Let the first version with θ_t be model *m1* and the second with $P_{t,t+1}$ be *m2.*

We begin by writing the linearized version of each model in structural form:

$$\mathbf{A^m y_t^m} = \mathbf{B^m x_t} + \mathbf{e_t^m}$$

$$t = 1, \dots, T$$

$$m \in \{m1, m2\}$$

Where $\mathbf{y_t^{m1}} = [\tilde{\theta}_t \quad \widetilde{W}_t \quad \tilde{u}_t]'$, $\mathbf{y_t^{m2}} = \left[\tilde{P}_{t,t+1} \quad \widetilde{W}_t \quad \tilde{u}_t\right]'$, and $\mathbf{x_t} = \left[\tilde{\varrho}_t \quad \tilde{\delta}_{t,t+1}\right]'$. The structural model matrices are given as follows:

$$\mathbf{A^{m1}} = \begin{bmatrix} \frac{c}{\mu}\varrho\theta^\sigma(r+\delta)\sigma & W & 0 \\ -\xi\varrho c\theta & W & 0 \\ \frac{(1-\sigma)\mu\theta^{1-\sigma}}{(G^L-1)+\delta+\mu\theta^{1-\sigma}} & 0 & 1 \end{bmatrix}$$

$$\mathbf{B^{m1}} = \begin{bmatrix} (1-\alpha)\varrho - \frac{c}{\mu}\varrho\theta^{\sigma}(r+\delta) & -\frac{c}{\mu}\varrho\theta^{\sigma}\delta \\ \xi\varrho\big((1-\alpha)+\theta c\big) & 0 \\ 0 & \frac{\mu\theta^{1-\sigma}}{\big(\delta+(G^L-1)\big)\big((G^L-1)+\delta+\mu\theta^{1-\sigma}\big)} \end{bmatrix}$$

And:

$$\mathbf{A^{m2}} = \begin{bmatrix} c\mu^{\frac{1}{\sigma-1}}\varrho P^{\frac{\sigma}{1-\sigma}}(r+\delta)\frac{\sigma}{1-\sigma} & W & 0 \\ -\xi\varrho c\mu^{\frac{1}{\sigma-1}}P^{\frac{1}{1-\sigma}}\frac{1}{1-\sigma} & W & 0 \\ \frac{P}{(G^L-1)+\delta+P} & 0 & 1 \end{bmatrix}$$

$$\mathbf{B^{m2}} = \begin{bmatrix} (1-\alpha)\varrho - c\mu^{\frac{1}{\sigma-1}}\varrho P^{\frac{\sigma}{1-\sigma}}(r+\delta) & -c\mu^{\frac{1}{\sigma-1}}\varrho P^{\frac{\sigma}{1-\sigma}}\delta \\ \xi\varrho(1-\alpha)+\xi\varrho c\mu^{\frac{1}{\sigma-1}}P^{\frac{1}{1-\sigma}} & 0 \\ 0 & \frac{\delta P}{\big(\delta+(G^L-1)\big)\big(\delta+(G^L-1)+P\ \big)} \end{bmatrix}$$

$\theta, P, W, \varrho, \delta, G^L$ in the model matrices above are the steady-state counterparts to the variables in the model. They are set to their mean values, which can be found in the second column of table 3.2 in the main text.

$\mathbf{e_t^m}$ is a *3*x*1* vector of structural errors. The first step in indirect inference testing is to find the structural errors, given the structural model matrices $\mathbf{A^m}$ and $\mathbf{B^m}$ and given the real world data.

$$\mathbf{e_t^m} = \mathbf{A^m y_t^m} - \mathbf{B^m\ x_t}$$

$$t = 1, \dots, T$$

The structural errors are assumed to be independent AR(1) processes. For each of the *m* structural errors of $\mathbf{e_t^m}$, univariate AR(1)s are estimated and the residuals $\hat{\varepsilon}_{k,t}^m$ determined:

$$e_{k,t}^m = \hat{\rho}_k^m e_{k,t-1}^m + \hat{\varepsilon}_{k,t}^m$$

$$k \in \{1,2,3\}$$

Here k refers to the number of structural equation in the order given above. Therefore $k = 1$ refers to the structural error for the job creation condition, $k = 2$ is the structural error for the wage equation, and $k = 3$ that for the steady-state unemployment condition.

We also estimate independent AR(1) equations on each of the two exogenous variables.

$$\tilde{\varrho}_t = \hat{\rho}_\varrho \tilde{\varrho}_{t-1} + \hat{\varepsilon}_{p,t}$$

$$\tilde{\delta}_{t,t+1} = \hat{\rho}_\delta \tilde{\delta}_{t-1,t} + \hat{\varepsilon}_{t,\delta}$$

Estimating these AR(1) regressions will give five vectors of residual innovations per model, which are grouped together ready for bootstrapping into a *Tx5* matrix as follows:

$$\mathbf{BOOT^m\_IN} = [\hat{\boldsymbol{\varepsilon}}_\mathbf{1}^\mathbf{m} \quad \hat{\boldsymbol{\varepsilon}}_\mathbf{2}^\mathbf{m} \quad \hat{\boldsymbol{\varepsilon}}_\mathbf{3}^\mathbf{m} \quad \hat{\boldsymbol{\varepsilon}}_\varrho \quad \hat{\boldsymbol{\varepsilon}}_\delta]$$

To draw a single bootstrap sample of length *T*, one simply makes *T* draws of rows of $\mathbf{BOOT^m\_IN}$ at random with replacement. Indirect inference testing in fact requires multiple samples to be drawn. we draw 2000 samples. The result can be written as 2000 *Tx5* matrices:

$$\mathbf{BOOT^m\_OUT^j} = \begin{bmatrix}\hat{\hat{\boldsymbol{\varepsilon}}}_1^{\mathbf{m,j}} & \hat{\hat{\boldsymbol{\varepsilon}}}_2^{\mathbf{m,j}} & \hat{\hat{\boldsymbol{\varepsilon}}}_3^{\mathbf{m,j}} & \hat{\hat{\boldsymbol{\varepsilon}}}_{\varrho}^{\mathbf{m,j}} & \hat{\hat{\boldsymbol{\varepsilon}}}_{\delta}^{\mathbf{m,j}}\end{bmatrix}$$

$$j = 1, \ldots, 2000$$

These 2000 samples of length *T* represent simulated innovations. As discussed in the main text, drawing a vector bootstrap in this way serves to preserve any correlation between To calculate simulated structural errors and simulated samples of exogenous variables, one uses the estimates $\hat{\rho}_1^m$, $\hat{\rho}_2^m$, $\hat{\rho}_3^m$, $\hat{\rho}_\varrho$ and $\hat{\rho}_\delta$ from the AR(1) regressions. One also requires initial values for the innovation. For these we simply use the first row of $\mathbf{BOOT^m\_OUT^j}$**.** Note that we use a “hat” to denote an estimated value, and a “double-hat” to denote a simulated value. (All of the bootstrapped residuals we count as simulated because they have been re-sampled.) The structural errors and exogenous variables can therefore be written as follows:

$$\hat{\hat{e}}_{k,t}^{m,j} = \hat{\rho}_k^m \hat{\hat{e}}_{k,t-1}^{m,j} + \hat{\hat{\varepsilon}}_{k,t}^{m,j}$$

$$k \in \{1,2,3\}$$

$$\hat{\hat{\varrho}}_t^j = \hat{\rho}_\varrho \hat{\hat{\varrho}}_{t-1}^j + \hat{\hat{\varepsilon}}_{\varrho,t}^j$$

$$\hat{\hat{\delta}}_{t,t+1}^j = \hat{\rho}_\delta \hat{\hat{\delta}}_{t-1,t}^j + \hat{\hat{\varepsilon}}_{\delta,t}$$

$$t = 1, \ldots, T$$
$$j = 1, \ldots, 2000$$
$$m \in \{m1, m2\}$$

As the structural model contains no expectations, it is simple to solve for the simulated endogenous variables, given the simulated exogenous variables and simulated structural errors. We compute 2000 samples of simulated endogenous variables as follows.

Let:

$$\hat{\hat{\mathbf{x}}}_\mathbf{t}^\mathbf{j} = \begin{bmatrix}\hat{\hat{\varrho}}_t^j \\ \hat{\hat{\delta}}_{t,t+1}^j\end{bmatrix}$$

$$\hat{\hat{\mathbf{e}}}_\mathbf{t}^\mathbf{m,j} = \begin{bmatrix}\hat{\hat{\varepsilon}}_{1,t}^{m,j} \\ \hat{\hat{\varepsilon}}_{2,t}^{m,j} \\ \hat{\hat{\varepsilon}}_{3,t}^{m,j}\end{bmatrix}$$

$$t = 1, \ldots, T$$
$$j = 1, \ldots, 1000$$
$$m \in \{m1, m2\}$$

Then the simulated endogenous variables are given by:

$$\widehat{\hat{\mathbf{y}}}_{\mathbf{t}}^{\mathbf{m,j}} = [\mathbf{A^m}]^{-1}\mathbf{B^m}\widehat{\hat{\mathbf{x}}}_{\mathbf{t}}^{\mathbf{j}} + [\mathbf{A^m}]^{-1}\widehat{\hat{\mathbf{e}}}_{\mathbf{t}}^{\mathbf{m,j}}$$

$$t = 1, \dots, T$$
$$j = 1, \dots, 2000$$
$$m \in \{m1, m2\}$$

It is useful to represent the simulated data as full-sample vectors:

$$\widehat{\hat{\mathbf{Y}}}^{\mathbf{m,j}} = \begin{bmatrix} \widehat{\hat{\mathbf{y}}}_{\mathbf{1}}^{\mathbf{m,j}\prime} \\ \dots \\ \widehat{\hat{\mathbf{y}}}_{\mathbf{T}}^{\mathbf{m,j}\prime} \end{bmatrix}$$

$$j = 1, \dots, 2000$$
$$m \in \{m1, m2\}$$

Also let:

$$\mathbf{Y^{m1}} = \begin{bmatrix} \theta_1 & W_1 & u_1 \\ \dots & \dots & \dots \\ \theta_T & W_T & u_T \end{bmatrix}$$

$$\mathbf{Y^{m2}} = \begin{bmatrix} P_{1,2} & W_1 & u_1 \\ \dots & \dots & \dots \\ P_{T,T+1} & W_T & u_T \end{bmatrix}$$

The auxiliary model of choice may then be estimated on each of the samples of $\widehat{\hat{\mathbf{Y}}}^{\mathbf{m1,j}}$ or $\widehat{\hat{\mathbf{Y}}}^{\mathbf{m2,j}}$, and on the true data $\mathbf{Y^{m1}}$ or $\mathbf{Y^{m2}}$.

Chapter 4: Testing the Yashiv (2006) extension of the Mortensen-Pissarides model using indirect inference

4.1 Introduction

Although a useful first step in showing how the method of indirect inference can be applied to aggregate search and matching models, the representation of the Mortensen Pissarides model used in Chapter 3 contained simplifications compared to those typically used in evaluation. The most important of these is the lack of dynamics in the model – meaning the use of steady-state unemployment in every period as opposed to the dynamic form of the unemployment equation. It also contained static expectations, contrary to most of the literature that uses rational expectations in aggregate search models. In this chapter we amend the analysis to deal with these issues.

Secondly, the model estimated is a relatively old-fashioned aggregate search theoretic sort of model, of the type criticised by Shimer (2005) and Costain and Reiter (2007). As was described in Chapter 1, Shimer argued that these models failed to generate realistic volatility in labour market variables, essentially because labour was unrealistically successful in absorbing gains in productivity into wages. There is now a lengthy literature that suggests amendments to these models – also described in Chapter 1.

In this chapter we apply the method of indirect inference to a model created and evaluated by Professor Eran Yashiv, in a paper entitled "Evaluating the Performance of the Search and Matching Model." The paper makes several amendments to the standard search and matching model, and Yashiv calibrates and simulates the model, and compares the simulated moments to those of U.S data. The main amendment to the standard model, which Yashiv credits with the model's superior empirical performance, is the incorporation of non-linear search costs facing the firm in vacancy creation and in hiring. Specifically, he makes search costs a cubic function of a linear combination of hires and vacancies, going against the standard assumption that search costs are linear in vacancies which was also used in Chapter 3.

The analysis of this chapter is a departure from the standard model in two important ways. Firstly, we use Yashiv's non-linear search cost specification. Secondly, we evaluate the model using indirect inference. Chapter 1 mentioned different ideas that have been suggested to resolve the Shimer puzzle, such as fixed firing costs, non-Nash bargaining. Here, in focussing on the search cost function faced by the firm, we explore just one of these ideas. The results of this augmented specification have already been explored in a *calibration* framework by Yashiv. However, we are changing the method of analysis in using indirect inference, which is a departure from the normal calibration-based method of analysis. Therefore, we argue that it makes sense to look at one just major change to the basic model at a time. We leave for further work the evaluation of alternative suggested resolutions to the Shimer puzzle under Indirect Inference.

The assumption of non-linear search costs is not arbitrary. Yashiv builds on his previous econometric work which attempts to directly estimate a structural model of search costs, albeit in a different setting (and using different techniques to those used here.)

In Yashiv (2000a) he models firm hiring for the Israeli economy as an attempt of a representative firm to maximize expected lifetime profits, by choosing a hiring level for each period. His primary aim is to use the very comprehensive Israeli Employment Service dataset on firm hiring between 1975 and 1989 in order to search for best fitting hiring-cost function for the model. The set of permissible functions are, in the notation used in Chapter 3 – (see Yashiv 2000a p495):

Table 4.1: Search cost specifications in Yashiv (2000a)

Quadratic	$\Gamma_t = \frac{\Theta}{2}\left(\frac{M_t}{N_t}\right)^2 F_t$
Generalized Power	$\Gamma_t = \frac{\Theta}{\gamma}\left(\frac{M_t}{N_t}\right)^\gamma F_t$
Polynomial	$\Gamma_t = \sum_{i=1}^{d} \frac{\Theta_i}{i}\left(\frac{M_t}{N_t}\right)^i F_t$ $d = 2,3,4$

Notes: Θ is a scale parameter, which plays a similar role to c in Chapter 3. M_t is number of period *t* hires of the firm. F_t is the firm's output level. N_t is the firm's number of employees in period *t.*

Yashiv uses General Method of Moments (GMM) to try to estimate the best cost function specification, as well as the Cobb-Douglas parameter α and the unemployment benefit parameter. With respect to hiring costs, Yashiv finds the best fit (based on the J- test of over-identifying restrictions) – with a generalised exponential search cost function of the form:

$$\Gamma_t = \frac{\Theta}{\gamma}\left(\frac{M_t}{N_t}\right)^\gamma F_t$$

So that under the best specification, search costs are modelled so as to depend on the scale parameter Θ and the exponent parameter γ. The GMM procedure precisely estimates γ and the range of estimates that are not rejected based on the J-test is [4.69, 4.94], indicating that the estimates are stable with respect to the choice of timing and instruments. However the estimates for the scale parameter Θ show a greater dispersion between specifications.

Yashiv attempts to corroborate these results by decomposing the firm's inter-temporal condition into terms related to the expected value of the marginal cost and the variance of the marginal costs of hiring, respectively. The marginal cost terms are interpreted as the asset value to the firm of a filled job, via the equilibrium condition of the model. These asset value terms are decomposed in terms of sample covariances for transformations of other variables appearing in the firm's inter-temporal condition – that is, separation rates, real interest rates and marginal profit rates. The postulated marginal cost function is the functional form selected by the J-test criterion in GMM, so again the task is to estimate Θ and γ. By choosing a range of the scale parameter, Yashiv compares the resulting exponent parameter arising when the sample moments are also fed into the expected value and variance terms of the asset. To pin down the value of Θ, he uses that which arises when γ is chosen so that both parameters are as close as possible to his acceptable GMM results. These parameters used in conjunction with the other data also allow for the calculation of mean asset values, which are not observed in the real world.

Yashiv (2000b) extends the analysis in Yashiv (2000a) in several ways. The dynamic demand for labour is again estimated using GMM, however rather than looking for the best-fitting function of the hiring rate, he looks for the best fitting function of a weighted sum of the vacancy rate and the hiring rate, relative to employment. The re-specification is to allow for the fact that firms may incur costs of maintaining (or indeed posting) vacancies, as well as when workers are matched.

In practise Yashiv's estimate of the weight on vacancies is not significantly different from zero (p1309)[33], and he restricts the weight to be zero for the rest of the paper. Note however that he brings back the weighted function of vacancies in Yashiv (2006) – justifiably- since the latter paper is based on U.S aggregate data rather than the Israeli Employment Service dataset used in Yashiv (2000a) and (2000b). The analysis presented here is an attempt to build upon the results of the latter of the three papers, therefore it also allows vacancies to affect search costs.

Yashiv (2000b) also estimates a behavioural equation for workers. This is facilitated by the fact that the Israeli Employment Service data contains a plausible proxy for the search intensity of the unemployed. During the 1975-1989 sample period, the Employment Service recorded the number days of visits of job-seeking unemployed workers. Yashiv reports that the employment service was at this time the only legal employment exchange in the Israeli labour market for those job-seekers that did not have a university degree. (pp1303-1304). This allows him to justify using the

[33] The point estimate is however 0.3, a value used in Yashiv (2000c) and in the analysis below (table 4.3), both of which model the U.S rather than the Israeli labour market.

frequency of visits to the Employment Service as a proxy for worker search intensity – a variable that goes unobserved in most settings.

In (2000b) Yashiv estimates the firm's and the worker's first-order conditions separately using GMM. The results for the firm are consistent with those of (2000a), and the estimates of the power function parameter γ in particular are almost exactly the same at around 4.7 with a small standard error (pp503-504 Yashiv (2000a) and p1309 Yashiv (2000b).[34] The estimates for the scale parameter Θ have relatively large standard errors and are not robust to specification changes in both papers. The firm's estimated asset value (that is, the ratio of the firm's estimated marginal cost of search to output-per-worker) are similar for some values of Θ. (The instability of estimates of Θ translates into unstable estimates for the firm's asset value). The similarity between papers is hardly surprising since the only innovation to the modelling of the firm in the latter paper is the previously mentioned weight on vacancies in the matching function. The results are reported to be consistent with the small number of existing studies of hiring costs faced by firms, including Hamermesh (1993).

In Yashiv (2006), the cubic hiring cost function is used to simulate a version of the aggregate search and matching model, with the objective of matching U.S data. Yashiv shows that the cubic hiring cost model generates simulated moments that are closer to the data for the period 1976Q1-2003Q3 than the version with linear search costs. Clearly, it is not the case that, just because the model fits Israeli data, it is applicable to the US over a different sample period. However, Yashiv identifies a general mechanism by which the cubic hiring cost mechanism may help to resolve the Shimer puzzle. He also calibrates the model to US values. The mechanism is most clearly explained in a working paper version of the same paper, Yashiv (2005), which shows that the non-linear search costs make the dynamic adjustment of vacancies slower and increases their persistence. The persistence of vacancies feeds through to other parts of the model. The greater volatility of the simulated variables of the model when search costs are non-linear is argued to be a direct consequence of the increased persistence (Yashiv (2005) p39, footnote 18). If the

[34] Yashiv (2000a) is "Hiring as Investment Behaviour" and Yashiv (2000b) "The Determinants of Equilibrium Unemployment."

model does turn out to be inappropriate for the US, indirect inference should allow us to reject the model based on a poor fit with the data.

4.2 Model

Yashiv's (2006) model in its most general form begins with an infinitely lived representative firm, maximizing expected future profit:

$$\max \;\; E_0 \sum_{t=1}^{\infty} \left(\prod_{j=0}^{t} \beta_j \right) [F_t - W_t N_t - \Gamma_t] \qquad (4.1)(\mathrm{a})$$

Note however that, $W_t = W_t(N_t)$ so that wages can respond to the level of employment. F_t is the firm's period *t* real output, W_t is the real wage and N_t is its employment level. β_t is the firm's one period discount factor.
Furthermore, the vacancy cost function now becomes (see Yashiv p913):

$$\Gamma_t = \frac{\Theta}{1+\gamma} \left(\frac{\phi V_t + (1-\phi) Q_t V_t}{N_t} \right)^{\gamma+1} F_t \qquad (4.1)(\mathrm{b})$$

With derivatives (Yashiv p913, footnote 6):

$$\frac{\partial \Gamma_t}{\partial V_t} = \Theta(\phi + (1-\phi)Q_t) \left(\frac{\phi V_t + (1-\phi) Q_t V_t}{N_t} \right)^{\gamma} \frac{F_t}{N_t}$$

$$(4.2)$$

$$\frac{\partial \Gamma_t}{\partial N_t} = \Theta \left(\frac{\phi V_t + (1-\phi) Q_t V_t}{N_t} \right)^{\gamma} \frac{F_t}{N_t} \left[\frac{1-\alpha}{1+\gamma} - 1 \right]$$

As before the matching function is given by the constant returns to scale (CRS), Cobb-Douglas form:

$$M_{t,t+1} = \mu U_t^{\sigma} V_t^{1-\sigma} \qquad (4.3)(\mathrm{a})$$

Or, normalizing by the labour force L_t:

$$m_{t,t+1} = \mu u_t^{\sigma} v_t^{1-\sigma} \qquad (4.3)(b)$$

where lower case denotes rates out of the labour-force.

The job-finding probabilities for workers and for firms are the same as in Chapter 1:

$$P_{t,t+1} = \frac{m_{t,t+1}}{u_t} = \mu \left(\frac{V_t}{U_t}\right)^{1-\sigma} = \mu \left(\frac{v_t}{u_t}\right)^{1-\sigma} \qquad (4.3)(c)$$

$$Q_{t,t+1} = \frac{m_{t,t+1}}{v_t} = \mu \left(\frac{V_t}{U_t}\right)^{-\sigma} = \mu \left(\frac{v_t}{u_t}\right)^{-\sigma} \qquad (4.3)(d)$$

and the production function is also CRS:

$$F_t = A_t K_t^{\alpha} N_t^{1-\alpha} \qquad (4.4)$$

The firm maximizes (4.1)(a) with respect to vacancies V_t and employment in the next period N_{t+1}, subject to the following law of motion for employment, N_t:

$$N_{t+1} = \left(1 - \delta_{t,t+1}\right) N_t + Q_{t,t+1} V_t \qquad (4.5)$$

Note that the firm's matching function in (4.1)(b) is, for $\gamma > 1$, a convex function of a linear combination of vacancies and hires, expressed relative to the stock of employment. It is also linear in output, as is the standard assumption in the literature.

$\left(1 - \delta_{t,t+1}\right)$ is the proportion of employees out of N_t that do not separate, as $\delta_{t,t+1}$ is the separation rate. $Q_{t,t+1}$ is the job-finding probability, V_t is the stock of vacancies, and so $Q_{t,t+1} V_t$ is the number of new hires between *t* and *t+1* since:

$$Q_{t,t+1} V_t = \mu \left(\frac{V_t}{U_t}\right)^{-\sigma} V_t = \mu U_t^{\sigma} V_t^{1-\sigma} = M_{t,t+1}$$

The first-order conditions are as follows. $(4.6)(a)$ is the first-order condition for V_t. $(4.6)(b)$ is the first-order condition for N_{t+1}. $(4.6)(c)$ is the first-order condition for the Lagrange multiplier Λ_t. $(4.6)(d)$is the transversality condition.

$$\frac{\partial \Gamma_t}{\partial V_t} = Q_{t,t+1}\Lambda_t \qquad (4.6)(\mathrm{a})$$

$$\Lambda_t = \mathrm{E}_t\beta_{t+1}\left[\frac{\partial F_{t+1}}{\partial N_{t+1}} - W_{t+1} - N_{t+1}\frac{\partial W_{t+1}}{\partial N_{t+1}} - \frac{\partial \Gamma_{t+1}}{\partial N_{t+1}}\right] + E_t\beta_{t+1}(1-\delta_{t+1,t+2})\Lambda_{t+1} \qquad (4.6)(\mathrm{b})$$

$$N_{t+1} = (1-\delta_{t,t+1})N_t + Q_{t,t+1}V_t \qquad (4.6)(\mathrm{c})$$

$$lim_{T\to\infty}E_t\left[\left(\prod_{j=0}^{T-1}\beta_j\right)\left\{\frac{\partial F_T}{\partial N_T} - W_T - N_T\frac{\partial W_T}{\partial N_T} - \frac{\partial \Gamma_T}{\partial N_T}\right\}\right] = 0 \qquad (4.6)(d)$$

Finally, firm and worker behaviour is assumed to correspond to the maximization of utility or profit, according to the value functions that describe the net benefit of being in each state:

$$J_t^F = \frac{\partial F_t}{\partial N_t} - W_t - N_t\frac{\partial W_t}{\partial N_t} - \frac{\partial \Gamma_t}{\partial N_t} + E_t\beta_{t+1}\left[\delta_{t,t+1}J_{t+1}^V + (1-\delta_{t,t+1})J_{t+1}^F\right] \qquad (4.7)(\mathrm{a})$$

$$J_t^V = -\frac{\partial \Gamma_\mathrm{t}}{\partial V_t} + E_t\beta_{t+1}\left[Q_{t,t+1}J_{t+1}^F + (1-Q_{t,t+1})J_{t+1}^V\right] \qquad (4.7)(\mathrm{b})$$

$$J_t^U = b_t + E_t\beta_{t+1}\left[P_{t,t+1}J_{t+1}^N + (1-P_{t,t+1})J_{t+1}^U\right] \qquad (4.7)(\mathrm{c})$$

$$J_t^N = W_t + E_t\beta_{t+1}\left[(1-\delta_{t,t+1})J_{t+1}^N + \delta_{t,t+1}J_{t+1}^U\right] \qquad (4.7)(\mathrm{d})$$

As is usually the case in Mortensen-Pissarides-style models, it is assumed that firms enter until the benefit from an additional vacancy is zero.

$$J_t^V = 0 \qquad (4.8)$$

As in Chapter 3, and in general for Mortensen-Pissarides style models, wages are determined by Nash-bargaining over the surplus value from the matched state. ξ is the bargaining power of labour.

$$W_t = \text{argmax}\ (J_t^N - J_t^U)^{\xi}(J_t^F - J_t^V)^{1-\xi} \qquad (4.9)$$

So far we have followed Yashiv's model entirely. However, we test a representation of the model that condenses into three linear equations, the endogenous variables of which are similar to those in Chapter 3.

After some manipulation, one can derive the following equations of the pre-linearized model. Of the variables and parameters that have not yet been mentioned or defined, $\lambda_t = \frac{\Lambda_t}{F_t/N_t}$ is the Lagrange multiplier normalized by output-per worker, $s_t = \frac{W_t N_t}{F_t}$ is the labour share of output, ξ is the Nash-bargaining power of labour, $G_{t+1}^X = \frac{F_{t+1}/N_{t+1}}{F_t/N_t}$ is the (gross) quarter-on-quarter growth rate of output-per-worker. η is a parameter that depends on the bargaining power of labour, ξ and also the replacement rate of income of unemployment benefits paid to unemployed workers: letting b_t denote the level of this benefit implies that $b_t = \tau W_t$. η is then given by:

$$\eta = \frac{\xi}{1-(1-\xi)\tau}$$

G^L is the gross quarter-on-quarter growth rate of the labour force, averaged over the sample and hence used as a parameter.

$$\frac{\lambda_t}{G_{t+1}^X} = \mathrm{E}_t\beta_{t+1}\left[\frac{1-\eta}{\eta}s_{t+1} - P_{t+1,t+2}\lambda_{t+1}\right] + E_t\beta_{t+1}\left(1-\delta_{t+1,t+2}\right)\lambda_{t+1} \qquad (4.10)(a)$$

$$\Theta(\phi + (1-\phi)Q_t)\left(\frac{\phi V_t + (1-\phi)Q_t V_t}{N_t}\right)^{\gamma} = Q_{t,t+1}\lambda_t \qquad (4.10)(b)$$

$$P_{t,t+1} = \mu\left(\frac{v_t}{1-n_t}\right)^{1-\sigma} \qquad (4.10)(c)$$

$$s_t = \eta\left[(1-\alpha)\left(\left[\frac{1}{1-\alpha\xi}\right] + \Theta\left(\frac{\phi v_t + (1-\phi)Q_t v_t}{n_t}\right)^{\gamma+1}\left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\right]\right) + P_{t,t+1}\lambda_t\right] \quad (4.10)(d)$$

$$G^L n_{t+1} = \left(1-\delta_{t,t+1}\right)n_t + P_{t,t+1}(1-n_t) \quad (4.10)(e)$$

In Appendix 4.2(a)-4.2(e) we show how each of these equations is derived from the first-order conditions, and the other search theoretic assumptions that come from the Mortensen-Pissarides framework. We first solve for the wages (that is we solve the problem in (4.9) to give W_t) in Appendix 4.1, as this is necessary to derive equation (4.10)(a).

Equation (4.10)(a) is an expression of the firm's inter-temporal first order condition. It says that the appropriately scaled Lagrange multiplier is equal to the expected discounted value of the marginal profit from a filled job, plus the continuation value of that filled job. To see this, note firstly that marginal profits are defined by:

$$\pi_t = \frac{\frac{\partial F}{\partial N} - W_t - \frac{\partial \Gamma_t}{\partial N_t} - N_t\frac{\partial W_t(N_t)}{\partial N_t}}{F_t/N_t}$$

and that the firm's first order inter-temporal condition $(4.6)b$ can therefore be written[35]

$$\frac{\lambda_t}{G_{t+1}^X} = \beta_{t+1}\mathrm{E}_t[\pi_{t+1}] + E_t\beta_{t+1}\left(1-\delta_{t+1,t+2}\right)\lambda_{t+1}$$

which implies that $\pi_{t+1} = \frac{1-\eta}{\eta}s_{t+1} - P_{t+1,t+2}\lambda_{t+1}$. Equation (4.10)(b) is the firm's intra-temporal condition, which says the firm equates the marginal cost of a vacancy with its expected marginal benefit, (the value of an extra employee in the following period, equal to the current period Lagrange multiplier). (4.10)(c) is the job-finding probability, expressed in terms of the vacancy and employment rates (note that it could just as easily be expressed in terms of the vacancy unemployment ratio). The functional form arises from the Cobb-Douglas matching function specification. (4.10)(d) is the solution for the labour share, which is shown to be increasing in both marginal search costs and in $P_{t,t+1}$ multiplied by the marginal value of a filled job to

[35] Note that this is spelled out clearly only in Yashiv's technical appendix to the paper, p6.

the firm. The whole expression is multiplied by η, which as demonstrated above is increasing both in the labour bargaining-power parameter ξ and the benefit or non-work income replacement ratio τ. These variables reflect the familiar intuition behind the Nash-bargaining assumption, which is that workers use their bargaining power to capture a proportion of the rents that accrue to workers and firms as a result of being in a "matched" state. (4.10)(e) is an equation describing the evolution of the employment rate, which is determined by workers that do not separate from one period to the next $[(1-\delta_{t,t+1})n_t]$ and by those unemployed workers that find jobs $[P_{t,t+1}(1-n_t)]$.

4.3 Data sources

Table 4.2 gives details of the data sources used in this chapter. There are some differences between the data used in Chapter 3.

Firstly, we have dispensed with the use of the vacancy-unemployment ratio θ_t as a measure of market tightness entirely. (Recall that in Chapter 3, both $P_{t,t+1}$and θ_t were used.) The reason is mainly for the sake of being concise, however it must also be remembered that $P_{t,t+1}$ may be measured independently of the vacancy rate v_t whereas $\theta_t$may not. Following Yashiv's argument that available measures of v_t conflate vacancies filled by job-to-job movers with those filled by the unemployed, we reason that $P_{t,t+1}$ is empirically a more appropriate variable to use (as in this model, as in Chapter 3, θ_t may be expressed in terms of $P_{t,t+1}$ by a simple monotonic transformation.) Furthermore, the results of Chapter 3 give us no reason to favour θ_t over $P_{t,t+1}$.

Table 4.2: Data definitions and sources.

Variable	*Variable definition in model*	*Source*
Endogenous variables:		
$P_{t,t+1}$	Job-finding probability among workers. $P_{t,t+1} = \frac{m_{t,t+1}}{u_t}$	The job finding-probability for workers is calculated as the ratio of our series for the hiring rate $m_{t,t+1}$ to our series for the unemployment rate u_t.
s_t	Labour share of output in nominal terms. $s_t = \frac{W_t N_t}{NGDP_t}$	For estimates of $W_t N_t$ (i.e. nominal payments to labour) we follow Yashiv (2006) in using the quarterly, seasonally-adjusted BEA series "Compensation of Employees, Paid" (code COE). We divide this by the quarterly seasonally-adjusted BEA series for (nominal) GDP (code GDP). Data downloaded from the St. Louis Fed. http://research.stlouisfed.org/fred2/.
n_t	Employment rate expressed as a proportion of the labour force.	Employment is BLS series LNS12000000Q divided by our series for L_t. The former is quarterly seasonally adjusted U.S unemployment, and was downloaded from http://www.bls.gov/cps/. For details of L_t see below in this table.
Exogenous variables:		
G^x_{t+1}	Quarter-on-quarter growth rate in real GDP per worker. $G^x_{t+1} = \frac{F_{t+1}/N_{t+1}}{F_t/N_t}$	We first calculate quarterly, seasonally-adjusted series of real output per worker. Output is BEA real quarterly U.S gross domestic product, (code GDPC1). We divide this by our series for n_t (see below). G^x_t is then simply the gross quarter-on-quarter growth-rate of the resulting series. BEA data downloaded from the St. Louis Fed. http://research.stlouisfed.org/fred2.
$\delta_{t,t+1}$	Separation probability of workers from employment into unemployment.	We create estimates of quarterly flows by matching individuals in consecutive months using the U.S Current Population survey, and observing and recording any changes in their labour market status between months. To do this we adapted code which has been made public by Robert Shimer. See the main text for details. This method gives a quarterly seasonally adjusted series of estimates for movements from employment to unemployment. To generate the separation probability $\delta_{t,t+1}$ we divide by N_t.
β_t	Firm's discount factor.	We use the quarterly Moody's Seasoned BAA Corporate Bond Yield (BEA code BAA). The quarterly figures are averages of daily yields. We seasonally adjust the series using quarterly dummy variables. Data downloaded from the St. Louis Fed. http://research.stlouisfed.org/fred2.

Other variables (used as means or in the calculation of exogenous and endogenous variables):		
$m_{t,t+1}$		We create estimates of quarterly flows by matching individuals in consecutive months using the U.S Current Population survey, and observing and recording any changes in their labour market status between months. To do this we adapted code which has been made public by Robert Shimer. This method gives a quarterly seasonally adjusted series of estimates for movements from unemployment to employment, $M_{t,t+1}$. To get the rate $m_{t,t+1}$ we divide by the labour force L_t (see below).
L_t	Labour force. Note that the unemployment rate u_t, and where required the vacancy rate v_t and matching rate m_t are all expressed as proportions of the labour force.	For the labour force we use BLS series LNS11000000Q. This is the quarterly seasonally adjusted U.S civilian labour force. The data was downloaded from http://www.bls.gov/cps/.
u_t	Unemployment rate.	Calculated as $1 - n_t$ (see above).
G_t^L	Quarterly (gross) growth rate of the labour force. $G_t^L = \frac{L_{t+1}}{L_t}$ Note that we only actually use the mean of this value for the sample period, so that G^L takes the role of a parameter.	Calculated as the gross quarter on quarter growth of L_t.

A second difference in the choice of variables between those used here and in Chapter 3 is in regards to the use of the labour share of nominal GDP, s_t, as opposed to real wages. Yashiv himself uses this transformation rather than real wages. The reason is essentially that Yashiv chooses to express labour productivity (output per worker) in terms of its quarterly growth rate, and uses this as a driving force in his model. Quarterly growth in output per worker has the benefit of being stationary. Imposing this transformation requires us to normalize the job-creation condition by output-per-worker. Dividing wages by output per worker gives the labour share in GDP. Thus, re-specifying productivity also requires that we recast wages in this way. We also follow Yashiv in using growth in output-per-worker, rather than the level as in Chapter 3 (See appendices 4.21 and 4.24 for details.)

Another important difference is that the labour market dynamics are represented by a difference equation in the employment rate rather than the unemployment rate – again so as to follow Yashiv as closely as possible for the sake of comparability. Note however that due to the fact that only two labour market states are modelled (that is, employment and unemployment), the two dynamic equations of Chapter 3 (in the unemployment rate) and Chapter 4 (in the employment rate) are formally equivalent.

Finally, there is an additional exogenous variable compared with the analysis in Chapter 3. Yashiv uses the discount factor β_t as an exogenous variable, representing the firm's financing costs. For data Yashiv uses a weighted average of average debt and equity financing costs as a measure of the cost of capital. This is an attempt to capture the range of means of financing that are available to different firms across the US economy. Unfortunately, we are unable to obtain a comparable measure of equity financing costs. For data on β_t we therefore use only a measure of the interest rate on BAA rated U.S corporate debt.

4.4 Parameter values and the steady state:

The steady state is more important in Chapter 4 than in the model used in Chapter 3. There are two main reasons for this. Firstly, Yashiv makes a subset of the model parameters functions of a smaller subset of steady state parameters. There are nine steady-state conditions which determine these parameter values (see (4.11)(a)-(i)). Twelve parameters have their values determined by Yashiv – these appear in the first panel of table 4.3. The remaining nine parameters in the bottom panel of the table are determined by the steady-state conditions. The second reason that the steady-state is important for Yashiv's model is that some of the parameters represent the steady-states of variables which appear as exogenous or endogenous variables in the model. Yashiv's original model is written in terms of log-deviations from these parameter values, allowing for the representation of the model in its log-linearized form. The use of these parameters as steady-state values allows Yashiv to avoid the use of an HP-filter or other time-series de-trending method. One could argue that this is preferable in the sense that the steady-state around which the variables are postulated to deviate is part of the model, rather than imposed a-theoretically. There is however a practical problem with the assumption of the constant steady-state for the whole of the sample period, which we discuss below.

Ultimately, this means that we use the HP-filter to de-trend our variables, rather than Yashiv's steady state values.

Table 4.3: Steady state values for different model specifications - comparison with Yashiv (2006)

		(i)	(ii)	(iii)	(iv)
		Yashiv (2006)		*Nanton*	
	Pool:	0	1	0	0
	Search costs:	*γ=2*		*γ=2*	*γ=0*
Production	$1-\alpha$	0.68		0.68	0.68
Matching	σ	0.4		0.4	0.4
Hiring (convexity)	γ	2		2	0
Hiring (vacancy weight)	ϕ	0.3		0.3	0.3
Productivity growth	Gx-1	0.003536		0.0037328	0.0037328
Labour force growth	Gl-1	0.004296	0.004199	0.0039294	0.0039294
Discount factor	β	0.9929		0.9081	0.9081
Separation rate	δ	0.0404		0.0315	0.0315
Unemployment	u	0.063	0.104	0.0634	0.0634
Labour share	s	0.579		0.575	0.575
Vacancy matching rate	Q	0.9		0.9	0.9
Benefit replacement rate	τ	0.25		0.25	0.25
Matching scale param.	μ	0.8	0.85	0.724	0.724
Hiring scale param.	Θ	465	82	715	2
Wage bargaining param.	ξ	0.37	0.41	0.40	0.29
Wage parameter	η	0.44	0.48	0.47	0.35
Vacancy rate	v	0.047	0.089	0.037	0.037
Market tightness	v/u	0.74	0.86	0.58	0.58
Workers' bargaining power	P	0.67	0.77	0.52	0.52
Profits	π	0.05	0.07	0.13	0.21
Asset value	λ	1.02	0.73	0.99	1.62

Notes: Columns (iii) and (iv) entitled "Nanton" refers to our own steady state results in this paper using system (4.11). All of our analysis is based upon "Pool 0" – the official BLS definition of unemployment. Yashiv, by contrast presents results for two different Pools – one of which is Pool 0, the other, Pool 1 includes a subset of the economically inactive. See the main text for details. Note that the first 12 rows of this table corresponds to calibrated values, the remaining 9 are the solutions to the steady-state system.

Many of the parameter values in the top panel are justified with reference to previous econometric studies. Merz and Yashiv (2006) give a point-estimate the labour elasticity of Cobb-Douglas output for the U.S, $1-\alpha$ as 0.68. The important calibrated value for the hiring convexity parameter γ – a value of around 2 - comes from the same study. The calibrated value of 0.3 for the relative weight on vacancies as opposed to new hires in the vacancy cost function, ϕ comes from the GMM estimate in Yashiv (2000b). The elasticity of hires with respect to unemployment, σ, set to 0.4 is based on Blanchard and Diamond's (1989) estimate for the U.S manufacturing sector. The steady-state value for the job-filling probability for firms Q, of 0.9 comes from a U.S study by Burdett and Cunningham (1982). The benefit replacement rate τ is set to 25/99 based on a study by Anderson and Meyer (1997), although Yashiv reports finding his simulation results scarcely affected by different values. Yashiv (2006) p926 is the reference for these values. These parameter values are summarized in the top panel of table 4.1, in columns (i) and (ii).

The aim of this chapter is to evaluate Yashiv's model under these parameter values using the method of indirect inference. We also wish to use indirect inference to examine how the fit of the model is affected by the use of cubic rather than linear search costs. For this reason, we keep the values of the parameters mentioned in the preceding paragraph the same in our analysis, with the exception of γ, which we set alternately to 2 and to 0, for cubic and linear search costs respectively (see table 4.3 columns (iii) and (iv), top panel. The parameters that appear in the top panel of table 4.3 which we have not yet mentioned are sample means of Yashiv's data. These are, productivity growth $G^x - 1$, labour force growth $G^L - 1$, the discount rate β, the separation rate δ, the unemployment rate u and the labour share of output s. Yashiv analyses the model using different 'pools' (i.e. measures) of unemployment, the most empirically successful of which turn out to be the conventional measure of unemployment that accords with the official CPS definition ('pool 0') and the official unemployment measure, *plus* those that are officially economically inactive members of the working-age population, but who say that they 'want a job now' ('pool 1') (Yashiv (2006), p919). It is for this reason that there are two figures for mean unemployment and for mean labour force growth in the first panel of table 4.3- the figures for pool 0 are in column (i) and those for pool 1 are in column (ii).[36] Our dataset differs from Yashiv's to varying degrees among some variables. Therefore, for those parameters which are set equal to sample means, we use the sample means of our own dataset rather than those of Yashiv's. Our values for mean

[36] Note that we differ from Yashiv in that we only consider the official pool of unemployment – Pool 0 - in the evaluation of our model. Columns (iii) and (iv) of table 4.3 correspond not to different definitions of unemployment, but to different values of the search cost parameter γ.

productivity growth $G^x - 1$ are somewhat higher, and those for labour force growth $G^L - 1$ somewhat lower. Also lower are the mean separation rates δ and the labour share s. The biggest difference is the between the values for the mean discount factor β, which is lower by roughly 8 percentage points in our version. The reason is that our data for the discount factor is very different from that constructed by Yashiv, see table 4.2.

The nine parameters of the model in the bottom panel of table 4.3 are solved-for using nine steady-state conditions, which are listed below in equations $(4.11)(\mathrm{a})$ – $(4.11)(\mathrm{i})$. These are solved jointly as a non-linear system. Details of the derivation of these equations is given in Appendix A.4.3. Note that the solutions for the parameter values in the bottom panels of column (i) and column (iii) of table 4.2 are similar. (These columns have both the same definition of unemployment (that is, the unemployment pool is pool 0) and are based on the assumption of cubic search costs: $\gamma = 2$.) Differences which do arise are the result of the different values for $G^x - 1$, $G^L - 1$, β, δ and s, and also from the fact that the solution is actually a non-linear approximation, which may have been implemented somewhat differently in Yashiv's paper. The greatest differences between parameters are for the hiring scale parameter Θ (which, as can be seen by comparison with columns (ii) and (iv), is sensitive to changes in the initial parameter values), and the mean marginal profit rate π, for which our parameterisation results in a value of more than double the comparable figure in Yashiv's study (0.13 versus 0.05).

$$v = \left(\frac{\delta + G^L - 1}{Q}\right)(1 - u) \qquad (4.11)(\mathrm{a})$$

$$\frac{v}{u} = \left(\frac{\delta + G^L - 1}{Q}\right)\frac{(1 - u)}{u} \qquad (4.11)(\mathrm{b})$$

$$P = \left[\frac{\delta}{u} + \frac{(G^L - 1)}{u} - \delta - (G^L - 1)\right] \qquad (4.11)(\mathrm{c})$$

$$n = \left[\frac{Qv}{(G^L - 1) + \delta}\right] \tag{4.11)(d}$$

$$\mu = Q\left(\frac{v}{u}\right)^{\sigma} \tag{4.11)(e}$$

$$\eta = \left[\frac{\xi}{1 - (1 - \xi)\tau}\right] \tag{4.11)(f}$$

$$\lambda = \left[\frac{G^X\beta}{1 - G^X\beta(1 - \delta)}\right]\pi \tag{4.11)(g}$$

$$\begin{aligned}\pi = \left(\frac{1}{1 + \eta P\frac{G^X\beta}{[1 - G^X\beta(1 - \delta)]}}\right)\Bigg[(1 - \eta)(1 - \alpha)\Bigg(\left[\frac{1}{1 - \alpha\xi}\right] \\ + \Theta\Bigg(\phi\frac{(\delta + G^L - 1)}{Q} \\ + (1 - \phi)(\delta + G^L - 1)\Bigg)^{\gamma+1}\left[\frac{\alpha + \gamma}{(1 + \gamma)(1 - \alpha)\big(1 - \xi(1 + \alpha + \gamma)\big)}\right]\Bigg)\Bigg]\end{aligned} \tag{4.11)(h}$$

$$\begin{aligned}s = \eta\Bigg[(1 - \alpha)\Bigg(\left[\frac{1}{1 - \alpha\xi}\right] \\ + \Theta\Bigg(\phi\frac{(\delta + G^L - 1)}{Q} \\ + (1 - \phi)(\delta + G^L \\ - 1)\Bigg)^{\gamma+1}\left[\frac{\alpha + \gamma}{(1 + \gamma)(1 - \alpha)\big(1 - \xi(1 + \alpha + \gamma)\big)}\right]\Bigg) + P\lambda\Bigg]\end{aligned} \tag{4.11)(i}$$

Table 4.3 also reveals the effect on the steady-state parameters of changing the search costs from cubic ($\gamma = 2$) to linear ($\gamma = 0$). This can be seen by comparing columns (iii) and (iv). Only four out of nine of the steady-state conditions are affected: these are ((4.11)(f), (g), (h) and (i)). The steady-state marginal profit rate π

can be seen to roughly double – this is because there is large increase in steady-state vacancy costs, which is not offset by a large fall in the scale of matching Θ in equation (4.11)(h). The marginal profit increase is also helped by a fall in labour's bargaining power ξ of around one-quarter, and consequently in η which is increasing in ξ as can be seen in (4.11)(f). This fall in bargaining power is required to maintain the steady state condition (4.11)(i), as the condition requires that there be no increase in the labour share which has been calibrated to the sample mean, and is therefore invariant to γ. Equation (4.11)(g) says that the value to the firm of a marginal employed worker (normalized by the average product of labour), λ is equal to the firm's marginal profit rate π multiplied by a discount factor. The discount factor is also invariant to changes in γ, which means that the value to the firm of a marginal employed worker increases when search costs are changed from being cubic to linear.

In summary, our data-set differs from Yashiv's in several important ways due to data availability. This is an issue because we don't have access to Yashiv's original dataset, and are therefore forced to reconstruct the data as closely as possible using the variable descriptions given in Yashiv's paper. Some of the data are not accessible to us, and some appear to have been subject to revisions. These data differences imply somewhat different means, and hence different steady-states for a subset of our variables (the first twelve rows of table 4.3). We use the same steady-state *conditions* as Yashiv to derive the steady-states for the remaining nine variables. But as the parameters to the steady-state system are precisely those twelve variable means at the top of table 4.3, the resulting steady-states for the remaining nine variables are also different. We argue that it makes more sense to use the actual means of our data for the original twelve variables, than it does to borrow the numerical means wholesale from Yashiv, when the latter are not the means of the data that we are actually using. The remaining nine steady-states that are solved from the system are therefore numerically different from those of Yashiv, but are derived according to the same logic.

Ideally it would be better for the purposes of evaluating the model under indirect inference if we were to have exactly the same dataset. However, this is not an option for us, and we are forced to compromise somewhat in the comparability of our results.

Note that below we also explore an alternative possibility for de-trending the data. That is, as an alternative de-trending method we use also use a log deviation around an HP(1600) filtered trend for each of our variables (as in Chapter 3). Although we explore the model under both treatments, we end up using the HP-de-trended data for indirect inference testing[37]. The reason is that it turns out that using log deviations around our constant mean values for the sample period fails to make our data stationary (see table 4.9 below). Again, here we deviate from Yashiv's original paper.

However, we note that there is nothing in Yashiv's proposed mechanism (increased persistence of vacancies) that relies on de-trending the variables around a constant steady-state. Indeed, the original Shimer (2005) puzzle was cast in terms of HP-de-trended data. Therefore, we maintain that HP-de-trending rather than steady-state de-trending should not be a gross distortion of the Yashiv model. De-trending using an HP-filter will of course result in different data generated and model generated moments, (since it will result in different bootstrap-residuals), as we document below.

4.5 The linearized system of equations

We combine equations (4.10)(a)- (4.10)(e) so as to eliminate λ_t , λ_{t+1}, $Q_{t,t+1}$ and v_t. The resulting system is given by (4.12)(a)-(c), in three equations, with three endogenous variables $\tilde{P}_{t,t+1}$, $\tilde{s}_t$, $\tilde{n}_{t+1}$ (the –tilde notation expresses log deviations from an assumed steady state) and exogenous variables $\tilde{G}_t^X$, $\tilde{\delta}_t^X$ and $\tilde{\beta}_t$. $\tilde{n}_t$ is of course pre-determined by period *$(t-1)$*. To recap: $\tilde{P}_{t,t+1}$ is the workers' job finding probability, $\tilde{s}_t$ is the labour share of output, $\tilde{n}_{t+1}$ is the employment rate in period $t+1$.

(4.12)(a) is the combined inter-temporal and intra-temporal condition. (4.12)(b) is the equation for the labour-share. (4.12)(c) is the log-linearized equation of motion for employment.

In anticipation of the econometrics to follow, we note that equations (4.12)(a)-(4.12)(c) are in structural form – they are conditions that have been derived from the

[37] Note that even when we de-trend the data using the HP-filter, the steady-state values are still used in the analysis, as they become parameters in the structural matrices of the model.

assumptions of the model (by way of log-linearization from an assumed steady state). No effort has yet been made to solve them. We have attached a structural error to the right-hand-side of each equation, given by $e_{p,t}$, $e_{s,t}$ and $e_{n,t}$ respectively. These play an important role in the indirect inference testing procedure below.

$$\begin{aligned}
&\frac{\lambda}{G^X}\left[[-\sigma(\Omega+\gamma\Omega-1)+\gamma]\frac{1}{1-\sigma}\right]\tilde{P}_{t,t+1}-\frac{\lambda}{G^X}\left[\frac{\gamma}{1-n}\right]\tilde{n}_t-\frac{\lambda}{G^X}\rho_{G^x}\tilde{G}^X_t+\beta\lambda\delta\rho_\delta\tilde{\delta}_{t,t+1}\\
&\quad-\beta\frac{1-\eta}{\eta}s\mathrm{E}_t[\tilde{s}_{t+1}]-\left(\beta\frac{1-\eta}{\eta}s+\beta\lambda(1-\delta-P)\right)\rho_\beta\tilde{\beta}_t\\
&\quad-\left\{(\beta(1-\delta)\lambda-\beta P\lambda)\left[[-\sigma(\Omega+\gamma\Omega-1)+\gamma]\frac{1}{1-\sigma}\right]\right.\\
&\quad\left.-\beta P\lambda\right\}E_t\left[\tilde{P}_{t+1,t+2}\right]+\beta(1-\delta-P)\lambda\left[\frac{\gamma}{1-n}\right]E_t[\tilde{n}_{t+1}]=e_{P,t}
\end{aligned}\qquad(4.12)(\mathrm{a})$$

$$\begin{aligned}
&s\tilde{s}_t-\left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\eta(1\right.\\
&\quad-\alpha)\Theta\left(\frac{\phi v+(1-\phi)Qv}{n}\right)^{\gamma+1}\frac{(\gamma+1)}{1-\sigma}[1+\sigma(1-\Omega)]\\
&\quad\left.+\frac{\eta P\lambda}{1-\sigma}[1-\sigma(\Omega+\gamma\Omega)+\gamma]\right]\tilde{P}_{t,t+1}\\
&\quad+\left(\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\eta(1\right.\\
&\quad\left.-\alpha)\Theta\left(\frac{\phi v+(1-\phi)Qv}{n}\right)^{\gamma+1}\frac{(\gamma+1)}{1-n}+\frac{\eta P\lambda\gamma}{1-n}\right)\tilde{n}_t=e_{s,t}
\end{aligned}\qquad(4.12)(\mathrm{b})$$

$$G^L\tilde{n}_{t+1}-(1-\delta-P)\tilde{n}_t+\delta\tilde{\delta}_{t,t+1}-\frac{P(1-n)}{n}\tilde{P}_{t,t+1}=e_{n,t}\qquad(4.12)(\mathrm{c})$$

4.6 Methodology

The method and ideas behind indirect inference were explained in detail in Chapter 3. Here we reproduce the schema from that chapter, highlighting the way in which we adapt the method for this particular model.

(1). *Estimate the auxiliary model once on the data.*

The auxiliary models we will use (in separate tests) are the standard deviations of the endogenous variables (in which case $K = 3$) and VAR(1) coefficients from estimating a VAR on the endogenous variables (in which case $K = 9$*).*

This will give us *k* parameter estimates.

$$\hat{\Gamma}_k \text{ for } k = 1,..,K.$$

(2). *Using the structural model, compute the r structural errors in $\mathbf{\hat{e}_t}$ under $\boldsymbol{\hat{\theta}}$ using*

$$\mathbf{\hat{e}_t} = \mathbf{A}(\boldsymbol{\hat{\theta}}^{\mathbf{A}})\mathbf{y_t} - \mathbf{B}(\boldsymbol{\hat{\theta}}^{\mathbf{B}})\mathbf{x_t} - \mathbf{C}(\boldsymbol{\hat{\theta}}^{\mathbf{C}})E_t\mathbf{y_{t+1}}$$

and using actual data for $\mathbf{y_t}$, $\mathbf{x_t}$ and if necessary, using a VAR to obtain estimated data values for $E_t\mathbf{y_{t+1}}$.

In this model, $r = 3$, and the vectors are as follows:

$$\mathbf{\hat{e}_t} = [\hat{e}_{P,t} \quad \hat{e}_{s,t} \quad \hat{e}_{n,t}]^T$$

$$\mathbf{y_t} = [\tilde{P}_{t,t+1} \quad \tilde{s}_t \quad \tilde{n}_{t+1}]^T$$

$$\mathbf{x_t} = [\tilde{G}_t^x \quad \tilde{\delta}_{t,t+1} \quad \tilde{\beta}_t]^T$$

$\boldsymbol{\hat{\theta}}$ contains all of the parameters in table 4.3, partitioned as required into $\boldsymbol{\hat{\theta}}^{\mathbf{A}}$, $\boldsymbol{\hat{\theta}}^{\mathbf{B}}$ and $\boldsymbol{\hat{\theta}}^{\mathbf{C}}$.

To obtain the structural errors $\mathbf{e_t}$, we feed in data to $\mathbf{y_t}$, $\mathbf{x_t}$ and $E_t\mathbf{y_{t+1}}$, and we populate $\mathbf{A}(\boldsymbol{\hat{\theta}}^{\mathbf{A}})$, $\mathbf{B}(\boldsymbol{\hat{\theta}}^{\mathbf{B}})$ and $\mathbf{C}(\boldsymbol{\hat{\theta}}^{\mathbf{C}})$ with the requisite parameters from (4.12)(a), (4.12)(b) and (4.12)(c).

In Chapter 3, $\mathbf{C}(\boldsymbol{\hat{\theta}}^{\mathbf{C}})$ was a matrix of zeros due to the assumption of static expectations. Here this is no longer the case, as the maintained assumption following Yashiv is that expectations are rational. To obtain data on the expectations, we estimate a first order VAR on the log-linearized endogenous variables in $\mathbf{y_t}$ using our sample of data. We construct fitted values from the regression and lead those values by one period.

(3). *Estimate univariate processes for each of the r structural errors in $\mathbf{e_t}$ so as to obtain white noise residuals for each. Estimate univariate processes for the exogenous variables as well, and obtain the residuals which are also white noise. And the end of step* (2) *one will have parameters $\hat{\rho}_1, \dots, \hat{\rho}_r; \hat{\varpi}_1, \dots, \hat{\varpi}_n$ and residuals $\hat{\varepsilon}_{1t}, \dots, \hat{\varepsilon}_{rt}; \hat{\zeta}_{1t}, \dots, \hat{\zeta}_{nt}$ $\forall t$ in the notation above.*

That is, we estimate the following univariate AR(1) equations:

$$e_{P,t} = \hat{\rho}_P e_{P,t-1} + \hat{\varepsilon}_{P,t}$$
$$e_{s,t} = \hat{\rho}_s e_{s,t-1} + \hat{\varepsilon}_{s,t}$$
$$e_{n,t} = \hat{\rho}_n e_{n,t-1} + \hat{\varepsilon}_{n,t}$$

$$\tilde{G}_t^X = \hat{\varpi}_{G^X} \tilde{G}_{t-1}^X + \hat{\zeta}_{G^X,t}$$
$$\tilde{\delta}_{t,t+1} = \hat{\varpi}_\delta \tilde{\delta}_{t-1,t} + \hat{\zeta}_{\delta,t}$$
$$\tilde{\beta}_t = \hat{\varpi}_\beta \tilde{\beta}_{t-1} + \hat{\zeta}_{\beta,t}$$

(3). We draw 1000 bootstrap pseudo-samples of the estimated residuals, $[\hat{\varepsilon}_{P,t} \quad \hat{\varepsilon}_{s,t} \quad \hat{\varepsilon}_{n,t} \quad \hat{\zeta}_{G^X,t} \quad \hat{\zeta}_{\delta.t,} \quad \hat{\zeta}_{\beta,t}]$, and the estimated univariate AR(1) equations to create 1000 pseudo-samples of exogenous variables, of structural errors and of exogenous variables. The re-sampling is done as a block bootstrap, as detailed in Chapter 3 so as to preserve any correlation between the residuals in the original time series.

$$\hat{\hat{e}}^{j}_{P,t} = \hat{\rho}_P \hat{\hat{e}}^{j}_{P,t-1} + \hat{\hat{\varepsilon}}^{j}_{P,t}$$
$$\hat{\hat{e}}^{j}_{s,t} = \hat{\rho}_s \hat{\hat{e}}^{j}_{s,t-1} + \hat{\hat{\varepsilon}}^{j}_{s,t}$$
$$\hat{\hat{e}}^{j}_{n,t} = \hat{\rho}_n \hat{\hat{e}}^{j}_{n,t-1} + \hat{\hat{\varepsilon}}^{j}_{n,t}$$
$$j = 1, \ldots, 1000$$
$$\widehat{\widehat{\tilde{G}^{X}}}^{j}_{t} = \hat{\rho}_{G^X} \widehat{\widehat{\tilde{G}^{X}}}^{j}_{t-1} + \hat{\hat{\zeta}}^{j}_{G^X,t}$$
$$\widehat{\widehat{\tilde{\delta}^{j}_{t,t+1}}} = \hat{\rho}_{\delta} \widehat{\widehat{\tilde{\delta}^{j}_{t-1,t}}} + \hat{\hat{\zeta}}^{j}_{\delta,t}$$
$$\widehat{\widehat{\tilde{\beta}^{j}_{t}}} = \hat{\rho}_{\beta} \widehat{\widehat{\tilde{\beta}^{j}_{t-1}}} + \hat{\hat{\zeta}}^{j}_{\beta,t}$$

(5). *Solve the model under $\widehat{\boldsymbol{\theta}}$ for $\mathbf{y_t}$, in terms of the structural errors and exogenous variables.*

Solving the model is somewhat more difficult than in Chapter 3 due to the presence of the expectations of endogenous variables in the structural equations. We use the equation solver in the MATLAB package Dynare to compute the reduced form of the model. The reduced form allows us to simulate the model 1000 times using the 1000 bootstrap samples of exogenous variables and error processes for the three structural equations.

$$\hat{\mathbf{y}}^{j}_{\mathbf{t}} = \boldsymbol{\Upsilon}_{\mathbf{1}}(\widehat{\boldsymbol{\theta}}) \begin{bmatrix} \hat{\hat{\mathbf{y}}}^{j}_{\mathbf{t-1}} \\ \hat{\hat{\mathbf{x}}}^{j}_{\mathbf{t-1}} \end{bmatrix} + \boldsymbol{\Upsilon}_{\mathbf{2}}(\widehat{\boldsymbol{\theta}}) \begin{bmatrix} \hat{\hat{\varepsilon}}^{j}_{P,t} \\ \hat{\hat{\varepsilon}}^{j}_{s,t} \\ \hat{\hat{\varepsilon}}^{j}_{n,t} \\ \hat{\hat{\zeta}}^{j}_{G^X,t} \\ \hat{\hat{\zeta}}^{j}_{\delta,t} \\ \hat{\hat{\zeta}}^{j}_{\beta,t} \end{bmatrix}$$

$$j = 1, \ldots, 1000$$

Where $\boldsymbol{\Upsilon}_{\mathbf{1}}$ and $\boldsymbol{\Upsilon}_{\mathbf{2}}$ are 3x6 matrices both of which depend on the underlying parameter vector $\widehat{\boldsymbol{\theta}}$.

(5). *Estimate the chosen K-parameter auxiliary model on each of the 1000 simulated samples.*

We use the notation $\hat{\Gamma}^{j}_{k}(\widehat{\boldsymbol{\theta}})$ for $j = 1, \ldots, 1000$ and $k = 1, \ldots, K$ to denote the kth parameter of the chosen auxiliary model, estimated on the jth pseudo-sample of exogenous and endogenous variables: $\{\hat{\mathbf{y}}^{j}_{\mathbf{t}}, \hat{\mathbf{x}}^{j}_{\mathbf{t}}\}^{T}_{t=1}$. As in Chapter 3, the auxiliary models are standard deviations of the model's endogenous variables (for which the parameter vector is a three-vector of standard deviation parameters, hence $K = 3$), and the VAR coefficients from a first-order VAR on the endogenous variables (in which case $K = 9$).

Calculating the Wald statistic for the model under $\widehat{\boldsymbol{\theta}}$ requires the mean of each of the auxiliary model parameters across the 1000 simulations, we therefore calculate:

$$\bar{\Gamma}_{k}(\widehat{\boldsymbol{\theta}}) = \frac{1}{1000} \sum_{j}^{1000} \hat{\Gamma}^{j}_{k}(\widehat{\boldsymbol{\theta}}) \text{ for } \quad k = 1, \ldots, K.$$

(7). *Compute the Wald statistic*

For this we require $\mathbf{v}$ and $\mathbf{W}(\widehat{\boldsymbol{\theta}})$. The Wald is given by $\mathbf{v}^{\mathbf{T}}\mathbf{W}(\widehat{\boldsymbol{\theta}})\mathbf{v}$.

$$\mathbf{v} = \begin{bmatrix} \hat{\Gamma}_1 - \bar{\Gamma}_1(\widehat{\boldsymbol{\theta}}) & \ldots & \hat{\Gamma}_K - \bar{\Gamma}_K(\widehat{\boldsymbol{\theta}}) \end{bmatrix}$$

$$\mathbf{W}(\widehat{\boldsymbol{\theta}}) = \begin{bmatrix} \operatorname{var}\left(\hat{\phi}_1^j(\widehat{\boldsymbol{\theta}}) - \bar{\phi}_1(\widehat{\boldsymbol{\theta}})\right) & \operatorname{cov}\left(\hat{\phi}_1^j(\widehat{\boldsymbol{\theta}}) - \bar{\phi}_1(\widehat{\boldsymbol{\theta}})\right)\left(\hat{\phi}_2^j(\widehat{\boldsymbol{\theta}}) - \bar{\phi}_2(\widehat{\boldsymbol{\theta}})\right) & \ldots \\ \operatorname{cov}\left(\hat{\phi}_1^j(\widehat{\boldsymbol{\theta}}) - \bar{\phi}_1(\widehat{\boldsymbol{\theta}})\right)\left(\hat{\phi}_2^j(\widehat{\boldsymbol{\theta}}) - \bar{\phi}_2(\widehat{\boldsymbol{\theta}})\right) & \operatorname{var}\left(\hat{\phi}_2^j(\widehat{\boldsymbol{\theta}}) - \bar{\phi}_2(\widehat{\boldsymbol{\theta}})\right) & \ldots \\ \ldots & \ldots & \ldots \end{bmatrix}^{-1}$$

Which makes $\mathbf{W}(\widehat{\boldsymbol{\theta}})$ a $K \times K$ matrix.

(8). *Compute the distribution of the Wald under the null as follows:* $WALD_0^j = \mathbf{v_0^j}'\mathbf{W}(\widehat{\boldsymbol{\theta}})\mathbf{v_0^j}$ for $j = 1, \ldots 1000$, where:

$$\mathbf{v_0^j} = \left[\hat{\Gamma}_1^j(\widehat{\boldsymbol{\theta}}) - \bar{\Gamma}_1(\widehat{\boldsymbol{\theta}}) \quad \ldots \quad \hat{\Gamma}_K^j(\widehat{\boldsymbol{\theta}}) - \bar{\Gamma}_K(\widehat{\boldsymbol{\theta}})\right]$$

For inference we use this finite-sample simulated distribution that can be constructed by ordering the 1000 values of $WALD_0^j$ by increasing magnitude.

(9). *The Wald statistic from step (7) can then be compared to the distribution of Wald under the null hypothesis from the previous step. The model under* $\widehat{\boldsymbol{\theta}}$ *is rejected if the Wald statistic lies in the tail outside of 95% of the values of the finite sample Wald distribution constructed in step* (8).

In contrast to the analysis in Chapter 3, we do not attempt to estimate the parameters of Yashiv's model. Our aim is simply to test the model under parameters that are purported to be successful in Yashiv (2006).

4.7 Differences between Yashiv (2006) and this paper

This chapter is an attempt to test the model of Yashiv (2006) using indirect inference, instead of using the simple comparison of moments that Yashiv employs. The model we have used is however not a perfect replica of the one used in the original paper. Our aim in this section is firstly to enumerate the differences between the models and secondly to show that these differences are not important – that the two models produce broadly the same results when we use a similar method of evaluation on each (that is, Yashiv's comparison of moments). This will allow us to make a valid comparison between Yashiv's and our results when we switch to using indirect inference, below. In other words, it will allow us to be reasonably sure that any differences in the model evaluation arise from the econometric procedures used, and not from fundamental differences between the models.

The first major difference between our model and that used by Yashiv, is that our model is smaller. We condense the model into three linearized equations ((4.12)(a)-

(c)), so as to remove all unobservables. This is necessary for the implementation of indirect inference, because one requires data on all of the variables in the structural equations to derive the structural errors (step (2), above). Removing unobservables is however not a requirement of Yashiv's moment comparison method, since he is free to simulate the whole model –both observable and unobservable parts – and may then simply compute the moments on which he has comparable data.

Removing unobservables from our version of the model means eliminating expressions of the (normalized by average labour productivity) Lagrange multiplier λ_t at any time index. This is easily done, since by the first order condition (4.6)(a) λ_t is equal to to the (normalized) marginal cost of a vacancy divided by the firm's hiring probability.

There are however more unobservables that must be removed. Since we ultimately agree with Yashiv's contention that the Help-Wanted-Index does not capture the relevant concept of vacancies (recall from Chapter 3 that the problem is that the Help-Wanted Index conflates vacancies which are eventually filled by job-to-job movers with those filled by the unemployed) and since we have no ready alternative, we must also eliminate the vacancy rate, v_t, and the firm's hiring probability $Q_{t,t+1}$, as the latter must be calculated using a measure of vacancies. Fortunately, as shown in Appendix 4.4, both of these variables may be eliminated entirely, by writing them in terms of employment n_t and the job-finding-probability for workers, $P_{t,t+1}$. As we have argued in Chapter 3, the latter can be measured in a way that does not conflate job-to-job movers with flows from unemployment to employment, as it is the ratio of hires to unemployment, and we can use the CPS data to measure hires that are strictly from unemployment.

The other major difference is in the way Yashiv chooses to model the three exogenous variables $\tilde{G}_t^X, \tilde{\delta}_{t,t+1}, \tilde{\beta}$. He uses a reduced form VAR(1) estimated on his data for these variables, and treats the resulting reduced-form shocks as the source of shocks to his model. This is the *only* source of shocks to his model.

In our version of the model, there is something of a dilemma that we face with regards to the modelling of shocks. The usual procedure with indirect inference is to assign one structural shock to each structural equation, as we have done in

(4.12)(a)-(4.12)(c). The model solution will specify the appropriate linear combination of the structural shocks required to convert the structural error to a reduced form error, enabling simulation of the endogenous variables. The exogenous variables *and* the structural errors are modelled as univariate AR(1) processes, but the residuals from each are bootstrapped as a block, with the aim of preserving any correlation between the residuals found in the data. The problem here is that this method implies block-bootstrapping residuals for six variables (or more generally, the number of structural shocks plus the number of exogenous variables), whereas Yashiv's model requires the modelling of just three exogenous shocks (or in general one shock per exogenous variable). This issue turns out to be a crucial one for this paper - we later argue based on our indirect inference results that Yashiv's calibration only produces simulated data that matches the data moments if one ignores the error inherent in the structural equations. For the moment, however, we are concerned with matching Yashiv's results as closely as possible. For this, we assume that the three structural equations (4.12)(a)-(4.12)(c) hold exactly – that is we force $e_{P,t}$, $e_{s,t}$ and $e_{n,t}$ to equal zero. We then implement the block bootstrap procedure using only the estimated residuals to the AR(1) processes of the exogenous variables. Effectively, this means amending steps (2)-(4) (above) as follows:

(3). *Estimate univariate processes for the exogenous variables, and obtain the residuals which are also white noise. And the end of step* (2) *one will have parameters* $\widehat{\varpi}_1, \dots, \widehat{\varpi}_n$ *and residuals;* $\hat{\zeta}_{1t}, \dots, \hat{\zeta}_{nt}\ \forall t$ *in the notation above.*

That is, we estimate the following univariate AR(1) equations:

$$\tilde{G}_t^X = \widehat{\varpi}_{G^X} \tilde{G}_{t-1}^X + \hat{\zeta}_{G^X,t}$$
$$\tilde{\delta}_{t,t+1} = \widehat{\varpi}_{\delta} \tilde{\delta}_{t-1,t} + \hat{\zeta}_{\delta,t}$$
$$\tilde{\beta}_t = \widehat{\varpi}_{\beta} \tilde{\beta}_{t-1} + \hat{\zeta}_{\beta,t}$$

(4). *Sample the residuals in order to create 1000 vector-bootstrapped samples of innovations. This means using each of the estimated univariate processes to generate 1000 simulated samples of structural errors, and 1000 samples of exogenous variables.*

We draw 1000 bootstrap pseudo-samples of the estimated residuals, $[\hat{\zeta}_{G^X,t} \quad \hat{\zeta}_{\delta.t,} \quad \hat{\zeta}_{\beta,t}]$, and the estimated univariate AR(1) equations to create 1000 pseudo-samples of exogenous variables, of structural errors and of exogenous variables. The re-sampling is done as a block bootstrap, as detailed in Chapter 3 so as to preserve any correlation between the residuals in the original time series.

$$\widehat{\widetilde{G^X}}_t^j = \hat{\rho}_{G^X}\widehat{\widetilde{G^X}}_{t-1}^j + \hat{\zeta}_{G^X,t}^j$$
$$\widehat{\tilde{\delta}_{t,t+1}^j} = \hat{\rho}_\delta \widehat{\tilde{\delta}_{t-1,t}^j} + \hat{\zeta}_{\delta,t}^j$$
$$\widehat{\tilde{\beta}_t^j} = \hat{\rho}_\beta \widehat{\tilde{\beta}_{t-1}^j} + \hat{\zeta}_{\beta,t}^j \qquad j = 1, \dots, 1000$$

(5). *Solve the model under $\hat{\boldsymbol{\theta}}$, in terms of the exogenous variables. Use the simulated data to generate 1000 samples of endogenous variables.*

$$\hat{\mathbf{y}}_{\mathbf{t}}^j = \boldsymbol{\Upsilon}_1(\hat{\boldsymbol{\theta}})\begin{bmatrix}\hat{\mathbf{y}}_{\mathbf{t-1}}^j \\ \hat{\mathbf{x}}_{\mathbf{t-1}}^j\end{bmatrix} + \boldsymbol{\Upsilon}_2(\hat{\boldsymbol{\theta}})\begin{bmatrix}\hat{\zeta}_{G^X,t}^j \\ \hat{\zeta}_{\delta,t}^j \\ \hat{\zeta}_{\beta,t}^j\end{bmatrix}$$

$$j = 1, \dots, 1000$$

Where $\boldsymbol{\Upsilon}_1$ is a 3×6 matrix and $\boldsymbol{\Upsilon}_2$ is a 3×3 matrix.

The three equations in step(3) used with a block-bootstrapped vector of residuals $[\hat{\zeta}_{G^X,t} \quad \hat{\zeta}_{\delta.t,} \quad \hat{\zeta}_{\beta,t}]$ are comparable with the reduced-form VAR(1) Yashiv uses to simulate his equations (p916):

$$\begin{bmatrix}\tilde{G}_{t+1} \\ \tilde{\delta}_{t+1,t+2} \\ \tilde{\beta}_{t+1}\end{bmatrix} = \boldsymbol{\Pi}\begin{bmatrix}\tilde{G}_t \\ \tilde{\delta}_{t,t+1} \\ \tilde{\beta}_{t+1}\end{bmatrix} + \boldsymbol{\Sigma}$$

where $\boldsymbol{\Pi}$ and $\boldsymbol{\Sigma}$ are both 3x3 matrices. Note however that Yashiv's specification is somewhat more general than our univariate equation setup in that it allows for dependencies between lagged exogenous variables (since $\boldsymbol{\Pi}$ is not forced to be diagonal), as well as for contemporaneous reduced form shocks in $\boldsymbol{\Sigma}$. On the other hand, $\boldsymbol{\Sigma}$ in Yashiv's analysis should capture the same shocks as our block bootstrap vector $[\hat{\zeta}_{G^X,t} \quad \hat{\zeta}_{\delta,t} \quad \hat{\zeta}_{\beta,t}]$.

Yashiv simulates a log-linearized version of his model, in which variables are represented as percentage differences from their steady state values. In order to compare the simulated series with the data, the data must be made stationary in the same way. The steady-state values are given in table 4.3, columns (i) and (ii) (recall that Yashiv's data and our data imply different steady states). I begin by using our own steady-state values (columns (iii) and (iv)) to de-trend our variables.

There is however an econometric problem here. The variables in the model are assumed to have been made stationary. We find evidence however that de-trending our sample of data in this way does not produce stationary data. Table 4.4 presents Dickey-Fuller tests for unit roots on the individual data series used in our study. In column (i) of the table, the series have been de-trended by taking log-differences from their respective steady-state values given in table 4.3. For all variables with the exception of quarterly labour productivity growth $\tilde{G}_{t,t+1}$, we fail to reject the null hypothesis that each series has a unit root. More prosaically, since many of the variables are trended and do not have a constant mean, it is not at all surprising that log-differencing from a constant does not counteract the non-stationarity.

Table 4.4: Dickey-fuller test for a unit root: McKinnon approximate P-values

	(i)	(ii)
	log-difference from steady state	*HP(1600)-filter*
$\tilde{P}_{t,t+1}$	0.3268	0.006
$\tilde{s}_t$	0.2346	0.0007
$\tilde{n}_t$	0.1610	0.0012
$\tilde{G}_{t,t+1}$	0.0000	0.0000
$\tilde{\delta}_{t,t+1}$	0.3917	0.001
$\tilde{\beta}_t$	0.6507	0.0007

P-values for the Dickey fuller test with 1 lag. The sample size is 109 in all cases. Results are similar for 0 and 2 lags.

A more standard approach to making variables stationary from the macroeconomics literature is to log-difference the data from their HP-filtered trend. Column (ii) of table 4.4 suggests that this works – after filtering and log-differencing the variables in this way we are able to reject the null hypothesis of a unit root for each at the 1% level of significance. Note that this use of the HP-filter does deviate from the methodology of Yashiv (2006). One could also object to using an a-theoretical time series process to treat the variables, although it is hard to see why log-differencing from an arbitrary sample mean is better in this respect.

The graphs in figure 4.1 shows graphs each variable both HP-filtered and expressed as a log-difference from its steady state. The graphs confirm the results of figure 4.4 visually. Simply log-differencing the variables around the steady-state appears to allow the variables to retain a time-varying mean, in other words, it does not de-trend them properly. This is not the case for the HP-filtered variables.

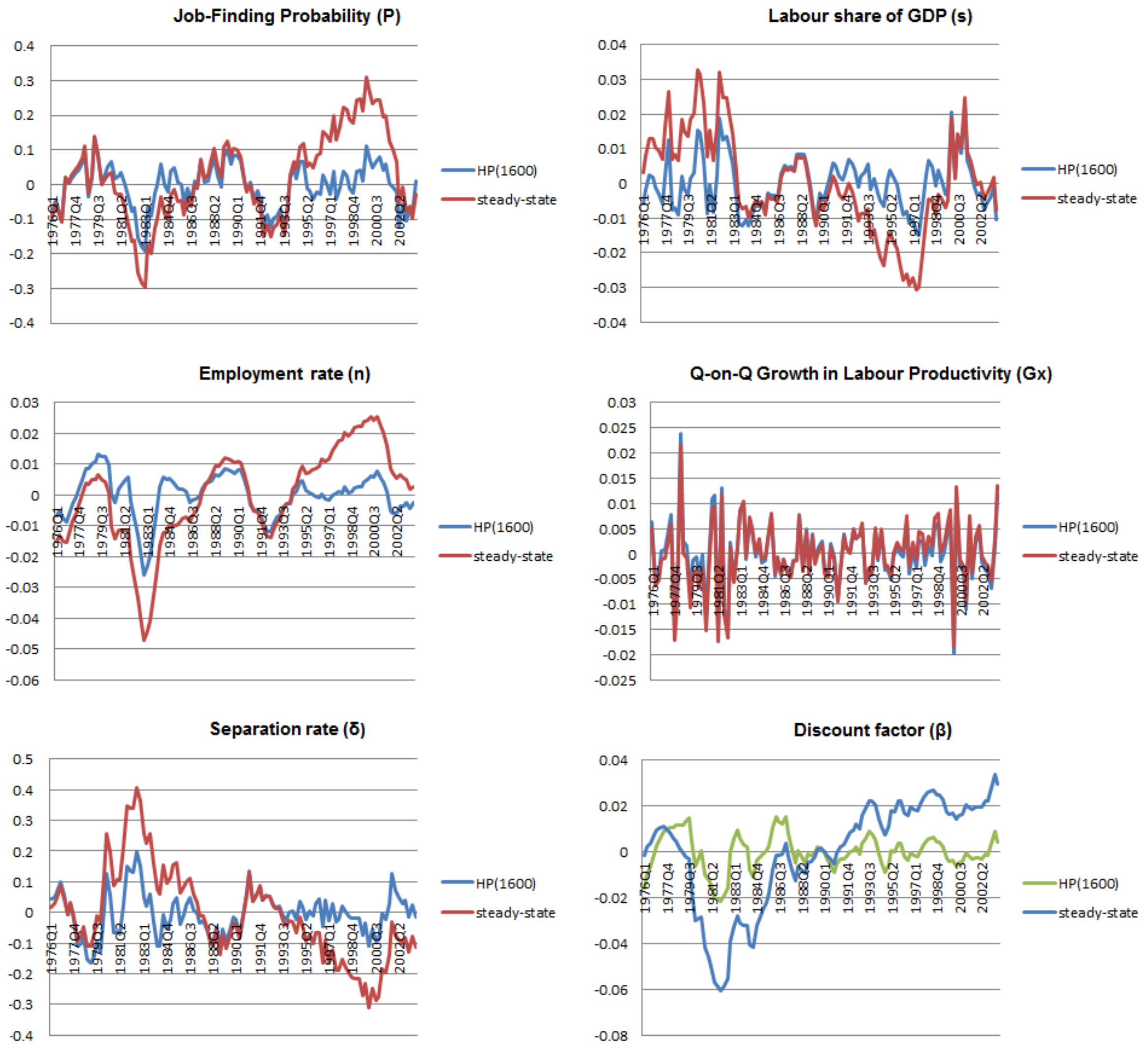

Figure 4.1. Steady-state and HP(1600) filter de-trended variables plotted for the sample period, 1976Q1-2003Q3.

The strategy we use in the remainder of the paper is as follows. In the next two tables (tables 4.5 and 4.6) we present summary statistics on different versions of the model, with a view to comparing our reproduction of Yashiv's (2006) model with the original. Thereafter, we focus on the HP-filtered version of the model, since we require stationary residuals for our indirect inference procedure.

Table 4.5: Comparison of summary statistics from Yashiv's (2006) analysis with those from our versions of the model, with shocks to the exogenous variables only.

	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
	Yashiv (2006)			*Nanton*					
Pool:	1			0					
	Data	*Model*		*Data*	*Model*		*Data*	*Model*	
Shocks:		*Exog. only*	*Exog. only*		*Exog. only*	*Exog. only*		*Exog. only*	*Exog. only*
Search costs:		*γ=2*	*γ=0*		*γ=2*	*γ=0*		*γ=2*	*γ=0*
Data treatment:	*Steady-state de-trending*			*Steady-state de-trended*			*HP-filter(1600) de-trended*		
AR1($\tilde{u}_t$)	0.97	0.98	0.88	0.97	0.99	0.99	0.92	0.96	0.80
AR1($\tilde{m}_{t,t+1}$)	0.85	0.99	0.47	0.89	0.94	0.95	0.56	0.66	0.74
AR1($\tilde{s}_t$)	0.88	0.98	0.59	0.65	0.99	0.99	0.20	0.96	0.80
Standard dev ($\tilde{n}_t$)	0.022	0.021	0.015	0.015	0.011	0.005	0.007	0.003	0.002
Standard dev ($\tilde{u}_t$)	0.188	0.183	0.126	0.222	0.159	0.080	0.094	0.049	0.022
Standard dev ($\tilde{m}_{t,t+1}$)	0.085	0.052	0.09	0.117	0.061	0.045	0.055	0.023	0.016
Standard dev ($\tilde{s}_t$)	0.016	0.056	0.068	0.014	0.015	0.034	0.008	0.006	0.010
corr($\tilde{u}_t, \tilde{m}_{t,t+1}$)	0.81	0.997	0.89	0.92	0.93	0.78	0.75	0.83	0.70
corr($\tilde{u}_t, \tilde{P}_{t,t+1}$)	-0.93	-1.000	-0.74	-0.93	-0.98	-0.79	-0.83	-0.92	-0.70
corr($\tilde{n}_t, \tilde{s}_t$)	-0.16	0.997	0.94	-0.33	0.63	0.78	-0.14	0.49	0.68
corr($\tilde{m}_{t,t+1}, \tilde{s}_t$)	0.45	-0.99	-0.99	0.32	-0.35	-0.27	0.03	0.04	0.02
corr($\tilde{m}_{t,t+1}, \tilde{\delta}_{t,t+1}$)	0.91	0.86	0.60	0.88	0.98	0.95	0.67	0.98	0.83

Notes: AR1($\tilde{x}_t$) stands for the univariate AR(1) coefficient of the variable $\tilde{x}_t$. Standard dev ($\tilde{x}_t$) is the standard deviation of variable $\tilde{x}_t$. corr($\tilde{x}_t, \tilde{y}_t$) stands for the correlation between variables $\tilde{x}_t$ and $\tilde{y}_t$. Columns (iii) and (iv) entitled "Nanton" refers to our own steady state results in this paper using system (4.11). All of our analysis is based upon "Pool 0" – the official BLS definition of unemployment. Yashiv's results are however based upon Pool 1, which includes a subset of the economically inactive. "Exog only" means that the only shocks applied to the model are to the known exogenous variables: $\tilde{G}_t^X$,$\tilde{\delta}_{t,t+1}$and $\tilde{\beta}_t$.

Table 4.5 compares Yashiv's results with our own results when we allow for shocks only to the exogenous variables in the way we outlined above.[38] The results are not Wald statistics, but comparisons of model and data moments. (Yashiv's results are columns (i)-(iii), and our results are in columns (iv)-(ix)). We present two separate analyses of our own results – in the first of which (columns (iv)-(vi)), we have attempted to make the data stationary by differencing from the steady state (though table 4.4 suggests that this is unsuccessful), and the second of which ((vii)-(ix)) we instead use the HP-filter for the same purpose. Thus, columns (v) and (vi) are results from our reconstruction of Yashiv's model using data that has been log-differenced around *our* steady-state values. Columns (viii) and (ix) are results from our reconstruction of Yashiv's model using data that has been log differenced around the HP(1600)-filtered trend.

For each analysis we also present the dataset which is most comparable with the analysis (using the same data treatment method as for the simulated data). The idea is to show that our model is a good representation of that of Yashiv, because the moments that arise from simulations of our model have a similar relationship to the relevant data moments as they do in Yashiv's model, especially with respect to changes in the search cost parameter, γ.

We begin with the AR(1) coefficients for unemployment, $\tilde{u}_t$ hires, $\tilde{m}_{t,t+1}$ and the labour share $\tilde{s}_t$. Note that in Yashiv's analysis in column (i), and in our analysis in column (iv) where the data are log-differenced from the steady state, unemployment and hires are highly persistent in the data. The labour share is also highly persistent in the data by Yashiv's reckoning, and moderately persistent in the data according to ours. In simulations the persistence is maintained, whether search costs are assumed to be linear or cubic is essentially irrelevant. Thus our results for these variables in columns (v) and (vi) are similar to Yashiv's in (ii) and (iii). The results are somewhat different when we use an HP filter on the variables (columns (vii)-

[38] Table 4.5 augments panel (a) of table 7 on p931 of Yashiv (2006). We choose this table because here Yashiv shows the effect of changing the search cost parameter γ from $\gamma = 2$ to $\gamma = 0$ on chosen moments from his simulation. His chosen moments do not correspond neatly to the variables we have chosen to use in our version of his model. Specifically, $\tilde{m}_{t,t+1}$ and $\tilde{u}_t$ do not feature in our model although they are present in the table. However, we do have data on both of these variables and for simulated values they are easily obtained from the variables we do choose to model. Since $u_t = 1 - n_t$ it follows that $\tilde{u}_t \approx -\frac{n}{1-n}\tilde{n}_t$ and since $P_{t,t+1} = \frac{m_{t,t+1}}{u_t}$ it follows that $\tilde{m}_{t,t+1} = \tilde{P}_{t,t+1} + \tilde{u}_t = \tilde{P}_{t,t+1} - \frac{n}{1-n}\tilde{n}_t$.

(ix)).[39] In particular, the persistence of hires in the data falls to 0.56 and that of the labour share falls to just 0.20. Nevertheless, the *simulated* persistences in this case are reasonably high, and the effects of moving from cubic to linear search costs are fairly small.

Moving on to examine the standard deviations, a comparison between columns (i) and (iv) reveals that there are differences between Yashiv's data moments and our own – even when the data are treated in the same way (that is, log-differencing from the steady state values). The employment rate is nearly one-third less volatile in our dataset than in Yashiv's and the unemployment rate is nearly 20% more volatile. The hiring rate is nearly 40% more volatile in our dataset. The labour share is 12.5% less volatile in our dataset than in Yashiv's. The reported data correlations all have same signs and are in most cases reasonably similar between datasets – the notable exception is the correlation between employment and the labour share which is -0.16 in Yashiv's dataset and -0.33 in our (log-differenced) dataset.

There are many factors that could explain these differences. Most likely is the fact that we use the block-bootstrap re-sampling procedure along with univariate AR(1) processes to simulate the exogenous shocks, whereas Yashiv uses his reduced form VAR. These are not guaranteed to produce the same results. Differences may also arise from our raw data samples. Nevertheless, we would like to draw attention to the relationships between columns (i) (ii) and (iii), and columns (iv), (v) and (vi) in table 4.5 respectively. Consider the simulations in columns (ii) and (v) (which are for $\gamma = 2$), in particular the results for the rows containing standard deviations. It can be verified that the difference between the model and simulated volatilities both have the same sign in both Yashiv and our results. (That is, the simulated variables are less volatile than the data for all reported variables except for the labour share $\tilde{s}_t$). The same can be said of the reported covariances, with the exception of $cov(\tilde{m}_{t,t+1}, \tilde{\delta}_{t,t+1})$. In the cases of some moments (in particular $std(\tilde{m}_{t,t+1}), corr(\tilde{u}_t, \tilde{P}_{t,t+1})$ and $corr(\tilde{n}_t, \tilde{s}_t)$) , the differences between simulation and data are of comparable magnitude, and in others they are not. Now consider columns (iii) and (vi). In the case of the standard deviations and covariances, changing search costs so that the function is linear rather than cubic (that is, moving

[39] Why does the data treatment procedures (log deviations around the steady-state value or the HP-filter) affect our simulation results? The reason is that the treated exogenous variables are used in the block bootstrap, which is in turn used to determine the shocks that drive the simulations.

from $\gamma = 2$ to $\gamma = 0$) serves in the case of most reported moments to move the simulated moments in the same direction in our simulation as in Yashiv's. An examination of columns (vii)-(ix), which shows moments of data and simulated variables which have been log-differenced from an HP trend shows similar relationships, indicating that our reproduction of Yashiv's model is robust to the data-treatment procedure.

4.8 Results

In this section we present results for indirect inference testing of our version of Yashiv's model, with a full set of shocks. However, we make one more small change to the method. Rather than simulating exogenous variables *and* structural shocks as AR(1) processes, we choose to model the structural shocks only, so that the exogenous parts of the model are subsumed within the structural errors. To see how this works, note that we can collect the exogenous parts of the model (terms in $\tilde{G}_t^X$, $\tilde{\delta}_{t,t+1}$ and $\tilde{\beta}_t$) in (4.12) into the respective structural error terms, so that we can write:

$$\begin{aligned}\frac{\lambda}{G^X}\left[[-\sigma(\Omega+\gamma\Omega-1)+\gamma]\frac{1}{1-\sigma}\right]\tilde{P}_{t,t+1} - \frac{\lambda}{G^X}\left[\frac{\gamma}{1-n}\right]\tilde{n}_t - \beta\frac{1-\eta}{\eta}s\mathrm{E}_t[\tilde{s}_{t+1}] \\ -\left\{(\beta(1-\delta)\lambda-\beta P\lambda)\left[[-\sigma(\Omega+\gamma\Omega-1)+\gamma]\frac{1}{1-\sigma}\right]\right. \\ \left. -\beta P\lambda\right\}E_t\left[\tilde{P}_{t+1,t+2}\right] + \beta(1-\delta-P)\lambda\left[\frac{\gamma}{1-n}\right]E_t[\tilde{n}_{t+1}] = e_{P,t}^*\end{aligned} \quad (4.13)(a)$$

$$\begin{aligned}s\tilde{s}_t - \Bigg[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\eta(1 \\ -\alpha)\Theta\left(\frac{\phi v+(1-\phi)Qv}{n}\right)^{\gamma+1}\frac{(\gamma+1)}{1-\sigma}[1+\sigma(1-\Omega)] \\ +\frac{\eta P\lambda}{1-\sigma}[1-\sigma(\Omega+\gamma\Omega)+\gamma]\Bigg]\tilde{P}_{t,t+1} \\ +\Bigg(\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\eta(1 \\ -\alpha)\Theta\left(\frac{\phi v+(1-\phi)Qv}{n}\right)^{\gamma+1}\frac{(\gamma+1)}{1-n}+\frac{\eta P\lambda\gamma}{1-n}\Bigg)\tilde{n}_t = e_{s,t}^*\end{aligned} \quad (4.13)(b)$$

$$G^L\tilde{n}_{t+1} - (1-\delta-P)\tilde{n}_t - \frac{P(1-n)}{n}\tilde{P}_{t,t+1} = e^*_{n,t} \quad (4.13)(c)$$

Where:

$$e^*_{P,t} = e_{P,t} + \frac{\lambda}{G^X}\rho_{G^x}\tilde{G}^X_t - \beta\lambda\delta\rho_\delta\tilde{\delta}_{t,t+1} + \left(\beta\frac{1-\eta}{\eta}s + \beta\lambda(1-\delta-P)\right)\rho_\beta\tilde{\beta}_t$$

$$e^*_{s,t} = e_{s,t} \quad (4.14)$$

$$e^*_{n,t} = e_{n,t} - \delta\tilde{\delta}_{t,t+1}$$

Steps (2)-(5) of the indirect inference methodology are now amended as follows:

(2). *Compute the r structural errors in $\mathbf{e_t}$ under $\widehat{\boldsymbol{\theta}}$ using*

$$\mathbf{e}^*_\mathbf{t} = \mathbf{A}(\widehat{\boldsymbol{\theta}}^\mathbf{A})\mathbf{y_t} - \mathbf{C}(\widehat{\boldsymbol{\theta}}^\mathbf{C})E_t\mathbf{y_{t+1}}$$

using actual data for $\mathbf{y_t}$, $\mathbf{x_t}$ and if necessary, using a VAR to obtain $E_t\mathbf{y_{t+1}}$.

Now we have the error vector:

$$\mathbf{e}^*_\mathbf{t} = [e^*_{P,t} \quad e^*_{s,t} \quad e^*_{n,t}]^T$$

where $e^*_{P,t}$, $e^*_{s,t}$ and $e^*_{n,t}$ are equal to the expressions given(4.14). Note however, we do not *construct* them according to equation (4.14) – they are calculated by feeding data into $\mathbf{y_t}$ and $E_t\mathbf{y_{t+1}}$ and parameter values into $\mathbf{A}(\boldsymbol{\theta}^\mathbf{A})$ and $\mathbf{C}(\boldsymbol{\theta}^\mathbf{C})$ and performing the matrix subtraction.

As before, $r = 3$, and:

$$\mathbf{y_t} = \left[\tilde{P}_{t,t+1} \quad \tilde{s}_t \quad \tilde{n}_{t+1}\right]^T$$

(2). *Estimate univariate processes for each of the r structural errors in $\mathbf{e}^*_\mathbf{t}$ so as to obtain white noise residuals for each. At the end of step* (2) *one will have parameters $\hat{\rho}_1, \dots, \hat{\rho}_r$; and residuals $\hat{\varepsilon}_{1t}, \dots, \hat{\varepsilon}_{rt}$; $\forall t$ in the notation above.*

That is, we estimate the following univariate AR(1) equations:

$$e^*_{P,t} = \hat{\rho}_P e^*_{P,t} + \hat{\varepsilon}^*_{P,t}$$
$$e^*_{s,t} = \hat{\rho}_s e^*_{s,t-1} + \hat{\varepsilon}^*_{s,t}$$
$$e^*_{n,t} = \hat{\rho}_n e^*_{n,t-1} + \hat{\varepsilon}^*_{n,t}$$

(3). *Resample the residuals in order to create 1000 vector-bootstrapped samples of innovations. This means using each of the estimated univariate processes to generate 1000 simulated* *samples of structural errors, and 1000 samples of exogenous variables.*

We draw 1000 bootstrap pseudo-samples of the estimated residuals, $[\hat{\varepsilon}^{*}_{P,t} \quad \hat{\varepsilon}^{*}_{s,t} \quad \hat{\varepsilon}^{*}_{n,t}]$, and the estimated univariate AR(1) equations to create 1000 pseudo-samples of exogenous variables, of structural errors and of exogenous variables. The re-sampling is done as a block bootstrap, as detailed in Chapter 3 so as to preserve any correlation between the residuals in the original time series.

$$\hat{e}^{*j}_{P,t} = \hat{\rho}_P \hat{e}^{*j}_{P,t-1} + \hat{\varepsilon}^{*j}_{P,t}$$
$$\hat{e}^{*j}_{s,t} = \hat{\rho}_s \hat{e}^{*j}_{s,t-1} + \hat{\varepsilon}^{*j}_{s,t}$$
$$\hat{e}^{*j}_{n,t} = \hat{\rho}_n \hat{e}^{*j}_{n,t-1} + \hat{\varepsilon}^{*j}_{n,t}$$
$$j = 1, \dots, 1000$$

(4). *Solve the model under* $\widehat{\boldsymbol{\theta}}$, *in terms of the structural errors and exogenous variables. Use the simulated data to generate 1000 samples of endogenous variables.*

The model solution now has the form:

$$\hat{\mathbf{y}}^{j}_{\mathbf{t}} = \boldsymbol{\Upsilon}_{\mathbf{1}}(\widehat{\boldsymbol{\theta}})[\hat{\mathbf{y}}^{j}_{\mathbf{t-1}}] + \boldsymbol{\Upsilon}_{\mathbf{2}}(\widehat{\boldsymbol{\theta}})\begin{bmatrix}\hat{\varepsilon}^{*j}_{P,t}\\ \hat{\varepsilon}^{j*}_{s,t}\\ \hat{\varepsilon}^{*j}_{n,t}\end{bmatrix}$$

$$j = 1, \dots, 1000$$

Where $\boldsymbol{\Upsilon}_{\mathbf{1}}$ and $\boldsymbol{\Upsilon}_{\mathbf{2}}$ are both 3x3 matrices.

We then follow the remaining steps to implement the indirect inference procedure, under the chosen auxiliary models. Before we present the indirect inference results, however, we first present the analogous table to 4.5, table 4.6 which differs from table 4.5 in that the shocks applied to the model are the bootstrapped structural shocks $\mathbf{e}^{*}_{\mathbf{t}} = [e^{*}_{P,t} \quad e^{*}_{s,t} \quad e^{*}_{n,t}]^{T}$ as opposed to only the shocks to the exogenous variables that were used as an approximation to Yashiv's method.

Table 4.6: Comparison of summary statistics from Yashiv's (2006) analysis with those from our versions of the model, with shocks to the structural equations only.

	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
Pool:	0			1					
	Yashiv			*Nanton*					
	Data	*Model*		*Data*	*Model*		*Data*	*Model*	
Shocks:		*Exog. only*	*Exog. only*	*Search costs:*	*Endog. only*	*Endog. only*	*Search costs*	*Endog. only*	*Endog. only*
Search costs:		*γ=2*	*γ=0*	*Shocks:*	*γ=2*	*γ=0*	*Shocks:*	*γ=2*	*γ=0*
Data treatment:	*Steady-state de-trending*			*Steady-state de-trended*			*HP-filter(1600) de-trended*		
AR1($\tilde{u}_t$)	0.97	0.98	0.88	0.97	0.99	0.97	0.92	0.97	0.90
AR1($\tilde{m}_{t,t+1}$)	0.85	0.99	0.46	0.89	0.81	0.72	0.56	0.58	0.65
AR1($\tilde{s}_t$)	0.88	0.98	0.59	0.65	0.99	0.97	0.20	0.97	0.90
Standard dev($\tilde{n}_t$)	0.022	0.021	0.015	0.015	0.192	0.024	0.007	0.204	0.007
Standard dev ($\tilde{u}_t$)	0.188	0.183	0.126	0.014	2.829	0.353	0.094	3.011	0.104
Standard dev ($\tilde{m}_{t,t+1}$)	0.085	0.052	0.09	0.117	0.513	0.160	0.055	0.830	0.075
Standard dev ($\tilde{s}_t$)	0.016	0.056	0.068	0.014	1.837	0.318	0.008	2.372	0.092
corr($\tilde{u}_t, \tilde{m}_{t,t+1}$)	0.81	0.997	0.89	0.92	-0.51	-0.13	0.75	-0.41	0.44
corr($\tilde{u}_t, \tilde{P}_{t,t+1}$)	-0.93	-1.000	-0.74	-0.93	-0.99	-0.91	-0.83	-0.97	-0.71
corr($\tilde{n}_t, \tilde{s}_t$)	-0.16	0.997	0.94	-0.33	0.89	0.88	-0.14	0.79	0.58
corr($\tilde{m}_{t,t+1}, \tilde{s}_t$)	0.45	-0.99	-0.99	0.32	0.84	0.49	0.03	0.88	0.38
corr($\tilde{m}_{t,t+1}, \tilde{\delta}_{t,t+1}$)	0.91	0.86	0.60	0.88	-	-	0.67	-	-

Notes: AR1($\tilde{x}_t$) stands for the univariate AR(1) coefficient of the variable $\tilde{x}_t$. Standard dev ($\tilde{x}_t$) is the standard deviation of variable $\tilde{x}_t$. corr($\tilde{x}_t, \tilde{y}_t$) stands for the correlation between variables $\tilde{x}_t$ and $\tilde{y}_t$. Columns (iii) and (iv) entitled "Nanton" refers to our own steady state results in this paper using system (4.11). All of our analysis is based upon "Pool 0" – the official BLS definition of unemployment. Yashiv's results are however based upon Pool 1, which includes a subset of the economically inactive. "Endog only" means that the only shocks applied to the model are structural shocks – one for each equation. These are defined such that they include any changes in exogenous variables which may affect the system. See equation (4.14).

It is clear from looking at table 4.6 that applying shocks to the endogenous parts of the three structural equations of the model in general does not produce similar moments as applying shocks to only the exogenous variables, and so does not produce moments that are in any way comparable to the results of Yashiv. In particular, the imposition of cubic search costs $(\gamma = 2)$ massively increases the standard deviations of the reported variables, to magnitudes of many times their sample counterparts. (Table 4.6, columns (v) and (viii)). The volatility is reduced when linear search costs are once again imposed $(\gamma = 0)$. In the case where the variables are de-trended by the steady state, the model volatilities of $\tilde{u}_t$ and $\tilde{s}_t$ are still far from those in the data even when $\gamma = 0$. (table 4.6, column (vi)). The model volatilities are close to those in the data for $(\gamma = 0)$ in the HP-filtered version of them model (column (ix)). However, the results suggest the opposite finding to Yashiv (2006) – that imposing cubic search costs make the model's performance worse rather than better.

The issue appears to be that the structural equations (4.13)(a)-(4.13)(c) fit the data poorly, producing large structural errors, and large estimated innovations to those errors. These produce volatile series of endogenous variables when the model is solved and simulated.

The results are similarly disappointing for the model covariances. For example $\mathrm{corr}(\tilde{u}_t, \tilde{m}_t)$ and $\mathrm{corr}(\tilde{n}_t, \tilde{s}_t)$both have the wrong sign under $(\gamma = 2)$. Only $\mathrm{corr}\left(\tilde{u}_t, \tilde{P}_{t,t+1}\right)$ is reasonably close to its data counterpart.

We now present more detailed results on the model's performance, using the method of indirect inference. As discussed above, table 4.4 suggests that taking the log-difference from Yashiv's steady state is not sufficient to induce stationarity. We therefore report results only for data de-trended using the HP-filter. It is somewhat regrettable, since it would have useful to follow Yashiv's data treatment procedures more exactly throughout the analysis. However, table 4.6 presents evidence that the steady-state de-trended and HP-filtered model behave in broadly the same way with respect to changes in the search cost parameter, γ.

Table 4.7 shows results using the standard deviations of the endogenous variables as the auxiliary model. We test both the version of the model with exogenous shocks only and the version with only shocks to the structural equations. Results are provided for versions of the model with cubic search costs ($\gamma = 2$) and with linear search costs ($\gamma = 0$).

For the data we report the standard deviations of endogenous variables (column (i)), and for the simulations we report the mean and the 95% distribution of the standard deviations. We also report the Wald statistic for the null hypothesis that the true data was generated by the model, and the 95th percentile of the Wald under the assumption that the null is true. At the bottom of table 4.7 we also report the equivalent t-statistic, which is calculated so that it would equal 1.645 at the 95th percentile of the Wald.[40]

We note first that for each version of the model the Wald is far outside the 95% critical value. For both treatments of shocks the assumption that $\gamma = 2$ does somewhat better than the assumption that $\gamma = 0$ in the sense of having a smaller Wald statistic relative to the critical value, As noted above, when $\gamma = 2$ (when search costs are cubic) the model standard deviations are far closer to the data when shocks are only applied to the exogenous variables. (Compare columns (ii) and (iv) with column (i).) Out of all the specifications the only data standard deviation that fits into the 95% distribution of model standard deviations is that of employment $\tilde{n}_t$when $\gamma = 0$ in the specification with shocks to each structural equation (column (iv)). Note however that *overall* this model still does worse than the same shock specification with $\gamma = 2$ (column (iii)).

[40] The formula for this is $t = 1.645 \times \left(\frac{\sqrt{2\times WALD} - \sqrt{2\times(K-1)}}{\sqrt{2\times WALD_{95}} - \sqrt{2\times(K-1)}}\right)$, where $WALD$ is the Wald statistic, $WALD_{95}$ is the 95th percentile value of the Wald statistic under the null hypothesis that the data was generated by the model and K is the number of auxiliary model parameters.

Table 4.7: Indirect inference testing results under Yashiv's parameters using standard deviations of endogenous variables as the auxiliary model.

		(i)	(ii)	(iii)	(iv)	(v)
		Data:	*Model*			
	Shocks:		*Exog. only*		*Endog. only*	
	Search costs:		*γ=2*	*γ=0*	*γ=2*	*γ=0*
			HP-filter de-trended			
$std(\tilde{P}_{t,t+1})$	*Mean*	0.064	0.0326	0.0156	3.4387	0.0969
	95% dist.	-	[0.028, 0.037]	[0.012,0.020]	[2.510, 4.529]	[0.077, 0.120]
$std(\tilde{s}_t)$	*Mean*	0.0075	0.0055	0.0101	2.3717	0.0919
	95% dist.	-	[0.005, 0.006]	[0.008, 0.013]	[1.920, 2.908]	[0.079, 0.106]
$std(\tilde{n}_t)$	*Mean*	0.0071	0.0033	0.0015	0.2038	0.007
	95% dist.	-	[0.003, 0.004]	[0.001, 0.002]	[0.144, 0.274]	**[0.005, 0.009]
Wald:		-	***391	***1,366,400	***118	***343
P-value:		-	0.000	0.000	0.000	0.000
t-stat.:		-	22.2	1412	11.3	19.9

Notes: *** means statistically significant at the 1% level. $std(\tilde{x}_t)$ denotes the standard deviation of variable $\tilde{x}_t$. "Exog only" means that the only shocks applied to the model are to the known exogenous variables: $\tilde{G}_t^X$,$\tilde{\delta}_{t,t+1}$and $\tilde{\beta}_t$. "Endog only" means that the only shocks applied to the model are structural shocks – one for each equation. These are defined such that they include any changes in exogenous variables which may affect the system. See equation (4.14). The rows entitled "Mean" in columns (ii)-(v) denote the mean standard deviation of the given variable across 1000 simulations. In column (i) it simply denotes the standard deviation of the data. The rows entitled "95% dist" are the 95% confidence intervals of model standard deviations for the requisite variable. ** denotes that the model standard deviation (column (i)) falls inside the model's 95% confidence interval for the same statistic.

To summarize table 4.7, all of the models are rejected by the data when the auxiliary model is the vector of standard deviations of endogenous variables. It is noteworthy that whereas the 95th percentiles of the Wald statistics under the null are of similar magnitude, the Walds themselves are wildly different. We return to this issue in the discussion section.

Table 4.8 contains indirect inference results based on an auxiliary model of VAR(1) coefficients estimated on the endogenous variables of the model. In contrast to the analysis in table 4.7, we use only the version of the model with shocks to the structural equations, since applying shocks to the exogenous variables only results in simulated samples of endogenous variables which result in singular or near-singular VARs when the auxiliary model is applied to them.

The results are disappointing. Both versions of the model with respect to search costs $(\gamma = 2)$ and $(\gamma = 0)$ are rejected, although the Wald is smaller for the case in which $\gamma = 0$, that is, for linear search costs.
Interestingly, for $\gamma = 2$, many of the data coefficients fall within the 95% model bounds, despite the overall rejection of the model. For $\gamma = 0$ which corresponds to the case of linear search costs, many of the 95% bounds of the auxiliary model's parameters are tighter, so that more estimated data parameters lie outside of these bounds.

Table 4.8: Indirect inference testing results under Yashiv's parameters using the VAR(1) coefficients of the endogenous variables as the auxiliary model.

		(i)	(ii)	(iii)
		Data	*Model*	
Shocks:			*Endog. only*	
Search costs:			*γ=2*	*γ=0*
		HP-filter de-trended		
$\mathrm{var}(\tilde{P},\tilde{P})$	*Mean*	0.28	0.27	1.5831
	95% dist		**[-4.9, 5.66]	[1.22, 1.97]
$\mathrm{var}(\tilde{P},\tilde{s})$	*Mean*	-0.03	0.52	-1.0822
	95% dist		**[-2.29, 3.29]	[-1.42, -0.81]
$\mathrm{var}(\tilde{P},\tilde{n})$	*Mean*	4.65	5.6	0.9148
	95% dist		**[-55.83, 65.29]	[-1.47, 3.42]
$\mathrm{var}(\tilde{s},\tilde{P})$	*Mean*	0	0.3	0.9029
	95% dist		[-5.57, 6.38]	[0.46, 1.39]
$\mathrm{var}(\tilde{s},\tilde{s})$	*Mean*	0.65	0.6	-0.463
	95% dist		**[-2.63, 3.71]	[-0.91, -0.11]
$\mathrm{var}(\tilde{s},\tilde{n})$	*Mean*	0.14	-3.53	1.0415
	95% dist		**[-73.82, 64.5]	**[-1.82, 3.95]
$\mathrm{var}(\tilde{n},\tilde{P})$	*Mean*	0.01	-0.03	0.0373
	95% dist		**[-0.21, 0.16]	[0.02, 0.06]
$\mathrm{var}(\tilde{n},\tilde{s})$	*Mean*	-0.1	0.04	-0.0392
	95% dist		**[-0.06, 0.13]	**[-0.06, -0.02]
$\mathrm{var}(\tilde{n},\tilde{n})$	*Mean*	0.79	1.06	0.8236
	95% dist		**[-1.09, 3.19]	**[0.71, 0.93]
Wald:		-	***484,920	***1005
P-value:		-	0.000	0.000
t-stat.:		-	807	33.6

Notes: In the notation of this table, $\mathrm{var}(\tilde{y},\tilde{x})$ denotes the VAR(1) coefficient that associates the variable $\tilde{y}$ with the first lag of variable $\tilde{x}$. "Endog only" means that the only shocks applied to the model are structural shocks – one for each equation. These are defined such that they include any changes in exogenous variables which may affect the system. See equation (4.14). The rows entitled "Mean" in columns (ii)-(v) denote the mean VAR(1) coefficient of the relevant variables across 1000 simulations. In column (i) it simply denotes the same VAR(1) coefficient of the data. The rows entitled "95% dist" are the 95% confidence intervals of model VAR(1) coefficients for the requisite variable. ** denotes that the model standard deviation (column (i)) falls inside the model's 95% confidence interval for the same statistic. ***Indicates that the Wald statistics are significantly different at the 1% level.

As the fit of both models is very poor under both type of auxiliary model, we do not bother to use a third auxiliary model combining both the standard deviations of endogenous variables and their VAR(1) coefficients, as we did in Chapter 3. Instead we move straight away to the discussion section.

4.9 Discussion

Chapter 1 reviewed literature on aggregate search and matching models, of the type developed by Mortensen and Pissarides. We located Yashiv's (2006) paper in the part of the literature that attempts to resolve the shortcomings of the standard model by modifying the way in which search costs are incurred by firms. The main relevant findings of this literature with respect to Yashiv's paper are as follows:

(i) By reallocating search costs away from vacancy costs and towards a fixed hiring cost per worker, the elasticity of key labour market variables (market tightness, vacancies, unemployment and the job finding probability) with respect to productivity increases significantly. (Pissarides (2009)). This is because the fixed hiring cost reduces the dampening effect of a higher vacancy-unemployment ratio through the expected vacancy duration channel.

(ii) For the channel to be operative, the costs must be sunk. If the costs are not sunk then they are merely equivalent to a reduction in productivity, some of which is passed from the firm to newly hired workers in the form of lower wages. This does not have a large effect on the elasticity of the labour market variables with respect to productivity. (Silva and Toledo (2013)).

We now consider Yashiv's stated mechanism for inducing the required volatility via search costs.

First of all, we must note that Yashiv's model contains both a vacancy cost and a hiring cost. Recall the cost function:

$$\Gamma_t = \frac{\Theta}{1+\gamma}\left(\frac{\phi V_t + (1-\phi)Q_t V_t}{N_t}\right)^{\gamma+1} F_t \qquad (4.1)(\mathrm{b})$$

The term $\frac{(1-\phi)Q_t V_t}{N_t} = \frac{(1-\phi)m_t}{n_t}$ represents a hiring cost. The term $\frac{\phi V_t}{N_t}$ is a more standard vacancy cost.

Yashiv's search costs are indeed sunk. When search costs are not sunk, they are available for bargaining over, and the result is that they enter the threat point of the firm and reducing wages of new entrants. (Silva and Toledo (2013, p838)). Conversely, when they are sunk they implicitly raise the bargaining power of the worker, and so appear with a positive sign in the equilibrium wage equation. In Yashiv's model however, search costs do not appear in the Nash bargaining equation (equation (4.9)), and these clearly raise the equilibrium labour share (see equation (4.10)(d)). The search costs in Yashiv's model are definitely sunk, as in the usual case.

However, the hiring costs are definitely not fixed. Instead they fluctuate pro-cyclically, although the effect is somewhat dampened due to the fact that employment is in the denominator. (Note that hires are 7.8 times more volatile than employment in our data and between 5.5 and 9 times more volatile in our model with only exogenous variation- see columns (iv), (v) and (vi) of table 4.5. This confirms that hiring costs are still pro-cyclical despite the countervailing effect of employment in the search cost function.) Furthermore, by the same reasoning, vacancy costs are non-linearly increasing with vacancies.[41] This suggests that the dampening effect of the expected duration channel should be particularly severe following a positive productivity shock.

The channel by which Yashiv purports to solve the Shimer puzzle is therefore different from that chosen by Pissarides and Silva and Toledo. As Yashiv reports, the mechanism is that the convexity of both vacancy and hiring costs when $\gamma = 2$ makes firms slow to adjust their vacancy stock in response to shocks. This induces extra persistence in vacancies, and raises their volatility due to the standard relationship between persistence and standard deviation. The mechanism works because Yashiv models vacancies as a control variable, and persistence and volatility is transmitted to the rest of the model by similar means. (Yashiv (2006) p39, footnote 18).

Table 4.9 below shows univariate AR(1) coefficients for the HP-filtered data (column (i)) and for the different shock-specifications and search cost parameter values used in the model (columns (ii)-(iv)). The table shows that raising the search cost

[41] Here we do not have vacancy data but the Help-Wanted Index suggests that they are many times more volatile than employment. For example Shimer (2005, p28) puts the quarterly standard deviation of the HP-de-trended Help-Wanted Index at 0.202 over the period 1951-2003.

parameter γ from $\gamma = 0$ to $\gamma = 2$ does indeed raise the persistence of key variables. (The exception to this is hires, which somewhat decrease in persistence when search costs are raised). However, unemployment and the labour share both increase in persistence when γ is raised. This suggests that the mechanism in our model works as it should.

Table 4.9: AR(1) coefficients

	(i)	(ii)	(iii)	(iv)	(v)
	Data	*Model*		*Model*	
		Exog. Only		*Endog. Only*	
		γ=2	*γ=0*	*γ=2*	*γ=0*
			HP Filter De-trended		
$AR1(\tilde{u})$	0.9181	0.9578	0.8085	0.9648	0.869
$AR1(\tilde{m})$	0.6965	0.6357	0.7553	0.5786	0.6505
$AR1(\tilde{s})$	0.6282	0.9578	0.8085	0.9648	0.896

Notes: $AR1(\tilde{x})$ is the univariate AR(1) coefficient of variable $\tilde{x}$. "Exog only" means that the only shocks applied to the model are to the known exogenous variables: $\tilde{G}_t^X$, $\tilde{\delta}_{t,t+1}$ and $\tilde{\beta}_t$. "Endog only" means that the only shocks applied to the model are structural shocks – one for each equation. These are defined such that they include any changes in exogenous variables which may affect the system. See equation (4.14).

Nonetheless, we have shown that the presence of the increased persistence is not sufficient to allow the model to fit the data based on our indirect inference testing procedure, even when – following Yashiv – only shocks to the known exogenous variables are included. Yashiv's structural parameters cannot produce a fit of his model to data for even the smallest of our auxiliary models - the standard deviations of endogenous variables. Results are no better with respect to the dynamics – as expected since the power of the Wald statistic increases with the number of auxiliary model parameters. (Le, Meenagh, Minford and Wickens, p17).

What is the problem? Firstly, we consider the inability of the model to fit even the data standard deviations, when only the shocks to the exogenous variables were considered. None of the data standard deviations fall inside the 95% model distributions in columns (ii) and (iii) of table 4.7. The results suggest that more volatility is needed to fit the data, but the results are still close enough that a different set of parameter values might allow the model to fit based on the same auxiliary model. As they are, the data standard deviations clearly lie on the fringes of our model's joint distribution of parameters. There are however some differences between our model and that of Yashiv. The model used in this chapter is somewhat smaller, we reduce the model down to three equations in which all the featuring variables are observable. We also HP-filter the data, rather than log-differencing it from Yashiv's steady state values, in order to achieve stationarity. Finally, there are unavoidable differences in some of the data used, due to issues of availability. These discrepancies may mean that a different set of parameters are needed to produce the same results, given the same modelled shocks. From this point of view there is a good chance that indirect inference estimation might be successful in finding a better fitting set of parameters according to the Wald criterion, which offset some of these differences.

However, the rest of our results show deeper problems. When we generate stochastic variation in the model using bootstrapped structural errors which we assume to include the exogenous variable shocks, we find that the structural equations in fact fit the data very poorly. The result is very large structural errors and large estimated structural innovations, which in turn generates large amounts of volatility when the structural errors are re-sampled. This is then transferred to the model during simulation. The result is massive standard deviations in the simulated endogenous variables as well as poorly fitting VAR coefficients. In Chapter 3 we argued that the Shimer puzzle related literature was an example of the 'puzzles' methodology, as described by Le, Minford and Wickens (2010). Yashiv's paper is no different in this case. The discrepancy between these results of ours and Yashiv's highlights an important difference between the puzzles methodology and the methodology of indirect inference. In the latter, one treats the structural error in the model as a factor generating variation in the endogenous variables, hence it can be subject to bootstrapping and simulation. The structural error can also be redefined (as it was in our case) to include variation in chosen exogenous variables. The model then reverts to the internal responses of the endogenous variables. Seen in this light, the puzzles method of applying shocks only to the chosen exogenous variables and ignoring the structural error term looks like an arbitrary restriction – it is equivalent to assuming that the structural equations hold exactly. That the volatility of the structural shocks turned out to be so large in our case suggests that this is not a good assumption.

There is still a potential reconciliation however, which we leave for further work. The obvious next step in our procedure would be to estimate Yashiv's model with unrestricted structural parameters, to see whether chosen data moments are consistent with the joint distribution of moments implied by the model, since it is still the case that a different set of parameters could improve the fit of the structural equations greatly.

4.10 Conclusion

We have shown in Chapters 3 and 4 that the Mortensen Pissarides model has a hard time fitting the data using indirect inference. In Chapter 3, we find as expected that the basic model fails to fit the data for *any* auxiliary model under Shimer's (2005) parameters. The three candidate auxiliary models are standard deviations of the endogenous variables, VAR(1) coefficients for the endogenous variables and a combination of these. Although the model is rejected by the Wald test, our bootstrapped simulated version of the model does display far more volatility than Shimer's. However, we also show that this is due to the extra volatility that results from treating the structural errors as shocks, which is a standard part of our indirect inference procedure. Upon estimating the basic model, we obtain values which appear to fit the data based on an auxiliary model of standard deviations only. The estimated results are however corner solutions, and therefore at the edge of our theoretically imposed constraints. We take this as a warning that there are indeed serious specification issues with this basic model, as has already been frequently suggested in the literature.

Moving on to Chapter 4, we test Yashiv's (2006) model with the indirect inference procedure, using Yashiv's chosen steady-state parameters. We choose Yashiv's model, as it is an example of a discrete time model which tries to introduce more volatility to the labour market variables than is allowed for by the canonical Mortensen-Pissarides model. The proposed cure is by the introduction of a cubic (rather than linear) search cost function faced by firms, which depends both on vacancies and on hires.

Again, our results show that whether the structural error shocks are used, or whether we simply shock the chosen exogenous variables (as Yashiv does) is an important issue. In the latter case, the properties of our reproduction of Yashiv's model are similar to the original. In the former case however, we derive large structural errors which cause large amounts of extra volatility, pushing the average simulated volatilities far from the data. Most importantly however, in both cases the model is rejected based on the Wald test for each auxiliary model used.

We have shown that the indirect inference method allows us to simulate the joint distribution of moments of interest, and that this is a very different from the standard moment-by-moment comparison approach that is widely taken by the literature. Because the joint distribution is derived from a simulation that takes into account correlation between the structural innovations, fitting the data based on moments is potentially challenging. The Wald test removes the ambiguity that arises from simply examining the simulated and target moments, and deciding whether they are "close". With many suggested solutions to the Shimer puzzle, we suggest that in general the wider adoption of indirect inference testing could sharpen the empirical analysis by allowing for far greater discrimination between models.

Note: As we have made clear throughout Chapter 4, the model presented here is essentially the same as that presented in Yashiv (2006), with a few adjustments to better adapt it for indirect inference. Mathematical details of the original model are available in the technical appendix to Yashiv (2006), which is available on request from Professor Yashiv himself.

Appendix 4.1: Deriving an expression for W_t *- solving* (4.9)*:*

$$W_t = \text{argmax}\ (J_t^N - J_t^U)^{\xi}(J_t^F - J_t^V)^{1-\xi} \tag{4.9}$$

The first-order condition is:

$$\xi \frac{\frac{\partial(J_t^N - J_t^U)}{\partial W_t}}{J_t^N - J_t^U} + (1-\xi)\frac{\frac{\partial(J_t^F - J_t^V)}{\partial W_t}}{J_t^F - J_t^V} = 0$$

Take the required differences:

$$J_t^N - J_t^U = W_t - b_t + E_t\beta_{t+1}\left[\left(1 - \delta_{t,t+1} - P_{t,t+1}\right)(J_{t+1}^N - J_{t+1}^U)\right]$$

Note therefore, that:

$$\frac{\partial(J_t^N - J_t^U)}{\partial W_t} = 1$$

Since, due to the free entry condition:

$$J_t^V = 0$$

It follows that:

$$J_t^F - J_t^V = J_t^F = \frac{\partial F_t}{\partial N_t} - W_t - N_t\frac{\partial W_t}{\partial N_t} - \frac{\partial \Gamma_t}{\partial N_t} + E_t\beta_{t+1}\left[\delta_{t,t+1}J_{t+1}^V + \left(1 - \delta_{t,t+1}\right)J_{t+1}^F\right]$$

And:

$$\frac{\partial(J_t^F - J_t^V)}{\partial W_t} = \frac{\partial J_t^F}{\partial W_t} = -1$$

Start by substituting the derivatives only into the first-order condition:

$$\xi\frac{1}{J_t^N - J_t^U} - (1-\xi)\frac{1}{J_t^F - J_t^V} = 0$$

Rearranging

$$\xi(J_t^F - J_t^V) = (1-\xi)(J_t^N - J_t^U)$$

Letting $J_t^V = 0$:

$$\xi J_t^F = (1-\xi)(J_t^N - J_t^U)$$

$$\xi\left[\frac{\partial F_t}{\partial N_t} - W_t - N_t\frac{\partial W_t}{\partial N_t} - \frac{\partial \Gamma_t}{\partial N_t} + E_t\beta_{t+1}\{(1-\delta_{t,t+1})J_{t+1}^F + \delta_{t,t+1}J_{t+1}^V\}\right]$$
$$= (1-\xi)\left[W_t - b_t + E_t\beta_{t+1}\left[(1-\delta_{t,t+1} - P_{t,t+1})(J_{t+1}^N - J_{t+1}^U)\right]\right]$$

Of course $J_{t+1}^V = 0$

$$\xi\left[\frac{\partial F_t}{\partial N_t} - N_t\frac{\partial W_t}{\partial N_t} - \frac{\partial \Gamma_t}{\partial N_t} + E_t\beta_{t+1}\{(1-\delta_{t,t+1})J_{t+1}^F\}\right]$$
$$= W_t + (1-\xi)\left[-b_t + E_t\beta_{t+1}\left[(1-\delta_{t,t+1} - P_{t,t+1})(J_{t+1}^N - J_{t+1}^U)\right]\right]$$

Use the first order condition to eliminate J_{t+1}^F. Since $J_{t+1}^V = 0$ by free entry:

$$\xi J_{t+1}^F = (1-\xi)(J_{t+1}^N - J_{t+1}^U)$$

Substituting into the above can be shown to give:

$$W_t + \xi N_t\frac{\partial W_t}{\partial N_t} = \xi\left[\frac{\partial F_t}{\partial N_t} - \frac{\partial \Gamma_t}{\partial N_t}\right] + (1-\xi)b_t + E_t\beta_{t+1}\left[P_{t,t+1}(1-\xi)(J_{t+1}^N - J_{t+1}^U)\right]$$

Now use the first order condition once again, to eliminate $(1-\xi)(J_{t+1}^N - J_{t+1}^U)$:

$$\xi J_{t+1}^F = (1-\xi)(J_{t+1}^N - J_{t+1}^U)$$

$$W_t + \xi N_t\frac{\partial W_t}{\partial N_t} = \xi\left[\frac{\partial F_t}{\partial N_t} - \frac{\partial \Gamma_t}{\partial N_t}\right] + (1-\xi)b_t + \xi E_t\beta_{t+1}\left[P_{t,t+1}J_{t+1}^F\right]$$

$$W_t + \xi N_t\frac{\partial W_t}{\partial N_t} = \xi\left[\frac{\partial F_t}{\partial N_t} - \frac{\partial \Gamma_t}{\partial N_t} + E_t\beta_{t+1}\left[P_{t,t+1}J_{t+1}^F\right]\right] + (1-\xi)b_t$$

The next step is to solve for $E_t\beta_{t+1}\left[P_{t,t+1}J_{t+1}^F\right]$ in the wage equation. From equation (4.7)(a) for J_t^F:

$$J_{t+1}^F = \frac{\partial F_{t+1}}{\partial N_{t+1}} - W_{t+1} - N_{t+1}\frac{\partial W_{t+1}}{\partial N_{t+1}} - \frac{\partial \Gamma_{t+1}}{\partial N_{t+1}} + E_{t+1}\beta_{t+2}\left[\delta_{t+1,t+2}J_{t+2}^V + (1-\delta_{t+1,t+2})J_{t+2}^F\right]$$

And from the first order condition for employment:

$$\Lambda_t = \mathrm{E}_t\beta_{t+1}\left[\frac{\partial F_{t+1}}{\partial N_{t+1}} - W_{t+1} - N_{t+1}\frac{\partial W_{t+1}}{\partial N_{t+1}} - \frac{\partial \Gamma_{t+1}}{\partial N_{t+1}}\right]$$
$$+ E_t\beta_{t+1}(1-\delta_{t+1,t+2})\Lambda_{t+1} \qquad (4.6)(\mathrm{b})$$

It can be shown that:

$$\Lambda_t = \mathrm{E}_t\beta_{t+1}J_{t+1}^F$$

Hence the wage equation may be written as:

$$W_t + \xi N_t \frac{\partial W_t}{\partial N_t} = \xi \left[\frac{\partial F_t}{\partial N_t} - \frac{\partial \Gamma_t}{\partial N_t} + P_{t,t+1}\Lambda_t\right] + (1-\xi) b_t$$

Or:

$$W_t = \xi \left[\frac{\partial F_t}{\partial N_t} - \frac{\partial \Gamma_t}{\partial N_t} - \xi N_t \frac{\partial W_t}{\partial N_t} + P_{t,t+1}\Lambda_t\right] + (1-\xi) b_t$$

Note that this is a differential equation in W_t and $\frac{\partial W_t}{\partial N_t}$. The differential equation arises from having wages dependent on employment. We do not provide the details of the solution here, but we have them and they are available on request. The result is the following equation for wages:

$$W_t = \xi \left[(1-\alpha) A \left(\frac{K}{N}\right)^{\alpha} \left(\left[\frac{1}{1-\alpha\xi}\right] + \Theta\left(\frac{\phi V_t + (1-\phi) Q_t V_t}{N_t}\right)^{\gamma+1} \left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\right]\right) + P_{t,t+1}\Lambda_t\right] + (1-\xi) b_t$$

Appendix 4.2: Solving for equations (4.10)(a)-(4.10)(c).

$$\frac{\lambda_t}{G_{t+1}^X} = \mathrm{E}_t \beta_{t+1} \left[\frac{1-\eta}{\eta} s_{t+1} - P_{t+1,t+2}\lambda_{t+1}\right] + E_t \beta_{t+1} \big(1 - \delta_{t+1,t+2}\big) \lambda_{t+1} \tag{4.10)(a}$$

$$\Theta(\phi + (1-\phi) Q_t) \left(\frac{\phi V_t + (1-\phi) Q_t V_t}{N_t}\right)^{\gamma} = Q_{t,t+1}\lambda_t \qquad (4.10)(b)$$

$$P_{t,t+1} = \mu \left(\frac{v_t}{1-n_t}\right)^{1-\sigma} \qquad (4.10)(c)$$

(Since $1 - n_t = u_t$ this follows from (4.3)).

$$s_t = \eta\left[(1-\alpha)\left(\left[\frac{1}{1-\alpha\xi}\right] + \Theta\left(\frac{\phi v_t + (1-\phi)Q_t v_t}{n_t}\right)^{\gamma+1}\left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\right]\right) + P_{t,t+1}\lambda_t\right] \quad (4.10)(\mathrm{d})$$

$$n_{t+1} = \left(1-\delta_{t,t+1}\right)n_t + P_{t,t+1}(1-n_t) \quad (4.10)(\mathrm{e})$$

A.4.2.1

Solving for:

$$\frac{\lambda_t}{G^X_{t+1}} = \mathrm{E}_t\beta_{t+1}\left[\frac{1-\eta}{\eta}s_{t+1} - P_{t+1,t+2}\lambda_{t+1}\right] + E_t\beta_{t+1}\left(1-\delta_{t+1,t+2}\right)\lambda_{t+1} \quad (4.10)(\mathrm{a})$$

Begin with the inter-temporal condition:

$$\Lambda_t = \mathrm{E}_t\beta_{t+1}\left[\frac{\partial F_{t+1}}{\partial N_{t+1}} - W_{t+1} - N_{t+1}\frac{\partial W_{t+1}}{\partial N_{t+1}} - \frac{\partial \Gamma_{t+1}}{\partial N_{t+1}}\right] + E_t\beta_{t+1}\left(1-\delta_{t+1,t+2}\right)\Lambda_{t+1} \quad (4.6)(\mathrm{b})$$

In Appendix 4.1 it was shown that:

$$W_t = \xi\left[\frac{\partial F_t}{\partial N_t} - \frac{\partial \Gamma_t}{\partial N_t} - \xi N_t \frac{\partial W_t}{\partial N_t} + P_{t,t+1}\Lambda_t\right] + (1-\xi)b_t \tag{4.6)(b}$$

We now assume that benefits are given by $b_t = \tau W_t$ - in other words τ is the replacement rate out wages that are paid in case of unemployment. This gives:

$$W_t = \xi\left[\frac{\partial F_t}{\partial N_t} - \frac{\partial \Gamma_t}{\partial N_t} - \xi N_t \frac{\partial W_t}{\partial N_t} + P_{t,t+1}\Lambda_t\right] + (1-\xi)\tau W_t$$

$$W_t = \frac{\xi}{1-\tau(1-\xi)}\left[\frac{\partial F_t}{\partial N_t} - \frac{\partial \Gamma_t}{\partial N_t} - \xi N_t \frac{\partial W_t}{\partial N_t} + P_{t,t+1}\Lambda_t\right]$$

$$W_t = \eta\left[\frac{\partial F_t}{\partial N_t} - \frac{\partial \Gamma_t}{\partial N_t} - \xi N_t \frac{\partial W_t}{\partial N_t} + P_{t,t+1}\Lambda_t\right]$$

Where:

$$\eta = \frac{\xi}{1-\tau(1-\xi)}$$

This implies that:

$$\frac{\partial F_t}{\partial N_t} - \frac{\partial \Gamma_t}{\partial N_t} - \xi N_t \frac{\partial W_t}{\partial N_t} = \frac{W_t}{\eta} - P_{t,t+1}\Lambda_t$$

Leading by one period and substituting into the intra-temporal condition gives:

$$\Lambda_t = \mathrm{E}_t\beta_{t+1}\left[\frac{W_{t+1}}{\eta} - P_{t+1,t+2}\Lambda_{t+1} - W_{t+1}\right] + E_t\beta_{t+1}\left(1-\delta_{t+1,t+2}\right)\Lambda_{t+1}$$

$$\Lambda_t = \mathrm{E}_t\beta_{t+1}\left[\left(\frac{1-\eta}{\eta}\right)W_{t+1} - P_{t+1,t+2}\Lambda_{t+1}\right] + E_t\beta_{t+1}\left(1-\delta_{t+1,t+2}\right)\Lambda_{t+1} \tag{4.6)(b}$$

Divide through by F_{t+1}/N_{t+1}

$$\frac{\Lambda_t}{F_{t+1}/N_{t+1}} = \mathrm{E}_t\beta_{t+1}\left[\frac{1-\eta}{\eta}\frac{W_{t+1}}{F_{t+1}/N_{t+1}} - P_{t+1,t+2}\frac{\Lambda_{t+1}}{F_{t+1}/N_{t+1}}\right] + E_t\beta_{t+1}\left(1-\delta_{t+1,t+2}\right)\frac{\Lambda_{t+1}}{F_{t+1}/N_{t+1}}$$

$$\frac{\Lambda_t}{F_t/N_t}\frac{F_t/N_t}{F_{t+1}/N_{t+1}} = \mathrm{E}_t\beta_{t+1}\left[\frac{1-\eta}{\eta}\frac{W_{t+1}}{F_{t+1}/N_{t+1}} - P_{t+1,t+2}\lambda_{t+1}\right] + E_t\beta_{t+1}\left(1-\delta_{t+1,t+2}\right)\frac{\Lambda_{t+1}}{F_{t+1}/N_{t+1}}$$

Noting that $\lambda_t = \frac{\Lambda_t}{F_t/N_t}$ is the normalized Lagrange multiplier, $\frac{W_{t+1}}{F_{t+1}/N_{t+1}} = \frac{W_{t+1}N_{t+1}}{F_{t+1}} = s_{t+1}$ is the labour share, $\frac{F_t/N_t}{F_{t+1}/N_{t+1}} = \frac{1}{G_{t+1}^X}$ is the growth rate of output-per-worker, the result is:

$$\frac{\lambda_t}{G_{t+1}^X} = \mathrm{E}_t\beta_{t+1}\left[\frac{1-\eta}{\eta}s_{t+1} - P_{t+1,t+2}\lambda_{t+1}\right] + E_t\beta_{t+1}\left(1-\delta_{t+1,t+2}\right)\lambda_{t+1} \qquad (4.10)(\mathrm{a})$$

as required.

A.4.2.2

Solving for:

$$\Theta(\phi + (1-\phi)Q_t)\left(\frac{\phi V_t + (1-\phi)Q_t V_t}{N_t}\right)^{\gamma} = Q_{t,t+1}\lambda_t \qquad (4.10)(\mathrm{b})$$

Begin with the intra-temporal condition in$(4.6)(\mathrm{a})$:

$$\frac{\partial \Gamma_t}{\partial V_t} = Q_{t,t+1}\Lambda_t \qquad (4.6)(\mathrm{a})$$

Also recall that:

$$\frac{\partial \mathit{\Gamma}_t}{\partial V_t} = \Theta(\phi + (1-\phi)Q_t)\left(\frac{\phi V_t + (1-\phi)Q_t V_t}{N_t}\right)^{\gamma}\frac{F_t}{N_t} \qquad (4.2)$$

Substituting in for $\frac{\partial \mathit{\Gamma}_t}{\partial V_t}$:

$$Q_{t,t+1}\Lambda_t = \Theta(\phi + (1-\phi)Q_t)\left(\frac{\phi V_t + (1-\phi)Q_t V_t}{N_t}\right)^{\gamma}\frac{F_t}{N_t}$$

Now divide through by $\frac{F_t}{N_t}$:

$$Q_{t,t+1}\frac{\Lambda_t}{F_t/N_t} = \Theta(\phi + (1-\phi)Q_t)\left(\frac{\phi V_t + (1-\phi)Q_t V_t}{N_t}\right)^{\gamma}$$

Since $\lambda_t = \frac{\Lambda_t}{F_t/N_t}$ it follows that:

$$\Theta(\phi + (1-\phi)Q_t)\left(\frac{\phi V_t + (1-\phi)Q_t V_t}{N_t}\right)^{\gamma} = Q_{t,t+1}\lambda_t$$

As required.

A.4.2.3

Solving for:

$$P_{t,t+1} = \mu\left(\frac{v_t}{1-n_t}\right)^{1-\sigma} \tag{4.10)(c}$$

This is trivial since $P_{t,t+1} = \mu\left(\frac{v_t}{u_t}\right)^{1-\sigma}$ by (4.3), and $u = 1 - n_t$.

A.4.2.4

Solving for:

$$\begin{aligned} s_t = \eta\Bigg[(1-\alpha)\Bigg(\left[\frac{1}{1-\alpha\xi}\right] \\ &+ \Theta\left(\frac{\phi v_t + (1-\phi)Q_t v_t}{n_t}\right)^{\gamma+1}\left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\right]\Bigg) \\ &+ P_{t,t+1}\lambda_t\Bigg] \end{aligned} \qquad (4.10)(\mathrm{d})$$

From Appendix (4.1):

$$W_t = \xi \left[(1-\alpha) A \left(\frac{K}{N}\right)^{\alpha} \left(\left[\frac{1}{1-\alpha\xi}\right] + \Theta \left(\frac{\phi V_t + (1-\phi) Q_t V_t}{N_t}\right)^{\gamma+1} \left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\right] \right) + P_{t,t+1}\Lambda_t \right] + (1-\xi) b_t$$

Using the assumption that $b_t = \tau W_t$, this can be written:

$$W_t = \eta \left[(1-\alpha) A \left(\frac{K}{N}\right)^{\alpha} \left(\left[\frac{1}{1-\alpha\xi}\right] + \Theta \left(\frac{\phi V_t + (1-\phi) Q_t V_t}{N_t}\right)^{\gamma+1} \left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\right] \right) + P_{t,t+1}\Lambda_t \right]$$

Where:

$$\eta = \frac{\xi}{1-(1-\xi)\tau}$$

Noting that $A\left(\frac{K}{N}\right)^{\alpha} = \frac{F_t}{N_t}$ that the labour share is given by $\frac{W_t}{F_t/N_t} = \frac{W_t N_t}{F_t} = s_t$, and that $\lambda_t = \frac{\Lambda_t}{F_t/N_t}$ the result is:

$$s_t = \eta \left[(1-\alpha) \left(\left[\frac{1}{1-\alpha\xi}\right] + \Theta \left(\frac{\phi v_t + (1-\phi) Q_t v_t}{n_t}\right)^{\gamma+1} \left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\right] \right) + P_{t,t+1}\lambda_t \right] \qquad (4.10)(\mathrm{d})$$

As required.

A.4.2.5

Solving for:

$$n_{t+1} = \big(1-\delta_{t,t+1}\big) n_t + P_{t,t+1}(1-n_t) \qquad (4.10)(\mathrm{e})$$

Beginning with the equation-of-motion for employment:

$$N_{t+1} = \big(1-\delta_{t,t+1}\big) N_t + Q_{t,t+1} V_t \qquad (4.5)$$

Dividing through by the labour-force L_t:

$$\frac{N_{t+1}}{L_t} = \big(1-\delta_{t,t+1}\big)\frac{N_t}{L_t} + Q_{t,t+1}\frac{V_t}{L_t}$$

Since we can write $\frac{N_{t+1}}{L_t} = \frac{N_{t+1}}{L_{t+1}}\frac{L_{t+1}}{L_t} = \frac{N_{t+1}}{L_{t+1}}G^L$ we have:

$$n_{t+1}G^L = \left(1-\delta_{t,t+1}\right)n_t + Q_{t,t+1}v_t$$

Where as usual, we use lower-case letters to denote rates out of the labour-force. Finally, note that by (4.3):

$$Q_{t,t+1}v_t = m_{t,t+1} = P_{t,t+1}u_t = P_{t,t+1}(1-n_t)$$

Substituting in gives:

$$n_{t+1}G^L = \left(1-\delta_{t,t+1}\right)n_t + P_{t,t+1}(1-n_t) \tag{4.10)(e}$$

As required.

Appendix 4.3: Deriving the system of equations for steady-state variables and parameters (4.11)(a) – (4.11)(i)*:*

$$v = \left(\frac{\delta + G^L - 1}{Q}\right)(1-u) \tag{4.11)(a}$$

$$\frac{v}{u} = \left(\frac{\delta + G^L - 1}{Q}\right)\frac{(1-u)}{u} \tag{4.11)(b}$$

$$P = \left[\frac{\delta}{u} + \frac{(G^L-1)}{u} - \delta + (G^L - 1)\right] \tag{4.11)(c}$$

$$n = \left[\frac{Qv}{(G^L-1)+\delta}\right] \tag{4.11)(d}$$

$$\mu = Q\left(\frac{v}{u}\right)^{\sigma} \tag{4.11(e)}$$

$$\eta = \left[\frac{\xi}{1-(1-\xi)\tau}\right] \tag{4.11(f)}$$

$$\lambda = \left[\frac{G^X\beta}{1-G^X\beta(1-\delta)}\right]\pi \tag{4.11(g)}$$

$$\begin{aligned}\pi = \left(\frac{1}{1+\eta P\frac{G^X\beta}{[1-G^X\beta(1-\delta)]}}\right)\Bigg[(1-\eta)(1-\alpha)\Bigg(\left[\frac{1}{1-\alpha\xi}\right] \\ +\Theta\Bigg(\phi\frac{(\delta+G^L-1)}{Q} \\ +(1-\phi)(\delta+G^L-1)\Bigg)^{\gamma+1}\left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\right]\Bigg)\Bigg]\end{aligned} \tag{4.11(h)}$$

$$\begin{aligned}s = \eta\Bigg[(1-\alpha)\Bigg(\left[\frac{1}{1-\alpha\xi}\right] \\ +\Theta\Bigg(\phi\frac{(\delta+G^L-1)}{Q} \\ +(1-\phi)(\delta+G^L-1)\Bigg)^{\gamma+1}\left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\right]\Bigg) \\ +P\lambda\Bigg]\end{aligned} \tag{4.11(i)}$$

A.4.3.1

Solving for:

$$v = \left(\frac{\delta+G^L-1}{Q}\right)(1-u) \tag{4.11(a)}$$

Start with the equation of motion for employment

$$N_{t+1} = (1 - \delta_{t,t+1})N_t + Q_{t,t+1}V_t \tag{4.5}$$

Appendix *A.4.2.5* we know can be written as:

$$n_{t+1}G^L = (1 - \delta_{t,t+1})n_t + Q_{t,t+1}v_t$$

Applying the steady state:

$$nG^L = (1 - \delta)n + Qv$$

Rearranging and substituting in $(1 - u) = n$ gives:

$$v = \left(\frac{\delta + G^L - 1}{Q}\right)(1 - u) \qquad (4.11)(\mathrm{a})$$

As required.

A.4.3.2

Solving for:

$$\frac{v}{u} = \left(\frac{\delta + G^L - 1}{Q}\right)\frac{(1 - u)}{u} \qquad (4.11)(\mathrm{b})$$

Simply dividing $(4.11)a$ through by u gives the required result.

A.4.4.3

Solving for:

$$P = \left[\frac{\delta}{u} + \frac{(G^L - 1)}{u} - \delta - (G^L - 1)\right] \qquad (4.11)(\mathrm{c})$$

Start with:

$$n_{t+1}G^L = (1 - \delta_{t,t+1})n_t + P_{t,t+1}(1 - n_t) \qquad (4.10)(\mathrm{e})$$

Set the values equal to their steady state values and rearrange:

$$nG^L = (1 - \delta)n + P(1 - n)$$

$$n(G^L - 1 + \delta) = P(1 - n)$$

And since $n = 1 - u$

$$(1-u)(G^L - 1 + \delta) = Pu$$

$$P = \frac{(1-u)}{u}(G^L - 1 + \delta) = \left[\frac{G^L - 1 + \delta}{u} - (G^L - 1) - \delta\right]$$

Hence:

$$P = \left[\frac{\delta}{u} + \frac{(G^L - 1)}{u} - \delta - (G^L - 1)\right] \tag{4.11(c)}$$

As required.

A.4.3.4

Solving for:

$$n = \left[\frac{Qv}{(G^L - 1) + \delta}\right] \tag{4.11(d)}$$

Start with the equation of motion for employment

$$N_{t+1} = \left(1 - \delta_{t,t+1}\right)N_t + Q_{t,t+1}V_t \tag{4.5}$$

Appendix *A.4.2.5* we know can be written as:

$$n_{t+1}G^L = \left(1 - \delta_{t,t+1}\right)n_t + Q_{t,t+1}v_t$$

Applying the steady state:

$$nG^L = (1-\delta)n + Qv$$

Collecting terms in n and solving gives:

$$n = \left[\frac{Qv}{(G^L - 1) + \delta}\right] \tag{4.11(d)}$$

As required.

A.4.3.5

Solving for:

$$\mu = Q\left(\frac{v}{u}\right)^{\sigma} \qquad (4.11)(e)$$

This is trivial since $Q_{t,t+1} = \mu\left(\frac{v_t}{u_t}\right)^{-\sigma}$ by equation (4.3).

A.4.3.6

Solving for:

$$\eta = \left[\frac{\xi}{1-(1-\xi)\tau}\right] \qquad (4.11)(f)$$

See *A.4.2.1* where the inter-temporal condition was solved-for. When the benefit level b_t is assumed to equal the product of the wage and the replacement ratio, so that $b_t = \tau W_t$ the equation for η above comes naturally out of the wage-equation.

A.4.3.7

Solving for:

$$\lambda = \left[\frac{G^X\beta}{1-G^X\beta(1-\delta)}\right]\pi \qquad (4.11)(g)$$

Begin with the inter-temporal condition:

$$\Lambda_t = \mathrm{E}_t\beta_{t+1}\left[\frac{\partial F_{t+1}}{\partial N_{t+1}} - W_{t+1} - N_{t+1}\frac{\partial W_{t+1}}{\partial N_{t+1}} - \frac{\partial \Gamma_{t+1}}{\partial N_{t+1}}\right] + E_t\beta_{t+1}\left(1-\delta_{t+1,t+2}\right)\Lambda_{t+1} \qquad (4.6)(b)$$

Divide through by average labour productivity one period ahead, F_{t+1}/N_{t+1}:

$$\frac{\Lambda_t}{F_{t+1}/N_{t+1}} = \mathrm{E}_t\beta_{t+1}\left[\frac{\frac{\partial F_{t+1}}{\partial N_{t+1}} - W_{t+1} - N_{t+1}\frac{\partial W_{t+1}}{\partial N_{t+1}} - \frac{\partial \Gamma_{t+1}}{\partial N_{t+1}}}{F_{t+1}/N_{t+1}}\right] + E_t\beta_{t+1}\left(1-\delta_{t+1,t+2}\right)\frac{\Lambda_{t+1}}{F_{t+1}/N_{t+1}}$$

Recall that $\lambda_{t+1} = \frac{\Lambda_{t+1}}{F_{t+1}/N_{t+1}}$, and that we can write $\frac{\Lambda_t}{F_{t+1}/N_{t+1}} = \frac{\Lambda_t}{F_{t+1}/N_{t+1}}\frac{F_t/N_t}{F_t/N_t} = \frac{\Lambda_t}{F_t/N_t}\frac{F_t/N_t}{F_{t+1}/N_{t+1}} = \frac{\lambda_t}{G^X_{t+1}}$

Now define marginal profit at time *t+1* as:

$$\pi_{t+1} = \left[\frac{\frac{\partial F_{t+1}}{\partial N_{t+1}} - W_{t+1} - N_{t+1}\frac{\partial W_{t+1}}{\partial N_{t+1}} - \frac{\partial \Gamma_{t+1}}{\partial N_{t+1}}}{F_{t+1}/N_{t+1}}\right]$$

Substituting everything in gives:

$$\frac{\lambda_t}{G^X_{t+1}} = \mathrm{E}_t\beta_{t+1}\pi_{t+1} + E_t\beta_{t+1}\left(1-\delta_{t+1,t+2}\right)\lambda_{t+1}$$

Finally, set everything in the equation to its steady-state value.

$$\frac{\lambda}{G^X} = \beta\pi + \beta(1-\delta)\lambda$$

Solving for λ gives:

$$\lambda = \left[\frac{G^X\beta}{1-G^X\beta(1-\delta)}\right]\pi \tag{4.11)(g}$$

As required.

A.4.3.8

Solving for:

$$\pi = \left(\frac{1}{1+\eta P\frac{G^X\beta}{[1-G^X\beta(1-\delta)]}}\right)\Bigg[(1-\eta)(1-\alpha)\Bigg(\left[\frac{1}{1-\alpha\xi}\right] + \Theta\Bigg(\phi\frac{(\delta+G^L-1)}{Q} + (1-\phi)(\delta+G^L-1)\Bigg)^{\gamma+1}\left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\left(1-\xi(1+\alpha+\gamma)\right)}\right]\Bigg)\Bigg] \tag{4.11)(h}$$

Defining marginal profit – as in *A.4.3.7*, as:

$$\pi_t = \frac{\left[\frac{\partial F_t}{\partial N_t} - \frac{\partial \Gamma_\mathrm{t}}{\partial N_t} - N_t\frac{\partial W_t}{\partial N_t} - W_t\right]}{F_t/N_t}$$

The wage solution, from *A.4.2.1*, is:

$$W_t = \eta\left[\frac{\partial F_t}{\partial N_t} - \frac{\partial \Gamma_t}{\partial N_t} - N_t \frac{\partial W_t}{\partial N_t}\right] + \eta P_{t,t+1}\Lambda_t$$

Recall the relationship between wages and the labour share:

$$W_t = s_t \frac{F_t}{N_t}$$

By eliminating $W_t$from the previous two equations it can be shown that:

$$\frac{\left[\frac{\partial F_t}{\partial N_t} - \frac{\partial \Gamma_t}{\partial N_t} - N_t \frac{\partial W_t}{\partial N_t}\right]}{F_t/N_t} = \frac{1}{\eta}\left(s_t - \eta P_{t,t+1}\lambda_t\right)$$

You will note that the right hand side is simply π_t.

$$\pi_t = \frac{s_t}{\eta} - \eta P_{t,t+1}\lambda_t$$

From $(4.10)(\mathrm{d})$ we have the following expression for s_t

$$\begin{aligned} s_t = \eta\Bigg[(1-\alpha)\Bigg(\left[\frac{1}{1-\alpha\xi}\right] \\ + \Theta\left(\frac{\phi v_t + (1-\phi)Q_t v_t}{n_t}\right)^{\gamma+1}\left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\right]\Bigg) \\ + P_{t,t+1}\lambda_t\Bigg] \end{aligned} \tag{4.10)(d}$$

Substituting into the expression for profits gives:

$$\begin{aligned} \pi_t = \Bigg[(1-\alpha)\Bigg(\left[\frac{1}{1-\alpha\xi}\right] \\ + \Theta\left(\frac{\phi v_t + (1-\phi)Q_t v_t}{n_t}\right)^{\gamma+1}\left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\right]\Bigg)\Bigg] \\ + (1-\eta)P_{t,t+1}\lambda_t \end{aligned}$$

Set all of the variables to their steady-state values:

$$\pi = \left[(1-\alpha)\left(\left[\frac{1}{1-\alpha\xi}\right] + \Theta\left(\frac{\phi v + (1-\phi)Qv}{n}\right)^{\gamma+1}\left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\right]\right)\right] + (1-\eta)P\lambda$$

Recall $(4.11)(\mathrm{g})$.

$$\lambda = \left[\frac{G^X\beta}{1-G^X\beta(1-\delta)}\right]\pi \qquad (4.11)(\mathrm{g})$$

Substituting $(4.11)(\mathrm{g})$ into the equation for π, and solving for π gives:

$$\pi = \left(\frac{1}{1+\eta P\frac{G^X\beta}{[1-G^X\beta(1-\delta)]}}\right)\left[(1-\eta)(1-\alpha)\left(\left[\frac{1}{1-\alpha\xi}\right] + \Theta\left(\phi\frac{(\delta+G^L-1)}{Q} + (1-\phi)(\delta+G^L-1)\right)^{\gamma+1}\left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\right]\right)\right] \qquad (4.11)(\mathrm{h})$$

As required.

A.4.3.9

Solving for:

$$(4.11)(\mathrm{i})$$

$$s = \eta\left[(1-\alpha)\left(\left[\frac{1}{1-\alpha\xi}\right] + \Theta\left(\phi\frac{(\delta+G^L-1)}{Q} + (1-\phi)(\delta+G^L-1)\right)^{\gamma+1}\left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\right]\right) + P\lambda\right]$$

The equation comes directly from setting $(4.10)(\mathrm{d})$ equal to its steady-state value.

Appendix 4.4: Deriving and linearizing the estimating equations $(4.12)(\mathrm{a}) - (4.12)(\mathrm{c})$.

$$\frac{\lambda}{G^X}\left[[-\sigma(\Omega+\gamma\Omega-1)+\gamma]\frac{1}{1-\sigma}\right]\tilde{P}_{t,t+1} - \frac{\lambda}{G^X}\left[\frac{\gamma}{1-n}\right]\tilde{n}_t - \frac{\lambda}{G^X}\rho_{G^X}\tilde{G}_t^X + \beta\lambda\delta\rho_\delta\tilde{\delta}_{t,t+1} - \beta\frac{1-\eta}{\eta}s\mathrm{E}_t[\tilde{s}_{t+1}] - \left(\beta\frac{1-\eta}{\eta}s + \beta\lambda(1-\delta-P)\right)\rho_\beta\tilde{\beta}_t - \left\{(\beta(1-\delta)\lambda - \beta P\lambda)\left[[-\sigma(\Omega+\gamma\Omega-1)+\gamma]\frac{1}{1-\sigma}\right] - \beta P\lambda\right\}E_t\left[\tilde{P}_{t+1,t+2}\right] + \beta(1-\delta-P)\lambda\left[\frac{\gamma}{1-n}\right]E_t[\tilde{n}_{t+1}] \approx e_t^p \qquad (4.12)(\mathrm{a})$$

$$-\left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\eta(1-\alpha)\Theta\left(\frac{\phi v+(1-\phi)Qv}{n}\right)^{\gamma+1}\frac{(\gamma+1)}{1-\sigma}[1+\sigma(1-\Omega)] + \frac{\eta P\lambda}{1-\sigma}[1-\sigma(\Omega+\gamma\Omega)+\gamma]\right]\tilde{P}_{t,t+1} + s\tilde{s}_t \approx -\left(\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\eta(1-\alpha)\Theta\left(\frac{\phi v+(1-\phi)Qv}{n}\right)^{\gamma+1}\frac{(\gamma+1)}{1-n} + \frac{\eta P\lambda\gamma}{1-n}\right)\tilde{n}_t \qquad (4.12)(\mathrm{b})$$

$$G^L\tilde{n}_{t+1} \approx (1-\delta-P)\tilde{n}_t - \delta\tilde{\delta}_{t,t+1} + \frac{P(1-n)}{n}\tilde{P}_{t,t+1} \qquad (4.12)(\mathrm{c})$$

A.4.4.1

Solving for:

$$\frac{\lambda}{G^X}\left[[-\sigma(\Omega+\gamma\Omega-1)+\gamma]\frac{1}{1-\sigma}\right]\tilde{P}_{t,t+1}-\frac{\lambda}{G^X}\left[\frac{\gamma}{1-n}\right]\tilde{n}_t-\frac{\lambda}{G^X}\rho_{G^X}\tilde{G}_t^X$$
$$+\beta\lambda\delta E_t\left[\tilde{\delta}_{t+1,t+2}\right]$$
$$\approx\beta\frac{1-\eta}{\eta}s\mathrm{E}_t[\tilde{s}_{t+1}]+\left(\beta\frac{1-\eta}{\eta}s+\beta\lambda(1-\delta-P)\right)\rho_\beta\tilde{\beta}_t \qquad (4.12)(a)$$
$$+\left\{(\beta(1-\delta)\lambda-\beta P\lambda)\left[[-\sigma(\Omega+\gamma\Omega-1)+\gamma]\frac{1}{1-\sigma}\right]\right.$$
$$\left.-\beta P\lambda\right\}E_t\left[\tilde{P}_{t+1,t+2}\right]-\beta(1-\delta-P)\lambda\left[\frac{\gamma}{1-n}\right]E_t[\tilde{n}_{t+1}]$$

In *A.4.2.1* it was shown that:

$$\Lambda_t=\mathrm{E}_t\beta_{t+1}\left[\frac{W_{t+1}}{\eta}-P_{t+1,t+2}\Lambda_{t+1}-W_{t+1}\right]+E_t\beta_{t+1}\left(1-\delta_{t+1,t+2}\right)\Lambda_{t+1}$$

$$\Lambda_t=\mathrm{E}_t\beta_{t+1}\left[\left(\frac{1-\eta}{\eta}\right)W_{t+1}-P_{t+1,t+2}\Lambda_{t+1}\right]+E_t\beta_{t+1}\left(1-\delta_{t+1,t+2}\right)\Lambda_{t+1}$$

Recall that $\lambda_{t+1}=\frac{\Lambda_{t+1}}{F_{t+1}/N_{t+1}}$, and that we can write $\frac{\Lambda_t}{F_{t+1}/N_{t+1}}=\frac{\Lambda_t}{F_{t+1}/N_{t+1}}\frac{F_t/N_t}{F_t/N_t}=\frac{\Lambda_t}{F_t/N_t}\frac{F_t/N_t}{F_{t+1}/N_{t+1}}=\frac{\lambda_t}{G_{t+1}^X}$. Substituting these in gives:

$$\frac{\lambda_t}{G_{t+1}^X}=\mathrm{E}_t\beta_{t+1}\left[\frac{1-\eta}{\eta}s_{t+1}-P_{t+1,t+2}\lambda_{t+1}\right]+E_t\beta_{t+1}\left(1-\delta_{t+1,t+2}\right)\lambda_{t+1}$$

Here we log-linearize the inter-temporal condition:

$$\frac{\lambda}{G^X}\left(1+\tilde{\lambda}_t\right)\left(1-\tilde{G}_{t+1}^X\right)=\mathrm{E}_t\beta_{t+1}\left[\frac{1-\eta}{\eta}s_{t+1}-P_{t+1,t+2}\lambda_{t+1}\right]+E_t\beta_{t+1}\left(1-\delta_{t+1,t+2}\right)\lambda_{t+1}$$

$$\frac{\lambda}{G^X}\left(1+\tilde{\lambda}_{\mathrm{t}}\right)\left(1-\tilde{G}_{t+1}^X\right)$$
$$\begin{aligned}
&= \beta\left(1+E_t\left[\tilde{\beta}_{t+1}\right]\right)\frac{1-\eta}{\eta}s\left(1+\mathrm{E}_t[\tilde{s}_{t+1}]\right) \\
&- \beta\left(1+E_t\left[\tilde{\beta}_{t+1}\right]\right)P\lambda\left(1+\mathrm{E}_t\left[\tilde{P}_{t+1,t+2}\right]\right)\left(1+\mathrm{E}_t\left[\tilde{\lambda}_{t+1}\right]\right) \\
&+ \beta\left(1+E_t\left[\tilde{\beta}_{t+1}\right]\right)(1-\delta)\lambda\left(1+\mathrm{E}_t\left[\left(1-\widetilde{\delta_{t+1,t+2}}\right)\right]\right)\left(1+\mathrm{E}_t\left[\tilde{\lambda}_{t+1}\right]\right)
\end{aligned}$$

Note that:

$$\left(1-\widetilde{\delta_{t+1,t+2}}\right) = -\tilde{\delta}_{t+1,t+2}\frac{\delta}{1-\delta}$$

$$\frac{\lambda}{G^X}\left(1+\tilde{\lambda}_{\mathrm{t}}\right)\left(1-\tilde{G}_{t+1}^X\right)$$
$$\begin{aligned}
&= \beta\left(1+E_t\left[\tilde{\beta}_{t+1}\right]\right)\frac{1-\eta}{\eta}s\left(1+\mathrm{E}_t[\tilde{s}_{t+1}]\right) \\
&- \beta\left(1+E_t\left[\tilde{\beta}_{t+1}\right]\right)P\lambda\left(1+\mathrm{E}_t\left[\tilde{P}_{t+1,t+2}\right]\right)\left(1+\mathrm{E}_t\left[\tilde{\lambda}_{t+1}\right]\right) \\
&+ \beta\left(1+E_t\left[\tilde{\beta}_{t+1}\right]\right)(1-\delta)\lambda\left(1-E_t\left[\tilde{\delta}_{t+1,t+2}\right]\frac{\delta}{1-\delta}\right)\left(1+\mathrm{E}_t\left[\tilde{\lambda}_{t+1}\right]\right)
\end{aligned}$$

$$\frac{\lambda}{G^X}\left(1+\tilde{\lambda}_{\mathrm{t}}\right)\left(1-\tilde{G}_{t+1}^X\right)$$
$$\begin{aligned}
&= \beta\frac{1-\eta}{\eta}s\left(1+\mathrm{E}_t[\tilde{s}_{t+1}]\right)\left(1+E_t\left[\tilde{\beta}_{t+1}\right]\right) \\
&- \beta P\lambda\left(1+\mathrm{E}_t\left[\tilde{P}_{t+1,t+2}\right]\right)\left(1+\mathrm{E}_t\left[\tilde{\lambda}_{t+1}\right]\right)\left(1+E_t\left[\tilde{\beta}_{t+1}\right]\right) \\
&+ \beta(1-\delta)\lambda\left(1+E_t\left[\tilde{\beta}_{t+1}\right]\right)\left(1-E_t\left[\tilde{\delta}_{t+1,t+2}\right]\frac{\delta}{1-\delta}\right)\left(1+\mathrm{E}_t\left[\tilde{\lambda}_{t+1}\right]\right)
\end{aligned}$$

$$\frac{\lambda}{G^X}\left(1+\tilde{\lambda}_{\mathrm{t}}-\tilde{G}_{t+1}^X\right)$$
$$\begin{aligned}
&= \beta\frac{1-\eta}{\eta}s\left(1+\mathrm{E}_t[\tilde{s}_{t+1}]+E_t\left[\tilde{\beta}_{t+1}\right]\right) \\
&- \beta P\lambda\left(1+\mathrm{E}_t\left[\tilde{P}_{t+1,t+2}\right]+\mathrm{E}_t\left[\tilde{\lambda}_{t+1}\right]+E_t\left[\tilde{\beta}_{t+1}\right]\right) \\
&+ \beta(1-\delta)\lambda\left(1+E_t\left[\tilde{\beta}_{t+1}\right]-E_t\left[\tilde{\delta}_{t+1,t+2}\right]\frac{\delta}{1-\delta}+\mathrm{E}_t\left[\tilde{\lambda}_{t+1}\right]\right)
\end{aligned}$$

Subtracting out the steady state:

$$\begin{aligned}
\frac{\lambda}{G^X}\tilde{\lambda}_{\mathrm{t}}-\frac{\lambda}{G^X}\tilde{G}_{t+1}^X \approx\ &\beta\frac{1-\eta}{\eta}s\left(\mathrm{E}_t[\tilde{s}_{t+1}]+E_t\left[\tilde{\beta}_{t+1}\right]\right) \\
&- \beta P\lambda\left(\mathrm{E}_t\left[\tilde{P}_{t+1,t+2}\right]+\mathrm{E}_t\left[\tilde{\lambda}_{t+1}\right]+E_t\left[\tilde{\beta}_{t+1}\right]\right) \\
&+ \beta(1-\delta)\lambda\left(E_t\left[\tilde{\beta}_{t+1}\right]-E_t\left[\tilde{\delta}_{t+1,t+2}\right]\frac{\delta}{1-\delta}+\mathrm{E}_t\left[\tilde{\lambda}_{t+1}\right]\right)
\end{aligned}$$

Collecting terms gives:

$$\begin{aligned}
\frac{\lambda}{G^X}\tilde{\lambda}_t-\frac{\lambda}{G^X}\tilde{G}_{t+1}^X \approx\ &\beta\frac{1-\eta}{\eta}s\mathrm{E}_t[\tilde{s}_{t+1}]+\left(\beta\frac{1-\eta}{\eta}s+\beta(1-\delta)\lambda-\beta P\lambda\right)E_t\left[\tilde{\beta}_{t+1}\right] \\
&- \beta\lambda\delta E_t\left[\tilde{\delta}_{t+1,t+2}\right]-\beta P\lambda\mathrm{E}_t\left[\tilde{P}_{t+1,t+2}\right]+\left(\beta(1-\delta)\lambda-\beta P\lambda\right)\mathrm{E}_t\left[\tilde{\lambda}_{t+1}\right]
\end{aligned}$$

Note that we still need to eliminate $\tilde{\lambda}_t$ and $\mathrm{E}_t[\tilde{\lambda}_{t+1}]$ from the expression obtained so far. We do this following Yashiv's technical appendix. We begin with the version of the intra-temporal condition given in (4.10)(b)

$$\Theta(\phi + (1-\phi)Q_{t,t+1})\left(\frac{\phi V_t + (1-\phi)Q_{t,t+1}V_t}{N_t}\right)^{\gamma} = Q_{t,t+1}\lambda_t \tag{4.10)(b}$$

For brevity of notation, define:

$$\bar{Q}_{t,t+1} = \phi + (1-\phi)Q_{t,t+1}$$

Which allows one to write the intra-temporal condition as:

$$\Theta\bar{Q}_{t,t+1}\left(\frac{\bar{Q}_{t,t+1}V_t}{N_t}\right)^{\gamma} = Q_{t,t+1}\lambda_t$$

Note that one can normalize V_t and N_t by the labour force which will not change anything since the ratio of the variables appears.

$$\Theta\bar{Q}_{t,t+1}\left(\frac{\bar{Q}_{t,t+1}\frac{V_t}{L_t}}{\frac{N_t}{L_t}}\right)^{\gamma} = \Theta\bar{Q}_{t,t+1}\left(\frac{\bar{Q}_{t,t+1}v_t}{n_t}\right)^{\gamma} = Q_{t,t+1}\lambda_t$$

The steady state of the intra-temporal condition is given by:

$$\Theta\bar{Q}\left(\frac{\bar{Q}v}{n}\right)^{\gamma} = Q\lambda$$

Log-linearizing around the steady-steady values:

$$\Theta\bar{Q}(1+\tilde{\bar{Q}}_{t,t+1})\left(\frac{\bar{Q}v}{n}\right)^{\gamma}(1+\gamma\tilde{\bar{Q}}_{t,t+1}+\gamma\tilde{v}_t-\gamma\tilde{n}_t) = Q\lambda(1+\tilde{Q}_{t,t+1})(1+\tilde{\lambda}_t)$$

$$\Theta\bar{Q}(1+\tilde{\bar{Q}}_{t,t+1})\left(\frac{\bar{Q}v}{n}\right)^{\gamma}(1+\gamma\tilde{\bar{Q}}_{t,t+1}+\gamma\tilde{v}_t-\gamma\tilde{n}_t) = Q\lambda(1+\tilde{Q}_{t,t+1})(1+\tilde{\lambda}_t)$$

$$(\Theta\bar{Q}+\Theta\bar{Q}\tilde{\bar{Q}}_{t,t+1})\left(\frac{\bar{Q}v}{n}\right)^{\gamma} + \Theta\bar{Q}\left(\frac{\bar{Q}v}{n}\right)^{\gamma}(\gamma\tilde{\bar{Q}}_{t,t+1}+\gamma\tilde{v}_t-\gamma\tilde{n}_t) = Q\lambda + Q\lambda(\tilde{Q}_{t,t+1}+\tilde{\lambda}_t)$$

$$\Theta\bar{Q}\left(\frac{\bar{Q}v}{n}\right)^{\gamma} + \Theta\bar{Q}\tilde{\bar{Q}}_{t,t+1}\left(\frac{\bar{Q}v}{n}\right)^{\gamma} + \Theta\bar{Q}\left(\frac{\bar{Q}v}{n}\right)^{\gamma}(\gamma\tilde{\bar{Q}}_{t,t+1}+\gamma\tilde{v}_t-\gamma\tilde{n}_t)$$
$$= Q\lambda + Q\lambda(\tilde{Q}_{t,t+1}+\tilde{\lambda}_t)$$

Subtracting out the steady-state gives:

$$\Theta\bar{Q}\left(\frac{\bar{Q}v}{n}\right)^{\gamma}\left(\tilde{\bar{Q}}_{t,t+1}+\gamma\tilde{\bar{Q}}_{t,t+1}+\gamma\tilde{v}_t-\gamma\tilde{n}_t\right)=Q\lambda\left(\tilde{Q}_{t,t+1}+\tilde{\lambda}_t\right)$$

Dividing out the remaining steady state values gives:

$$\tilde{\bar{Q}}_{t,t+1}+\gamma\tilde{\bar{Q}}_{t,t+1}+\gamma\tilde{v}_t-\gamma\tilde{n}_t=\tilde{Q}_{t,t+1}+\tilde{\lambda}_t$$

The challenge is now to derive an expression for $\tilde{\bar{Q}}_{t,t+1}$. Recall that $\bar{Q}_{t,t+1}=\phi+(1-\phi)Q_{t,t+1}$ so we will first find an expression for $\tilde{Q}_{t,t+1}$. Use equation (4.3):

$$Q_{t,t+1}=\mu\left(\frac{v_t}{u_t}\right)^{-\sigma}$$

The fact that $u_t=1-n_t$ allows me to write:

$$Q_{t,t+1}=\mu\left(\frac{v_t}{1-n_t}\right)^{-\sigma}$$

Hence, log-linearizing $Q_{t,t+1}$ around its steady state:

$$\begin{aligned}
\tilde{Q}_{t,t+1}&=-\sigma[ln(\mu)+ln(v_t)-ln(1-n_t)-ln(\mu)-ln(v)+ln(1-n)]\\
&=-\sigma[ln(v_t)-ln(v)-ln(1-n_t)+ln(1-n)]\\
&=-\sigma[ln(v_t)-ln(v)-ln(1-n_t)+ln(1-n)]\\
&=-\sigma\left[\tilde{v}_t-\frac{1-n_t-1+n}{1+n}\right]
\end{aligned}$$

$$=-\sigma\left[\tilde{v}_t-\frac{1-n_t-1+n}{1-n}\right]$$

$$=-\sigma\left[\tilde{v}_t-\frac{-(n_t-n)}{n}\frac{n}{1-n}\right]$$

$$\tilde{Q}_{t,t+1}=-\sigma\left[\tilde{v}_t+\frac{n}{1-n}\tilde{n}_t\right]$$

From here one can derive an expression for $\tilde{\bar{Q}}_{t,t+1}$:

Since $\bar{Q}_{t,t+1}=\phi+(1-\phi)Q_{t,t+1}$, it follows that:

$$\tilde{\bar{Q}}_{t,t+1}=\frac{\phi+(1-\phi)Q_{t,t+1}-\phi-(1-\phi)Q}{\phi+(1-\phi)Q}=\frac{(1-\phi)\left(Q_{t,t+1}-Q\right)}{\phi+(1-\phi)Q}$$

$$=\frac{(1-\phi)Q}{\phi+(1-\phi)Q}\left[\frac{Q_{t,t+1}-Q}{Q}\right]$$

$$= \frac{(1-\phi)Q}{\phi + (1-\phi)Q}\tilde{Q}_{t,t+1}$$

$$= \Omega\tilde{Q}_{t,t+1}$$

Where

$$\Omega = \frac{(1-\phi)Q}{\phi + (1-\phi)Q}$$

Substituting these expressions into the linearized intra-temporal condition above and rearranging gives:

$$\tilde{\tilde{Q}}_{t,t+1} + \gamma\tilde{\tilde{Q}}_{t,t+1} + \gamma\tilde{v}_t - \gamma\tilde{n}_t = \tilde{Q}_{t,t+1} + \tilde{\lambda}_t$$

$$\Omega\tilde{Q}_{t,t+1} + \gamma\Omega\tilde{Q}_{t,t+1} - \tilde{Q}_{t,t+1} + \gamma\tilde{v}_t = \gamma\tilde{n}_t + \tilde{\lambda}_t$$

$$\tilde{Q}_{t,t+1}(\Omega + \gamma\Omega - 1) + \gamma\tilde{v}_t = \gamma\tilde{n}_t + \tilde{\lambda}_t$$

$$-\sigma\left(\tilde{v}_t + \frac{n}{1-n}\tilde{n}_t\right)(\Omega + \gamma\Omega - 1) + \gamma\tilde{v}_t = \gamma\tilde{n}_t + \tilde{\lambda}_t$$

$$[-\sigma(\Omega + \gamma\Omega - 1) + \gamma]\tilde{v}_t = \left[\gamma + \sigma\frac{n}{1-n}(\Omega + \gamma\Omega - 1)\right]\tilde{n}_t + \tilde{\lambda}_t$$

We now have a linearized expression for the intra-temporal condition. However, we do not have data on $\tilde{v}_t$ and so wish to express the equation in terms of $\tilde{P}_{t,t+1}$. From equation (4.3) we have:

$$P_{t,t+1} = \mu\left(\frac{v_t}{u_t}\right)^{1-\sigma} = \mu\left(\frac{v_t}{1-n_t}\right)^{1-\sigma}$$

Log linearizing around an assumed steady-state gives:

$$P\left(1 + \tilde{P}_{t,t+1}\right) \approx \mu\left(\frac{v}{1-n}\right)^{1-\sigma}(1 + (1-\sigma)\tilde{v}_t)\left(1 - (1-\sigma)(\widetilde{1-n_t})\right)$$

$$(\widetilde{1-n_t}) = \frac{(1-n_t)-(1-n)}{1-n} = \frac{-(n_t-n)}{1-n} = \frac{-(n_t-n)}{1-n}\frac{n}{n} = \frac{-(n_t-n)}{n}\frac{n}{1-n}$$
$$= -\tilde{n}_t\frac{n}{1-n}$$

$$P\left(1+\tilde{P}_{t,t+1}\right) \approx \mu\left(\frac{v}{1-n}\right)^{1-\sigma}\left(1+(1-\sigma)\tilde{v}_t\right)\left(1+\frac{(1-\sigma)n}{1-n}\tilde{n}_t\right)$$

$$P\left(1+\tilde{P}_{t,t+1}\right) \approx \mu\left(\frac{v}{1-n}\right)^{1-\sigma}\left(1+(1-\sigma)\tilde{v}_t+\frac{(1-\sigma)n}{1-n}\tilde{n}_t\right)$$

Subtracting out the steady-state:

$$P\tilde{P}_{t,t+1} \approx \mu\left(\frac{v}{1-n}\right)^{1-\sigma}\left((1-\sigma)\tilde{v}_t+\frac{(1-\sigma)n}{1-n}\tilde{n}_t\right)$$

Divide out the steady state:

$$\tilde{P}_{t,t+1} \approx (1-\sigma)\tilde{v}_t+\frac{(1-\sigma)n}{1-n}\tilde{n}_t$$

Now solve to get an expression for $\tilde{v}_t$:

$$\tilde{P}_{t,t+1}-\frac{(1-\sigma)n}{1-n}\tilde{n}_t \approx (1-\sigma)\tilde{v}_t$$

$$\tilde{v}_t \approx \frac{1}{1-\sigma}\tilde{P}_{t,t+1}-\frac{n}{1-n}\tilde{n}_t$$

Returning to the linearized intra-temporal condition

$$[-\sigma(\Omega+\gamma\Omega-1)+\gamma]\tilde{v}_t = \left[\gamma+\sigma\frac{n}{1-n}(\Omega+\gamma\Omega-1)\right]\tilde{n}_t+\tilde{\lambda}_t$$

Substituting to eliminate $\tilde{v}_t$ gives:

$$[-\sigma(\Omega+\gamma\Omega-1)+\gamma]\left(\frac{1}{1-\sigma}\tilde{P}_{t,t+1}-\frac{n}{1-n}\tilde{n}_t\right) \approx \left[\gamma+\sigma\frac{n}{1-n}(\Omega+\gamma\Omega-1)\right]\tilde{n}_t+\tilde{\lambda}_t$$

Rearranging gives the desired form of the linearized intra-temporal condition.

$$\tilde{\lambda}_t \approx \left[[-\sigma(\Omega+\gamma\Omega-1)+\gamma]\frac{1}{1-\sigma}\right]\tilde{P}_{t,t+1}-\left[\frac{\gamma}{1-n}\right]\tilde{n}_t$$

The linearized intra-temporal condition can be substituted back into the inter-temporal condition, which was:

$$\begin{aligned}\frac{\lambda}{G^X}\tilde{\lambda}_t-\frac{\lambda}{G^X}\tilde{G}^X_{t+1} \approx \beta\frac{1-\eta}{\eta}s\mathrm{E}_t[\tilde{s}_{t+1}]+\left(\beta\frac{1-\eta}{\eta}s+\beta(1-\delta)\lambda-\beta P\lambda\right)E_t\left[\tilde{\beta}_{t+1}\right]\\ -\beta\lambda\delta E_t\left[\tilde{\delta}_{t+1,t+2}\right]-\beta P\lambda\mathrm{E}_t\left[\tilde{P}_{t+1,t+2}\right]+(\beta(1-\delta)\lambda-\beta P\lambda)\mathrm{E}_t\left[\tilde{\lambda}_{t+1}\right]\end{aligned}$$

Substituting to eliminate $\tilde{\lambda}_t$ and $\mathrm{E}_t\left[\tilde{\lambda}_{t+1}\right]$ and rearranging gives:

$$
\begin{aligned}
&\frac{\lambda}{G^X}\left[[-\sigma(\Omega+\gamma\Omega-1)+\gamma]\frac{1}{1-\sigma}\right]\tilde{P}_{t,t+1}-\frac{\lambda}{G^X}\left[\frac{\gamma}{1-n}\right]\tilde{n}_t-\frac{\lambda}{G^X}\tilde{G}^X_{t+1}+\beta\lambda\delta E_t\left[\tilde{\delta}_{t+1,t+2}\right]\\
&\approx\beta\frac{1-\eta}{\eta}s\mathrm{E}_t[\tilde{s}_{t+1}]+\left(\beta\frac{1-\eta}{\eta}s+\beta\lambda(1-\delta-P)\right)E_t\left[\tilde{\beta}_{t+1}\right]\\
&+\left\{(\beta(1-\delta)\lambda-\beta P\lambda)\left[[-\sigma(\Omega+\gamma\Omega-1)+\gamma]\frac{1}{1-\sigma}\right]\right.\\
&\left.-\beta P\lambda\right\}E_t\left[\tilde{P}_{t+1,t+2}\right]-\beta(1-\delta-P)\lambda\left[\frac{\gamma}{1-n}\right]E_t[\tilde{n}_{t+1}]
\end{aligned}
$$

Take expectations at time *t*.

$$
\begin{aligned}
&\frac{\lambda}{G^X}\left[[-\sigma(\Omega+\gamma\Omega-1)+\gamma]\frac{1}{1-\sigma}\right]\tilde{P}_{t,t+1}-\frac{\lambda}{G^X}\left[\frac{\gamma}{1-n}\right]\tilde{n}_t-\frac{\lambda}{G^X}E_t\left[\tilde{G}^X_{t+1}\right]+\beta\lambda\delta E_t\left[\tilde{\delta}_{t+1,t+2}\right]\\
&\approx\beta\frac{1-\eta}{\eta}s\mathrm{E}_t[\tilde{s}_{t+1}]+\left(\beta\frac{1-\eta}{\eta}s+\beta\lambda(1-\delta-P)\right)E_t\left[\tilde{\beta}_{t+1}\right]\\
&+\left\{(\beta(1-\delta)\lambda-\beta P\lambda)\left[[-\sigma(\Omega+\gamma\Omega-1)+\gamma]\frac{1}{1-\sigma}\right]\right.\\
&\left.-\beta P\lambda\right\}E_t\left[\tilde{P}_{t+1,t+2}\right]-\beta(1-\delta-P)\lambda\left[\frac{\gamma}{1-n}\right]E_t[\tilde{n}_{t+1}]
\end{aligned}
$$

Finally, assume that the expectations of the exogenous variables are formed using the univariate AR(1) coefficient. That is: $E_t\left[\tilde{G}^X_{t+1}\right]=\rho_{G^x}\tilde{G}^X_t$, $E_t\left[\tilde{\delta}_{t+1,t+2}\right]=\rho_\delta\tilde{\delta}_{t,t+1}$. $E_t\left[\tilde{\beta}_{t+1}\right]=\rho_\beta\tilde{\beta}_t$. Substituting in gives:

$$
\begin{aligned}
&\frac{\lambda}{G^X}\left[[-\sigma(\Omega+\gamma\Omega-1)+\gamma]\frac{1}{1-\sigma}\right]\tilde{P}_{t,t+1}-\frac{\lambda}{G^X}\left[\frac{\gamma}{1-n}\right]\tilde{n}_t-\frac{\lambda}{G^X}\rho_{G^x}\tilde{G}^X_t+\beta\lambda\delta\rho_\delta\tilde{\delta}_{t,t+1}\\
&\approx\beta\frac{1-\eta}{\eta}s\mathrm{E}_t[\tilde{s}_{t+1}]+\left(\beta\frac{1-\eta}{\eta}s+\beta\lambda(1-\delta-P)\right)\rho_\beta\tilde{\beta}_t\\
&+\left\{(\beta(1-\delta)\lambda-\beta P\lambda)\left[[-\sigma(\Omega+\gamma\Omega-1)+\gamma]\frac{1}{1-\sigma}\right]\right.\\
&\left.-\beta P\lambda\right\}E_t\left[\tilde{P}_{t+1,t+2}\right]-\beta(1-\delta-P)\lambda\left[\frac{\gamma}{1-n}\right]E_t[\tilde{n}_{t+1}]
\end{aligned}
$$

$$\frac{\lambda}{G^X}\left[[-\sigma(\Omega+\gamma\Omega-1)+\gamma]\frac{1}{1-\sigma}\right]\tilde{P}_{t,t+1}-\frac{\lambda}{G^X}\left[\frac{\gamma}{1-n}\right]\tilde{n}_t-\frac{\lambda}{G^X}\rho_{G^x}\tilde{G}_t^X+\beta\lambda\delta\rho_\delta\tilde{\delta}_{t,t+1}$$
$$\approx\beta\frac{1-\eta}{\eta}s\mathrm{E}_t[\tilde{s}_{t+1}]+\left(\beta\frac{1-\eta}{\eta}s+\beta\lambda(1-\delta-P)\right)\rho_\beta\tilde{\beta}_t$$
$$+\left\{(\beta(1-\delta)\lambda-\beta P\lambda)\left[[-\sigma(\Omega+\gamma\Omega-1)+\gamma]\frac{1}{1-\sigma}\right]\right.$$
$$\left.-\beta P\lambda\right\}E_t\left[\tilde{P}_{t+1,t+2}\right]-\beta(1-\delta-P)\lambda\left[\frac{\gamma}{1-n}\right]E_t[\tilde{n}_{t+1}]$$

$$\frac{\lambda}{G^X}\left[[-\sigma(\Omega+\gamma\Omega-1)+\gamma]\frac{1}{1-\sigma}\right]\tilde{P}_{t,t+1}-\frac{\lambda}{G^X}\left[\frac{\gamma}{1-n}\right]\tilde{n}_t-\frac{\lambda}{G^X}\rho_{G^x}\tilde{G}_t^X+\beta\lambda\delta\rho_\delta\tilde{\delta}_{t,t+1}$$
$$-\beta\frac{1-\eta}{\eta}s\mathrm{E}_t[\tilde{s}_{t+1}]-\left(\beta\frac{1-\eta}{\eta}s+\beta\lambda(1-\delta-P)\right)\rho_\beta\tilde{\beta}_t \qquad (4.12)(a)$$
$$-\left\{(\beta(1-\delta)\lambda-\beta P\lambda)\left[[-\sigma(\Omega+\gamma\Omega-1)+\gamma]\frac{1}{1-\sigma}\right]\right.$$
$$\left.-\beta P\lambda\right\}E_t\left[\tilde{P}_{t+1,t+2}\right]+\beta(1-\delta-P)\lambda\left[\frac{\gamma}{1-n}\right]E_t[\tilde{n}_{t+1}]\approx e_t^p$$

A.4.4.2

Solving for:

$$s\tilde{s}_t - \left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\eta(1-\alpha)\Theta\left(\frac{\phi v+(1-\phi)Qv}{n}\right)^{\gamma+1}\frac{(\gamma+1)}{1-\sigma}[1 + \sigma(1-\Omega)] + \frac{\eta P\lambda}{1-\sigma}[1-\sigma(\Omega+\gamma\Omega)+\gamma]\right]\tilde{P}_{t,t+1} + \left(\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\eta(1-\alpha)\Theta\left(\frac{\phi v+(1-\phi)Qv}{n}\right)^{\gamma+1}\frac{(\gamma+1)}{1-n}+\frac{\eta P\lambda\gamma}{1-n}\right)\tilde{n}_t \approx e_t^s \qquad (4.12)(b)$$

Beginning with the labour share equation from $(4.10)(\mathrm{d})$

$$s_t = \eta\left[(1-\alpha)\left(\left[\frac{1}{1-\alpha\xi}\right] + \Theta\left(\frac{\phi v_t+(1-\phi)Q_t v_t}{n_t}\right)^{\gamma+1}\left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\right]\right) + P_{t,t+1}\lambda_t\right] \qquad (4.10)(\mathrm{d})$$

Log-linearizing around the steady-state:

$$s(1+\tilde{s}_t) = \eta\left[(1-\alpha)\left(\left[\frac{1}{1-\alpha\xi}\right] + \Theta\left(\frac{\phi v+(1-\phi)Qv}{n}\right)^{\gamma+1}\left(1 + (\gamma+1)\left(\widetilde{\frac{\phi v_t+(1-\phi)Q_t v_t}{n_t}}\right)\right)\left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\right]\right) + P\lambda\big(1+\tilde{P}_{t,t+1}\big)\big(1+\tilde{\lambda}_t\big)\right]$$

First we will log-linearize $\left(\frac{\phi v+(1-\phi)Qv}{n}\right)$:

$$\left(\widetilde{\frac{\phi v_t+(1-\phi)Q_t v_t}{n_t}}\right) = \phi n_t^{-1}v_t + (1-\phi)n_t^{-1}Q_t v_t$$

$$= \frac{\phi n_t^{-1}v_t + (1-\phi)n_t^{-1}Q_t v_t - \phi n^{-1}Vv - (1-\phi)n^{-1}Qv}{\phi n^{-1}v + (1-\phi)n^{-1}Qv}$$

$$= \frac{\phi n_t^{-1}v_t - \phi n^{-1}v + (1-\phi)n_t^{-1}Q_t v_t - (1-\phi)n^{-1}Qv}{\phi n^{-1}v + (1-\phi)n^{-1}Qv}$$

$$= \frac{\phi n_t^{-1} v_t - \phi n^{-1} v}{\phi n^{-1} v + (1-\phi) n^{-1} Q v} + \frac{(1-\phi) n_t^{-1} Q_t v_t - (1-\phi) n^{-1} Q v}{\phi n^{-1} v + (1-\phi) n^{-1} Q v}$$

$$= \frac{\phi n_t^{-1} v_t - \phi n^{-1} v}{\phi n^{-1} v}\left[\frac{\phi n^{-1} v}{\phi n^{-1} v + (1-\phi) n^{-1} Q v}\right]$$
$$+ \frac{(1-\phi) n_t^{-1} Q_t v - (1-\phi) n^{-1} Q v}{(1-\phi) n^{-1} Q v}\left[\frac{(1-\phi) n^{-1} Q v}{\phi n^{-1} v + (1-\phi) n^{-1} Q v}\right]$$

$$= \phi \widetilde{n_t^{-1} v_t}\left[\frac{\phi n^{-1} v}{\phi n^{-1} v + (1-\phi) n^{-1} Q v}\right] + (1-\phi) \widetilde{n_t^{-1} Q_t} v_t \left[\frac{(1-\phi) n^{-1} Q v}{\phi n^{-1} v + (1-\phi) n^{-1} Q v}\right]$$

$$= \phi \widetilde{n_t^{-1} v_t}\left[\frac{\phi}{\phi + (1-\phi) Q}\right] + (1-\phi) \widetilde{n_t^{-1} Q_t} v_t \left[\frac{(1-\phi) Q}{\phi + (1-\phi) Q}\right]$$

$$= \left(ln(\phi n_t^{-1} v_t) - ln(\phi n^{-1} v)\right)\left[\frac{\phi}{\phi + (1-\phi) Q}\right]$$
$$+ \left(ln\left((1-\phi) n_t^{-1} Q_t v_t\right) - ln\left((1-\phi) n^{-1} Q v\right)\right)\left[\frac{(1-\phi) Q}{\phi + (1-\phi) Q}\right]$$

$$= (\tilde{v}_t - \tilde{n}_t)\left[\frac{\phi}{\phi + (1-\phi) Q}\right] + \left(\tilde{v}_t + \tilde{Q}_{t,t+1} - \tilde{n}_t\right)\left[\frac{(1-\phi) Q}{\phi + (1-\phi) Q}\right]$$

Noting that $\left[\frac{(1-\phi)Q}{\phi+(1-\phi)Q}\right] = \Omega$ and $\left[\frac{\phi}{\phi+(1-\phi)Q}\right] = 1 - \Omega$:

$$\left(\frac{\phi v_t + \widetilde{(1-\phi)} Q_t v_t}{n_t}\right) = (\tilde{v}_t - \tilde{n}_t)\Omega + \left(\tilde{v}_t + \tilde{Q}_{t,t+1} - \tilde{n}_t\right)[1-\Omega]$$

$$= (\tilde{v}_t - \tilde{n}_t)\Omega + \left(\tilde{v}_t + \tilde{Q}_{t,t+1} - \tilde{n}_t\right)[1-\Omega]$$

$$= \left(\tilde{v}_t - \tilde{n}_t + \tilde{Q}_{t,t+1}[1-\Omega]\right)$$

Substituting into the expression for $\tilde{s}_t$:

$$s(1+\tilde{s}_t)=\eta\left[(1-\alpha)\left(\left[\frac{1}{1-\alpha\xi}\right]\right.\right.$$
$$+\Theta\left(\frac{\phi v+(1-\phi)Qv}{n}\right)^{\gamma+1}\Big(1$$
$$+(\gamma+1)\big(\tilde{v}_t-\tilde{n}_t$$
$$\left.+\tilde{Q}_{t,t+1}[1-\Omega]\big)\Big)\left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\right]\right)$$
$$\left.+P\lambda\big(1+\tilde{P}_{t,t+1}\big)\big(1+\widetilde{\lambda_t}\big)\right]$$

Subtracting out the steady-state gives:

$$s\tilde{s}_t\approx\eta(1-\alpha)\left(\Theta\left(\frac{\phi v+(1-\phi)Qv}{n}\right)^{\gamma+1}(\gamma+1)\big(\tilde{v}_t-\tilde{n}_t\right.$$
$$\left.+\tilde{Q}_{t,t+1}[1-\Omega]\big)\left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\right]\right)$$
$$+\eta P\lambda\big(\tilde{P}_{t,t+1}+\widetilde{\lambda_t}\big)$$

Rearranging the terms:

$$s\tilde{s}_t\approx\left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\eta(1-\alpha)\Theta\left(\frac{\phi v+(1-\phi)Qv}{n}\right)^{\gamma+1}(\gamma+1)\right]\tilde{v}_t$$
$$-\left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\eta(1\right.$$
$$\left.-\alpha)\Theta\left(\frac{\phi v+(1-\phi)Qv}{n}\right)^{\gamma+1}(\gamma+1)\right]\tilde{n}_t$$
$$+\left[\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\eta(1\right.$$
$$\left.-\alpha)\Theta\left(\frac{\phi v+(1-\phi)Qv}{n}\right)^{\gamma+1}(\gamma+1)[1-\Omega]\right]\tilde{Q}_{t,t+1}+[\eta P\lambda]\tilde{P}_{t,t+1}$$
$$+[\eta P\lambda]\widetilde{\lambda_t}$$

It remains to remove the terms in $\tilde{v}_t$, $\widetilde{\lambda_t}$ and $\tilde{Q}_{t,t+1}$, using the expressions derived so far.

$$\tilde{v}_t \approx \left(\frac{1}{1-\sigma}\tilde{P}_{t,t+1} - \frac{n}{1-n}\tilde{n}_t\right)$$

$$\tilde{\lambda}_t \approx \left[[-\sigma(\Omega+\gamma\Omega-1)+\gamma]\frac{1}{1-\sigma}\right]\tilde{P}_{t,t+1} - \left[\frac{\gamma}{1-n}\right]\tilde{n}_t$$

$$\tilde{Q}_{t,t+1} \approx \frac{\sigma}{1-\sigma}\tilde{P}_{t,t+1}$$

After eliminating $\tilde{v}_t$, $\widetilde{\lambda_t}$ and $\tilde{Q}_{t,t+1}$ in the expression for $\tilde{s}_t$ gives:

$$\begin{aligned} s\tilde{s}_t - \Bigg[&\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\eta(1-\alpha)\Theta\left(\frac{\phi v+(1-\phi)Qv}{n}\right)^{\gamma+1}\frac{(\gamma+1)}{1-\sigma}[1 \\ &+\sigma(1-\Omega)] + \frac{\eta P\lambda}{1-\sigma}[1-\sigma(\Omega+\gamma\Omega)+\gamma]\Bigg]\tilde{P}_{t,t+1} \\ &+\Bigg(\frac{\alpha+\gamma}{(1+\gamma)(1-\alpha)\big(1-\xi(1+\alpha+\gamma)\big)}\eta(1 \\ &-\alpha)\Theta\left(\frac{\phi v+(1-\phi)Qv}{n}\right)^{\gamma+1}\frac{(\gamma+1)}{1-n}+\frac{\eta P\lambda\gamma}{1-n}\Bigg)\tilde{n}_t \approx e_t^s \end{aligned} \qquad (4.12)(b)$$

As required.

A.4.4.3

Solving for:

$$G^L\tilde{n}_{t+1} \approx (1-\delta-P)\tilde{n}_t - \delta\tilde{\delta}_{t,t+1} + \frac{P(1-n)}{n}\tilde{P}_{t,t+1} \qquad (4.12)(c)$$

Beginning with:

$$G^L n_{t+1} = \big(1-\delta_{t,t+1}\big)n_t + P_{t,t+1}(1-n_t) \qquad (4.10)(e)$$

Log-linearize around the steady-state:

$$G^L n(1+\tilde{n}_{t+1}) = (1-\delta)\left(1+\left(\widetilde{1-\delta_{t,t+1}}\right)\right)n(1+\tilde{n}_t)$$
$$+P\left(1+\tilde{P}_{t,t+1}\right)(1-n)\left(1+\left(\widetilde{1-n_t}\right)\right)$$

$$G^L n(1+\tilde{n}_{t+1}) = (1-\delta)n\left(1+\left(\widetilde{1-\delta_{t,t+1}}\right)\right)(1+\tilde{n}_t)$$
$$+P(1-n)\left(1+\tilde{P}_{t,t+1}\right)\left(1+\left(\widetilde{1-n_t}\right)\right)$$

$$G^L n(1+\tilde{n}_{t+1}) \approx (1-\delta)n\left(1+\tilde{n}_t+\left(\widetilde{1-\delta_{t,t+1}}\right)\right)+P(1-n)\left(1+\tilde{P}_{t,t+1}+\left(\widetilde{1-n_t}\right)\right)$$

Note that:

$$\left(\widetilde{1-\delta_{t,t+1}}\right) = \frac{\left(1-\delta_{t,t+1}\right)-(1-\delta)}{1-\delta} = -\frac{\left(\delta_{t,t+1}-\delta\right)}{1-\delta} = -\frac{\left(\delta_{t,t+1}-\delta\right)}{\delta}\frac{\delta}{1-\delta}$$
$$= -\tilde{\delta}_{t,t+1}\frac{\delta}{1-\delta}$$

$$\left(\widetilde{1-n_t}\right) = \frac{(1-n_t)-(1-n)}{1-n} = -\frac{(n_t-n)}{1-n} = -\frac{(n_t-n)}{n}\frac{n}{1-n} = -\tilde{n}_t\frac{n}{1-n}$$

Substituting in gives:

$$G^L n(1+\tilde{n}_{t+1}) \approx (1-\delta)n\left(1+\tilde{n}_t-\tilde{\delta}_{t,t+1}\frac{\delta}{1-\delta}\right)+P(1-n)\left(1+\tilde{P}_{t,t+1}-\tilde{n}_t\frac{n}{1-n}\right)$$

Subtracting out the steady-state gives:

$$G^L\tilde{n}_{t+1} \approx (1-\delta)\left(\tilde{n}_t-\tilde{\delta}_{t,t+1}\frac{\delta}{1-\delta}\right)+P\left(\frac{1-n}{n}\right)\left(\tilde{P}_{t,t+1}-\tilde{n}_t\frac{n}{1-n}\right)$$
$$G^L\tilde{n}_{t+1} \approx (1-\delta)\left(\tilde{n}_t-\tilde{\delta}_{t,t+1}\frac{\delta}{1-\delta}\right)+P\left(\frac{1-n}{n}\right)\left(\tilde{P}_{t,t+1}-\tilde{n}_t\frac{n}{1-n}\right)$$

Rearranging gives:

$$G^L\tilde{n}_{t+1} \approx (1-\delta-P)\tilde{n}_t-\delta\tilde{\delta}_{t,t+1}+\frac{P(1-n)}{n}\tilde{P}_{t,t+1} \qquad (4.12)(c)$$

As required.

Conclusion

In Chapter 2 of this thesis we have used the Petrongolo and Pissarides method on UK Labour Force Survey panel data, and attempted to apply it to quarterly gross flows in the Welsh labour market for the period 1997Q2-2010Q4. For comparison purposes, we have also updated the results of Petrongolo and Pissarides, by re-estimating the gross flow shares for the whole of the UK over our more recent sample period.

We have found that, under steady-state assumptions and with smoothed data, gross flow hazard shares "into" and "out of" unemployment are approximately equal in Wales, whereas they are weighted towards "outs" in the UK and the area outside of Wales, (specifically the UE hazards). The conclusions for the UK are similar to those obtained by Petrongolo and Pissarides from their earlier sample.
However, once we relax the smoothing and/or steady state assumptions, and when we change the sample period so as to exclude the period of recession in the UK that began in 2008Q2, the results are shown to be fragile and highly contingent on the particular assumptions made. Our results are least stable when we impose neither the assumption of the steady state nor data smoothing. Most worryingly, the results from unsmoothed data do *not* have the property that the results for the whole of the UK are a weighted-average of those for Wales and the area Outside-of-Wales – although this property is present when the data is smoothed. We suggest that there is a particular problem with small panel sample sizes for Wales, giving noisy estimates of hazard shares without smoothing.

How might this problem be addressed? Firstly, it is clear that we need bigger sample sizes to make progress with decomposing unemployment for Wales. It is clear both from visual inspections of the derived gross flows (figure 2.3) and from the tables (2.2 and 2.3) that the series are much more volatile than those for the whole of the UK and for Outside-of-Wales. There is an extent to which this is natural: we would expect the flows pertaining to a smaller geographical area to be smaller and hence naturally more volatile. However, our average gross flow figures are close to the ONS minimum of 17,000. It also telling that we cannot identify flows of any type *between* geographical regions from the panel data. The fact that the panel grosses its sample estimates to population level using sample weights makes the zero values for migration figures potentially problematic. If the sample were to pick up even one or two cross-border changes of labour market state per quarter, the sampling weights would gross these up to presumably much larger number. But with zero values, the sampling weights have nothing to multiply, so that the results are a likely source of downwards bias.

This suggests that the sample coverage is insufficient, and that future work should seek to use a larger sample to obtain more robust conclusions. An exercise which would also be interesting is to check the method using a larger geographical sub-region of the UK. Examining the gross flows within and between London and the rest of the UK, for example, would make for a more precise analysis if it were subject to careful robustness checks on the boundary between "London" and "Outside-of-London".

Another idea to improve the analysis is to change the construction of unemployment. There is a very sharp dichotomy in our analysis between steady-state unemployment, which we decompose without any error term, and observed unemployment, which we decompose with an error term. Rather than model steady-state unemployment as Petrongolo and Pissarides do, we could perhaps follow Shimer (2005(a)) in the simulation of unemployment data. That is, having estimated the hazard rates from the panel data, we would specify an initial value for unemployment at the beginning of the simulated series. We could then simulate unemployment *within* quarters based on the assumption of the constant hazard rates estimated for the quarter from the data (mapping time within each quarter to the [0,1] interval, $\tau \in [0,1]$ as described in Shimer (2005(a))). We could then read off the simulated value for unemployment at end of each quarter ($\tau = 1$), and treat this as the series to be decomposed. This is ought to result in a less volatile series for unemployment, without the need for extra smoothing. It might also have a smaller error term in the decomposition than the (non-smoothed) analyses presented here.

Finally, we note that there may be scope to use search and matching theory, of the type covered in chapters 3 and 4, to improve our knowledge of the properties of gross-flows decomposition. In this thesis, we report results for different sample periods, one of which includes a deep recession. Our results show notable differences in shares attributable to hazard rates, even in our estimates for Outside-of-Wales and for the UK with large sample sizes.

It would be interesting to explore the properties of the gross-flows variance decomposition procedure to changes in the way the economy is believed to work. One could conduct a variety of Monte-Carlo simulations using theoretical models (specified with different sorts of shocks of various magnitudes), different assumptions about search and matching, and the potential for structural breaks. One could then conduct variance-decompositions on the simulated data. This could be one way to get a more comprehensive understanding of the effects of smoothing, and the error term properties for the unemployment decompositions. The analysis would presumably require a finer differentiation between different types of flows (for example – separations would need to be decomposed into separations and layoffs, in the manner of Elsby, Michaels and Solon (2007)). In light of this analysis, one could then adapt the variance decomposition according to evidence on the circumstances of the economy for the sample period.

This brings us back to our indirect – inference assessment of a part of search and matching theory. Our contribution has been to test a version of the Mortensen-Pissarides model, under a set of parameters suggested by Shimer, using indirect inference rather than calibration and simple moment comparison. Shimer showed that the data fails to match essential moments of the model, due to the sensitivity of wages with respect to changes in productivity, a result that arises from the Nash bargaining assumption over wages.

Our results in Chapter 3 affirm the finding that under Shimer's original parameter values, the model fits poorly. (In indirect inference terms, we reject the hypothesis that the data were generated by the model). The auxiliary model coefficients from the data – be they standard deviations of endogenous variables or VAR(1) coefficients - fall far outside of 95% of model outcomes. We then estimate the model using indirect inference estimation, and find that the model is not rejected by the data (really that the data is not rejected by the model) when the auxiliary model used to fit the data consists of standard deviations of the endogenous variables (the vacancy-unemployment ratio, wages and unemployment). However, the estimated primitive parameters that generate the “closest” standard deviations are corner solutions – they are at the boundary of the restricted parameter space, which we impose both to economize of parameter search and for theoretical reasons. When we try to fit the data based on VAR coefficients, we cannot get the model to fit the data at all.

Our results broadly support the idea that the Mortensen Pissarides model generates moments that are inconsistent with the data. Although the results under indirect inference for volatilities of endogenous variables look different, the differences can be at least partially resolved by decomposing shocks to the exogenous variables (which our model has in common with the traditional Shimer puzzle approach) and the endogenous variables (which are a result of our bootstrapping procedure and are not present in Shimer’s analysis). However, one can criticise our work here along certain dimensions, which could be the subject of further work.

Firstly, we have linearized the standard model, so that we lose second or higher order effects in the model. One particular implication of this is that we lose part of the original Nash bargaining wage equation, for example non-labour incomes or the benefit replacement ratio drops out of the model under linearization. We also lose all of the non-linearity in the search cost specification term in the job-creation condition. Proponents of the original search and matching approach would be right to question the robustness of our work to this assumption. Further work could see if the same conclusions can be reached if we keep the non-linearities of the original model.

A second potential criticism relates to the treatment of time, and is similar to our comment on our gross flows decomposition procedure. In Chapter 3 we use a steady-state approximation to the unemployment rate. A better approach would be to use a difference equation in the manner of Chapter 4. Better still however would be to follow Shimer (2005(a)) and model the continuously evolving intra-period unemployment rate in the manner described above, mapping time within each quarter to the [0,1] interval, $\tau \in [0,1]$ and recording the quarterly unemployment rate when $\tau = 1$. Doing this would re-emphasize the role of hazard rates as in the original formulation, and would arguably be a more faithful representation to the original model.

In Chapter 4 we reproduce Yashiv's version of the Mortensen-Pissarides model, which employs cubic search costs to attempt to resolve Shimer's volatility puzzle. The Yashiv model has some extra features compared to the model in Chapter 3, such as rational rather than static expectations and an unemployment difference equation that moves dynamically (rather than assuming steady state unemployment in every period). The search cost function incorporates new hires as well as vacancies into the search cost function. Yashiv's putative mechanism to induce more volatility in the labour market variables is to increase the persistence of vacancies, reducing the speed at which they respond to productivity shocks. This has a positive effect on the volatility of vacancies, which feeds through into other labour market variables.

Yashiv's paper shows that the model is much more successful at fitting US time series data moment-for-moment than the original, when a moment comparison-based calibration methodology is used. We recreate Yashiv's model as closely as possible, subject to constraints on our ability to recreate Yashiv's data set. We show that our version of the model behaves in a similar way to that of Yashiv's with respect to variations in the search cost function, subject to a few exceptions. Finally we subject each version of the model (with the standard linear search cost and Yashiv's cubic search cost specification, respectively) to indirect inference testing. We find using an auxiliary model comprised of standard deviations of endogenous variables, that the standard deviations fit better under cubic search costs, although both versions of the model – linear and cubic search costs - are rejected by the data.

Our main interpretation of this result is –as in Chapter 3 - that shocks from the unexplained variation in the model are important. Our method of indirect inference includes this unknown variation in the way that the traditional calibration method (simulating the shocks to the chosen exogenous variables in the model only) does not, since it involves drawing errors from the distance between data and structural model fitted values, and simulating their serially independent components. The main point is that including this source of variation results in very different model behaviour compared to the standard formulation in which one simulates shocks to the exogenous variables only. Because the structural errors – the distance between the data and the fitted values – are so large, the effect is to drive our results further away from those of Yashiv. Indirect inference also constitutes a stringent test of the "closeness" of the chosen auxiliary model parameters to those of the data, because it is based on the *joint* distribution of model generated auxiliary parameters. Moment-by-moment comparisons are not, strictly speaking, a statistical test at all, but they are an analogue of a marginal distribution. The failure of the model to fit using an auxiliary model based on the standard deviations of the three endogenous variables is perhaps surprising. But as the power of indirect inference testing increases with the number of auxiliary model parameters, it is not surprising that the model does not fit when we use a VAR as an auxiliary model.

There are in principle several infidelities to Yashiv's model in our analysis. In the final analysis we use an HP-filter to de-trend our variables, rather than log differencing the data from steady-state values as Yashiv does, as we find that the latter fails to make most of the series stationary. There are also differences between our original data and Yashiv's - for example our construction of the discount factor faced by the firm, which constitutes one of the model's exogenous variables, has a different construction to the original, and also has very different statistical properties. There is in truth no way to be sure that this does not affect the result. Other unavoidable data discrepancies could also make a difference, however they are less severe and their effect is presumably smaller.

A broader question to consider is whether Yashiv's original formulation is a good one, aside from the specification of search costs. For example, as described in Chapter 1, recent work suggests that separations are an important part of the evolution of cyclical unemployment. It may therefore be better to try to build upon the original Mortensen Pissarides model, which had an endogenous separation rate. Progress could perhaps be made by focussing on the distinction between fires and quits, a strategy followed in the more recent empirical work by Elsby, Michaels and Solon (2007). As with Chapter 3, it may also be worthwhile to consider keeping the model non-linear, rather than linearizing it, for the sake of the second order (and higher) moments.

Finally we note that in Chapter 4 we have focussed on only one proposed mechanism for resolving the Shimer puzzle. Many others have been suggested, notably the re-specification of shocks (Barnichon) and alterations to the assumptions of Nash bargaining (for example Hall (2005a), Shimer (2004)). The incorporation of job-to-job flows could also improve the results, and would avoid the problem of using the vacancy data encountered in Yashiv's paper and in our own work. There is no reason why these models cannot be explored using the indirect inference tools set out in this thesis.

Bibliography:

ABRAHAM, K. G. & KATZ, L. F. 1986. Cyclical Unemployment: Sectoral Shifts or Aggregate Disturbances? *Journal of Political Economy,* 94.

ANDERSON, P. M. & MEYER, B. D. 1997. Unemployment Insurance Takeup Rates and the After-Tax Value of Benefits. *The Quarterly Journal of Economics,* 112**,** 913-937.

ANDOLFATTO, D. 1996. Business Cycles and Labor-Market Search. *American Economic Review,* 86**,** 112-132.

BELL, B. & SMITH, J. 2002. On gross worker flows in the United Kingdom: evidence from the Labour Force Survey. Bank of England.

BLANCHARD, O. J. & DIAMOND, P. 1989. The Beveridge Curve. *Brookings Papers on Economic Activity,* 20.

BURDETT, K. & CUNNINGHAM, E. J. 1998. Towards a theory of vacancies. *Journal of Labor Economics,* 16**,** 445-478.

COGLEY, T. 2006. Data Filters, *New York University* https://files.nyu.edu/tc60/public/datafilters.pdf

COLE, H. L. & ROGERSON, R. 1999. Can the Mortensen-Pissarides Matching Model Match the Business-Cycle Facts? *International Economic Review,* 40**,** 933-959.

COLES, M. G. & SMITH, E. 1996. Cross-Section Estimation of the Matching Function: Evidence from England and Wales. *Economica,* 63**,** 589-597.

COSTAIN, J. S. & REITER, M. 2008. Business Cycles, Unemployment Insurance and the Calibration of Matching Models. *Journal of Economic Dynamics and Control,* 32**,** 1120-1155.

DAVIS, S. J. & HALTIWANGER, J. 1992. Gross Job Creation, Gross Job Destruction and Employment Reallocation. *The Quarterly Journal of Economics,* 107**,** 819-863.

DEN HAAN, W. J. & KALTENBRUNNER, G. 2009. Anticipated Growth and Business Cycles in Matching Models. Journal of Monetary Economics, 56, 309-327.

DIAMOND, P. A. 1982. Aggregate Demand Management in Search Equilibrium. *Journal of Political Economy,* 90**,** 881-894.

ELSBY, M. W., MICHAELS, R. L. & SOLON, G. 2007. The Ins and Outs of Cyclical Unemployment. NBER Working Papers.

ELSBY, M. W. L., SMITH, J. C. & WADSWORTH, J. 2011. The role of worker flows in the dynamics and distribution of UK unemployment. *Oxford Review of Economic Policy,* 27**,** 338-363.

FALLICK, B. & A., F. C. 2004. Employer-to-Employer Flows in the U.S. Labour Market: The Complete Picture of Gross Worker Flows. Federal Reserve: Finance and Economics Discussion Series.

FUJITA, S. & RAMEY, G. 2007. Job matching and propagation. *Journal of Economic Dynamics and Control,* 31**,** 3671-3698.

FUJITA, S. & RAMEY, G. 2009. The Cyclicality of Separation and Job Finding Rates. *International Economic Review,* 50**,** 415-430.

GOMES, P. 2009. Labour market flows: facts from the United Kingdom. Bank of England.

GOTTSCHALK, P. & MOFFITT, R. A. 2000. Job Instability and Insecurity for Males and Females in the 1980s and 1990s. *In:* N. (ed.) *On the Job: Is Long-term Employment a Thing of the Past?* New York: Russell Sage Foundation.

GOURIEROUX, C. S. & MONFORT, A. 1993. Indirect Inference. *Journal of Applied Econometrics,* 8**,** S85-S118.

HAGERDORN, M. & MANOVSKII, I. 2008. The Cyclical Behaviour of Unemployment and Vacancies Revisited. *American Economic Review,* 98**,** 1692.

HALL, R. E. 2005a. Employment Fluctuations with Equilibrium Wage Stickiness. *American Economic Review,* 95**,** 50-65.

HALL, R. E. 2005b. Job Loss, Job Finding, and Unemployment in the U.S. Economy over the Past Fifty Years. *NBER Macroeconomics Annual.* MIT Press.

HALL, R. E. & MILGROM, P. R. 2008. The Limited Influence of Unemployment on the Wage Bargain. *American Economic Review,* 98**,** 1653-1674.

HAMERMESH, D. S. 1993. *Labor demand*, Princeton University Press.

HANSEN, G. D. 1985. Indivisible labor and the business cycle. *Journal of Monetary Economics,* 16**,** 309-327.

JAEGER, D. A. & STEVENS, A. H. 2000. Is Job Stability in the United States Falling? Reconciling Trends in the Current Population Survey and the Panel Study of Income Dynamics. *In:* N. (ed.) *On the Job: Is Long-Term Employment a Thing of the Past?* New York: Russell Sage Foundation.

KRAUSE, M. U. & LUBIK, T. 2007. The (Ir)relevance of Real Wage Rigidity in the New Keynesian Model with Search Frictions. *Journal of Monetary Economics,* 54**,** 706-727.

KYDLAND, F. E. & PRESCOTT, E. C. 1982. Time to Build and Aggregate Fluctuations. *Econometrica,* 50**,** 1345-1370.

LE, V. P. M. & MEENAGH, D. 2013. Testing and Estimating Models using Indirect Inference. Cardiff Business School, Economics Section.

LE, V. P. M., MEENAGH, D., MINFORD, P. & WICKENS, M. 2012. Testing DSGE models by indirect inference and other methods: some Monte Carlo experiments. Cardiff Business School, Economics Section.

LE, V. P. M., MINFORD, P. & WICKENS, M. 2010. The 'Puzzles' methodology: Enroute to Indirect Inference? *Economic Modelling,* 27**,** 1417–1428.

LILIEN, D. M. 1982. Sectoral Shifts and Cyclical Unemployment. *The Journal of Political Economy,* 90**,** 777-793.

MERZ, M. 1995. Search in the labor market and the real business cycle. *Journal of Monetary Economics,* 36**,** 269-300.

MORTENSEN, D., T. & NAGYPÁL, É. 2007. More on Unemployment and Vacancy Fluctuations. *Review of Economic Dynamics,* 10**,** 327-347.

MORTENSEN, D. T. & PISSARIDES, C. A. 1994. Job Creation and Job Destruction in the Theory of Unemployment. *The Review of Economic Studies,* 61**,** 397-415.

ONS 2012. Labour Force Survey User Guide: LFS Two-Quarter and Five-Quarter Longitudinal Datasets - 2011. Office for National Statistics.

PETRONGOLO, B. & PISSARIDES, C., A. 2001. Looking into the Black Box: A Survey of the Matching Function. *Journal of Economic Literature,* 39**,** 390-431.

PETRONGOLO, B. & PISSARIDES, C. A. 2008. The Ins and Outs of European Unemployment. *American Economic Review,* 98**,** 256-262.

PISSARIDES, C. A. 1986. Unemployment and vacancies in Britain. *Economic Policy,* 1**,** 499-559.

PISSARIDES, C. A. 2000. *Equilibrium Unemployment Theory*, MIT Press.

PISSARIDES, C. A. 2009. The Unemployment Volatility Puzzle: Is Wage Stickiness the Answer? *Econometrica,* 77**,** 1339-1369.

SHIMER, R. 2004. Consequences of Rigid Wages in Search Models. *Journal of the European Economic Association,* 2**,** 469-479.

SHIMER, R. 2005a. Reassessing the Ins and Outs of Unemployment. University of Chicago.

SHIMER, R. 2005b. The Cyclical Behaviour of Equilibrium Unemployment and Vacancies. *American Economic Review,* 95**,** 25-49.

SILVA, J. I. & TOLEDO, M. 2009. Labor Turnover Costs and the Cyclical Behaviour of Vacancies and Unemployment. *Macroeconomic Dynamics,* 13**,** 76-96.

SILVA, J. I. & TOLEDO, M. 2013. The Unemployment Volatility Puzzle: The Role of Matching Costs Revisited. *Economic Inquiry,* 51**,** 836-843.

SMITH, J. C. 2011. The Ins and Outs of UK Unemployment. *The Economic Journal,* 121**,** 402-444.

YASHIV, E. 2000a. Hiring as Investment Behavior. *Review of Economic Dynamics,* 3**,** 486-522.

YASHIV, E. 2000b. The Determinants of Equilibrium Unemployment. *American Economic Review,* 90**,** 1297-1322.

YASHIV, E. 2006. Evaluating the Performance of the Search and Matching Model. *European Economic Review,* 50**,** 909-936.